Doctoral theses at NTNU, 2016:253

Leif Ove Hansen

Geometric Aspects of Quantum Entanglement

ISBN 978-82-326-1844-6 (printed ver.)
ISBN 978-82-326-1845-3 (electronic ver.)
ISSN 1503-8181

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of
Science and Technology
Thesis for the Degree of
Philosophiae Doctor
Faculty of Natural Sciences and Technology
Department of Physics

Doctoral theses at NTNU, 2016:253

NTNU

Leif Ove Hansen

Geometric Aspects of Quantum Entanglement

Thesis for the Degree of Philosophiae Doctor

Trondheim, September 2016

Norwegian University of Science and Technology
Faculty of Natural Sciences and Technology
Department of Physics

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

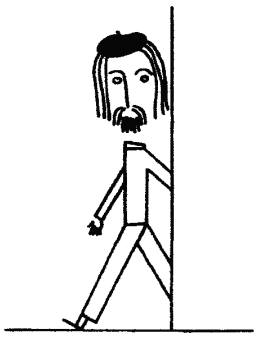
Thesis for the Degree of Philosophiae Doctor

Faculty of Natural Sciences and Technology
Department of Physics

ISBN 978-82-326-1844-6 (printed ver.)
ISBN 978-82-326-1845-3 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2016:253

Printed by NTNU Grafisk senter



Contents

Acknowledgements

The present work is a dissertation (or thesis) for the degree Doctor of Philosophy (Ph.D) in theoretical physics at the Norwegian University of Science and Technology (NTNU).

I am eternally grateful to my *doktorvater* professor Jan Myrheim for taking me as his Ph.D-student, it has been a privilege to work together with such an insightful theoretical physicist. His mild and generous nature has been instrumental towards the completion of this work.

I would also like to thank Per Øyvind Sollid and Jon Magne Leinaas from the University of Oslo for a very rewarding collaboration, and for the hospitality shown during visits to Oslo.

On the same note I also want to give tribute to Andreas Hauge, who passed away in 2015. He made essential contributions to the numerical study and theoretical understanding of the extremal entanglement witnesses. He also wrote the first draft of Paper II, which is a natural continuation of his master thesis.

I am also indebted to Børge Irgens for his contribution with regards to programming issues, especially in the use and development of the simulated annealing routines, but also for suggesting the notion of genuine SL-symmetry.

Due credit is given to the Faculty of Natural Sciences and Technology (NT-Faculty) at NTNU for financing this project. I am also indebted to the many kind and helpful people in the administrative staff at NTNU for their valuable assistance, cooperation and patience.

I would also like to thank my family for their lasting acknowledgement and support, and especially dedicate this thesis to my father, who passed away during the course of this work.

Thesis

Introduction

The work in this thesis is an attempt to understand entanglement and entangled states in quantum physics, with an emphasis on the underlying geometric structures. Entanglement is a strange feature contained within the quantum mechanical framework, which was first observed by Erwin Schrödinger and Albert Einstein, and eventually led to a seminal paper by Einstein, Podolsky and Rosen, published in Physical Review in 1935. The next important contributions to the theory of quantum entanglement was given by Bell in 1960s, and an experimental demonstration of entanglement, *i.e.* a realization of the *gedanken experiment* proposed by EPR in 1935, was finally performed by Aspect *et al.* in 1982. A considerable growth of the interest in quantum entanglement has occurred since the early 1990s, and much of the reason for this is that the tools of quantum optics and modern instrumentation have made it possible to measure and manipulate entangled quantum systems in a way which was not possible before.

Motivation

An inherent attribute of entanglement is non-locality, and in the continuation of this, several non-reductionist features, which are conceptually challenging and philosophically interesting. There are however several other motivating factors for the study of entanglement. At the very early stage, entanglement was considered to be merely a puzzling feature of quantum mechanics, mostly confined to the philosophical realm. Today, entanglement is considered a resource in quantum information theory, and it has found several applications within this field. These include quantum computers and several areas in quantum communications. A problem with entanglement is that it is a fragile resource, in the sense that it has a tendency to decay. In any entangled system which interacts with an environment, the entanglement generally will diminish as a result of this interaction. In order to deal with this loss, methods such as entanglement distillation have been invented in order to distill less entangled states into more entangled ones. There exists two qualitatively different types of entanglement, called free and bound entanglement. The type referred to as bound entanglement has been shown to be impossible to distill in a complete fashion, and they are thus much less useful in applications.

The separability problem

A fundamental problem related to entanglement, is the *separability problem*. This is the problem of determining whether a given quantum state is separable or entangled,

as it cannot be none or both. For pure quantum states (states of rank one) the problem is completely solved by *Schmidt decomposition*, but for mixed quantum states this is not the case, and the problem is complicated even for the low dimensional cases. Both from a practical point of view and as part of the theoretical framework of quantum mechanics, it is an important problem to address.

In 1996 the PPT criterion for determining whether or not a state is entangled, was presented by Peres. This criterion is based on the partial transposition map. It appeared to clearly distinguish separable and entangled states, in the sense that it formulated an easily testable condition which separable states satisfies, but which entangled states do not. The year after however, Pawel Horodecki provided an example of an entangled state that also satisfies the condition. The PPT criterion essentially separates all quantum states into PPT and non-PPT (NPT) states, and all quantum states belong to exactly one of the two groups, with the separable states as PPT states for all dimensions $N_A \times N_B$. An important fact is that testing whether a state is PPT or NPT is easily achieved using a computer.

It is known that for the 2×2 and 2×3 systems, the PPT states are identical to the separable states, and that the NPT states are identical to the entangled states, so the Peres criterion is always conclusive. But for larger dimensions, such as for instance the 3×3 system, there always exist PPT states that are entangled, so for this case the PPT criterion can only be used in a contrapositive way, *i.e.* to prove that a state is *not* separable. Since entangled PPT states in this sense seem to have great significance, their characteristics are worth studying. A natural choice is the 3×3 system, which is the next logical step from the 2×2 and 2×3 systems.

As mentioned, the PPT criterion introduced by Peres relied on the use of the transposition map on one of the two subsystems of a composite quantum system, *i.e.* partial transposition. Transposition is an example of a *positive map*. A positive map essentially maps all quantum states into quantum states, while a non-positive map will not do so, *i.e.* it will map at least one quantum state into a non-quantum state. It is possible to use positive maps to formulate a much more general criterion for the separability problem. The PPT criterion then becomes a special case of a larger and more general formulation of the separability problem. This was done by Michal, Pawel and Ryszard Horodecki in 1996, and is naturally referred to as the Horodecki criterion. There is a one-to-one correspondence known as the Choi-Jamiolkowski isomorphism between positive maps and a subset of Hermitian operators called *entanglement witnesses*. The Horodecki criterion states that a quantum state is separable if and only if the expectation value of *all* entanglement witnesses in the state are non-negative. If the expectation value of one (or possibly several) entanglement witnesses in a given quantum state is negative, then the state is always entangled. One might say that these entanglement witnesses reveal or are "witnesses" to the entanglement of this state. This criterion is both necessary and sufficient. The problem is that in order to implement this criterion we potentially must check for all entanglement witnesses. In order to achieve this, it is required that we must have knowledge and understanding of the structure of the set of entanglement witnesses, and since this set is convex, the extremal entanglement witnesses are particularly important.

Geometry and experimental mathematics

The approach in this thesis is through geometry. The geometrical aspects of entanglement are interesting in themselves, and might lead to insights which is impossible or very difficult to arrive at by other methods. Instead of developing explicit examples, it is then possible to obtain a larger picture by studying the entire set of generic objects, and also to gain an understanding of the various subsets with specific properties. Of fundamental importance to our studies is the convex nature of the sets in question. We study these objects as matrices in convex spaces, and thus knowledge and understanding of the extremal points in these sets is very relevant. Also, since convex sets have well defined faces, and since all extremal points of the total convex set always lie on one of these faces, the study of the faces of these sets is essential.

Undertaking completely analytic studies of the geometry of entanglement is in general very hard, especially as the dimension of the systems grow. It is therefore very useful to develop and use numerical tools to aid in these studies. It becomes possible to gain new information, and to test the feasibility of conjectures and ideas that arise. In this sense we might call this activity *experimental mathematics*. It is however important that analytical developments run alongside the numerical investigations, for example by providing analytical starting points for further investigations.

Thesis structure

This thesis is organized into two parts. The first part consists of an introduction into, and a summary of essential features and results from the four papers, which are themselves presented in part two. Many of the details from the papers are left out, but some examples and illustrations which are not contained in the papers are found here. There is also a summary of some of the most important numerical methods used during the work presented in the papers, a feature which is not included in the papers themselves. The first part is organized as follows:

- In Chapter 1 we take an historical view of some important aspects of the development of quantum mechanics, and the role played by entanglement. It showcases some reasons why entanglement is interesting. The importance of the work of John Bell is highlighted, and the experimental contribution of Aspect *et al.* is described.

- Chapter 2 gives a short description of entanglement as a resource in several areas of quantum information theory. We define the nuts and bolts of LOCC and quantum operations, before illustrating quantum dense coding, quantum cryptography, quantum teleportation and quantum computing.

- In Chapter 3 we develop the mechanics of mixed composite quantum systems. We look at partial trace and decoherence, and use these considerations to exhibit some non-reductionist aspects of entanglement. We define partial transposition and the important sets of density matrices $\mathcal{D}$, separable states $\mathcal{S}$ and the PPT states $\mathcal{P}$. We also look at product vectors in different subspaces of $\mathbb{C}^N$, and use

this to state the range criterion, which is a criterion that any separable state must satisfy. Finally we discuss the important concept of product transformations, and symmetries under these.

- Chapter 4 contains a review of convex sets, and some of their important properties. The use of barycentric coordinates, Carathéodorys theorem and dual cones are some of the essential features. Some of the important concepts regarding faces on convex sets are discussed, and a special type of convex sets called simplexes are defined.

- Chapter 5 then takes us into the geometry of quantum entanglement. We look at the geometrical nature of the sets $\mathcal{D}$, $\mathcal{S}$ and $\mathcal{P}$, and use some of these features to make two-dimensional plots through these multi-dimensional structures. We then define some of the tools we need to investigate the structure of the set of PPT states $\mathcal{P}$, and use these to describe how we can perform perturbations in the set of PPT states. We end with a description of the faces of the three convex sets $\mathcal{D}$, $\mathcal{S}$ and $\mathcal{P}$.

- In Chapter 6 we discuss the structure of the set of PPT states $\mathcal{P}$ for dimensions 3×3. We look at the generic states of rank four and five, and describe three methods for tracing out surfaces of these states. While the extremal entangled PPT states of rank four show a relatively simple structure, this is not the case for such states of rank five. For the rank five PPT states there appears to exist a large number of non-generic types. A generic PPT state of rank five is extremal and has no product vectors in the kernel, but it is possible to construct non-generic PPT states of rank five with a non-zero number of product vectors in the kernel. This is the main theme of Paper 4, but here we discuss these matters quite briefly, and highlight the perhaps most interesting case, where the kernel contains four product vectors.

- Chapter 7 contains a description of the strategy pertaining to the use of positive maps and entanglement witnesses to solve the separability problem. We first define entanglement witnesses and then describe the one-to-one correspondence with positive maps. Different types of positive maps are then defined, such as completely positive maps and decomposable maps. Of profound importance to our research are the zeros of entanglement witness, and the constraints related to these zeros, which define two very important types of entanglement witnesses: quadratic and quartic witnesses. We use these zeros and the facial structure of the set of entanglement witnesses $\mathcal{S}^\circ$ to understand extremality on this set, and then to construct a method for producing such extremal witnesses. This method is formalised as an algorithm in Section 8.6. Some results from the investigations on entanglement witnesses are included. Since the set of separable states $\mathcal{S}$ and the set of entanglement witnesses $\mathcal{S}^\circ$ are dual to each other, it is possible to use the faces of $\mathcal{S}^\circ$ to say something about the faces of $\mathcal{S}$. We finally end with a brief review of the special set of entanglement witnesses called optimal witnesses,

including a discussion of the so-called SPA separability conjecture. Finally, we develop a method to transform an entanglement witness to a form which makes the corresponding positive map both unital and trace-preserving, and details regarding the numerical procedures related to this is presented in Section 8.7.

- In Chapter 8 we summarize the essential details of the most important numerical methods we have used in the four papers. A description of these methods is not included in the papers themselves, with the exception of the algorithms for finding extremal witnesses (Paper 2), and for transforming positive maps to unital and trace-preserving form (Paper 3).

- Chapter 9 contains short summaries of the four publications, with some added remarks on the outlook for further research.

Throughout the thesis attempts have been made to keep the notation as clear and pedagogical as possible. Even though some of the notation may appear superfluous to hardened and experienced quantum physicists, underlining and sometimes exaggerating points or notation is useful to readers without detailed knowledge of this field. Many of the theorems presented are given without formal proofs, though in some cases a line of reasoning is given. The papers (particularly paper II) contain more formal proofs.

Chapter 1

Historical background

An outline of the early developments of quantum mechanics and some of the epistemologically problematic sides of the theory is given. We then discuss the EPR-argument, and the insight given by John Bell, which effectively closed the possibility of including quantum mechanics in a local realistic hidden variable theory. The important Aspect experiment is then presented, and an important set of quantum states called Bell states are defined.

1.1 The birth of quantum mechanics

Quantum mechanics as a physical theory dates back to the beginning of the 20th century. Several theories existed that could describe the fundamental realms of nature known at the time. These included Isaac Newton's mechanics and theory of gravity, the laws of thermodynamics, and James Clerk Maxwell's elegant and successful summary of electromagnetism and optics. During the next 20 years the whole foundation of physics and science was to be revolutionized.

At the start of the 20th century there were several big puzzles in physics. It was widely believed, especially among the conservative part of the physics community, that these puzzles could be understood and described within the realm of the established models. Some of these puzzles were related to fundamental questions regarding the nature of light, which eventually gave rise to the theory of special relativity, and further down the line a completely new understanding of gravity. Two of the problems that gave way to quantum theory was the observed spectral lines of atoms and the so-called ultraviolet catastrophe. Spectral lines are among the many characteristics of atoms that are impossible to understand using the classical theory of electromagnetism, because they imply that the energy states of the atom is a discrete set rather than the continuous spectrum which is to be expected from classical electromagnetism. The structure of the atom was discovered by Ernest Rutherford in the early 20th century [1], and led Niels Bohr to develop the so-called Bohr model of the atom. Also, if the classical theories of thermodynamics and electromagnetism are used to calculate the amount of radiation emitted by a black body of temperature T, we get the Rayleigh-Jeans law [2]. The spectral radiance $B_\nu(T)$ summed over all frequencies ν should of course be finite, and

if we include the frequencies up to visible radiation the Rayleigh-Jeans law works quite well, but if we include larger frequencies it is a catastrophe, and in fact it diverges to infinity if we sum over all frequencies.

In 1900 Max Planck submitted two seminal papers [3, 4]. By *quantizing* oscillators interacting with the radiation field, he calculated the spectral radiance $B_\nu(T)$ and ended up with a new expression, today known as the Planck distribution. Planck himself considered the quantization of the interaction with the radiation field to be nothing more than a mathematical trick, that bore no specific physical significance. However, in the following years many of the phenomena that were completely unexplainable, could be understood by using the idea of quantization, including the divergence in the Rayleigh-Jeans law. The photoelectric effect, which gave Albert Einstein the Nobel prize, was another triumph for the idea of radiation quanta, *i.e.* photons [5]. In 1925, Werner Heisenberg and others invented matrix mechanics [6], and building on Louis de Broglie's work [7], Erwin Schrödinger shortly after invented wave mechanics [8]. Though the two were seemingly different descriptions, it was shown by Schrödinger that the two approaches are equivalent [9]. The formalized framework established during this period still defines the platform for how we use quantum mechanics today, but the number of applications has grown enormously.

While there are no doubts about the successes of quantum mechanics in giving predictions that are in amazing agreement with observations and experiments, it is clear that the theory contains many problematic aspects. Many of these problems have a philosophical origin, and challenge our intuitively based understanding of the nature of reality.

1.1.1 Determinism and hidden variables

The trade-off that quantum mechanics makes is that it describes outcomes of measurements by probabilities. This lack of certainty about an outcome is well known in many situations, *e.g.* the tossing of a coin. The fundamental difference between coins and electrons is that a coin is made up of a very large amount of subsystems for which it is impossible (or at least extremely difficult) to have complete knowledge, and this enforces a statistical description. We assume this is not the case for an electron, and claim that the quantum mechanical description contained in the state $|\psi\rangle$, is the most complete description we can have. Since the state vector $|\psi\rangle$ only contains information that gives us probabilities, we end up with a *non-deterministic* description. Claims that there is a more complete description containing parameters $\lambda_1, \ldots, \lambda_n$ that are hidden, so as to make a deterministic theory, are called hidden variable theories (HV). The existence of such hidden variables would effectively mean that quantum mechanics is an incomplete theory.

1.1.2 Measurement and realism

In most interpretations of quantum mechanics we are forced to reinterpret the meaning of the term measurement. If we perform a measurement on a ball to find its colour, and

find that the ball is red, then intuition tells us that the ball was red before we looked at it, and that the act of measurement simply revealed an already existing characteristic of the ball. Say we instead want to measure the spin of an electron, and we arrange it so that we have the two possible outcomes -1 or $+1$, corresponding to states $|\psi_-\rangle$ or $|\psi_+\rangle$. A *non-realistic* interpretation of quantum mechanics would then be that the state of the electron need not be $|\psi_-\rangle$ nor $|\psi_+\rangle$, but a superposition. In such a model the very act of measurement puts the electron in a state that corresponds to the result. This does not mean that the electron did not have a state before the measurement, it was simply in a state that did not correspond to a possible outcome of the specific measurement, and this is a highly non-classical feature.

1.1.3 Non-locality

Another feature of profound importance in quantum mechanics is *entanglement*. It was originally Schrödinger and Einstein who first recognized what appeared to be a "spooky" action at a distance present in the quantum mechanical description of certain systems. In the paper where Schrödinger introduces his famous cat [10], he also uses the word "Verschränkung" to describe a rather peculiar non-local feature, which he later translates to entanglement. Schrödinger observed the existence of global states $|\psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B$ of composite systems $\mathcal{H}_A$ and $\mathcal{H}_B$, which cannot be written as a product of states of the individual subsystems. The term entanglement underlines an intrinsic order of what appears to be statistical relations between subsystems of the compound quantum system. This introduces an inherent non-locality into the system. Earlier the same year Einstein, working at Princeton, submitted a paper with the two colleagues Boris Podolsky and Nathan Rosen, usually abbreviated the EPR-paper [11]. During a large part of the 1920s Einstein had intense debates with Bohr about the foundations of quantum mechanics. A debate where Einstein could not find a clinching argument to prove that quantum mechanics must be incomplete. After his emigration to the USA he kept thinking about these questions, and finally ended up with what he thought would settle the debate in his favour.

1.2 The EPR-argument

The EPR-argument is simply based on a thought experiment suggested by Einstein, Podolsky and Rosen to prove by pure logic, not by actually performing the experiment, that quantum mechanics is not a complete theory of nature. It is referred to as a "gedanken" experiment because the means to actually conduct such experiments were not available at the time. Today these experiments can be conducted by utilizing the many tools that quantum optics provides. The heart of the argument proposed by Einstein, Podolsky and Rosen was that entanglement is an effect that *must* exist for quantum mechanics to be a complete theory of nature, and since the physical implications of such a "spooky" action at a distance are preposterous, entanglement as a physical phenomenon cannot exist.

Before progressing further we must make the following clarifications

Locality
Results of measurements obtained at one location are independent of, *i.e.* not correlated to, any actions performed at the system at space-like separation.

Realism
Measurement results are determined by properties the system carries prior to, and independent of measurements.

This sort of realism is often referred to as *counterfactual definiteness*. It is worth noting that the demand for locality does not mean that results of measurements cannot be correlated. Since we know today that entanglement exists, the argument provided by Einstein, Podolsky and Rosen is often, and wrongfully, labelled as flawed. Their logic and argumentation still stand, it is the presumption of local realism which is flawed. The essential argument can be formulated as: if nature adheres to the principles of locality and realism, then quantum mechanics is not a complete theory. Most physicists at the time that were seeking to invalidate the argument never questioned locality as the wrongful assumption, but rather the assumption of realism. Realism had been questioned and dismissed by Bohr, Heisenberg and other supporters of the Copenhagen interpretation. While locality can be quantified through correlation functions from measurements, the case for realism is far more subtle and difficult, even though we are dealing with a relatively precise definition here. Clever attempts have been made to rule out realism as well [12]. The essential setup of the EPR experiment is shown below.

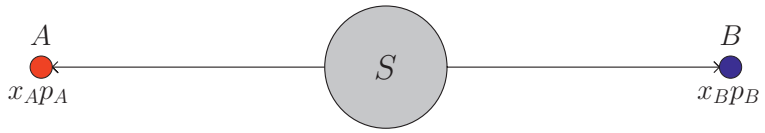


Figure 1.1: The basic setup of the EPR experiment.

EPR used measurement of the positions x_A and x_B along with momenta p_A and p_B of a pair of identical particles created from a source S. For the sake of the EPR argument this is sufficient, but if we want to develop these things further we must introduce spins (or polarizations). An important fact here is that the quantities x and p are incompatible observables according to quantum mechanics or $[x_A, p_A] \neq 0$, which means that a particle cannot have definite values of the same components of position and momentum at the same time.

It is clear from conservation that a measurement of x_A will also give the value for x_B, and likewise a measurement of p_A will give the value for p_B. Since EPR assume locality, a measurement on particle A cannot in any way influence the state of particle B. Since we can choose between measuring x_A or p_A and thus make this quantity an element of reality, and since this choice cannot influence the state of particle B in any way, we must assume that the quantities x_B and p_B were clearly defined, or

elements of reality, all along. This is in contradiction to quantum mechanics which states that particle B cannot attain sharply defined values of both these quantities at the same time. So, EPR claimed, this means that there exist elements of reality in this system which quantum mechanics cannot fully describe, ergo, it is an incomplete theory. So it follows with elegant inevitability that we cannot have nature to be local and realistic, and make quantum mechanics a complete description of it. Einsteins idea was that a fuller and better theory would be possible, and that this theory should *include* quantum mechanics but contain additional and hidden variables that make the theory deterministic, local and realistic. As we shall see this idea was demolished by Bell a little less than ten years after Einstein's death.

1.3 Bell inequalities

During the 1950s several attempts were made to construct hidden variable theories, following the proposals that Einstein had made. The currently best known hidden variable theory is due to the physicist and philosopher David Bohm. Originally published in 1952 [13], it is a non-local hidden variable theory. Bohm unknowingly rediscovered and extended the idea that de Broglie had proposed and abandoned in 1927 [14]. Hence this theory is commonly called the de Broglie-Bohm theory. Bohm postulated that any quantum particle, *e.g.* an electron, has a hidden "guiding wave" that governs its motion. Thus, in this theory electrons are quite clearly particles, so that when a double slit experiment is performed, the trajectory goes through one slit rather than the other. The trajectory is governed by the hidden guiding wave, and since this guiding wave is non-local it can produce the wave pattern which is observed.

In 1964 Bell proposed studying correlations between measurements of the spin components of two spin-½ particles, where the measurement axes for the two particles were independently oriented [15].

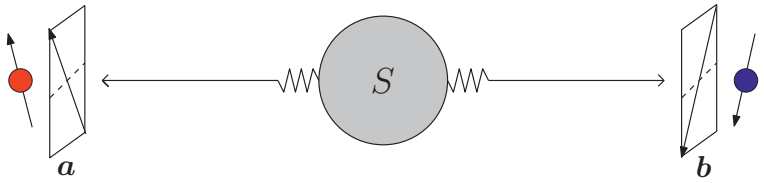


Figure 1.2: Illustration of the setup in a Bell experiment. Anti-correlated pairs of spin-½ particles pass through measurement devices in A and B. In A the spin along an axis represented by $\boldsymbol{a}$ is measured, and likewise $\boldsymbol{b}$ along B.

Suppose we have a source that produces pairs of anti-correlated spin-½ particles that shoot off in opposite directions before they pass through detectors. Assume also that the detector for particle A is oriented along an axis defined by the unit vector $\boldsymbol{a}$, and likewise $\boldsymbol{b}$ for particle B. We also introduce the locality principle when we assume that the detectors are sufficiently separated so that any action (or measurement) in A, cannot influence the outcome of an action (or measurement) in B.

We first treat this system using standard quantum mechanics. The wave function can be expressed as

$$|\psi\rangle = |\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\{|01\rangle - |10\rangle\} \tag{1.1}$$

with the states $|01\rangle = \{|0\rangle_A \otimes |1\rangle_B\} \in \mathcal{H}_{AB}$ and $|10\rangle = \{|1\rangle_A \otimes |0\rangle_B\} \in \mathcal{H}_{AB}$ as product states made up of basis states from each subspace $\mathcal{H}_A$ and $\mathcal{H}_B$. The component of spin along $\boldsymbol{a}$ is $S_A = \boldsymbol{S}_A\boldsymbol{a}$, and using $\boldsymbol{S} = (\hbar/2)\boldsymbol{\sigma}$ we find that a measurement of S_A gives $\pm\hbar/2$, with probabilities that depend on the orientation of $\boldsymbol{a}$. Since we are interested in correlations we must find the expectation value for the joint spin S_AS_B. Using (1.1) we get

$$\begin{aligned} E(\boldsymbol{a},\boldsymbol{b}) &= \langle\psi|S_aS_b|\psi\rangle \\ &= -\boldsymbol{a}\boldsymbol{b} = -\cos\,\phi \end{aligned} \tag{1.2}$$

where $\phi = \measuredangle(\boldsymbol{a},\boldsymbol{b})$. When $\phi = 0$ you get $E(\boldsymbol{a},\boldsymbol{b}) = -1$ which indicates complete anti-correlation. This type of correlation is easily reproduced classically, and an example is given by Bell with reference to the socks of monsieur Bertlmann [16]. The function E here, is nothing more mysterious than the quantum mechanical expression for the average of joint measurements of the two spins S_A and S_B.

We now turn to the local hidden variable (LHV) description given by Bell in [15]. We introduce a set of hidden variables $\lambda_1, \ldots, \lambda_n$ that completely describe the system in a deterministic way. For our purposes these hidden variables can without loss of generality be gathered into one parameter λ, and we make no more assumptions about this hidden variable λ other than that it satisfies a probability distribution $p(\lambda)$

$$\int p(\lambda)d\lambda = 1 \qquad\qquad p(\lambda) \geq 0 \tag{1.3}$$

Measurements of the spin components are now defined by

$$F_A(\boldsymbol{a},\lambda) = \pm 1 \qquad\qquad F_B(\boldsymbol{b},\lambda) \;\; = \pm 1 \tag{1.4}$$

Locality is ensured by the fact that F_A only depends on the setting of the measurement apparatus $\boldsymbol{a}$ in A and λ, and likewise for F_B on $\boldsymbol{b}$ and λ. Spin conservation, or total anti-correlation for $\phi = 0$ is ensured by

$$F_A(\boldsymbol{a},\lambda) = -F_B(\boldsymbol{a},\lambda) \tag{1.5}$$

The average of joint measurements of the spins is then $\xi(\boldsymbol{a},\boldsymbol{b})\,\hbar^2/4$ where the $\xi(\boldsymbol{a},\boldsymbol{b})$ is factorized with respect to the A and B subsystems, and then weighted by the probability distribution expressed by the hidden variable λ

$$\xi(\boldsymbol{a},\boldsymbol{b}) = \int p(\lambda)F_A(\boldsymbol{a},\lambda)F_B(\boldsymbol{b},\lambda)\,d\lambda \tag{1.6}$$

where as assumed, $E(\boldsymbol{a},\boldsymbol{a}) = \xi(\boldsymbol{a},\boldsymbol{a}) = -1$.

Now we derive so-called Bell inequalities from this, which are bounds on absolute values of the function $\xi(\boldsymbol{a},\boldsymbol{b})$. We start with

$$\begin{aligned}\xi(\boldsymbol{a},\boldsymbol{b}) - \xi(\boldsymbol{a},\boldsymbol{c}) &= \int p(\lambda)\big\{F_A(\boldsymbol{a},\lambda)F_B(\boldsymbol{b},\lambda) - F_A(\boldsymbol{a},\lambda)F_B(\boldsymbol{c},\lambda)\big\}\,d\lambda \\ &= -\int p(\lambda)\big\{F_A(\boldsymbol{a},\lambda)F_A(\boldsymbol{b},\lambda) + F_A(\boldsymbol{b},\lambda)F_B(\boldsymbol{c},\lambda)\big\}\,d\lambda\end{aligned} \tag{1.7}$$

where $\boldsymbol{c}$ is another angle in subsystem B. Since the functions F_A and F_B only take values ± 1 we get

$$\begin{aligned}|\xi(\boldsymbol{a},\boldsymbol{b}) - \xi(\boldsymbol{a},\boldsymbol{c})| &< \int p(\lambda)\big\{1 + F_A(\boldsymbol{b},\lambda)F_B(\boldsymbol{c},\lambda)\big\}\,d\lambda \\ \Rightarrow\ |\xi(\boldsymbol{a},\boldsymbol{b}) - \xi(\boldsymbol{a},\boldsymbol{c})| &< 1 + \xi(\boldsymbol{b},\boldsymbol{c})\end{aligned} \tag{1.8}$$

It should be emphasized that these Bell inequalities say nothing at all about quantum mechanics, they are derived from very general demands that any LHV theory must satisfy. If however quantum mechanics is to be included in such a LHV theory, then the function E in (1.2) must also satisfy (1.8). We can choose the directions $\boldsymbol{a}$, $\boldsymbol{b}$ and $\boldsymbol{c}$ to satisfy

$$\measuredangle(\boldsymbol{a},\boldsymbol{b}) = \pi/3 \qquad \measuredangle(\boldsymbol{b},\boldsymbol{c}) = \pi/3 \qquad \measuredangle(\boldsymbol{a},\boldsymbol{c}) = 2\pi/3 \tag{1.9}$$

Using (1.2) we then get

$$\begin{aligned}|E(\boldsymbol{a},\boldsymbol{b}) - E(\boldsymbol{a},\boldsymbol{c})| &= 1 \\ 1 + E(\boldsymbol{b},\boldsymbol{c}) &= 1/2\end{aligned} \tag{1.10}$$

so for the angles (1.9) clearly

$$|E(\boldsymbol{a},\boldsymbol{b}) - E(\boldsymbol{a},\boldsymbol{c})| > 1 + E(\boldsymbol{b},\boldsymbol{c}) \tag{1.11}$$

which is in clear contradiction to (1.8). This shows that Einstein's idea of incorporating quantum mechanics in a larger theory, which should be deterministic and obey local reality, is not possible. Hidden variable theories, *i.e.* deterministic descriptions, that includes quantum mechanics is still possible, but the hidden variables $\lambda_1, \ldots, \lambda_n$ will necessarily have to contain non-local features.

1.4 The Aspect experiments

The gradual transition from theoretical considerations to the laboratory began in the 1960s, but the actual realization of the EPR experiment as a convincing test of the Bell inequalities was not done until 1982, when a set of experiments was conducted by Aspect *et al.* [17], usually known as the Aspect experiments. They used entangled

photon pairs from a two-photon decay in calcium atoms. The schematic is shown in Figure 1.3.

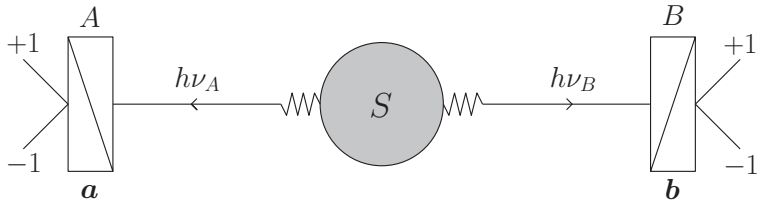


Figure 1.3: The setup for the Aspect experiments using entangled photon pairs which pass through polarization filters. The polarization of the photons are then measured along directions defined by $\boldsymbol{a}$ and $\boldsymbol{b}$ respectively.

Since these are entangled *photons* we use the Bell state

$$|\psi\rangle_{AB} = \frac{1}{\sqrt{2}}\{|01\rangle + |10\rangle\} \tag{1.12}$$

and we define the quantum mechanical correlation function $E(\boldsymbol{a}, \boldsymbol{b})$ by

$$E(\boldsymbol{a}, \boldsymbol{b}) = \frac{N_1(\boldsymbol{a}, \boldsymbol{b})}{N_2(\boldsymbol{a}, \boldsymbol{b})} \tag{1.13}$$

with the two coincidence rates

$$\begin{aligned} N_1(\boldsymbol{a}, \boldsymbol{b}) &= N_{++}(\boldsymbol{a}, \boldsymbol{b}) + N_{--}(\boldsymbol{a}, \boldsymbol{b}) - N_{-+}(\boldsymbol{a}, \boldsymbol{b}) - N_{+-}(\boldsymbol{a}, \boldsymbol{b}) \\ N_2(\boldsymbol{a}, \boldsymbol{b}) &= N_{++}(\boldsymbol{a}, \boldsymbol{b}) + N_{--}(\boldsymbol{a}, \boldsymbol{b}) + N_{-+}(\boldsymbol{a}, \boldsymbol{b}) + N_{+-}(\boldsymbol{a}, \boldsymbol{b}) \end{aligned} \tag{1.14}$$

where for example N_{++} represents, for fixed values of $\boldsymbol{a}$ and $\boldsymbol{b}$, the number of incidents where both measurements gave $+1$. If we calculate this using (1.12), we get in much the same way as for the spin-½ case used by Bell that

$$E(\boldsymbol{a}, \boldsymbol{b}) = \cos 2\phi \tag{1.15}$$

with $\phi = \measuredangle(\boldsymbol{a}, \boldsymbol{b})$. The correlation function for the LHV theory $\xi(\boldsymbol{a}, \boldsymbol{b})$ is found in the same way as for the fermion case

$$\xi(\boldsymbol{a}, \boldsymbol{b}) = \int p(\lambda) F_A(\boldsymbol{a}, \lambda) F_B(\boldsymbol{b}, \lambda)\, d\lambda \tag{1.16}$$

Several Bell inequalities can be calculated from this. Aspect *et al.* chose one based on

$$S = \xi(\boldsymbol{a}, \boldsymbol{b}) - \xi(\boldsymbol{a}, \boldsymbol{b'}) + \xi(\boldsymbol{a'}, \boldsymbol{b}) + \xi(\boldsymbol{a'}, \boldsymbol{b'}) \tag{1.17}$$

where $\boldsymbol{a}$ and $\boldsymbol{a'}$ are two possible orientations of the same polarization analyzer in A, and likewise $\boldsymbol{b}$ and $\boldsymbol{b'}$ in B. Now, provided that $F_A(\boldsymbol{a}, \lambda)$ and $F_B(\boldsymbol{b}, \lambda)$ both take values ± 1 we observe that

$$-2 \leq S(\phi) \leq 2 \tag{1.18}$$

The directions that were used in one of the Aspect experiments, and that goes into (1.17) are defined by

$$\begin{aligned} \boldsymbol{ab} = \boldsymbol{a'b} = \boldsymbol{a'b'} = \cos\phi \\ \boldsymbol{ab'} = \cos 3\phi \end{aligned} \tag{1.19}$$

The correlations for a large set of ϕ in the whole range from $\phi = 0°$ to $\phi = 180°$ were then measured. An illustrative depiction of this data is shown in Figure 1.4. The real set of data presented in [17] contained uncertainty bars which are not shown here. The blue points are the observed data points and the red curve is the function $S(\phi)$ with the quantum mechanical correlation function $E(\boldsymbol{a}, \boldsymbol{b}) = \cos 2\phi$ substituted for the LHV correlation function $\xi(\boldsymbol{a}, \boldsymbol{b})$ in (1.17). So, the correlations predicted by the LHV theory are limited to $-2 \leq S(\phi) \leq 2$, while the correlations predicted by quantum mechanics are given by the red curve. The Bell inequalities (1.18) are clearly not satisfied, while the red curve and the observed data points are in very high agreement.

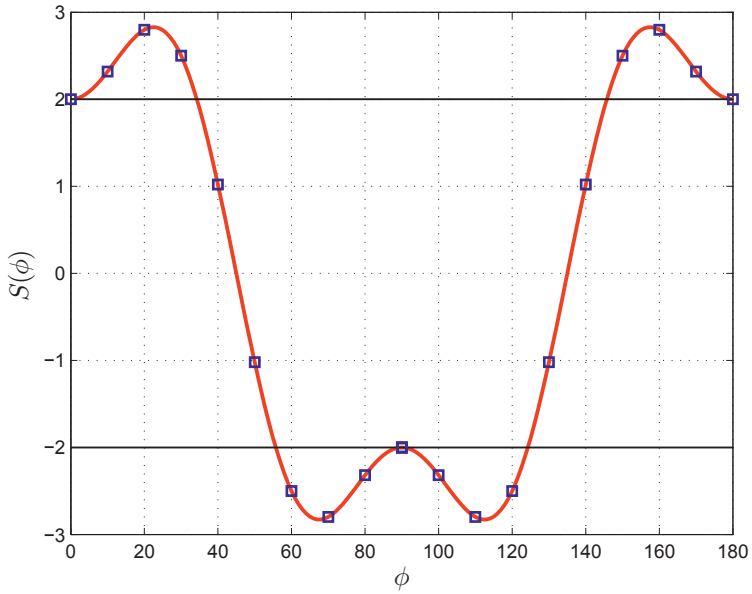


Figure 1.4: Illustrative depiction of data from one of the Aspect experiments. The blue points are observed data points, while the red curve is the quantum mechanical prediction of the correlations for different ϕ. The Bell inequality $-2 \leq S(\phi) \leq 2$ predicted by LHV theories is clearly broken.

It is also easy to see that for some values of ϕ the Bell inequalities *are* satisfied. This means that for these values the measurements give correlations that can be described by LHV models. This includes the special case for $\phi = 0°$ where the filters are completely

aligned, and these correlations can be reproduced by observing the colors of the socks of monsieur Bertlmann.

Following the Aspect experiments, many similar experiments producing non-local correlations have been made by several groups [18, 19], and all these experiments have confirmed the predictions made by quantum mechanics. There is also some evidence that entanglement between photons can be maintained over long distances, at least distances well exceeding 100 km [20]. The major problem using fiber optics at such distances is that transferring single photons with sufficient reliability becomes difficult, and therefore an effective strategy is to transfer these photons in straight laser beams through the atmosphere.

1.5 Bell states and entanglement

If we have a composite system of two subspaces $\mathcal{H}_A$ and $\mathcal{H}_B$ of dimension N_A and N_B respectively, this is in quantum mechanics described by a tensor product of the individual Hilbert spaces

$$\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B \tag{1.20}$$

The tensor product of Hilbert spaces is a way to extend the tensor product construction from vectors to spaces, so that the result of taking a tensor product of two Hilbert spaces of dimension N_A and N_B is another Hilbert space of dimension $N = N_A N_B$. Since a Hilbert space is an inner product space, the product space must also have a well defined inner product.

Set a basis for $\mathcal{H}_A$ as $|0\rangle_A, \ldots, |N_A - 1\rangle_A$ and for $\mathcal{H}_B$ as $|0\rangle_B, \ldots, |N_B - 1\rangle_B$. Then a basis for $\mathcal{H}_{AB}$ is $|i\rangle_A \otimes |j\rangle_B$ with $0 \leq i \leq N_A - 1$ and $0 \leq j \leq N_B - 1$. So a general state $|\psi\rangle_{AB} \in \mathcal{H}_{AB}$ can be expressed as

$$|\psi\rangle_{AB} = \sum_{ij} c_{ij} |ij\rangle \tag{1.21}$$

where $|ij\rangle$ is short notation for $|i\rangle_A \otimes |j\rangle_B$, and of course $c_{ij} \in \mathbb{C}$. A state $|\psi\rangle_{AB}$ that can be written as

$$|\psi\rangle_{AB} = |\phi\rangle_A \otimes |\chi\rangle_B \tag{1.22}$$

where $|\phi\rangle_A = \sum_i a_i |i\rangle_A \in \mathcal{H}_A$ and $|\chi\rangle_B = \sum_j b_j |j\rangle_B \in \mathcal{H}_B$ is called a product state, or *separable* state. The interesting thing, which was originally pointed out by Schrödinger [10], is that there are states in $\mathcal{H}_{AB}$ that cannot be expressed in the form (1.22). These are entangled states. It is then clear that a quantum state is either entangled or separable, it cannot be both or neither.

The states (1.1) and (1.12) are called Bell states. They are examples of entangled states for the case of lowest possible dimensions $N_A = N_B = 2$, which is often written as the 2×2 system. There are two more Bell states in addition to the two we have used so far. We define the four Bell states in the following way

Definition 1.1 (Bell states)**.**

$$\begin{aligned} |\alpha\pm\rangle &= \frac{1}{\sqrt{2}}\{|01\rangle \pm |10\rangle\} \\ |\beta\pm\rangle &= \frac{1}{\sqrt{2}}\{|00\rangle \pm |11\rangle\} \end{aligned} \tag{1.23}$$

The product states $|00\rangle$, $|01\rangle$, $|10\rangle$ and $|11\rangle$ form a basis for $\mathcal{H}_{AB}$, and so do the four non-product Bell states $|\alpha\pm\rangle$ and $|\beta\pm\rangle$. The product states $|ij\rangle$ are often called the computational basis. To demonstrate how the product structure arises for states in the 2×2 system we can observe that a general product state in this system can be written as

$$\begin{aligned} |\psi\rangle &= \{a_0|0\rangle_A + a_1|1\rangle_A\} \otimes \{b_0|0\rangle_B + b_1|1\rangle_B\} \\ &= a_0b_0\,|00\rangle + a_0b_1\,|01\rangle + a_1b_0\,|10\rangle + a_1b_1\,|11\rangle \end{aligned} \tag{1.24}$$

We see that the Bell states (1.23) cannot be written on this form for any choice of a_i, b_i, so they are entangled. Clearly, putting for example $a_1 = 0$ produces the separable state

$$|\psi\rangle = \frac{1}{\sqrt{2}}\{|00\rangle + |01\rangle\} \tag{1.25}$$

Due to the clear non-factorisable mathematical structure that is present in some quantum states, and taking into consideration the non-local correlations which are clearly seen in some measurements, there is a wide consensus that some systems in nature do not adhere to the locality principle. The orthodox Copenhagen interpretation also claims that nature violates realism. Hidden variable theories are by their very nature deterministic, but the hidden variables must be able reproduce the correlations that violate Bell inequalities, *i.e.* they must operate in a non-local manner.

Any quantum product state, *e.g.* a state of the form (1.22), can be described by a LHV theory. So a system described by such a state will only contain correlations of a local nature, which means by definition that they will satisfy *all* Bell inequalities. As long as we restrict ourselves to pure quantum states, by which we mean states that are described by a single Hilbert state vector, all entangled states will describe systems that contain non-local correlations, which means that there will be correlations which break Bell inequalities. Mixed quantum states, which we shall introduce in detail later, are states that are statistical ensembles of pure states. These states describe a more general situation than the pure quantum states, and are therefore very useful. For mixed states the picture is more complex. All separable mixed states are still local states, in that they can be described by LHV models, and cannot produce correlations that break Bell inequalities. But examples of entangled mixed states that do *not* violate Bell inequalities have been given by Werner [21], these states admit LHV models and do not contain non-local correlations. It appears that the link between entanglement and non-local correlations is not straightforward, at least not for mixed states.

Chapter 2

Entanglement as a resource

During the 1990s it had become clear that entanglement was not only the subject of philosophical debates, but could be considered a new resource for tasks that cannot be performed by classical means. The development of tools to exert finely tuned nanoscale control on quantum systems have been developed, such as for instance to identify, image and manipulate individual atoms. We here first discuss some basic definitions related to local operations and classical communications, including some important examples. We then describe three important applications in the area of quantum communication, and end with some very brief notes on quantum computers.

2.1 LOCC

Entanglement theory today owes its form in great measure to the discovery and development of entanglement manipulation. A natural class of operations suitable for manipulating entanglement is that of local operations and classical communications or LOCC [22].

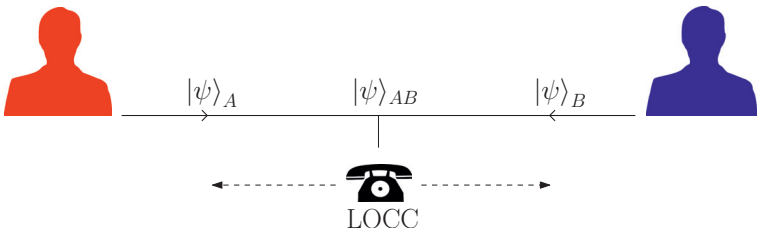


Figure 2.1: Local quantum operations are performed on a subsystem, and information regarding these operations are sent to other parts, where this information is used to decide further operations, often according to a predetermined protocol.

The paradigm of LOCC means that local operations are performed on parts of the system, and the results of these operations are communicated classically to another part, where other local operations, based on the classical information received, are performed. The information passed may be regarding what type of operations were performed, or in the case that the operations were measurements, what results were

obtained. Usually, complicated and lengthy operations that involve several actions and reactions on each subsystem are done according to a predetermined protocol.

A central understanding called the monotonicity condition, is that the total amount of entanglement in a state shared by system A and B cannot be increased by performing LOCC operations. If for instance the state $|\psi\rangle_{AB}$ in Figure 2.1 is made from two locally prepared states $|\phi\rangle = |\psi\rangle_A$ and $|\chi\rangle = |\psi\rangle_B$, it will always be separable. We have claimed that a quantum state is either entangled or separable, and although this is true there exist ways to quantify entanglement with the aid of so-called *entanglement measures*. There are many types of such measures with different strengths and weaknesses, but basically an entanglement measure could be described as a function $f : |\psi\rangle \mapsto [0, 1]$. Clearly all separable states would then have $f(|\psi\rangle) = f(|\phi\rangle \otimes |\chi\rangle) = 0$, while the Bell states (1.23) would have $f(|\alpha\pm\rangle) = f(|\beta\pm\rangle) = 1$, since they are known to be maximally entangled.

2.2 Quantum operations

We will now define some quantum operations for the 2×2 system. The first two examples are both local quantum operations, in the sense that they act only on qubits locally. But first we define the term *qubit*.

Definition 2.1 (Qubit)**.** *In a two-dimensional Hilbert space we define the quantum state*

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle \tag{2.1}$$

with $a_0, a_1 \in \mathbb{C}$, as a qubit or quantum bit.

We can use

$$|0\rangle = \begin{pmatrix}1\\0\end{pmatrix} \quad |1\rangle = \begin{pmatrix}0\\1\end{pmatrix} \quad \Rightarrow \quad |\psi\rangle = \begin{pmatrix}a_0\\a_1\end{pmatrix} \tag{2.2}$$

as a matrix representation for the qubit.

The operations X, iY and Z

We first define the operation NOT or $\mathbf{X}$. This operation works like

$$\mathbf{X}\begin{pmatrix}a_0\\a_1\end{pmatrix} = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}\begin{pmatrix}a_0\\a_1\end{pmatrix} = \begin{pmatrix}a_1\\a_0\end{pmatrix} \tag{2.3}$$

so that $\mathbf{X}|\psi\rangle = a_1|0\rangle + a_0|1\rangle$ and the $\mathbf{X}$ operation simply switches the coefficients for the two states $|0\rangle$ and $|1\rangle$, which also means for example that $\mathbf{X}|0\rangle = |1\rangle$. Similarly the $\mathbf{Z}$ operation is $\mathbf{Z}|\psi\rangle = a_0|0\rangle - a_1|1\rangle$. We can also use the operation $\mathrm{i}\mathbf{Y}$ defined by $\mathrm{i}\mathbf{Y}|\psi\rangle = a_1|0\rangle - a_0|1\rangle$. The observant reader will realize that the effects of the operations $\mathbf{X}, \mathrm{i}\mathbf{Y}$ and $\mathbf{Z}$ are defined by the Pauli matrices σ_x, σ_y and σ_z.

Hadamard

We now also define the Hadamard operation **H**

$$\mathbf{H}\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \frac{a_0}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} + \frac{a_1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \tag{2.4}$$

which effectively represents a split of the basis vectors

$$\begin{aligned} \mathbf{H}|0\rangle &= \frac{1}{\sqrt{2}}\{|0\rangle + |1\rangle\} = |+\rangle \\ \mathbf{H}|1\rangle &= \frac{1}{\sqrt{2}}\{|0\rangle - |1\rangle\} = |-\rangle \end{aligned} \tag{2.5}$$

Since each of these are performed on a single qubit, which could represent one half of an entangled state, they are local operations. If we use the **X** operation on the first qubit belonging to subsystem A of the Bell state $|\beta+\rangle$ we get

$$\mathbf{X}\,|\beta+\rangle = \mathbf{X}\left(\frac{1}{\sqrt{2}}\{|00\rangle + |11\rangle\}\right) = \frac{1}{\sqrt{2}}\{|10\rangle + |01\rangle\} = |\alpha+\rangle \tag{2.6}$$

which is another Bell state. So by performing a quantum operation on subsystem A we have changed our part of the entangled state and thus transformed the global state $|\beta+\rangle$ into another state $|\alpha+\rangle$ that still carries maximum entanglement.

CNOT

We now move on to another very important tool in the arsenal of quantum operations, namely the CNOT operation, which stands for controlled NOT operation. This is a non-local operation, which means that it defines a protocol that operates on both qubits in a state $|\psi\rangle_{AB}$. The CNOT operation works on product states in such a way that if the first qubit is $|0\rangle$, then it does nothing to the second qubit, but if the first qubit is $|1\rangle$ then it changes the other qubit from $|0\rangle \mapsto |1\rangle$ or alternatively from $|1\rangle \mapsto |0\rangle$. We realize the non-local character of this operation since we need to be able to read the first qubit as a control qubit, and in addition we need to be able to change the second qubit in accordance with this reading. In total for the CNOT operation

$$|00\rangle \mapsto |00\rangle \qquad |01\rangle \mapsto |01\rangle \qquad |10\rangle \mapsto |11\rangle \qquad |11\rangle \mapsto |10\rangle \tag{2.7}$$

If we step back again to the matrix representations (2.2) we can use the tensor product to calculate

$$|00\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \qquad |01\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \qquad |10\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} \qquad |11\rangle = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \tag{2.8}$$

and a matrix representation for the CNOT operation is then

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \tag{2.9}$$

which is a unitary matrix that cannot be written as a tensor product of local unitary operations. This is proof that the CNOT operation is a non-local operation. We can in fact see that the combined operations Hadamard and CNOT will transform a pure product state into a Bell state. First, Hadamard on system A gives

$$\mathbf{H}|00\rangle = \left(\frac{1}{\sqrt{2}}\{|0\rangle + |1\rangle\}\right) \otimes |0\rangle = \frac{1}{\sqrt{2}}\{|00\rangle + |10\rangle\} \tag{2.10}$$

which is still a separable state since Hadamard is a local operation. Then CNOT gives

$$\mathbf{C}\left(\frac{1}{\sqrt{2}}\{|00\rangle + |10\rangle\}\right) = \frac{1}{\sqrt{2}}\{|00\rangle + |11\rangle\} = |\beta+\rangle \tag{2.11}$$

It is worth observing that $[\mathbf{H}, \mathbf{C}] \neq 0$ since the reversed operation is $\mathbf{HC}|00\rangle = \mathbf{H}|00\rangle$, which is separable.

The quantum operations above can be realized physically in many different ways. Since the technology for manipulating photons is very well developed, most of the realizations of entanglement utilize photonic systems [23, 24].

2.3 Applications of entanglement

We will here describe some applications of entanglement. First, in order to illustrate some of the power of entanglement in communication tasks we will look at some examples of this. Then we try to give a (very) basic account of the ideas of quantum computing and quantum computers. The three most prominent applications in the area of quantum communication are quantum dense coding, quantum cryptography (or quantum key distribution) and quantum teleportation. All these examples involve the transmission of information in some quantum form. Quantum computing involves, as we shall see, the use and manipulation of quantum bits as defined in (2.1) as the tool for making computations, rather than the traditional bit which is either $|0\rangle$ *or* $|1\rangle$.

Quantum dense coding

In classical communication, if we want to send two bits of information, we can accomplish this by sending two physical bits to the receiver. By using entanglement we can send the same amount of information by transmitting only one quantum bit of information between the parties. This is known as quantum dense coding or superdense coding [25], and has had several experimental realizations [26]. The two parties Alice and Bob

must each control one component, or qubit, of a shared entangled state, say the Bell state

$$|\beta+\rangle = \frac{1}{\sqrt{2}}\{|00\rangle + |11\rangle\} \tag{2.12}$$

Alice is in possession of the qubit at A and Bob of the qubit at B. Now Alice want to communicate two pieces of information to Bob. She relays this information through her choice of quantum operation $\mathbf{U}$. She chooses one out of four alternative local operations on her qubit in A. Firstly she can do nothing, which means that the entangled state is unchanged $\mathbf{U}|\beta+\rangle = |\beta+\rangle$, or she can use the operations $\mathbf{X}$, i$\mathbf{Y}$ or $\mathbf{Z}$ that we defined earlier. In these cases we get $\mathbf{X}|\beta+\rangle = |\alpha+\rangle$, i$\mathbf{Y}|\beta+\rangle = |\alpha-\rangle$ and $\mathbf{Z}|\beta+\rangle = |\beta-\rangle$.

So the entangled state she shares with Bob will change according to what type of operation she chooses to do. If she now sends her qubit to Bob, he will be in possession of both qubits, and can perform measurements on both. If he does the measurement in the Bell state basis he faces four possible outcomes, corresponding to $|\alpha\pm\rangle$ and $|\beta\pm\rangle$. In this way he can determine which of the four operations Alice did on her qubit. The one information bit Alice sends is either 0 or 1, but it turns into two information bits which Bob can determine as either 00, 01, 10 or 11.

Quantum cryptography

Another use of entanglement is quantum cryptography, a form of communication where information is encoded into a stream of qubits, is such a way that it is extremely difficult, if not impossible, to eavesdrop. Entanglement contains a feature which is very useful in this matter, namely the fact that if two subsystems A and B are in a pure entangled state, then no other outside system can be made to share these correlations. The fact that entanglement represents correlations that cannot be shared by third parties is connected to *entanglement monogamy*, which expresses unshareability of entanglement, as developed by Terhal [27].

There are several quantum key distribution (QKD) protocols that do not use entanglement directly, such as the BB84 protocol [28]. A measurement on a quantum state will usually irrevocably disturb the state in a non-deterministic and irreversible manner. The BB84 protocol utilizes this, and the *no-clone theorem*. The short version of the no-clone theorem is that it is impossible to make a copy of an unknown quantum state $|\psi\rangle$, while keeping the original unchanged. Cloning an unknown state would make it possible to exactly measure all the properties of the unknown state simultaneously, including non-commuting ones, which quantum mechanics expressly forbids.

Closely related to the no-clone theorem is the fact that it is impossible to distinguish two non-orthogonal states in quantum mechanics. For example if we take the four states

$$\begin{aligned} |\psi_{00}\rangle &= |0\rangle & |\psi_{10}\rangle &= |1\rangle \\ |\psi_{01}\rangle &= \frac{1}{\sqrt{2}}\{|0\rangle + |1\rangle\} = |+\rangle & |\psi_{11}\rangle &= \frac{1}{\sqrt{2}}\{|0\rangle - |1\rangle\} = |-\rangle \end{aligned} \tag{2.13}$$

then it is impossible to distinguish between them with certainty by measurements, and the BB84 protocol uses this fact.

1. If Alice wants to send a private key to Bob using the BB84 protocol she begins with two strings of bits, a and b, each n bits long. She then encodes these two strings as a string of n qubits

$$|\psi\rangle = |\psi_{a_1 b_1}\rangle \otimes \ldots \otimes |\psi_{a_n b_n}\rangle = \bigotimes_{i=1}^{n} |\psi_{a_i b_i}\rangle \tag{2.14}$$

   While the bit a_i can be regarded as the information she wants to send, the bit b_i contains information about which basis a_i is encoded in, either $b_i = 0$ for $|0\rangle, |1\rangle$ or $b_i = 1$ for $|+\rangle$, $|-\rangle$. So, the strings a_i, b_i are effectively translated, or encoded, into a quantum state $|\psi\rangle$. Complete knowledge of the state $|\psi\rangle$ will then give the strings a_i, b_i. In fact, together $a_i b_i$ essentially give us an index into the four qubit states (2.13).

2. Alice sends $|\psi\rangle$ over a public quantum channel to Bob. In the middle, Eve is eavesdropping into the flow of qubits. After Bob receives the string of qubits, all three parties, namely Alice, Bob and Eve, have their own states.

3. However, since only Alice knows b_i, it is virtually impossible for either Bob or Eve to distinguish the states of the qubits. Also, after Bob has received the qubits, we know by the no-clone theorem that Eve cannot be in possession of a perfect copy of the qubits sent to Bob. If Eve makes measurements on qubit a_i she risks disturbing the qubit with a likelihood of 50%, since she do not know b_i. So in effect half of the qubits that Eve measures, and presumably sends on to Bob, will be wrong.

4. Bob proceeds to generate a string of random bits b' of the same length as b. Using b', he then measures the string he has received from Alice, with the results as a'. Bob then announces publicly that he has received Alice's transmission. Alice then knows she can safely announce b. Bob then communicates over a public channel with Alice which b_i and b'_i are not equal. Both Alice and Bob now discard the qubits a_i and a'_i where b_i and b'_i do not match.

5. From the remaining k bits where both Alice and Bob measured in the same basis, Alice randomly chooses $k/2$ bits and discloses her choices over the public channel. Both Alice and Bob announce these bits publicly and run a check to see if more than a certain number of them agree. If this check passes, Alice and Bob proceed to use privacy amplification and information reconciliation techniques to create some number of shared secret keys. Otherwise, they realize with a certain probability that they have been eavesdropped, so they cancel and start the process all over.

Today a large number of protocols to conduct QKD exist. A major problem in the construction of such protocols is the closing of loopholes that allow eavesdroppers to interfere. Hacking strategies target vulnerabilities in the operation of a QKD protocol

or deficiencies in the components of the physical devices used in construction of the QKD system, and the volume of possible attack strategies is massive.

The first public use of QKD happened in Vienna in 2004, when Bank Austria Creditanstalt on behalf of the city of Vienna, transferred €3000 to research groups at Universität Wien. Today there are several companies offering commercial QKD systems.

Quantum teleportation

We will now discuss the basic theory underlying quantum teleportation, which could be argued is the most amazing application of entanglement. We assume that Alice and Bob share a Bell state $|\beta+\rangle$, where as before Alice controls the A qubit and Bob controls the B qubit of the state

$$|\beta+\rangle = \frac{1}{\sqrt{2}}\{|00\rangle + |11\rangle\} \tag{2.15}$$

In addition to her half of the entangled state $|\beta+\rangle$, Alice also controls another qubit $|\psi\rangle = a_0|0\rangle + a_1|1\rangle$ which she hopes to send or "teleport" to Bob by using only classical communication, *e.g.* a telephone. If she in any way tries to measure the state $|\psi\rangle$ she will ruin or change it, so she wants to avoid that. We must remember that the complex factors a_0 and a_1 can take on any value.

Alice starts by letting her qubit $|\psi\rangle$ interact with her half of $|\beta+\rangle$ so that the entangled state that Alice and Bob shares becomes

$$|\gamma_0\rangle = \frac{1}{\sqrt{2}}\Big(a_0|0\rangle\{|00\rangle + |11\rangle\} + a_1|1\rangle\{|00\rangle + |11\rangle\}\Big) \tag{2.16}$$

where we use the convention that the first two qubits (from the left) belong to Alice and the third to Bob. Alice then sends her qubits through a CNOT gate (2.7), and we have

$$|\gamma_1\rangle = \frac{1}{\sqrt{2}}\Big(a_0|0\rangle\{|00\rangle + |11\rangle\} + a_1|1\rangle\{|10\rangle + |01\rangle\}\Big) \tag{2.17}$$

She then uses a Hadamard gate on the first qubit

$$|\gamma_2\rangle = \frac{1}{2}\Big(a_0\{|0\rangle + |1\rangle\}\{|00\rangle + |11\rangle\} + a_1\{|0\rangle - |1\rangle\}\{|10\rangle + |01\rangle\}\Big) \tag{2.18}$$

This state is now rewritten

$$\begin{aligned}|\gamma_2\rangle = \frac{1}{2}\Big(&|00\rangle\{a_0|0\rangle + a_1|1\rangle\} + |01\rangle\{a_0|1\rangle + a_1|0\rangle\} \\ &+ |10\rangle\{a_0|0\rangle - a_1|1\rangle\} + |11\rangle\{a_0|1\rangle - a_1|0\rangle\}\Big)\end{aligned} \tag{2.19}$$

Please note that the $|00\rangle$ in (2.18) and (2.19) is not the same. The expression above naturally breaks down into four terms. The first term for instance, has Alice's qubits as $|00\rangle$ and Bob's qubit as $a_0|0\rangle + a_1|1\rangle$ which is the original state $|\psi\rangle$.

Alice now performs a measurement on her two qubits in the state $|\gamma_2\rangle$ in the product basis $|ij\rangle$. If her result is for instance $|00\rangle$ she reports this to Bob. He then knows that his qubit is a perfect copy of the initial qubit $|\psi\rangle$, since $|\psi\rangle$ is perfectly correlated to $|00\rangle$. If Alice measures, say $|10\rangle$, then he knows his qubit is $a_0|0\rangle - a_1|1\rangle$. He can perform an appropriate local quantum operation $\mathbf{U}$ on his qubit to reproduce the state $|\psi\rangle$. So in total, for $|00\rangle$ he does nothing ($\mathbf{U} = \mathbf{I}$), for $|01\rangle$ he uses $\mathbf{U} = \mathbf{X}$, $|10\rangle$ means he uses $\mathbf{U} = \mathbf{Z}$ and finally if Alice reports $|11\rangle$ from her measurement, he uses the operation $\mathbf{U} = \mathbf{ZX}$.

What effectively happens when Alice performs her measurement is that all information about the state $|\psi\rangle$ is irreversibly destroyed in A but reappears in B. Hence the name quantum teleportation. The reason this works is due to the initial point where Alice interacts the qubit $|\psi\rangle$ with her qubit of the entangled state $|\beta+\rangle$, by doing this she "imprints" information about $|\psi\rangle$ into the entangled state $|\beta+\rangle$. It is worth noting that there is no instant transport of energy going on, which one could be led to believe from the term "teleportation", and also since Bob cannot know the state $|\psi\rangle$ before he receives classical information from Alice about the result of her measurement, there is no instant information transport.

The quantum copier is a bit like a copy machine that makes a perfect copy of an unopened letter, but changes or ruins the original in an unpredictable way.

Quantum computing

The idea behind quantum computers is the distinction between bits and qubits. A classical bit in a classical computer is somewhat like a light switch, it is either on or off. We have looked at qubits before, and we refresh from (2.1)

$$|\psi\rangle = a_0|0\rangle + a_1|1\rangle \tag{2.20}$$

A switch in the state $|\psi\rangle$ is neither in state $|0\rangle$ nor $|1\rangle$, but in a superposition with weights $|a_0|^2$ and $|a_1|^2$. Since a classical bit is either 0 or 1, a simulation of N two-level systems requires 2^N bits, while a quantum computer requires only N bits. For large N the difference can be quite substantial.

In general, a quantum computer that maintains a sequence of N qubits can be in an arbitrary superposition of up to 2^N different states simultaneously. This compares to a normal computer that can only be in *one* of these 2^N states at any one time. A quantum computer operates by setting the qubits in a controlled initial state that represents the problem at hand and by manipulating those qubits with a fixed sequence of quantum logic operations. The sequence of gates to be applied is called a quantum algorithm. The calculation ends with a measurement, collapsing the system of qubits into one of the 2^N pure states, where each qubit is purely zero or one. The outcome can therefore be at most N classical bits of information. Quantum algorithms are often non-deterministic, in that they provide the correct solution only with a certain known probability. However, by repeatedly initializing, running and measuring the quantum computer, the probability of getting the correct answer can (and must) be increased.

An example of a realized quantum computer is the NMR quantum computer, based on nuclear magnetic resonance [29]. The qubits in such a computer are the spin states of molecules. These can be manipulated by varying a magnetic field which interacts with the spins. Using a setup with seven qubits, such a system has been used to factor the number $15 = 3 \cdot 5$, an experimental demonstration of the principle of the quantum algorithm known as Shor's algorithm.

The big problem with quantum computers is the loss and degradation of entanglement. It is a result of the quantum state interacting with the environment, and is sometimes referred to as decoherence. In most applications in quantum computation, the states used are pure entangled states, but as they interact with the environment they will become more and more mixed. This puts limits not only on the operation of quantum computers, but also on most forms of quantum communications involving entanglement.

Chapter 3

Quantum mechanics of composite systems

We will first define the basic principles of quantum mechanics, and then move on the description of composite quantum systems, which involves the use of tensor products. For those who have suffered at least one course in quantum mechanics this should mostly be well known material. We also define the operations partial trace and partial transposition, and give some examples to underline their importance. The concept of product vectors in various subspaces is discussed, and also related concepts such as entangled subspaces and the range criterion. We end by showing the usefulness of product transformations in the description of these quantum systems.

3.1 Fundamental principles of quantum mechanics

We define the basic postulates of quantum mechanics, first for pure quantum states, and we then develop the formalism for mixed states and density matrices.

Pure quantum states

A pure state in quantum mechanics is usually represented by a vector in a complex inner product space, usually referred to as a Hilbert space $\mathcal{H}$. Since $\mathcal{H}$ is a vector space the superposition principle applies, which means that if $|\psi_1\rangle, |\psi_2\rangle \in \mathcal{H}$ then the sum $|\psi_1\rangle + |\psi_2\rangle \in \mathcal{H}$. A vector space must also be closed under multiplication, which means that also $c|\psi_1\rangle \in \mathcal{H}$ for $c \in \mathbb{C}$. The vectors $c|\psi_1\rangle$ and $|\psi_1\rangle$ represent the same physical state, so the space of physical states is a projective space of lines in $\mathcal{H}$. The only vector that cannot be projected onto a physical state is thus the zero vector.

Since the Hilbert space is an inner product space we also need an inner product. In the conventional Dirac notation $|\psi_1\rangle^\dagger = \langle\psi_1|$, and we put the inner product $\langle\psi_1, \psi_2\rangle$ of $|\psi_1\rangle$ and $|\psi_2\rangle$ as $\langle\psi_1|\psi_2\rangle = \langle\psi_2|\psi_1\rangle^*$. It is possible, and often practical to write the Hilbert space vector $|\psi\rangle$ simply as an $N \times 1$ matrix $\psi \in \mathbb{C}^N$. In this notation $\langle\psi| = \psi^\dagger$ and the inner product becomes $\langle\psi_1, \psi_2\rangle = \psi_1^\dagger\psi_2$, while the outer product $|\psi_1\rangle\langle\psi_2| = \psi_1\psi_2^\dagger$. Occasionally we use ψ_k to mean the kth basis vector in a basis $\psi_1, \ldots, \psi_n$, and occasionally we take ψ_k to be the kth component of a vector $\psi \in \mathbb{C}^N$,

i.e. a scalar. In short, context often dictates the meaning. In some cases, as illustrated in Chapters 1 and 2, we prefer to use the Dirac notation because it is more convenient and illustrates more clearly the main points.

To every observable there is associated an Hermitian operator $\mathbf{F}$, where Hermitian implies that $\mathbf{F}^\dagger = \mathbf{F}$. For the case of an N-dimensional Hilbert space, a representation of $\mathbf{F}$ could be an $N \times N$ complex matrix. The matrix $\mathbf{F}$ will have a set of eigenvectors e_n with corresponding real eigenvalues λ_n for $1 \leq n \leq N$

$$\mathbf{F}e_n = \lambda_n e_n \tag{3.1}$$

where the set λ_n may or may not be degenerate. For any matrix F that represents $\mathbf{F}$, the rank of F is equal to the number of eigenvalues $\lambda_m \neq 0$. For an eigenvector basis e_n we will find the elements of $F = F_{ij} = e_i^\dagger \mathbf{F} e_j$. If the system is in the state ψ, the results of a measurement of the observable $\mathbf{F}$ in the eigenvector basis e_n, will always be one of the eigenvalues λ_n with probability $e_n^\dagger \psi$, and a measurement will always leave the system in the state e_n corresponding to the result. The average value of measurements in the state ψ is given by the expectation value

$$\langle \mathbf{F} \rangle = \psi^\dagger \mathbf{F} \psi \tag{3.2}$$

Any matrix F for which $FF^\dagger = F^\dagger F$ is a *normal* matrix. It is easy to see that any Hermitian matrix is normal. A well known and important result for normal matrices is the spectral theorem, which states that any normal matrix F can be written in terms of its eigenvalues and eigenvectors

$$F = \sum_n \lambda_n \left(e_n e_n^\dagger \right) \tag{3.3}$$

where the eigenvectors form a complete set $\sum_n e_n e_n^\dagger = I$.

Mixed quantum states

For a system represented by a pure vector state $\psi \in \mathcal{H}$ the entropy is zero, but from statistical mechanics we know that any irreversible change that occurs in a system will increase entropy or create disorder. For a pure vector state the uncertainty about the outcome of any measurement is not due to the lack of information about the system. In standard interpretations of quantum mechanics a pure state $\psi = \lambda_1 e_1 + \lambda_2 e_2$ represents the maximum amount of information we can have. Any uncertainties are purely non-classical, and related to the statistical interpretation of the state vector ψ. If we perform a measurement on this system in a e_n basis we would get a result that corresponded either to e_1 or e_2, with probabilities $|\lambda_1|^2$ and $|\lambda_2|^2$. The system is in a pure quantum state, which is a *neither nor* state that is impossible for golf balls, but possible for electrons.

On the other hand, if we have a system which is described by an ensemble ψ_k, with p_k as the probability associated with the state ψ_k, we could regard this as a probability

distribution or statistical ensemble of several quantum mechanical states. This state is *either or*, a type of distribution which is possible for both golf balls and electrons. The average value of a measurement on an observable $\mathbf{F}$ is then

$$\langle \mathbf{F} \rangle = \sum_k p_k \, (\psi_k^\dagger \mathbf{F} \psi_k) = \sum_k p_k \langle \mathbf{F} \rangle_k \tag{3.4}$$

which is a sum of the average values from (3.2) weighted by the p_k. On this note we can introduce the concept of a density operator.

Definition 3.1 (Density operator)**.** *We define the probability distribution of projections $\psi_k \psi_k^\dagger$ onto the state vectors $\psi_k \in \mathcal{H}$*

$$\rho = \sum_k p_k \, (\psi_k \psi_k^\dagger) \qquad p_k > 0 \qquad \sum_k p_k = 1 \tag{3.5}$$

as a density operator or density matrix on $\mathcal{H}$.

The two conditions that $p_k > 0$ and $\sum_k p_k = 1$ can be inferred from the fact that the matrix ρ is considered to represent a probability distribution. The latter implies that $\mathrm{tr}(\rho) = 1$. We take the expectation value of a density operator in any vector ψ

$$\psi^\dagger \rho \psi = \sum_k p_k |\psi_k^\dagger \psi|^2 \geq 0 \tag{3.6}$$

The fact that $\psi^\dagger \rho \psi \geq 0$ for any state ψ means that ρ has all non-negative eigenvalues $\lambda_k \geq 0$, *i.e.* it is a *positive operator* and this is usually written $\rho \geq 0$. It follows from this and the spectral representation of ρ that $\psi^\dagger \rho \psi \geq 0$ if and only if $\rho \psi \geq 0$.

So, based on this we define an Hermitian positive semidefinite $N \times N$ matrix of unit trace which can be expressed in the form (3.5), to be a density matrix on the N-dimensional Hilbert space $\mathcal{H}$.

Definition 3.2 (The set $\mathcal{D}$)**.** *The set of matrices formed by the density matrices is usually written*

$$\mathcal{D} = \mathcal{D}_N = \{\rho \in H_N \,|\, \rho \geq 0, \, \mathrm{tr}(\rho) = 1\} \tag{3.7}$$

The natural structure of the set H_N of Hermitian $N \times N$ matrices is that of a real Hilbert space of dimension N^2, where for $X, Y \in H_N$

$$\langle X, Y \rangle = \mathrm{tr}(X^\dagger Y) = \mathrm{tr}(XY) \tag{3.8}$$

is a well defined inner product, called the *Hilbert-Schmidt* inner product. This gives us access in the matrix space to concepts such as distances $|X - Y|^2 = \mathrm{tr}[(X - Y)^2]$ and angles $\cos\theta = \mathrm{tr}(XY)$, and from the latter orthogonality as $\mathrm{tr}(XY) = 0$. Note that if the Hilbert-Schmidt norm of a matrix X is one, this means that $\mathrm{tr}(X^2) = 1$, which if $\mathrm{rank}(X) > 1$ in turn implies that $\mathrm{tr}(X) < 1$. Since we insist that $\mathrm{tr}(\rho) = 1$, the Hilbert-Schmidt norm has its main use in geometrical aspects, such as distances.

The average value of an operator $\mathbf{F}$ in terms of the density matrix is

$$\langle \mathbf{F} \rangle = \mathrm{tr}(\mathbf{F}\rho) = \sum_k p_k \, (\psi_k^\dagger \mathbf{F} \psi_k) \tag{3.9}$$

If the number of terms in the decomposition (3.5) is equal to one, we have only one state ψ_1 with probability one. The density matrix then represents a pure state $\rho_1 = \psi_1 \psi_1^\dagger$, a state of rank one. For the cases where $k > 1$ in (3.5) we have a genuinely mixed state, and since ρ is Hermitian (normal), we can utilize (3.3) to show that $\mathrm{tr}(\rho^2) = \sum_k p_k^2 < 1$. So only for a pure state do we have $\rho^2 = \rho$, which is then an idempotent projection operator onto a single pure state. So, the trace of the square of ρ is a measure of how mixed ρ is, where $\mathrm{tr}(\rho^2) = \mathrm{tr}(\rho) = 1$ means no mixing at all. The maximum amount of mixing for a state ρ on an N-dimensional Hilbert space is the maximally mixed state

$$\rho = I/N \tag{3.10}$$

with I as the $N \times N$ identity matrix. We realize that this state is the one which gives minimal knowledge about the system, since all the states in the ensemble ψ_k are equally weighted with $p_k = 1/N$.

It is a tempting and common fallacy to suppose that the ensemble of states ψ_k along with probability factors p_k, which is used to compose a density matrix has a special physical significance. This is however not the case. To illustrate this we look at the 2×2 density matrix

$$\rho = \frac{3}{4} |0\rangle\langle 0| + \frac{1}{4} |1\rangle\langle 1| \tag{3.11}$$

with $|0\rangle$ and $|1\rangle$ as for instance (2.2). Suppose we defined the two states

$$\begin{aligned} |a\rangle &= \sqrt{\frac{3}{4}} |0\rangle + \sqrt{\frac{1}{4}} |1\rangle \\ |b\rangle &= \sqrt{\frac{3}{4}} |0\rangle - \sqrt{\frac{1}{4}} |1\rangle \end{aligned} \tag{3.12}$$

then we see that

$$\rho = \frac{1}{2} |a\rangle\langle a| + \frac{1}{2} |b\rangle\langle b| = \frac{3}{4} |0\rangle\langle 0| + \frac{1}{4} |1\rangle\langle 1| \tag{3.13}$$

So the construction of mixed density matrices is not bijective, in the sense that there is no unique decomposition of ρ into an ensemble. For most cases there are infinitely many such decompositions, and the spectral decomposition (3.3) is always one of them. It therefore does not make any sense to ask out of which states a mixed state was originally constructed. For pure states, *i.e.* states of rank one, the decomposition will however always be unique.

3.2 Tensor product states

It is important to describe several subsystems (usually particles) with a single quantum state. These states are elements of a tensor product space. If we assume a bipartite quantum states with system A described by an N_A-dimensional Hilbert space $\mathcal{H}_A$, and likewise for B by the N_B-dimensional Hilbert space $\mathcal{H}_B$, then this Hilbert space will be $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, with the complex tensor product $\mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$. The tensor product is a way of putting vector spaces together to form larger vector spaces. The construction is crucial to understanding the quantum mechanics of multiparticle systems.

Pure tensor product states

Consider two quantum systems A and B. As above, $|\mathcal{H}_A| = N_A$ and $|\mathcal{H}_B| = N_B$. We start with the construction of a state of the composite system AB of the form $\psi = \phi \otimes \chi$ with $\phi \in \mathcal{H}_A$ and $\chi \in \mathcal{H}_B$, and call this a product state. A very important observation is that states of this product form, do not make up a vector space, because we realize that a combination of two product vectors

$$\psi = c_1 (\phi_1 \otimes \chi_1) + c_2 (\phi_2 \otimes \chi_2) \tag{3.14}$$

with $c_1, c_2 \in \mathbb{C}$, is not generally a product vector. The total set of product vectors is therefore called the *Cartesian product* of $\mathcal{H}_A$ and $\mathcal{H}_B$, *i.e.* the set of all ordered pairs consisting of a vector from $\mathcal{H}_A$ and a vector from $\mathcal{H}_B$.

Since a general principle of quantum mechanics is that systems are described by complex vector spaces, we must insist that the superposition principle also holds for composite quantum systems. This means that the appropriate state space for the composite system is not just the Cartesian product, but rather the entire vector space spanned by product vectors $\psi = \phi \otimes \chi$. The states that are not of product form $\psi = \phi \otimes \chi$ are then the entangled quantum states for the system. The vector space $\mathcal{H} = \mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ spanned by the product vectors, is called the tensor product of $\mathcal{H}_A$ and $\mathcal{H}_B$. The inner product of $\mathcal{H}$ is defined by

$$(\phi_1 \otimes \chi_1)^\dagger (\phi_2 \otimes \chi_2) = (\phi_1^\dagger \phi_2)(\chi_1^\dagger \chi_2) \tag{3.15}$$

Any vector $\psi \in \mathcal{H}$ can be written as a linear combination of product vectors. All the vectors in the products can be expanded in orthonormal bases ϕ_i with $i = 1, \ldots, N_A$ for A, and χ_j with $1, \ldots, N_B$ for B. For the product vectors $\phi_i \otimes \chi_j$ we can write in increasing order of omitting redundancies

$$|\phi_i\rangle \otimes |\chi_j\rangle \equiv |\phi_i\rangle|\chi_j\rangle \equiv |\phi_i, \chi_j\rangle \equiv |i, j\rangle \equiv |ij\rangle \tag{3.16}$$

For ease of notation we here use Dirac notation, so that the expansion of an arbitrary vector in $\mathcal{H}$ is then

$$|\psi\rangle = \sum_{ij} |ij\rangle\langle ij|\psi\rangle = \sum_{ij} c_{ij}|ij\rangle \tag{3.17}$$

which indeed makes the dimension of $|\mathcal{H}| = N = N_A N_B$. The coefficients $c_{ij} = \langle ij|\psi\rangle$ can be written as an $N_A N_B \times 1$ column vector, so we can effectively think of a vector in $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ as an N_A-dimensional vector whose components are N_B-dimensional vectors.

$$|\psi\rangle = \begin{pmatrix} \mathbf{c}_1 \\ \mathbf{c}_2 \\ \vdots \\ \mathbf{c}_{N_A} \end{pmatrix} \qquad \text{with} \qquad \mathbf{c}_k = \begin{pmatrix} c_{k1} \\ c_{k2} \\ \vdots \\ c_{kN_B} \end{pmatrix} \tag{3.18}$$

If the coefficients $c_{ij} = \langle ij|\psi\rangle$ factor as $c_{ij} = a_i b_j$ then

$$|\psi\rangle = \sum_{ij} a_i b_j |ij\rangle = \left(\sum_i a_i |i\rangle\right) \otimes \left(\sum_j b_j |j\rangle\right) = |\phi\rangle \otimes |\chi\rangle \tag{3.19}$$

which is a product state in $\mathcal{H}_N$. Contrasted with the general form (3.17), the demand that $c_{ij} = a_i b_j$ is not generally satisfied. It is for example easy to see that for the 2×2 system, where $|\psi\rangle$ would be a 4×1 vector with coefficients ψ_i, we would need to have $\psi_1\psi_4 = \psi_2\psi_3$ for this to be the case, or equivalently $c_{11}c_{22} = c_{12}c_{21}$. We see here the correspondence between the double index system ij and the indices for the $N_A N_B \times 1$ vector $|\psi\rangle$, in that we write the components of $|\psi\rangle$ as $\psi_I = \psi_{ij}$ where

$$\begin{gathered} I = 1, 2, \ldots, N_B, N_B + 1, N_B + 2, \ldots, N \\ \updownarrow \\ ij = 11, 12, \ldots, 1N_B, 21, 22, \ldots, N_A N_B \end{gathered} \tag{3.20}$$

Schmidt decomposition

For pure states the problem of determining whether a state is separable or entangled, is completely solved due to Schmidt decomposition [30], which is essentially a restatement of the *singular value decomposition* scheme. When formulated for our purposes this decomposition essentially means that for any vector $\psi \in \mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ there exists an integer r such that $1 \leq r \leq \min(N_A, N_B)$, a set of real numbers $c_i > 0$, and orthonormal sets of vectors $\phi_i \in \mathbb{C}^{N_A}$ and $\chi_i \in \mathbb{C}^{N_B}$ such that

$$\psi = \sum_{i=1}^{r} c_i \left(\phi_i \otimes \chi_i\right) \tag{3.21}$$

with $\sum_{i=1}^{r} c_i^2 = 1$. This scheme is possible because the orthonormal vectors ϕ_i and χ_i are chosen differently for each vector ψ, and are not fixed like a complete orthonormal basis would be. The minimum integer $r = r_{\min}$ allowed for the form (3.21) is called the *Schmidt rank* of ψ, and $r_{\min} = 1$ if and only if ψ is separable.

The important thing is that the task of finding the Schmidt decomposition of a general state ψ is numerically easy.

Tensor products of operators

We now move to operators acting on the tensor product space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. We have seen that we can use for instance $(\phi_1 \otimes \chi_1)(\phi_2 \otimes \chi_2)^\dagger = \phi_1\phi_2^\dagger \otimes \chi_1\chi_2^\dagger$, and this arranges the outer product for states in $\mathcal{H}$, as a tensor product of outer product operators for the individual systems A and B. We now let $\mathbf{F} = \mathbf{A} \otimes \mathbf{B}$ be a product operator on $\mathcal{H}$. Recalling (3.16) and (3.20) we then write $F_{IK} = F_{ij;kl}$ with $I \leftrightarrow ij$ and $K \leftrightarrow kl$ as

$$F = \left(\sum_{ik} A_{ik}|i\rangle\langle k|\right) \otimes \left(\sum_{jl} B_{jl}|j\rangle\langle l|\right) \tag{3.22}$$

so effectively we get

$$F = \sum_{ij;kl} A_{ik}B_{jl}|i\rangle\langle k| \otimes |j\rangle\langle l| = \sum_{ij;kl} A_{ik}B_{jl}|ij\rangle\langle kl| \tag{3.23}$$

The matrix F above is thus an $N_A \times N_A$ matrix, where each component is an $N_B \times N_B$ matrix. To observe of this works we write an operator $\mathbf{A} \otimes \mathbf{I}$ with $\mathbf{A}$ acting on $\mathcal{H}_A$ and the identity $\mathbf{I}$ acting on $\mathcal{H}_B$

$$A \otimes I = \sum_{ij;kl} A_{ik}|i\rangle\langle k| \otimes |j\rangle\langle j| = \sum_{ij;kl} A_{ik}|ij\rangle\langle kj| \tag{3.24}$$

where the l component becomes superfluous since the j component is sufficient to carry the identity operation. Also, a useful expression for the expectation value of an operator $\mathbf{F}$, which need not necessarily be a product operator of the form (3.22), in a product state $\psi = \phi \otimes \chi$ is then

$$(\phi \otimes \chi)^\dagger \mathbf{F} (\phi \otimes \chi) = \sum_{ij;kl} \phi_i^* \chi_j^* F_{ij;kl} \phi_k \chi_l \tag{3.25}$$

This is a bilinear form in the coordinates $\phi_i, \phi_k, \chi_j, \chi_l$ of ϕ and χ, with $i, k = 1, \ldots, N_A$ and $j, l = 1, \ldots, N_B$. Its significance will become very clear in Chapter 7. A product operator $\mathbf{F} = \mathbf{A} \otimes \mathbf{B}$ makes operations on each subsystem individually, and we can easily calculate $\langle \mathbf{A} \otimes \mathbf{B} \rangle$ in a product state

$$(\phi \otimes \chi)^\dagger (\mathbf{A} \otimes \mathbf{B})(\phi \otimes \chi) = (\phi^\dagger \mathbf{A} \phi) \otimes (\chi^\dagger \mathbf{B} \chi) = \langle A \rangle \langle B \rangle \tag{3.26}$$

so as expected, the expectation values factorize. Of course for an entangled state $\psi \neq \phi \otimes \chi$ this may very well not be the case.

Mixed tensor product states

We can now define density matrices $\rho = \rho_{AB}$ on a tensor product space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, and these will respect the real tensor product $H_N = H_{N_A} \otimes H_{N_B}$. If we consider the ensemble $\psi_k \in \mathcal{H}$ with the probability distribution p_k, this gives rise to a density matrix

$$\rho = \sum_k p_k \left(\psi_k \psi_k^\dagger\right) = \sum_k p_k \rho_k \tag{3.27}$$

where $\sum_k p_k = 1$ and all $p_k \geq 0$. So the mixed tensor product state ρ is a convex sum of outer products on $\mathcal{H}$. Since the density matrix $\rho = \sum_k p_k \rho_k$ is a convex sum of individual density matrices ρ_k of lesser rank, it could be viewed as an *ensemble of ensembles*, but we usually prefer to write $\rho_k = \psi_k \psi_k^\dagger$, so that $\text{rank}(\rho_k) = 1$ for all k. Further, we follow Werner [21], and make the definition

Definition 3.3 (Separable states)**.** *A density matrix ρ is separable if it can be expressed in the form*

$$\rho = \sum_k p_k \left(\phi_k \phi_k^\dagger \otimes \chi_k \chi_k^\dagger\right) = \sum_k p_k \left(\mu_k \otimes \tau_k\right) \tag{3.28}$$

with $\mu_k \in \mathcal{D}_{N_A}$ and $\tau_k \in \mathcal{D}_{N_B}$. The complete set of separable states is $\mathcal{S}$.

If this is not possible, the mixed state ρ is entangled. So for ρ to be separable there must exist an ensemble ψ_k, where all $\psi_k = \phi_k \otimes \chi_k$, along with a probability distribution p_k, such that the form (3.28) is possible.

As we have seen, a density matrix may be constructed from a whole range of different ensembles, so the defining property for a separable density matrix is that the form (3.28) is one of the *possible*. We saw a very simple example of this in (3.13), but purely to illustrate the complexity involved even for low dimensional cases such as the 2×2 system, we consider another example. In the $|ij\rangle$ basis we consider the state defined by

$$\begin{aligned}
\rho = &\frac{3}{24}|00\rangle\langle 00| + \frac{\sqrt{2}}{24}\mathrm{i}\,|00\rangle\langle 01| + \frac{\sqrt{2}}{12}\mathrm{i}\,|00\rangle\langle 11| \\
&- \frac{\sqrt{2}}{24}\mathrm{i}\,|01\rangle\langle 00| + \frac{1}{4}|01\rangle\langle 01| - \frac{\sqrt{2}}{12}\mathrm{i}\,|01\rangle\langle 10| \\
&+ \frac{\sqrt{2}}{12}\mathrm{i}\,|10\rangle\langle 01| + \frac{5}{24}|10\rangle\langle 10| - \frac{\sqrt{2}}{24}\mathrm{i}\,|10\rangle\langle 11| \\
&- \frac{\sqrt{2}}{12}\mathrm{i}\,|11\rangle\langle 00| + \frac{\sqrt{2}}{24}\mathrm{i}\,|11\rangle\langle 10| + \frac{5}{12}|11\rangle\langle 11|
\end{aligned} \tag{3.29}$$

If we employ the representation (2.8) we can write

$$\rho = \begin{pmatrix}
\frac{3}{24} & \frac{\sqrt{2}}{24}\mathrm{i} & 0 & \frac{\sqrt{2}}{12}\mathrm{i} \\
-\frac{\sqrt{2}}{24}\mathrm{i} & \frac{1}{4} & -\frac{\sqrt{2}}{12}\mathrm{i} & 0 \\
0 & \frac{\sqrt{2}}{12}\mathrm{i} & \frac{5}{24} & -\frac{\sqrt{2}}{24}\mathrm{i} \\
-\frac{\sqrt{2}}{12}\mathrm{i} & 0 & \frac{\sqrt{2}}{24}\mathrm{i} & \frac{5}{12}
\end{pmatrix} \tag{3.30}$$

To unveil that this matrix represents a separable state, *i.e.* it can be written in the form (3.28), is not trivial. Still, one possibility is

$$\rho = \frac{1}{2}\left(\mu_1 \otimes \tau_1\right) + \frac{1}{2}\left(\mu_2 \otimes \tau_2\right) \tag{3.31}$$

where

$$\begin{aligned} \mu_1 &= \frac{1}{2}\,|1\rangle\langle 1| + \frac{1}{2}\,|-\rangle\langle -| \qquad & \tau_1 &= |\gamma+\rangle\langle\gamma+| \\ \mu_2 &= \frac{1}{2}\,I/4 + \frac{1}{2}\,|+\rangle\langle +| \qquad & \tau_1 &= |\gamma-\rangle\langle\gamma-| \end{aligned} \tag{3.32}$$

where $|\pm\rangle$ are defined by (2.5), $|\gamma\pm\rangle = \frac{1}{\sqrt{3}}\,|0\rangle \pm \sqrt{\frac{2}{3}}\,\mathrm{i}\,|1\rangle$ and we recognize $I/4$ as the maximally mixed state.

For the pure separable state $\psi = \phi \otimes \chi$ we saw that $\langle \mathbf{A} \otimes \mathbf{B}\rangle = \langle \mathbf{A}\rangle\langle \mathbf{B}\rangle$ so that no quantum correlations could exist. Mixed separable states will in general contain correlations, but these correlations are purely classical. For example the state

$$\rho = \frac{1}{2}\,|00\rangle\langle 00| + \frac{1}{2}\,|11\rangle\langle 11| \tag{3.33}$$

is correlated, in the sense that any measurement in the basis $|0\rangle, |1\rangle$ in either subsystem will always yield the same result. It can be compared to the correlations that exist in the Bell state (1.12) when the angle between the axes of measurements of the polarizations is $\phi = 0°$.

3.3 Partial trace

For composite systems the partial trace is a trace operation with respect to a subset of the systems. In many ways this is the opposite operation to combining two Hilbert spaces.

Definition 3.4 (Partial trace)**.** *For a general $\rho = \rho_{ij;kl}$ on $\mathcal{H}_A \otimes \mathcal{H}_B$, we can define the partial trace with respect to system B by*

$$\rho_{ij;kl} \to \rho_{ik} = \sum_j \rho_{ij;kj} = tr_B(\rho) \tag{3.34}$$

where the indices ik represent system A and jl system B.

What actually happens is that we trace out, or average over all correlations that may or may not be present in the state. One may of course equivalently do the partial trace operation with respect to system A. The ranks r_A and r_B of the states $\rho_A = \mathrm{tr}_B(\rho)$ and $\rho_B = \mathrm{tr}_A(\rho)$ are called the local ranks of ρ. The physical interpretation of $\rho_A = \mathrm{tr}_B(\rho)$ is that it provides the correct measurement statistics for measurements that could be made on system A. For the case of the product state $\rho = \mu \otimes \tau$, we have

$$\mathrm{tr}_B(\mu \otimes \tau) = \mu\,\mathrm{tr}(\tau) = \mu \tag{3.35}$$

which could be expected. A less trivial example is the Bell state $|\beta+\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$. This state has a density operator

$$\begin{aligned}\rho &= |\beta+\rangle\langle\beta+| \\ &= \left(\frac{1}{\sqrt{2}}\{|00\rangle + |11\rangle\}\right)\left(\frac{1}{\sqrt{2}}\{\langle 00| + \langle 11|\}\right) \\ &= \frac{1}{2}\left(|00\rangle\langle 00| + |11\rangle\langle 00| + |00\rangle\langle 11| + |11\rangle\langle 11|\right)\end{aligned} \tag{3.36}$$

or in the standard matrix representation

$$\rho = \frac{1}{2}\left(\begin{array}{cc:cc} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ \hdashline 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{array}\right) \tag{3.37}$$

Tracing out the second qubit can be done in different ways. Observing that the partial trace $\mathrm{tr}_B(|ij\rangle\langle kl|)$ is the matrix $\langle l|j\rangle\,(|i\rangle\langle k|)$ and that the basis $|ij\rangle$ is orthonormal we can appeal to (3.36) to get

$$\rho_A = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = I/2 \tag{3.38}$$

The same answer could have been achieved by replacing the four 2×2 submatrices in (3.37) with their trace.

Since the state $I/2$ is a mixed state, this is a remarkable result. The state of the joint system of qubits is a pure state, that is a state of which we have complete knowledge or certainty. But the state of the first qubit (and by symmetry also the second) is a mixed state, of which we have less knowledge, and in this case even minimal knowledge since the state $I/2$ is the maximally mixed state. The strange property that we have more knowledge about the whole system than the parts of it, is a highly *non-reductionist* effect, which is entirely due to entanglement.

3.4 Decoherence

It is an easy mistake to assume that a convex mix of two density matrices for two entangled states, is again an entangled state. Let us again use the Bell states (1.23) to illustrate an important concept. The Bell states are known to be maximally entangled. We take an even mix of the two states $|\beta+\rangle$ and $|\beta-\rangle$

$$\rho = \frac{1}{2}|\beta+\rangle\langle\beta+| + \frac{1}{2}|\beta-\rangle\langle\beta-| \tag{3.39}$$

Again using the basis $|ij\rangle$ and drawing from (3.36) and (3.37) we observe that the density matrix for the state $|\beta\pm\rangle$ is $(|00\rangle\langle 00| \pm |11\rangle\langle 00| \pm |00\rangle\langle 11| + |11\rangle\langle 11|)/4$, so the two middle terms cancel when we make the sum (3.39), and we get the separable state

$$\rho = \frac{1}{2}|00\rangle\langle 00| + \frac{1}{2}|11\rangle\langle 11| \tag{3.40}$$

In general if a state evolves from $\rho_1 \to \rho_2$ in such a way that it becomes increasingly mixed, *i.e.* $\mathrm{tr}(\rho_1^2) > \mathrm{tr}(\rho_2^2)$ this is known as *decoherence* or *quantum noise*. As a general rule any type of mixing will destroy some entanglement in the total state. In many cases the state may stay entangled, but as a rule it will be *less entangled*. To understand this further we would need to develop a theory of how to quantify the amount of entanglement for a state ρ. As was mentioned in Section 2.1 we can employ the use of entanglement measures, but as the theory of entanglement measures is a formidable subject we will not dwell into it further.

The most common way of looking at decoherence is to imagine that we control two systems A and B described by $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, and that we wish to perform quantum operations on these. We now introduce a third system E, which represents the environment, and over which we have none or very limited control. We then effectively have a tripartite system

$$\mathcal{H} = \underbrace{\mathcal{H}_A \otimes \mathcal{H}_B}_{\text{Control}} \otimes \mathcal{H}_E \tag{3.41}$$

of which we can control only A and B. The interactions between the AB system and the environment E will in general involve adding terms to the density matrix ρ_{AB}. This will lead to a more mixed state, and inevitably to loss of entanglement, and an accompanied loss of non-local correlations between parts A and B of the system.

In many applications in quantum information this is a formidable problem, and since the desired working states for the AB system are pure states, any mixing with the environment is highly undesirable. The ability of an entangled state ρ to withstand mixing with the environment, which leads to loss of entanglement, is called *robustness of entanglement*, and is an area of considerable research activity. The biggest challenge today if one is hoping to build and operate quantum computers of a sufficient size and complexity, is decoherence.

Finally, a skill which is very practical for any quantum physicist in social situations, is to be able to answer questions about the before mentioned Schrödingers cat. The seemingly paradoxical situation is easily resolved by appealing to decoherence. One assumes that the cat should be in a superposition between dead and alive, or rather in an entangled state with the poison trap. Any such superposition would be immediately ruined by decoherence, so putting the cat into the wanted superposition is altogether impossible. One may argue that the cat is quite rightly in a quantum state, but in a mixed and incredibly complex one.

3.5 PPT states

We have defined and investigated two important sets of matrices on the Hilbert space $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, namely the set of density matrices $\mathcal{D}$, and the set of separable density matrices $\mathcal{S} \subset \mathcal{D}$. We will now define a third set which will also be a subset of $\mathcal{D}$, and to start us up we will define an operation called *partial transposition*.

3.5.1 Partial transposition

This is an operation where the part of a density matrix ρ on $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ that corresponds to system B, is transposed. The transpose map $\mathbf{T} : M_{ij} \mapsto M_{ji}$ is a positive map in the sense that if a matrix $M = M_{ij}$ is positive, meaning that all eigenvalues $\lambda_i \geq 0$, then so will the transpose $M^T = M_{ji}$. Since transposition also preserves trace, this means that a density matrix is mapped into a density matrix by transposition. For any matrix ρ acting on $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ partial transposition is defined by the product operator $\mathbf{I} \otimes \mathbf{T}$

$$\rho^P = (\mathbf{I} \otimes \mathbf{T})\rho \tag{3.42}$$

which may or may not be a positive matrix. The identity operator $\mathbf{I}$ and transposition operator $\mathbf{T}$ works on system A and B respectively. In matrix notation a more compact and more useful way of doing this is to simply interchange the indices referring to subsystem B.

Definition 3.5 (Partial transposition)**.** *If $\rho_{ij;kl}$ is on the Hilbert space $\mathcal{H}_A \otimes \mathcal{H}_B$, and the indices i, k represent system A and j, l system B, then the operation*

$$\rho^P_{ij;kl} = \rho_{il;kj} \tag{3.43}$$

is called partial transposition with respect to system B.

The affix P expresses that this is partial transposition, and we shall choose P to symbolize partial transposition as defined above, *i.e.* with respect to system B. This can also formally be expressed as $\rho^P = \rho^{T_B}$. It is obvious that $(\rho^P)^P = \rho$, and also that the eigenvalues of ρ^P will not necessarily be positive, so that we may have $\rho^P \notin \mathcal{D}$. It is also useful to observe that since $\mathrm{tr}(X^P Y^P) = \mathrm{tr}(XY)$, partial transposition is also an orthogonal transformation.

The partial transposition operation, or rather a good approximation of it, has in fact been realized experimentally [31].

3.5.2 PPT states and the Peres set

In general ρ^P will have different eigenvalues than ρ, but for some states they remain positive, so that $\rho \in \mathcal{D} \Leftrightarrow \rho^P \in \mathcal{D}$. The states with this property are called positive partial transpose states, or PPT states.

Definition 3.6 (PPT states and the Peres set)**.** *Any density matrix ρ such that ρ^P is a density matrix, we call a PPT state, and the set of PPT states $\mathcal{P} \subset \mathcal{D}$ is called the Peres set.*

The Peres set and PPT states was first introduced by Peres [32]. The PPT property is independent of whether we define partial transposition with respect to subsystem A or B. Since $(\rho^{T_A})^{T_B} = \rho^T \geq 0$ always holds, it follows that $\rho^{T_A} \geq 0$ if and only if $\rho^{T_B} \geq 0$. If $\rho \in \mathcal{P}$ with $\mathrm{rank}(\rho) = m$ and $\mathrm{rank}(\rho^P) = n$, we usually say that ρ is a PPT

state of rank (m, n). It is quite easy to realize that any separable state ρ will always be a PPT state since

$$(\mathbf{I}\otimes\mathbf{T})\rho = \rho^P = (\mathbf{I}\otimes\mathbf{T})\left(\sum_k p_k(\mu_k\otimes\tau_k)\right) = \sum_k p_k\big(\mu_k\otimes\mathbf{T}(\tau_k)\big) \tag{3.44}$$

and since $\mathbf{T}(\tau_k)$ is a density matrix, so will ρ^P. In fact $\rho^P \in \mathcal{S}$.

Theorem 3.1 (Peres-Horodecki criterion)**.** *Let $\rho \in \mathcal{D}$ and ρ^P be the partially transposed matrix of ρ. Then if $\rho^P \not\geq 0$, the state ρ is entangled.*

For the 2×2 and 2×3 systems the PPT property is exclusive to separable states, but for systems of dimension $N = N_A N_B \geq 8$ there exist PPT states that are entangled [33]. Thus for systems with $N \geq 8$ the PPT property can only be used in a contrapositive way: *i.e.*, if a state is NPT we know for any N that it is not separable. The first example of an entangled PPT state was given by Michal and Pawel Horodecki [34]. The general problem of distinguishing separable states from entangled states is a fundamental and unsolved problem in quantum information theory. For bipartite systems the problem is solved for the 2×2 and 2×3 systems, since checking the PPT property of a density matrix is a quick task numerically.

Since a PPT state is defined to be a state ρ for which the partial transpose ρ^P is a density matrix, we can also define the set of PPT states as

$$\mathcal{P} = \{\rho \,|\, \rho \in \mathcal{D} \cap \mathcal{D}^P\} \tag{3.45}$$

where $\mathcal{D}^P$ is the set of transposed density matrices. An attempt to visualize this is made in Figure 3.1

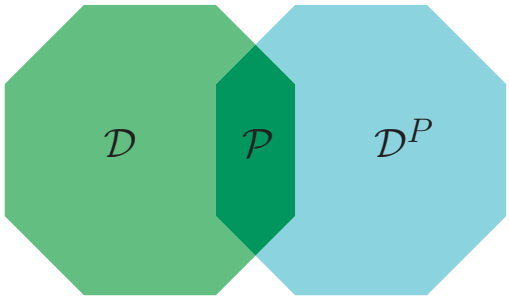


Figure 3.1: We see a visualization of the Peres set $\mathcal{P}$ as the intersection of the sets $\mathcal{D}$ and $\mathcal{D}^P$

The Werner states for the 2×2 system, initially published by Werner [21], which are a convex combination of the maximally entangled state $|\beta+\rangle$ and the maximally mixed state $I/4$, provide a magnificent example

$$\rho(x) = x\,|\beta+\rangle\langle\beta+| + (1-x)\,\frac{I}{4} \qquad\qquad x \in [0,1] \tag{3.46}$$

This gives the density matrix and its partial transpose as

$$\rho(x) = \frac{1}{4}\left(\begin{array}{cc:cc} 1+x & 0 & 0 & 2x \\ 0 & 1-x & 0 & 0 \\ \hdashline 0 & 0 & 1-x & 0 \\ 2x & 0 & 0 & 1+x \end{array}\right) \tag{3.47}$$

$$\rho^P(x) = \frac{1}{4}\left(\begin{array}{cc:cc} 1+x & 0 & 0 & 0 \\ 0 & 1-x & 2x & 0 \\ \hdashline 0 & 2x & 1-x & 0 \\ 0 & 0 & 0 & 1+x \end{array}\right) \tag{3.48}$$

The partial transpose of ρ with respect to system B is found by transposing each 2×2 submatrix of ρ. It is easy to verify that all the eigenvalues of ρ are positive for $x \in [0, 1]$. The eigenvalues of ρ^P are $\lambda_1 = (1 - 3x)/4$ and the remaining three are degenerate and equal to $\lambda_{2,3,4} = (1 + x)/4$. This means that $\rho^P(x)$ for the 2×2 system is positive definite for $x \leq 1/3$. So $\rho(x)$ is in fact entangled for $x > 1/3$ and separable for $x \leq 1/3$. It seems reasonable to think that the amount of entanglement in $\rho(x)$ will gradually approach zero as $x \to (1/3)^+$.

3.5.3 Bound and free entanglement

Entanglement distillation is the transformation of N copies of an arbitrary entangled state ρ into some number $n < N$ of approximately pure Bell pairs, using only LOCC operations. The main motivation behind these operations is decoherence, or loss of entanglement. Let us assume that we have an entanglement measure E which is scaled in such a way that $E(\rho_{\text{sep}}) = 0$ and $E(\rho_{\text{Bell}}) = 1$. Let us further assume that we have a system of $N = 100$ identical states ρ_0 for which we have $E(\rho_0) = 1/10$, so that the total amount of entanglement $E_0 = \sum_{N=1}^{100} E(\rho_0) = 10$. Assume that during a process each of these 100 states evolved such that $\rho_1 = U\rho_0 U^\dagger$ where $E(\rho_1) = 1/100 \Rightarrow E_1 = \sum_{N=1}^{100} E(\rho_1) = 1$. If one could transform all the ρ_1 states by LOCC operations into one Bell state without significant loss of entanglement, then the entanglement in the system would be considerably more concentrated. If the state ρ_1 allows such a scheme, the entanglement it contains is said to be *free entanglement*, in the opposite case it is called *bound entanglement*. The first attempt at describing this idea was by Bennett *et al.* [35]. Since this first work, many others have published on this matter, providing several different protocols for distilling entangled states.

There is a link between the PPT property of a state ρ and the type of entanglement it contains. It is known that any entangled PPT state has bound entanglement, and in the 3×3 system they form a small subset of $\mathcal{D}$ [36]. The entangled states with the PPT property have some characteristics that would appear to make them more "separable" than the NPT (non-PPT) entangled states. A legitimate question would then be whether an NPT state always contains free entanglement. This question has not been finally answered yet, but there is some evidence which suggests that NPT states with bound entanglement do exist [37, 38].

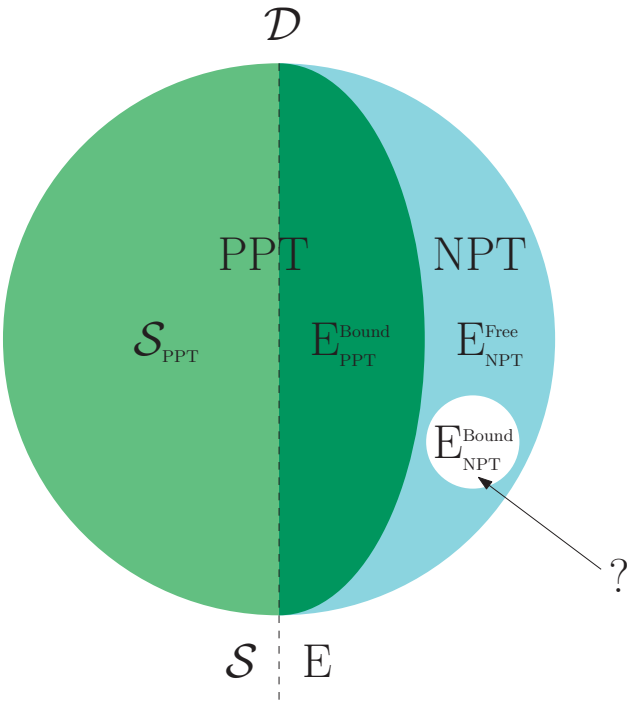


Figure 3.2: The set $\mathcal{D}$ with the subset $\mathcal{S}$ in light green, the set of entangled PPT states in dark green and the set of NPT states is in light blue. The left half are the separable states and the right half the entangled states. The small white circle represents the hypothetical NPT states with bound entanglement.

In Figure 3.2 the subset $\mathcal{S}$ is in light green, the set of entangled PPT states in dark green and the set of NPT states in light blue. The left half are the separable states and the right half the entangled states. The small white circle represents the hypothetical NPT states with bound entanglement.

3.6 Product vectors in subspaces

$\text{Img}\,\rho$, or the range of ρ, is defined as the set $R(\rho) = \{\psi \in \mathbb{C}^N \,|\, \exists\gamma \in \mathbb{C}^N\text{: } \rho\gamma = \psi\}$, while $\text{Ker}\,\rho$, or the null space of ρ, is the set $N(\rho) = \{\psi \in \mathbb{C}^N \,|\, \rho\psi = 0\}$.

An important characteristic in the classification of the properties of a density matrix ρ is the number of product vectors in the range $\text{Img}\,\rho$ and in the kernel $\text{Ker}\,\rho$. Let the number of product vectors in $\text{Img}\,\rho$ be n_{img} and likewise for $\text{Ker}\,\rho$ be n_{ker}. For purposes of studying PPT states we are also interested in the number of product vectors in $\text{Img}\,\rho^P$ which we put $\widetilde{n}_{\text{img}}$. From the identity

$$(a \otimes b)^\dagger \rho (c \otimes d) = (a \otimes d^*)^\dagger \rho^P (c \otimes b^*) \tag{3.49}$$

we can, by taking $\psi = \phi \otimes \chi$ and $\widetilde{\psi} = \phi \otimes \chi^*$, infer the general relation

$$\psi^\dagger \rho \psi = \widetilde{\psi}^\dagger \rho^P \widetilde{\psi} \tag{3.50}$$

Now, since ρ and by assumption ρ^P are positive matrices, we know that $\psi^\dagger\rho\psi \geq 0 \Leftrightarrow \rho\psi \geq 0$. That $\psi \in \mathrm{Ker}\,\rho$ is then equivalent to the condition that $\widetilde{\psi} \in \mathrm{Ker}\,\rho^P$. So in effect, we always have that $n_{\mathrm{ker}} = \widetilde{n}_{\mathrm{ker}}$. In summary $\{n_{\mathrm{img}}, \widetilde{n}_{\mathrm{img}}; n_{\mathrm{ker}}\}$ is a characterization of a state ρ with regards to the number of product vectors in both the range and kernel of ρ and ρ^P.

3.6.1 Generic cases

We have seen earlier that a vector $|\psi\rangle = \sum_{ij} c_{ij} |ij\rangle \in \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ can be written as a $N_A N_B \times 1$ matrix ψ with components $\psi_I = \psi_{ij}$, where

$$\begin{array}{c} I = 1, 2, \ldots, N_B, N_B+1, N_B+2, \ldots, N \\ \updownarrow \\ ij = 11, 12, \ldots, 1N_B, 21, 22, \ldots, N_A N_B \end{array} \tag{3.51}$$

A product vector $\psi = \phi \otimes \chi$ has components $\psi_{ij} = \phi_i \chi_j$, with ϕ_i, χ_j being the components of the vectors ϕ and χ. We observe that ψ is a product vector if and only if its components satisfy the quadratic equations

$$\psi_{ij}\psi_{kl} = \psi_{il}\psi_{kj} \tag{3.52}$$

These equations are not all independent, and the number of independent complex equations is

$$m = (N_A - 1)(N_B - 1) = N - N_A - N_B + 1 \tag{3.53}$$

For example, if $\psi_1 \neq 0$ we get a complete set of independent equations by taking $i = j = 1$, $k = 2, \ldots, N_A$ and $l = 2, \ldots, N_B$.

Since the equations are homogeneous, any solution $\psi \neq 0$ gives rise to a one parameter family of solutions $c\psi$ where $c \in \mathbb{C}$. A vector ψ in a subspace of dimension n has n independent complex components. Since the most general non-zero solution must contain at least one free complex parameter, we conclude that a generic subspace of dimension n will contain product vectors if and only if

$$n \geq m + 1\ . \tag{3.54}$$

The limiting dimension

$$n = m + 1 = N - N_A - N_B + 2 \tag{3.55}$$

is particularly interesting. In this special case, a non-zero solution will contain exactly one free parameter, which has to be a complex normalization constant. Thus up to proportionality there will exist a finite set of product vectors in a generic subspace of

this dimension, in fact it was shown by Hartshorne [39] that the number of product vectors is

$$p = \binom{N_A + N_B - 2}{N_A - 1} = \frac{(N_A + N_B - 2)!}{(N_A - 1)!(N_B - 1)!} \tag{3.56}$$

A generic subspace of lower dimension will contain no product vector, whereas any subspace of higher dimension will contain a continuous infinity of different product vectors (different in the sense that they are not proportional).

It is important to emphasize that these results hold only for *generic* subspaces. It is trivially clear that there exist non-generic subspaces with dimension $n < m + 1$ that contain product vectors.

3.6.2 Entangled subspaces

If the number of product vectors p in a given subspace $\mathcal{U} \subset \mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ is such that $0 \leq p < |\mathcal{U}|$, it is called an *entangled subspace*. If $p = 0$ it is called a completely entangled subspace. A theorem proved by Parthasarathy [40], that holds generally (and not only for generic subspaces), says that the maximal dimension of a completely entangled subspace for a bipartite system of dimensions $N = N_A N_B$, is given by (3.53). So it is possible to construct a subspace $\mathcal{V} \subset \mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ with $|\mathcal{V}| \leq m = N - N_A - N_B + 1$ that contains no product vectors, but if $|\mathcal{V}| \geq m + 1$ it would have to contain at least one product vector.

3.6.3 Product vectors in orthogonal subspaces

If we want product vectors $\psi = \phi \otimes \chi$ in two orthogonal subspaces, for instance in both the range and the kernel of some state ρ, then they must satisfy a set of orthogonality conditions. We put the product vectors in the range as

$$w_i = u_i \otimes v_i \qquad i = 1, \ldots, n_{\text{img}} \tag{3.57}$$

And likewise for $\operatorname{Ker} \rho$:

$$z_j = x_j \otimes y_j \qquad j = 1, \ldots, n_{\text{ker}} \tag{3.58}$$

Since the two subspaces are orthogonal, it is necessary that $w_i^\dagger z_j = 0$ for all i, j, hence for every pair i, j we must have either $u_i^\dagger x_j = 0$ or $v_i^\dagger y_j = 0$, or both.

3.6.4 The range criterion

An important result connected to product vectors and partial transposition is the range criterion developed by Horodecki [34]. It states

Theorem 3.2 (Range criterion)**.** *A density matrix ρ is separable only if there is a set of product vectors $w_i = u_i \otimes v_i$ that span the range of ρ, and the product vectors $\widetilde{w}_i = u_i \otimes v_i^*$ obtained by partial conjugation of the w_i, span the range of ρ^P.*

In order to convince ourselves of the validity of the range criterion we start with the density operator

$$\rho = \sum_k p_k(\psi_k\psi_k^\dagger) \tag{3.59}$$

and then study the vectors $\xi \in \mathrm{Ker}\,\rho$. These satisfy $\rho\xi = 0 \Leftrightarrow \xi^\dagger\rho\xi = 0$, and from (3.6) we have $\xi^\dagger\rho\xi = \sum_k p_k|\psi_k^\dagger\xi|^2$, from which we deduce that $\psi_k\xi = 0$ for all k. Since $\xi \in \mathrm{Ker}\,\rho$ by assumption, it follows that $\psi_k \perp \xi \Rightarrow \psi_k \in \mathrm{Img}\,\rho$ for all k. Since from (3.59) we have that k must be at least equal to $\mathrm{rank}(\rho)$, the set ψ_k span $\mathrm{Img}\,\rho$. We now assume that ρ is a separable state, so we may choose all $\psi_k = \phi_k \otimes \chi_k$, and further put $\widetilde{\psi}_k = \phi_k \otimes \chi_k^*$. We make use of the identity

$$\{(a \otimes b^*)(a \otimes b^*)^\dagger\}^P = (a \otimes b)(a \otimes b)^\dagger \tag{3.60}$$

from which it emerges that $(\widetilde{\psi}_k\widetilde{\psi}_k^\dagger)^P = \psi_k\psi_k^\dagger$. We then have

$$\rho^P = \sum_k p_k(\psi_k\psi_k^\dagger)^P = \sum_k p_k(\widetilde{\psi}_k\widetilde{\psi}_k^\dagger) \tag{3.61}$$

and using the same argument as for ρ, we find that the $\widetilde{\psi}_k$ span $\mathrm{Img}\,\rho^P$.

Since there are entangled states that satisfy the range criterion, this can only be used in a contrapositive way to prove that the state ρ is not separable. Please note that the range criterion demands that there should exist such a set of product vectors, not that *all* product vectors in the range should satisfy this. This is for instance relevant in the cases where there are an infinite number of product vectors in the range of ρ. On the same note we define the term *edge state*. These are states that break the range criterion in a strong way. That is to say that there exist *no* product vectors satisfying $w_i \in \mathrm{Img}\,\rho \Leftrightarrow \widetilde{w}_i \in \mathrm{Img}\,\rho^P$. Obviously, by the range criterion, all edge states must be entangled.

In addition to the PPT criterion and the range criterion there exist several alternative separability criteria. The reduction criterion [41, 42] and majorization criterion [43] both use the partial traces $\mathrm{tr}_A(\rho)$ and $\mathrm{tr}_B(\rho)$ to establish criteria for separability. We can also mention the entropy criterion [44] and the realignment criterion [45]. All these criteria are, like the PPT criterion, not conclusive in the sense that they can decide with certainty whether a state is separable. Most of them work only in a contrapositive way. There exist criteria based on positive maps, and thereof so-called entanglement witnesses, that work with complete certainty. We shall return to positive maps and entanglement witnesses in Chapter 7.

3.7 Product transformations

Using the properties of linear transformations and their actions on state vectors and density matrices greatly simplifies the efforts to characterize these quantum systems. Focusing on properties that are invariant under certain transformations can make the

study of these systems much easier. This is a desirable property because we can regard the two states as indistinguishable, or equivalent, up to a certain type of transformations. These transformations make an immediate appeal to the notion of *equivalence classes*.

Definition 3.7 (Equivalence under transformations)**.** *The state ρ' is said to be equivalent to the state ρ under the matrix transformation U if*

$$\rho' = U\rho U^\dagger \tag{3.62}$$

for some matrix U.

To ensure the symmetric property of this equivalence, an inverse U^{-1} must exist, so we insist that all transformations U have full rank. Several properties of our states will be invariant under U if the transformation carries a product structure, *i.e.* we insist that $U = U_A \otimes U_B$, where U_A is on $\mathcal{H}_A$ and U_B is on $\mathcal{H}_B$. If we want to ensure complete LOCC equivalence the operators U_A and U_B must be unitary, so that $U_A \in \mathrm{U}(N_A)$ and $U_B \in \mathrm{U}(N_B)$.

3.7.1 SLOCC

We may however extend the equivalence classes by allowing transformations that succeed with a non-zero probability. These are called *stochastic LOCC transformations*, or SLOCC. A linear product transformation, or SLOCC transformation has the form

$$\rho \mapsto \rho' = a\, V\rho V^\dagger \qquad\qquad V = V_A \otimes V_B \tag{3.63}$$

where $a > 0$ is a normalization factor and $V_A \in \mathrm{SL}(N_A, \mathbb{C})$, $V_B \in \mathrm{SL}(N_B, \mathbb{C})$. In total, we no longer require that the local operations are unitary, only that they are members of the groups of non-singular linear operators. There is some abuse of language here, as we refer to transformations such as (3.63) as $\mathrm{SL} \times \mathrm{SL}$-transformations. Even though $\det(V_A) = \det(V_B) = 1$, the total transformation $\rho \mapsto \rho'$ in fact belongs to the more general group $\mathrm{GL}(N_A, \mathbb{C}) \otimes \mathrm{GL}(N_B, \mathbb{C})$. Also, we call states that can be transformed into each other by transformations like (3.63), by the shorter term SL-equivalent, so the product structure of the transformations is often assumed. The $\mathrm{SL} \times \mathrm{SL}$-transformation $V = V_A \otimes V_B$ will also transform Hilbert space vectors as

$$\psi \mapsto \psi' = c\, V\psi \tag{3.64}$$

where $c \in \mathbb{C}$ is a normalization constant. The study of equivalence under deterministic transformations (LOCC) between pure states, was initiated by Lo and Popescu [46], and the extension towards SLOCC equivalence was proposed by Dür *et al.* [47].

Due to the linearity of SLOCC transformations the convexity of structures are preserved. Other important invariants for a state ρ under $\mathrm{SL} \times \mathrm{SL}$-transformations are rank and positivity of both ρ and ρ^P. If we insist on using only local (or product) operations, we also conserve separability and in addition also the number of product

vectors in the range and kernel of ρ and ρ^P. The rank and nullity are always preserved as long as we use invertible transformations. To see that SL $\times$ SL-transformations preserve separability, observe that if we transform a product vector $\psi = \phi \otimes \chi$ we get $\psi \mapsto \psi' = c\,V\psi = c\,(V_A \otimes V_B)(\phi \otimes \chi) = c\,(V_A\phi \otimes V_B\chi) = c\,(\phi' \otimes \chi')$ which is also a product vector. To show that positivity is preserved we accept that ρ can be expanded in a suitable ensemble ψ_k

$$\rho = \sum_k p_k\,(\psi_k\psi_k^\dagger) \tag{3.65}$$

and then take the expectation value of ρ' in any vector ψ

$$\begin{aligned}\psi^\dagger\rho'\psi &= \psi^\dagger(a\,V\rho V^\dagger)\psi \\ &= a\sum_k p_k(\psi^\dagger V\psi_k)(\psi_k^\dagger V\psi) \\ &= a\sum_k p_k\big|\psi^\dagger V\psi_k\big|^2 \geq 0\end{aligned} \tag{3.66}$$

because both $a, p_k > 0$. It is possible to show that

$$\rho' = a\,V\rho V^\dagger \qquad \Rightarrow \qquad \rho'^P = a\,\widetilde{V}\rho^P\widetilde{V}^\dagger \tag{3.67}$$

if $V = V_A \otimes V_B$ and $\widetilde{V} = V_A \otimes V_B^*$. Since we showed above that SL $\times$ SL-transformations preserve positivity, it follows from (3.67) that if ρ is a PPT state, then so is ρ'. The range and kernel of ρ and ρ^P transform in the following ways

$$\begin{aligned}\operatorname{Img}\rho' &= V\operatorname{Img}\rho & \operatorname{Ker}\rho' &= (V^\dagger)^{-1}\operatorname{Ker}\rho \\ \operatorname{Img}\rho'^P &= \widetilde{V}\operatorname{Img}\rho^P & \operatorname{Ker}\rho'^P &= (\widetilde{V}^\dagger)^{-1}\operatorname{Ker}\rho^P\end{aligned} \tag{3.68}$$

with $V = V_A \otimes V_B$ and $\widetilde{V} = V_A \otimes V_B^*$. All these transformations are of product form, and together with the fact that a product vector $\psi = \phi \otimes \chi$ is transformed by $V = V_A \otimes V_B$ into a product vector $\psi' = c\,(\phi' \otimes \chi')$ with $c \in \mathbb{C}$, we see that the number of product vectors in a given subspace will remain the same.

Since SL-transformations of product type $V = V_A \otimes V_B$ preserve the number of product vectors in a subspace, we see that in the case that there exists a product transformation $\rho \mapsto \rho^P = a\,V\rho V^\dagger$, this transformation must transform the set of n_{img} product vectors in the range of ρ to the set of $\widetilde{n}_{\text{img}} = n_{\text{img}}$ product vectors in the range of ρ^P. If the product vectors in $\operatorname{Img}\rho$ is $w_i = u_i \otimes v_i$ with $i = 1, \ldots, n_{\text{img}}$ and if in $\operatorname{Img}\rho^P$ we have $\widetilde{w}_i = \widetilde{u}_i \otimes \widetilde{v}_i$ with $i = 1, \ldots, \widetilde{n}_{\text{img}}$, then $V_A\,u_i = \widetilde{u}_i$ and $V_B\,v_i = \widetilde{v}_i$

Note that we need SL-transformations V_A and V_B on each subsystem that transform *all* vectors u_i and v_i respectively. Since the understanding of the relation between w and $\widetilde{w}$ is quite limited for entangled states, it is difficult to say much in general about what makes some states SL-symmetric and others not. But it is clear that the vectors $w_i = u_i \otimes v_i$ which are fixed for a given range of ρ, and the vectors $\widetilde{w}_i = \widetilde{u}_i \otimes \widetilde{v}_i$ which depend on the specific state ρ (or ρ^P) must have a structure that allows the existence of V_A and V_B to satisfy $V_A\,u_i = \widetilde{u}_i$ and $V_B\,v_i = \widetilde{v}_i$. For a generic state ρ there will exist no such V_A and V_B.

3.7.2 SL-symmetric states

We can now introduce two further, and stronger, conditions on the SL-equivalence relations related to a state ρ.

Definition 3.8 (SL-symmetry)**.** *If a state ρ is* SL*-equivalent with ρ^P, i.e. there exists a product transformation $V = V_A \otimes V_B$ such that*

$$\rho \mapsto \rho' = a\, V\rho V^\dagger = \rho^P \tag{3.69}$$

we say that the state ρ is SL*-symmetric.*

In these special cases ρ and ρ^P belong to the same equivalence class with respect to SL $\times$ SL-transformations. We also introduce an even more stringent property, namely *genuine* SL-equivalence:

Definition 3.9 (Genuine SL-symmetry)**.** *A state ρ is genuinely* SL*-symmetric if there exists a transformation such that*

$$\rho \mapsto \rho' = U\rho U^\dagger \qquad\qquad \rho' = \rho'^P \tag{3.70}$$

with $U = U_A \otimes U_B$.

The transformation of ρ implies that

$$\rho'^P = a\, \widetilde{U}\rho^P \widetilde{U}^\dagger \tag{3.71}$$

when we define $\widetilde{U} = U_A \otimes U_B^*$. Then assuming genuine SL-symmetry we get that

$$\widetilde{U}\rho^P \widetilde{U}^\dagger = U\rho U^\dagger \tag{3.72}$$

and hence

$$\rho^P = V\rho V^\dagger \qquad \text{with} \qquad V = \widetilde{U}^{-1}U = I \otimes V_B \tag{3.73}$$

and with $V_B = (U_B^*)^{-1}U_B$. This shows that genuine SL-symmetry implies SL-symmetry.

Note that the relation $V_B = (U_B^*)^{-1}U_B$ implies that $V_B^* = V_B^{-1}$, so that V_B is unitary if and only if it is symmetric. Since in general $\operatorname{Tr}\rho^P = \operatorname{Tr}\rho$, by (3.73) we require that V preserves the trace of ρ,

$$\operatorname{Tr}(V\rho V^\dagger) = \operatorname{Tr}(\rho V^\dagger V) = \operatorname{Tr}\rho \tag{3.74}$$

A sufficient, but perhaps not necessary condition is that V_B is unitary.

We conclude that for the state ρ to be genuinely SL-symmetric it must be SL-symmetric with a transformation of the form given in (3.73). For the 3×3 system V would have the form

$$V = \begin{pmatrix} V_B & 0 & 0 \\ 0 & V_B & 0 \\ 0 & 0 & V_B \end{pmatrix} \tag{3.75}$$

with $V_B \in \mathrm{SL}(3,\mathbb{C})$ and $V_B^* = V_B^{-1}$.

Since the transformation has the form $V = I \otimes V_B$ in the case of genuine SL-symmetry, the product vectors in $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$ will be related by the transformations $\widetilde{u}_i = u_i$ and $\widetilde{v}_i = V_B v_i$. This is a necessary condition for genuine SL-symmetry which may be tested as soon as we know the product vectors in $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$.

Assume that for a given PPT state ρ we find that ρ and ρ^P are related by a transformation of the form given in (3.73). Then a further problem to be solved is to find a transformation $U = U_A \otimes U_B$ that demonstrates explicitly the genuine SL-symmetry of ρ. Thus we have to solve the equation $V_B = (U_B^*)^{-1} U_B$ for U_B. Assume that λ is an eigenvalue of U_B, then $\mu = \lambda/\lambda^*$ is an eigenvalue of V_B. Hence $|\mu| = 1$, and we may assume that $\lambda = \mathrm{e}^{\mathrm{i}\alpha}$ where α is real. Then λ must be a solution of the equation $\mu = \lambda^2$, suggesting that we may try to take U_B as the matrix square root of V_B. We find in practice that it is possible to choose simply

$$U = I \otimes \sqrt{V_B} \tag{3.76}$$

Chapter 4

Convex sets

Since the set of density matrices $\mathcal{D}$, and the subsets $\mathcal{S}$ and $\mathcal{P}$, are convex sets, it is important to develop the tools for describing such sets. Especially the geometric features of convex sets is important to us, such as extremal points, dual sets and faces.

4.1 Convex combinations

From a geometrical viewpoint a mixture of two points ρ_1 and ρ_2 in a convex set X, is a point on the line segment between the points. For the set to be convex we insist that all points on all line segments also belong to the set. We often describe a convex set as an *affine* space, which means that there is no choice of origin, or absolute reference point.

Definition 4.1 (Convex combination)**.** *The straight line through the points ρ_1 and ρ_2 defined by*

$$\rho = x_1\rho_1 + x_2\rho_2 \qquad x_1 + x_2 = 1 \qquad x_1, x_2 \geq 0 \tag{4.1}$$

is a convex combination of ρ_1 and ρ_2.

The case (4.1) extends recursively to convex combinations of more than two points, so more generally we may write a point ρ in X as a

$$\rho = \sum_{i=1}^{k} x_i\rho_i \tag{4.2}$$

where ρ_i are points in X. If we relax the demand that $x_i \geq 0$, the real numbers x_i in (4.2) are called *barycentric* coordinates. The demands $x_i \geq 0$ and $\sum x_i = 1$ are strongly related to the fact that we use convexity to describe probability distributions. The expectation value of a random variable Y relates to convexity. If we assume that Y is a discrete variable taking values in some finite set of real numbers y_i, with probabilities p_i of the event $Y = y_i$, the expectation value of Y is $EY = \sum_i p_i y_i$. So EY is a convex combination of the set y_i. In Figure 4.1 we see the difference between a convex and a non-convex (or concave) set. Intuitively a convex set is a set such that one can always see the entire set from any inside vantage point.

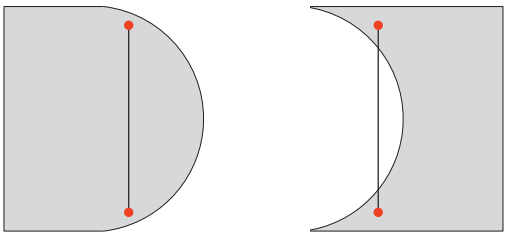


Figure 4.1: For a set X to be convex, we insist that all points on any line segment between points in X, also belong to the set.

Compact convex sets

A set X with a norm $|x|$ is *bounded* if there is a number $c \in \mathbb{R}$ such that $|\rho_1 - \rho_2| < c$ for any two points ρ_1 and ρ_2. Another important property is *closure*. In a topological space, a closed set can be defined as a set which contains all its limit points. But one may also use a norm again, and first define a boundary point of X as a point b where the set $\{y \,|\, |b - y| < \epsilon\}$ always contains an $y \notin X$ for all $\epsilon > 0$. The boundary of X is then the set of all such points b, and a closed set X will always contain its boundary, which is sometimes denoted by ∂X. All these notions demand that there is a space of similar objects outside the convex set, this space is sometimes referred to as the *container space* of the compact set X. For almost all our purposes this container space will be the space of Hermitian matrices, which is unbounded.

4.2 Extremal points

Any given set of points S can define a convex set. The *convex hull* of a set S is the set of all convex combinations of points in S. Thus if $S = \{\rho_1, \rho_2\}$, then the convex hull of S is the line segment (4.1). The convex hull of any set S is therefore always a convex set. Another, but equivalent way, is to define the convex hull of S as the smallest convex set that contains S. The convex set of a finite set of points is called a polytope. A very important concept here is *extremal points*:

Definition 4.2 (Extremal point)**.** *An extremal point of a convex set X is a point ρ that cannot be written as a convex combination of any other two points in the set.*

The hypercube $C_\nu = \{x \in \mathbb{R}^\nu \,|\, |x_i| \leq 1\}$ has 2^ν extremal points, namely the "corners" $(\pm 1, \pm 1, \ldots, \pm 1)$. The hyperball $B_\nu = \{x \in \mathbb{R}^\nu \,|\, |x| \leq 1\}$ has the entire spherical shell $|x| = 1$ as extremal points, which is an infinite set. We define the *convex rank* of a point ρ in a convex set as the minimal number of extremal points required in the convex combination (4.2). An extremal point of a convex set X thus always has convex rank one. The rank of a matrix, should not immediately be confused with the convex

rank in a set X. We shall see that in the set of density matrices $\mathcal{D}$, the convex rank is always equal to the matrix rank, but we shall also encounter matrices of higher matrix rank than one, which are extremal on the convex set $\mathcal{P}$.

A very essential and practical feature of convex sets is that any point in the set can be written as a convex combination of a number of extremal points. This is known as Carathéodorys theorem [48]. It can be formulated as

Theorem 4.1 (Carathéodory)**.** *If X is a convex set of dimension n, then any point $\rho \in X$ can be expressed as a convex combination of at most $n+1$ extremal points in X.*

The original theorem by Carathéodory do not use extremal points explicitly, but because of a theorem by Minkowski which states that any convex body is the convex hull of its extremal points, Theorem 4.1 can be formulated as above. Therefore, if we can identify the extremal points of a convex set, we essentially know the entire set. Further, a convex sum of k points all of convex rank one, can of course be made into a sum of points of higher rank, containing a smaller number of terms. The fact that we generally need $n+1$ extremal points for an n-dimensional set, is a consequence of our reluctance to define an origin. It should be noted that the extremal points of a convex set do not in any way form a convex basis, in the sense that the same set of $n+1$ points can be used to generate any point ρ.

There is an alternative way of defining convex sets rather than using extremal points. This is by using inequalities to define the interior points of the set, with equality on the boundary. This is especially useful for polytopes. For instance, it turns out to be rather easy to define both extremal points and inequalities for the full set of density matrices $\mathcal{D}$, while for the set of separable states $\mathcal{S}$ the extremal points are well known, while inequalities are difficult to obtain.

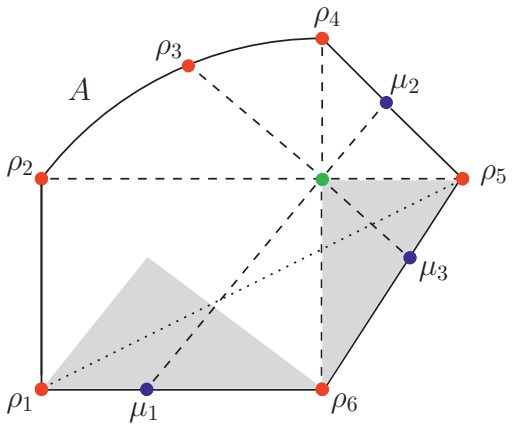


Figure 4.2: A convex set X, where extremal points are represented in red. All three blue points on the boundary ∂X and the green inner point have rank two. All points on the curved boundary segment A between ρ_2 and ρ_4 are extremal. The only points of rank three are those in the shaded areas, with the exception of those on the line from ρ_1 to ρ_5.

In Figure 4.2 we see a convex set, where the points in red are extremal points. In fact

all points along the curved boundary segment A, are extremal points. All the extremal points have convex rank one. Obviously all extremal points of a convex set X lies on the boundary ∂X, but not all points on ∂X are extremal points. All the blue points represent points that are convex combinations of two extremal points. If we observe the green point, it clearly can be constructed in several ways, by points of different ranks. We see, by the dashed lines, that it can be constructed by two extremal points (ρ_2, ρ_5), or three (ρ_3, μ_3) and four (μ_1, μ_2). But since we define the convex rank of a point ρ as the minimal number of extremal points required in the convex combination, the green point actually also has rank two, the same as the blue points. We observe that another possibility is (ρ_4, ρ_6), which also are two extremal points, showing that even if we restrict ourselves to the minimum number of extremal points, there are ambiguities in the construction of points.

It is clear that points in the (elliptic) arc like region defined by A and ρ_6 can be written as a convex combination of ρ_6 and any point on A. The same applies for the arc regions defined by A and ρ_1, and by A and ρ_5. Another such area is the one defined by A and the line segment from ρ_2 to ρ_4, where any point may be constructed from two points on A. On closer inspection, we see that the points in the grey triangles, with the exception of the ones on the line from ρ_1 to ρ_5, are the only ones of rank three.

The main point of this, is that we realize that even for a simple two-dimensional figure, the convex geometrical structure can be somewhat involved. For systems of higher dimensions, the complexity of the convex geometrical structures increases rapidly.

4.3 Convex and dual cones

If the demand $\sum_i x_i = 1$ that the convex coefficients x_i in (4.2) be summed to one is relaxed, we are able to define another construction, namely that of a *convex cone*.

Definition 4.3 (Convex cone)**.** *A set C is called a convex cone if $x_1\rho_1 + x_2\rho_2 \in C$ whenever $\rho_1, \rho_2 \in C$ and $x_1, x_2 \geq 0$.*

The zero point O is also contained in this set since $x_1 = x_2 = 0$ is a possibility. In addition to convex combinations, a convex cone is closed under multiplication by a non-negative real scalar, *i.e.* if $\rho \in C$ and $\alpha \in \mathbb{R}_+$, then $\alpha\rho \in C$, and the set $\alpha\rho$ is called a *ray* in the convex cone C, and is merely the real and positive multiples of the state ρ. We may denote the slice, or segment, of the cone corresponding to $\sum_{i=1}^{k} x_i = \alpha$ as C_α. Since α is arbitrarily high, the convex cone is unbounded upwards, but limited downwards by the zero point that corresponds to C_0. We find the entire structure in the convex set for which $\sum_{i=1}^{k} x_i = \alpha_0$ repeated for other values of α, so to study the structures of a convex set it is usually ample to restrict to the C_1 case.

The study of convex cones is nevertheless very useful, since many operations that do not preserve the norm take us into other areas of the cone, where $\alpha \neq 1$. Another important reason why we extend to the convex cone structure, is the very useful concept of *dual cones*. The concept of duality is equally valid for other sets, but it is easier to visualize if we define it for cones.

Definition 4.4 (Dual cone)**.** *The dual cone to the convex cone X is*

$$X^\circ = \{y \,|\, \langle x, y\rangle \geq 0 \;\; \forall\, x \in X\} \tag{4.3}$$

We observe that $(X^\circ)^\circ = X$, and that $\theta = \pi/2 \Leftrightarrow X = X^\circ$. If $X = X^\circ$ the convex cone is self dual. Since the inner product $\langle x, y\rangle$ can be given a geometrical interpretation by $\langle x, y\rangle = \cos\theta$, this can be visualized as in Figure 4.3.

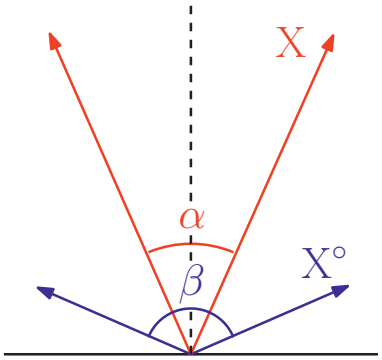


Figure 4.3: The cone angle for X is α, and likewise for X° is β. The angle represents the width or size of the convex cones.

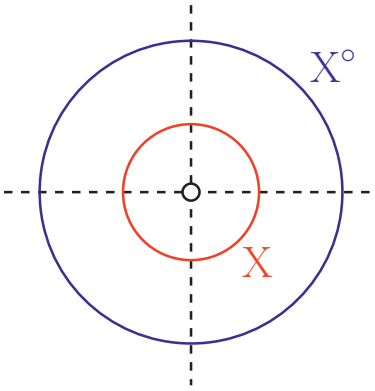


Figure 4.4: The cones from Figure 4.3 seen from above through a cut or section. It appears quite clear that $X \subset X^\circ$.

It must be remembered that the cones representing the density matrices of actual quantum systems are hypercones with a large number of dimensions, and a very complex structure. For instance, the set of density matrices $\mathcal{D}$ for the 3×3 system is an 80-dimensional space, and any two-dimensional cut or segment through this set will in general be different from the two concentric circles in Figure 4.4.

4.4 Faces of convex sets

We will here introduce a very important concept in connection with convex sets, namely that of a face. Further on, we will look at a special class of faces, known as exposed faces.

Definition 4.5 (Face)**.** *A face of a convex set X is a subset $\mathcal{F} \subseteq X$ such that if*

$$\rho = x_1\rho_1 + x_2\rho_2 \qquad x_1 + x_2 = 1 \qquad x_1, x_2 \geq 0 \tag{4.4}$$

then $\rho \in \mathcal{F}$ if and only if $\rho_1, \rho_2 \in \mathcal{F}$.

One might say that a face $\mathcal{F}$ is a convex subset of X which is convexly closed, or stable under convex mixing and purification. Since the set X is also formally a face on X, we say that any face $\mathcal{F}$ that is strictly smaller than X is a proper face. In Figure

4.5 we see this visualized. The convex set X consists of the points inside and on the square and the semicircle S. Since the only way of constructing points on the edges $\rho_1\rho_2$, $\rho_1\rho_4$ and $\rho_2\rho_3$ can be done by using points on these edges, they are faces of X. But the line $\rho_1\rho_3$ is not a face on X, because any point on $\rho_1\rho_3$ can also be written by using the two points μ_1 and μ_2, which are not on $\rho_1\rho_3$. The dotted line $\rho_3\rho_4$ is also obviously not a face on X, for the same reason as $\rho_1\rho_3$, but it is a face if S is omitted, *i.e.* of the square. If $|X| = n$ then any face of dimension $n-1$ is called a *facet*, while a one-dimensional face is an *edge*. All the points along the semicircle S in Figure 4.5 are faces of dimension zero, or 0-faces, and this is also the case for all the points ρ_i. In fact an extremal point for a convex set is by definition a face of dimension zero.

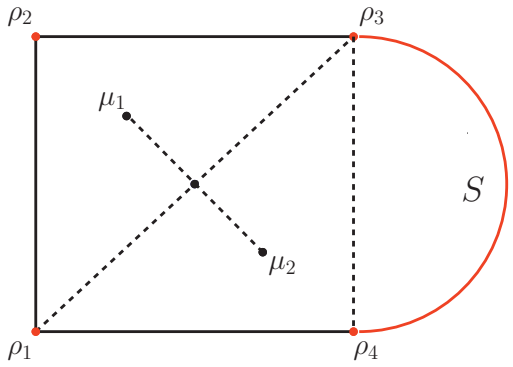


Figure 4.5: The convex set X consists of the points inside the square and the red semicircle S. The extremal points or 0-faces of X are the ρ_i and all points on S, and the additional faces are the edges $\rho_1\rho_2$, $\rho_1\rho_4$ and $\rho_2\rho_3$.

The hypercube C_ν has $3^\nu - 1$ faces, namely 2^ν extremal points, $\nu 2^{\nu-1}$ edges, $\binom{2}{\nu}2^{\nu-2}$ facial planes, all the way to $2\binom{\nu-1}{\nu}$ facets of dimension $\nu - 1$. The only faces on the hyperball B_ν are its extremal points.

A face $\mathcal{F}$ of a convex set X always lies on the topological boundary of X, and also any point on this boundary will always lie in at least one face $\mathcal{F} \subset X$. Also, if we assume that a face $\mathcal{F} \subset X$, and let another set $Y \subset \mathcal{F}$, then Y is a face of $\mathcal{F}$ if and only if it is a face of X. In this way we realize that the faces of a convex set X form hierarchies, in the sense that if $\mathcal{F}_{n-1}$ is a face on $\mathcal{F}_n$, and $\mathcal{F}_{n-2}$ is a face on $\mathcal{F}_{n-1}$, then $\mathcal{F}_{n-2}$ is a face on $\mathcal{F}_n$. Since a face $\mathcal{F}$ is a convex set in its own right, it will have extremal points. The extremal points in a hierarchy of faces are hereditary, so that the extremal points of a face $\mathcal{F}$ are extremal points in all subfaces of $\mathcal{F}$. Since the largest face on any convex set is the set itself, we understand that the extremal points on any proper face $\mathcal{F} \subset X$ are extremal points on X. So for any convex set we are able to descend a sequence of faces

$$X \supset \mathcal{F}_n \supset \mathcal{F}_{n-1} \supset \ldots \supset \mathcal{F}_1 \supset \mathcal{F}_0 \tag{4.5}$$

where $|\mathcal{F}_0| = 0$, so we end up in an extremal point of X. Theorem 4.1 by Carathéodory can be justified and understood by appealing to the structure of faces (4.5).

Exposed faces

Consider the sphere $B_3 = \{x \in \mathbb{R}^3 \,|\, |x| \leq 1\}$. Through a boundary point of B_3, *i.e.* a point with $|x| = 1$, we can place a tangent plane H. It consists of the points x satisfying $\langle a, x\rangle = 1$, where a is the boundary point. We say that H supports B_3 at $x = a$, meaning that all of B_3 lie on the same side of the plane, as all points in the sphere satisfy $\langle a, x\rangle \leq 1$.

It is sometimes easier to use the Euclidean vector space $\mathbb{R}^n$ rather than an affine description. We here utilize this to define supporting hyperplanes. We define the set $H = \{x \in \mathbb{R}^n \,|\, \langle a, x\rangle = c\}$, where $a \in \mathbb{R}^n$ and $c \in \mathbb{R}$, to be a hyperplane in $\mathbb{R}^n$ of dimension $n - 1$. We get parallel hyperplanes for various values of c. Each hyperplane divides the space into two halfspaces $H^+ = \{x \in \mathbb{R}^n \,|\, \langle a, x\rangle \geq c\}$ and $H^- = \{x \in \mathbb{R}^n \,|\, \langle a, x\rangle \leq c\}$. If now $X \subset \mathbb{R}^n$ is a convex set and is completely contained in one of the two half-spaces, and in addition $H \cap X$ is non-empty, we say that H is a *supporting hyperplane* of X, and that it supports X at $x \in H \cap X$. So a supporting hyperplane of X intersects the set only at the boundary ∂X. One can prove that there is a supporting hyperplane passing through every point on the boundary of a convex set, and that only convex sets has this property. So in fact supporting hyperplanes can be used to define convex sets.

We now use hyperplanes to introduce another distinction, that of an *exposed* face of a convex set. This is related to the notion of hierarchies of faces.

Definition 4.6 (Exposed face)**.** *Let H be a supporting hyperplane of the convex set X, then any intersection $H \cap X$ is an exposed face of X.*

Another way of formulating this is to say that a face $\mathcal{F}$ is an exposed face if there exists a supporting hyperplane H such that $\mathcal{F} \cap H = \mathcal{F}$. So the hyperplane cannot contain any points in X that is not in the face $\mathcal{F}$.

In Figure 4.6 the convex set X has essentially the same structure as in Figure 4.5, but with the addition of an alternative dotted elliptic boundary segment E.

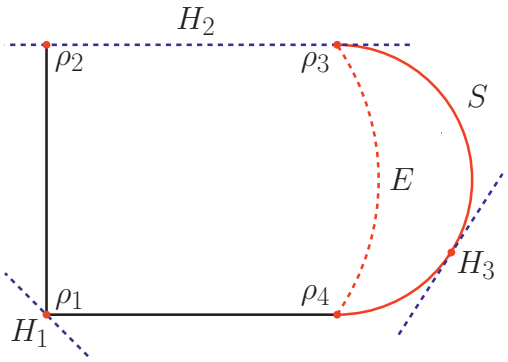


Figure 4.6: Three hyperplanes supporting three exposed faces: The edge $\rho_2\rho_3$, the extremal point ρ_1 and another extremal point on S. The points ρ_3 and ρ_4 are extremal points, *i.e.* 0-faces. If the curved part of the boundary of X is defined by S, they are not exposed faces. If however E defines the boundary, then ρ_3 and ρ_4 become exposed faces.

We see three hyperplanes H_1, H_2 and H_3 supporting three faces. In all three cases the faces are exposed, since the hyperplanes only intersect points in X that is contained in the faces. The point ρ_3 is an extremal point, *i.e.* a 0-face, but with S defining the boundary of X it is not an exposed face, because any supporting hyperplane in ρ_3 will also contain points in the edge face $\rho_2\rho_3$ or interior points of X. With E defining ∂X, the boundary is non-smooth in ρ_3, so that a supporting hyperplane which contains only ρ_3 can be constructed, *i.e.* ρ_3 is an exposed face of X.

It appears that unexposed faces are related to transitions, in one of several possible directions, where a line segment (of zero curvature) passes into and a curved segment, and where this transition is smooth.

4.5 Simplexes

A *simplex* Δ^n can be defined as the convex hull of a set of $n+1$ affinely independent points (or vectors) in $\mathbb{R}^n$. By Theorem 4.1 Δ^n will then be an n-dimensional set, and a polytope. More formally, suppose that we have $n+1$ affinely independent points $\rho_i \in \mathbb{R}^n$ where $i = 1, \ldots, n+1$. Then the simplex defined by them is the set of points

$$\Delta^n = \left\{ \sum_{i=1}^{n+1} x_i \rho_i \,\middle|\, x_i \geq 0, \sum_{i=1}^{n+1} x_i = 1 \right\} \tag{4.6}$$

In this way a 0-simplex is a point ρ_1, a 1-simplex is the line segment between two points ρ_1 and ρ_2, and a 2-simplex is the triangle of convex combinations of three points ρ_1, ρ_2 and ρ_3. The coefficients x_i are then convex coordinates defining a point in the simplex.

Geometrically a simplex Δ^n is the generalization of the triangle from two dimensions to n. In Figure 4.7 we see a 3-simplex with four affinely independent points. The 3-simplex Δ^3 has four facets (faces of dimension two) which are triangles, six edges (faces of dimension one) and four extremal points (faces of dimension zero).

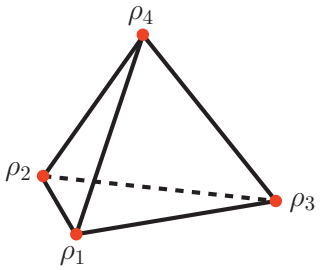


Figure 4.7: A 3-simplex Δ^3 with four facets, six edges and four extremal points. The 3-simplex is an extension of the triangle (or 2-simplex) by one dimension.

The triangle Δ^2 is a convex set which has no surplus extremal points, by way of affine independence. Since it is two-dimensional it can have maximally three affinely independent extremal points. All simplexes Δ^n have this property. Two examples of

convex sets that do not exhibit this property is the square and the circle, with four and infinitely many extremal points, but both two-dimensional. Another remarkable property of a simplex is that the lattice of faces is self dual, meaning that the set of $(n-k-1)$-dimensional faces is in one-to-one correspondence with set of k-dimensional faces. Specifically the extremal points $k = 0$ correspond to the set of facets, or $(n-1)$-dimensional faces.

Chapter 5

Geometry of density matrices

Building on Chapters 3 and 4, we now discuss the geometric properties of the set of density matrices $\mathcal{D}$, and some of the subsets, primarily $\mathcal{S}$ and $\mathcal{P}$. We illustrate the geometric properties and relations between these sets by constructing two-dimensional cross sections. A method for investigating the geometrical structure of $\mathcal{P}$ is to make restricted perturbations within the set. These perturbations can be used to construct an extremality test for $\mathcal{P}$, and to investigate substructures in $\mathcal{P}$. We also discuss the faces of the sets $\mathcal{D}$, $\mathcal{P}$ and $\mathcal{S}$, and to some extent the properties of the dual sets

5.1 Density matrices as simplexes

A density matrix acting on a Hilbert space $\mathcal{H}$ of dimension N, is characterized by N^2-1 free parameters. Since all Hermitian matrices are unitary diagonalizable, a general state ρ can be described by an $N \times N$ diagonal matrix D and a unitary transformation of determinant one. The maximum number of free parameters in a normalized form of D is $N-1$, and the dimension of $SU(N)$ is N^2-1, but the dimension of the subset in $SU(N)$ that commutes with diagonal matrices, and therefore do not change the diagonal form of D, is $N-1$. Using the eigenvectors e_k and eigenvalues λ_k of a density matrix ρ we can, according to Theorem 3.3, write

$$\rho = D = \sum_{k=1}^{r} \lambda_k \, (e_k e_k^\dagger) \tag{5.1}$$

where $r = |\mathrm{Img}\, \rho|$ is the rank of ρ. The e_k that correspond to eigenvalues $\lambda_k > 0$ span the range of ρ, and the e_k for which $\lambda_k = 0$ span the kernel of ρ. We have $\mathcal{H} = \mathrm{Img}\, \rho \oplus \mathrm{Ker}\, \rho$, $\mathrm{Img}\, \rho \perp \mathrm{Ker}\, \rho$ and thereby $N = r + |\mathrm{Ker}\, \rho|$.

Any density matrix ρ which is diagonal in e_k can be written as (5.1), but of course with different λ_k for different ρ. We clearly see that since all $\lambda_k \geq 0$ and $\sum_k \lambda_k = 1$, the set $e_k e_k^\dagger$ are vertices in a simplex, and the λ_k are convex coefficients that define density matrices lying in the simplex. All matrices defined as in (5.1), *i.e.* with the same set e_k, will then have the same range $\mathrm{Img}\, \rho$ and kernel $\mathrm{Ker}\, \rho$. The simplex defined above is the eigenensemble or the eigenvalue simplex of ρ, which also means that the set e_k will be an orthogonal set.

Any state ρ with rank r will then be found within an eigenensemble $(r-1)$-simplex, which is a special $(r-1)$-dimensional subset of $\mathcal{D}$. There are also simplexes with r non-orthogonal, but linearly independent vertices that can build the same state, but the projections will then mix the different dimensions contained in the range, so the values λ_k in the convex sum will not be the eigenvalues of ρ.

Such a simplex is a collection of many states built from the vertices. For example a rank five state will be found in the interior of a four-dimensional simplex Δ^4. States on the different faces of Δ^4 will then be states of lesser ranks. The vertices, which are the extremal points of the simplex will have rank one, the edges will contain rank two states and the facets will be made up of rank four states.

In Section 4.2 we defined the rank of a point in a convex set as the minimum number of extremal points needed in the convex combination (3.5). We observe that a Hermitian matrix of matrix rank r can be written as a convex sum of no less than r orthogonal projections onto the basis of $\mathcal{H}$, with the $r \leq N$ projections being onto $\operatorname{Img}\rho$. Hence the convex rank r of a density matrix in the convex set $\mathcal{D}$, is equal to the matrix rank of $\rho \in H_N$. The maximal convex rank of a mixed state in $\mathcal{D}_N$ is thus N, which is much less than the upper bound N^2 from Carathéodorys theorem. This is due to the fact that any density matrix ρ can be reached by unitary transformations from a diagonal density matrix D.

It is possible to construct the entire set of density matrices $\mathcal{D}_N$ by using $(N-1)$-simplexes. For example any density matrix $\rho \in \mathcal{D}_9$ can be obtained by special unitary transformations

$$\rho' \mapsto \rho = U\rho' U^\dagger \tag{5.2}$$

where $U \in \mathrm{SU}(9)$ and $\rho' \in \Delta^8$. Since the dimensions taken by Δ^8 is eight, there still remains 72 dimensions that we can perform unitary rotations in. So, any density matrix ρ of any rank will be found in at least one unitarily rotated simplex with extremal points $U(e_k e_k^\dagger)U^\dagger$, with e_k as a linearly independent ensemble in $\mathcal{H}$.

5.2 The convex sets $\mathcal{D}$, $\mathcal{S}$ and $\mathcal{P}$

The set of density matrices $\mathcal{D}$, the set of separable density matrices $\mathcal{S}$ and the Peres set $\mathcal{P}$ of density matrices with positive partial transpose, all manifest a convex structure. This is of fundamental importance to the study of the geometric properties of these sets.

The set of density matrices $\mathcal{D}$

That the set of density matrices $\mathcal{D}$ is a convex set follows more or less from Definition 3.1, as well as from (5.1). With $\psi_k \in \mathcal{H}$ we repeat (3.5)

$$\rho = \sum_k p_k \left(\psi_k \psi_k^\dagger\right) \qquad p_k > 0 \qquad \sum_k p_k = 1 \tag{5.3}$$

Since the sum above always can be constructed as a convex sum over pure states, it also follows that the extreme points of $\mathcal{D}$ are the pure states $\psi\psi^\dagger$. These pure states include both the product states $\psi = \phi \otimes \chi$ and the non-product states $\psi \neq \phi \otimes \chi$. The pure states are the only ones for which $\mathrm{tr}(\rho) = \mathrm{tr}(\rho^2)$, and these states have the maximum distance within $\mathcal{D}$, from the maximally mixed state I/N. Using the Hilbert-Schmidt metric (3.8) we get $|\psi\psi^\dagger - I/N| = \sqrt{N - 1/N}$.

We should expect that the dimension $N^2 - 1$ of the full set $\mathcal{D}$ is larger than the dimension of the pure states. A pure state is a one-dimensional projection onto a normalized vector $\psi_k \in \mathbb{C}^N$ with N complex components, which again gives $2N$ free real parameters. The normalization demand $\psi_k^\dagger \psi_k = 1$ removes one real factor, and the fact that we can multiply ψ with a phase factor $\psi \mapsto e^{i\alpha_k}\psi_k$ such that $\psi_k^\dagger \psi_k$ remains unchanged, removes another. So in total the dimension of the pure states is $2N - 2 < N^2 - 1$ for all $N > 1$. The boundary $\partial\mathcal{D}_N$ is always by definition $(N^2 - 2)$-dimensional. This means that for $N = 2$ the dimension of the boundary is equal to the dimension of the pure states, so the entire boundary is then pure states. To illustrate this, we observe that a density matrix in $\mathcal{D}_2$ has the form

$$\rho = \frac{1}{2}\begin{pmatrix} 1+z & x - \mathrm{i}y \\ x + \mathrm{i}y & 1 - z \end{pmatrix} \tag{5.4}$$

with $x, y, z \in \mathbb{R}$ and $x^2 + y^2 + z^2 \leq 1$. Thus $\mathcal{D}_2$ is a three-dimensional sphere, the Block sphere. The boundary states, with $x^2 + y^2 + z^2 = 1$, are the pure states. For $N > 2$ the boundary of $\mathcal{D}$ is no longer a sphere, but has a more complex structure. For instance for $N = 3$ we have $|\mathcal{D}_3| = 8$, and its boundary $\partial\mathcal{D}_3$ consists of a seven-dimensional set of rank two matrices and a four-dimensional set of pure states. Every state of rank two in $\mathcal{D}_3$ then lies in the interior of some Bloch sphere, which again has a boundary of pure states. This Bloch sphere is a three-dimensional face on $\mathcal{D}_3$, where the boundary points, or pure states, are extremal points for both this Bloch sphere and $\mathcal{D}_3$.

The traditional way of determining the boundary of $\mathcal{D}_N$ is by appealing to inequalities for the eigenvalues. For a density matrix $\rho \in \mathcal{D}_N$ we have effectively N inequalities $\lambda_k \geq 0$ for the eigenvalues λ_k of ρ. The boundary of $\mathcal{D}_N$ can then be defined as all the states ρ that satisfy all the N inequalities $\lambda_k \geq 0$, but for which at least one is satisfied with equality. So in effect, all the states $\rho \in \partial\mathcal{D}$ have $\det(\rho) = 0$, and for all the interior points we have $\det(\rho) > 0$. The description of $\partial\mathcal{D}$ by the set of inequalities $\lambda_k \geq 0$ relies on obtaining solutions of a complex Nth degree polynomial, but this can generally be done in a fast and effective way numerically.

The set of separable states $\mathcal{S}$

Just as for the set of density matrices $\mathcal{D}$ we can see from (3.28) that the set of separable states $\mathcal{S}$ is a convex set, where the pure extremal states are replaced by a subset, namely the pure product states. Thus the only extremal states for the set of separable states are the pure product states, *i.e.* one-dimensional projections onto a normalized product vector $\psi = \phi \otimes \chi \in \mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$.

For a general product vector $\psi = \phi \otimes \chi$, the vector ϕ has $2N_A$ real elements, and

likewise $2N_B$ for χ, in total $2(N_A + N_B)$. If we take both the vectors ϕ and χ to be normalized and in addition remove a phase from each, we are left with $2N_A + 2N_B - 4$ free parameters. It is a fact that $2N - 2 = 2N_AN_B - 2 > 2(N_A + N_B) - 4$ for all choices of $N_A, N_B \geq 2$. This indicates that the set of pure product states have a smaller dimension than the total set of pure states, which is compatible with the fact that a general Hilbert space vector $\psi \in \mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ will not be a product vector, and thus the pure state $\rho = \psi\psi^\dagger$ will not be a pure product state. In order for ψ to be a product vector its components must satisfy (3.52), and generically this is not the case.

It is however a fact that the dimension of the set of separable states $\mathcal{S}$ is the same as the set of density matrices $\mathcal{D}$. The easiest way to understand this is by looking at the *ball of separable states* around the maximally mixed state I/N, as defined by Gurvits and Barnum [49]. It is known that in the space of density matrices $\mathcal{D}_N$ there exists a hyperball B_{N^2-1} centered in I/N, such that all states within a certain distance $r_{\max}$, *i.e.* states that satisfies $|\rho - I/N| \leq r_{\max}$, are separable. In the Hilbert-Schmidt metric we have $r_{\max} = 1/\sqrt{N(N-1)}$. Since the dimension of the ball is $N^2 - 1$, which equals the dimension of $\mathcal{D}$, we see that $|\mathcal{S}| = |\mathcal{D}|$.

In our attempts to investigate the separable states, and thereby also the entangled states, we are almost entirely dependent on the extremal points of $\mathcal{S}$, *i.e.* the pure product states. The main reason is that it is not possible, or at least extremely difficult, to characterize the separable states according to the eigenvalues of the states.

It should be noted that even though the set of separable states is a convex set, this is not the case for the set of entangled states. A convex combination of two (or more) separable states always returns another separable state, since the product structure is preserved in the convex sum. But convex combinations of entangled states do not necessarily produce an entangled state. This is illustrated for the 2×2 system in (3.39), where two Bell states are mixed to produce a separable state, as an example of decoherence.

The Peres set $\mathcal{P}$

The set of partially transposed matrices $\mathcal{D}^P$ is obtained by taking the partial transpose of every point in $\mathcal{D}$. Since the partial transpose operator is a linear map, and by using (5.3), we see that the partial transpose of a state $\rho = \sum_k p_k\rho_k = \sum_k p_k\,(\psi_k\psi_k^\dagger) \in \mathcal{D}$ is a convex combination of partially transposed pure states

$$\rho^P = \sum_{k=1}^{r} p_k\,\rho_k^P = \sum_{k=1}^{r} p_k\,(\psi_k\psi_k^\dagger)^P \tag{5.5}$$

where $r = \mathrm{rank}(\rho)$. The pure states ρ_k are extremal in $\mathcal{D}$, but in general $\rho_k^P \notin \mathcal{D}$. The state ρ_k^P will in any case lie in the set $\mathcal{D}^P$, and since any convex combination of these ρ_k^P also lie in $\mathcal{D}^P$, we see that the set $\mathcal{D}^P$ of partially transposed density matrices is a convex set. For pure product states $\rho_k = \phi_k\phi_k^\dagger \otimes \chi_k\chi_k^\dagger$ we have $\rho_k^P = \phi_k\phi_k^\dagger \otimes (\chi_k\chi_k^\dagger)^T$, which is again a pure product state, and thus extremal in $\mathcal{S}$ and $\mathcal{D}$.

We have seen in Section 3.5.2 that the Peres set, or the set of density matrices that remain positive under partial transposition, can be understood as the intersection

$\mathcal{P} = \mathcal{D} \cap \mathcal{D}^P$. Since the intersection of two convex sets is always again a convex set, we realize that the Peres set $\mathcal{P}$ is also a convex set.

We know the extremal points of the set $\mathcal{D}$ to be the set of pure states $\psi\psi^\dagger$, and the extremal points of $\mathcal{S}$ to be the set of pure product states $\psi\psi^\dagger$ with $\psi = \phi \otimes \chi$. Since $\mathcal{S} \subseteq \mathcal{P}$ with equality for the 2×2 and 2×3 systems, all the extremal points of $\mathcal{S}$ must also be extremal points of $\mathcal{P}$. But we also know that for higher dimensions $N \geq 8$ entangled PPT states exist, so in order to construct convex combinations that represent these entangled PPT states, there must be a further set of extremal points for $\mathcal{P}$. Since convex combinations of product states only return separable states, these additional extremal points of $\mathcal{P}$ must be entangled.

5.3 The positive convex cone

The concept of convex cones is given in Definition 4.3 in Section 4.3. The set of normalized $N \times N$ Hermitian matrices H_N has the same dimension as the set of density matrices $\mathcal{D}$, namely $N^2 - 1$. The difference is that the set $\mathcal{D}$ is a compact set, as defined in Section 4.1. The compactness is entirely due to the demands that $\rho \geq 0$ and $\mathrm{tr}(\rho) = 1$. If the latter demand is cancelled, we get a convex cone of rays $z\rho$, where $\rho \in \mathcal{D}$. The set of all true density matrices with $\mathrm{tr}(\rho) = 1$ is the intersection in the container space of Hermitian matrices H_N, of the positive convex cone with a hyperplane parallel to the subspace of traceless Hermitian operators. We see this in Figure 5.1.

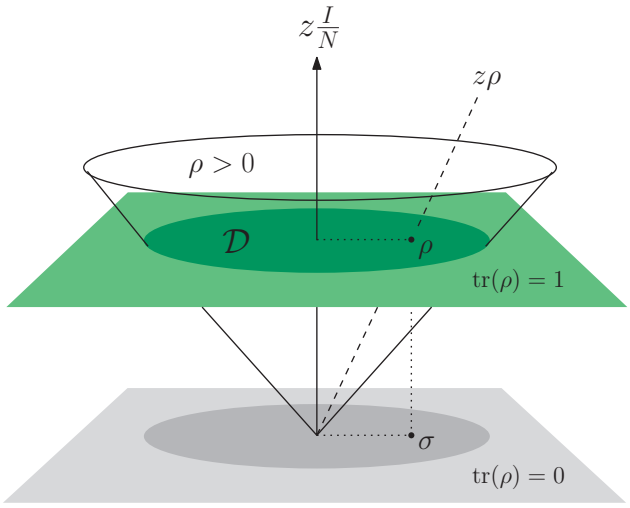


Figure 5.1: The set of all true density matrices with $\mathrm{tr}(\rho) = 1$ is the intersection in the space of Hermitian matrices H_N, of the positive convex cone with a hyperplane parallel to the subspace of traceless Hermitian operators. The cone consists of rays $z\rho$ where $\rho \in \mathcal{D}$. Any density matrix ρ can be projected onto the plane of traceless Hermitian operators, and uniquely represented by a Hermitian matrix σ with $\mathrm{tr}(\sigma) = 0$.

The inside of the cone consists of all Hermitian matrices with $\det(\rho) > 0$ and all eigenvalues $\lambda_k > 0$, while the edge of the cone is all Hermitian matrices with $\det(\rho) = 0$ and all eigenvalues $\lambda_k \geq 0$. In the center of the cone lies multipla of the identity $(z/N)I$, with the maximally mixed state I/N as the center point in the set of true density matrices $\mathcal{D}$.

The set of normalized separable states $\mathcal{S}$ and normalized PPT states $\mathcal{P}$ also define convex cones, which are then nested inside the full convex cone of density matrices $\mathcal{D}$. It is also possible to define the dual cones of the sets $\mathcal{D}$, $\mathcal{S}$ and $\mathcal{P}$, along the lines in Definition 4.4 and presented in Figures 4.3 and 4.4. This will be treated in some detail in Chapter 7.

5.4 Two-dimensional cross sections

Because of the rapid increase in the dimension $N^2 - 1$ of $\mathcal{D}_N$, as the dimension $|\mathcal{H}| = N$ of the underlying Hilbert space grows, the set of density matrices is in general difficult to visualize. For any set of points, once we have three points in the set, we can define a two-dimensional cut through the set. If we choose three density matrices ρ_0, ρ_1 and ρ_2, where usually $\rho_0 = I/N$, then the convex sum

$$\rho = \frac{I}{N} + p_1\rho_1 + p_2\rho_2 \tag{5.6}$$

is not a density matrix, since unless $p_1 = -p_2$ we have $\mathrm{tr}(\rho) > 1$. A way of mastering this is to write density matrices as convex combinations of I/N and a set of traceless Hermitian matrices. This set of traceless Hermitian matrices is the hyperplane depicted in Figure 5.1 as the grey plane (floor) in the convex cone. Since the set of Hermitian matrices has dimension N^2 we can write an arbitrary matrix $C \in H_N$ as

$$C = \sum_{k=1}^{N^2} c_k M_k \tag{5.7}$$

where $c_k \in \mathbb{R}$, and the matrices $M_k \in H_N$ are linearly independent. We can think of c_k as elements of an N^2-dimensional real vector in $\mathbb{R}^{N^2}$, so the Hermitian matrices of dimension $N \times N$ define a real N^2-dimensional vector space. We then perform additions, multiplications and inner products of the Hermitian matrices A and B by performing operations on their real representations in $\mathbb{R}^{N^2}$. Using this strategy, we now represent a density matrix ρ by

$$\rho = \frac{I}{N} + \sum_{k=1}^{N^2-1} d_k D_k \equiv \frac{I}{N} + \sigma(x) \tag{5.8}$$

where the D_k are a basis for the set of traceless Hermitian matrices, and the real coefficients d_k are chosen so that $\rho \geq 0$. This set of traceless $N \times N$ matrices form a real vector space of dimension $N^2 - 1$, *i.e.* it is closed under addition and multiplication. Since all information sits in the matrix σ, this procedure entirely defines the set $\mathcal{D}_N$.

We usually choose $\rho_0 = I/N$ to be the origin, and let the two states ρ_1 and ρ_2 define two directions in the two-dimensional section by the traceless matrices

$$\sigma_1 = a(\rho_1 - \rho_0) \qquad \sigma_2 = b(\rho_2 - \rho_0) \tag{5.9}$$

where $a, b \in \mathbb{R}$ are scaling parameters. A matrix

$$\rho = \rho_0 + x\sigma_1 + y\sigma_2 \tag{5.10}$$

where $x, y \in \mathbb{R}$, will then lie in the two-dimensional section of H_N that contains ρ_0, ρ_1 and ρ_2. A compact subset of this xy-plane will then give $\rho \geq 0$, so to make $\rho \in \mathcal{D}$. The boundary of $\mathcal{D}$ is defined to be the points where at least one of the eigenvalues of ρ is zero, and the rest are positive. The distance $r = \sqrt{x^2 + y^2}$ from the origin to $\partial\mathcal{D}$ will then depend on the direction θ, and can be calculated. This is seen in Figure 5.2.

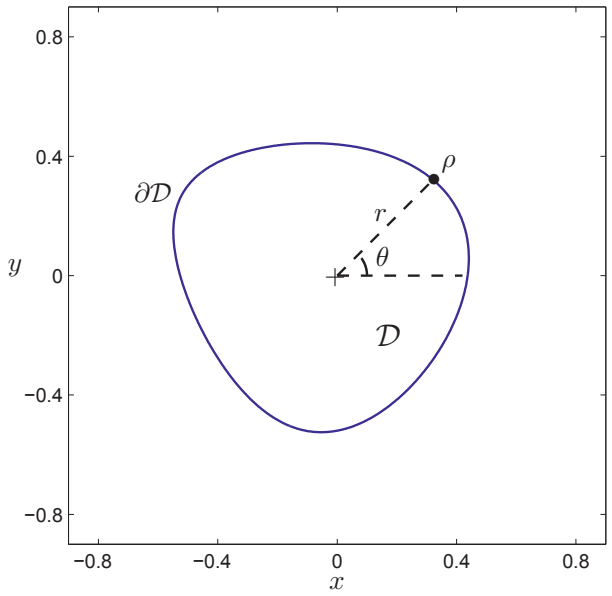


Figure 5.2: A two-dimensional section through the set of density matrices $\mathcal{D}$. The curve represents the boundary $\partial\mathcal{D}$ which contains the states of less than full rank. For each θ the distance r can be calculated. The state represented by " + " is the origin ρ_0, usually chosen as I/N, and the state ρ is here a state on the boundary $\partial\mathcal{D}$.

In general $\mathrm{tr}(\sigma_1\sigma_2) \neq 0$, so for these directions $\sigma_1 \not\perp \sigma_2$. To avoid skew coordinate axes we must rotate one of them, for instance σ_2 so that $\sigma_2 \mapsto \sigma_2 - \mathrm{tr}(\sigma_1\sigma_2)\sigma_1$, which ensures that $\sigma_1 \perp \sigma_2$. This is a matter of convenience, since σ_1 and the new σ_2 still define the same plane. Since we use the Hilbert-Schmidt metric, we have to make sure that the axes σ_1 and σ_2 are scaled accordingly, so that distances and angles are correctly represented. The constants a and b in (5.9) must then be chosen such that $\mathrm{tr}(\sigma_1^2) = \mathrm{tr}(\sigma_2^2) = 1$. We now call these two traceless, normalized and orthogonal directions σ_x and σ_y, and write any density matrix ρ in the cross section as

$$\rho = \frac{I}{N} + r\left[(\cos\theta)\sigma_x + (\sin\theta)\sigma_y\right] \tag{5.11}$$

Now, the partial transpose of the plane through $\mathcal{D}$ defined by ρ_0, ρ_1 and ρ_2 is

$$\rho^P = \rho_0^P + x\sigma_1^P + y\sigma_2^P = \frac{I}{N} + x\sigma_1^P + y\sigma_2^P \tag{5.12}$$

The points I/N, σ_1^P and σ_2^P now define a new two-dimensional section in the space of Hermitian matrices H_N, passing through $\mathcal{D}^P$. We can calculate the border $\partial\mathcal{D}^P$ for this section in the same way we found the border of $\mathcal{D}$ for the plane defined by I/N, σ_1 and σ_2. The border of $\partial\mathcal{D}^P$ is then found by appealing to the inequality $\rho^P \geq 0$, and consists of states for which $\det(\rho^P) = 0$. Since $\mathcal{P}$ consists of the states for which $\rho \geq 0$ and $\rho^P \geq 0$, the boundary of $\mathcal{P}$ can be found in the cross section between $\mathcal{D}$ and $\mathcal{D}^P$.

The borders of $\mathcal{D}$ and $\mathcal{D}^P$ can be visualized in different plots, but it is more informative to gather them in the same figure. We here choose the two-dimensional section defined by I/N, σ_1^P and σ_2^P. Again, to ensure orthogonality we must rotate $\sigma_2^P \mapsto \sigma_y$, and to make the system consistent we must also rotate the original $\sigma_2 \mapsto \sigma_y^P$, so as to ensure that it is indeed the partial transpose of σ_y.

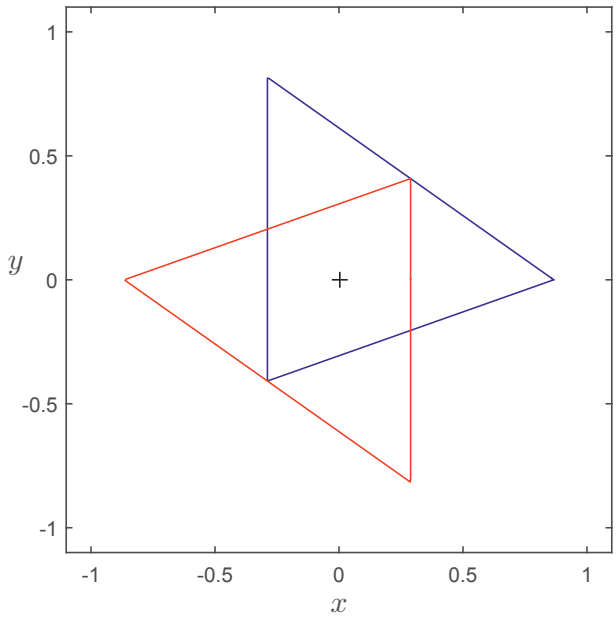


Figure 5.3: The blue curve is $\partial\mathcal{D}_4$ defined by a two-dimensional section that contains the two Bell states $\rho_1 = |\alpha+\rangle\langle\alpha+|$ and $\rho_2 = |\beta+\rangle\langle\beta+|$, and the maximally mixed state $I/4$, which is marked with a " + ". The red curve is the boundary of the partial transposes of states in the blue triangle. The points inside the blue/red parallelogram for which $\rho \geq 0$ and $\rho^P \geq 0$, make up the set $\mathcal{S} = \mathcal{P}$. The Bell states lie outside $\mathcal{P} = \mathcal{S}$, and since they are maximally entangled they both have maximum distance from $I/4$. Also observe that the sets $\mathcal{D}$ and $\mathcal{S} = \mathcal{P}$ are convex, while the set of entangled states is not.

In Figure 5.3 the blue curve is the boundary of $\mathcal{D}$ defined by the plane that consists of $I/4$ and the Bell states $\rho_1 = |\alpha+\rangle\langle\alpha+|$ and $\rho_2 = |\beta+\rangle\langle\beta+|$. The red triangle is the boundary of the partial transposes of states in the blue triangle, or more compactly $\partial\mathcal{D}^P$. The points inside the blue and red parallelogram in the middle, are states for which $\rho \geq 0$ and $\rho^P \geq 0$, or in other words the set of PPT states. Since we are looking at the 2×2 system all these states are separable, since we have $\mathcal{S} = \mathcal{P}$. The Bell states ρ_1 and ρ_2 are both entangled pure states, so they lie outside $\partial\mathcal{P}$, and since they are maximally entangled they both represent states with the maximum distance from $I/4$.

We also observe that since the Bell states commute, the curves that define $\partial\mathcal{D}$ are straight lines. For any convex combination $\rho = I/N + x\sigma_x + y\sigma_y$, the σ-matrices can then be diagonalized simultaneously, and so $\det(\rho)$ can be factored into a product of linear terms, which results in linear boundaries for $\mathcal{D}$. For non-commuting σ-matrices $\det(\rho)$ will in general be a polynomial of degree equal to the rank of the matrix.

In addition it is easy to see that the set of density matrices $\mathcal{D}$ and the set of separable states $\mathcal{S} = \mathcal{P}$ are convex sets, while the set of entangled states is clearly not.

In Figures 5.4 and 5.5 we see the same as in Figure 5.3, but with sections defined by $I/4$, the Bell state $\rho_1 = |\beta+\rangle\langle\beta+|$ and where ρ_2 is a random pure product state. In Figure 5.4 the pure product state ρ_2 is chosen so that it commutes with the Bell state ρ_1. The boundary of $\mathcal{D}$ are then straight lines, but since $[\rho_1^P, \rho_2^P] \neq 0$, parts of the boundary of $\mathcal{D}^P$ will be curved. In Figure 5.5 we choose ρ_2 so that it does not commute with the Bell state ρ_1, and so neither $\mathcal{D}$ nor $\mathcal{D}^P$ will be entirely straight lines. In both cases we observe that the pure product state ρ_2 is extremal in both $\mathcal{D}$ and $\mathcal{P} = \mathcal{S}$.

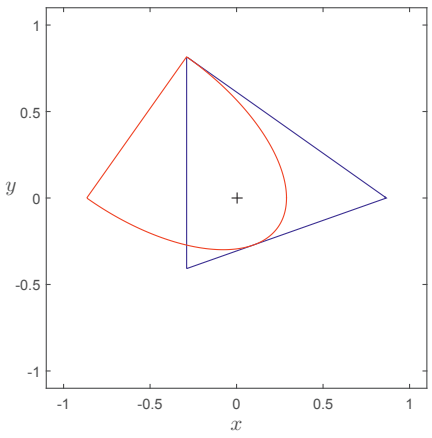


Figure 5.4: Sections are defined by $I/4$, the Bell state $\rho_1 = |\beta+\rangle\langle\beta+|$ and a random pure product state ρ_2, such that $[\rho_1, \rho_2] = 0$. Observe that the pure product state ρ_2 is the only state to be extremal in both $\mathcal{D}$ and $\mathcal{P} = \mathcal{S}$.

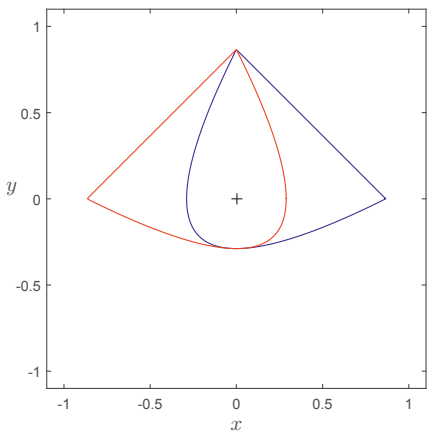


Figure 5.5: Sections are defined by $I/4$, the Bell state $\rho_1 = |\beta+\rangle\langle\beta+|$ and a pure product state ρ_2, such that $[\rho_1, \rho_2] \neq 0$. Observe that the pure product state ρ_2 is extremal in both $\mathcal{D}$ and the drop shaped set $\mathcal{P} = \mathcal{S}$.

In Figures 5.6 and 5.7 we are looking at states in 3×3 system. First in Figure 5.6 we have used a two-dimensional section consisting of $I/9$ and two extremal PPT states

of rank $(5, 5)$. Since there exists entangled PPT states for these dimensions, we can no longer easily identify the set of separable states. All that can be said with certainty is that the set $\mathcal{S}$ is constrained to the points that lie inside both $\partial\mathcal{D}$ and $\partial\mathcal{D}^P$, *i.e.* the Peres set. The entangled NPT states are exclusively the points that are outside $\partial\mathcal{D}^P$ but inside $\partial\mathcal{D}$, which on Figure 5.6 are areas that are relatively small compared to the corresponding area covered by the same class of states for the 2×2 system in Figure 5.3. In Figure 5.7 we use a two-dimensional section consisting of $I/9$ and two random NPT states of rank $(9, 9)$. Both partial transposes have only one relatively small negative eigenvalue, so both states are only marginally NPT. The area of the entangled NPT states is considerably larger than for the case of the extremal rank $(5, 5)$ PPT states. We can clearly see the convex shape of the sets $\mathcal{D}$ and $\mathcal{P}$, while at least the entangled NPT states have a concave (non-convex) shape.

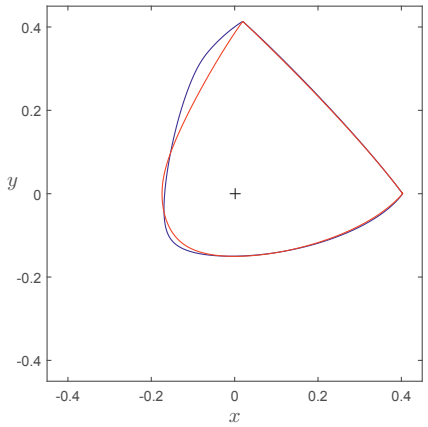


Figure 5.6: Section defined by $I/9$ and two extremal PPT states of rank $(5, 5)$. The entangled NPT states are the points outside $\partial\mathcal{D}^P$ but inside $\partial\mathcal{D}$, which are rather small areas. The Peres set is now the entangled PPT states in addition to $\mathcal{S}$.

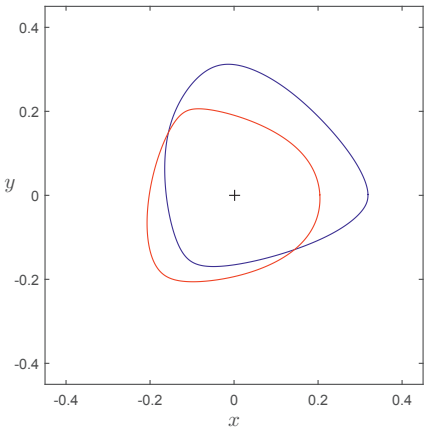


Figure 5.7: Section defined by $I/9$ and two entangled NPT states of rank $(9, 9)$. The entangled NPT states are the points that are outside $\partial\mathcal{D}^P$ but inside $\partial\mathcal{D}$, which is a larger area than in Figure 5.6. The Peres set is now the entangled PPT states in addition to $\mathcal{S}$.

5.5 Perturbations in $\mathcal{D}$ and $\mathcal{P}$

Because it is Hermitian, a density matrix ρ has a spectral representation (3.3) in terms of a complete set of orthonormal eigenvectors $e_k \in \mathbb{C}^N$ with real eigenvalues λ_k. We now define the Hermitian matrices

$$P = \sum_{k,\lambda_k\neq0} e_k e_k^\dagger \qquad Q = I - P = \sum_{k,\lambda_k=0} e_k e_k^\dagger \tag{5.13}$$

which project orthogonally onto the two complementary and orthogonal subspaces $\operatorname{Img}\rho \subseteq \mathbb{C}^N$ and $\operatorname{Ker}\rho \subset \mathbb{C}^N$ respectively. These matrices commute with any den-

sity matrix $[P, \rho] = [Q, \rho] = 0$, which also means that $P\rho P = \rho$ and $Q\rho Q = 0$. In the following we consider a perturbation of the density matrix ρ

$$\rho \mapsto \rho' = \rho + \epsilon\sigma \tag{5.14}$$

where $\sigma \neq 0$ is a traceless Hermitian matrix. The parameter $\epsilon \in \mathbb{R}$ may be finite or infinitesimal.

5.5.1 Finite perturbations

We observe that if $\operatorname{Img}\sigma \subset \operatorname{Img}\rho$, which is equivalent to $P\sigma P = \sigma$, then there will be a finite range of values $\Delta\epsilon = \{\epsilon \,|\, \epsilon_1 < \epsilon < \epsilon_2\}$ with $\epsilon_1 < 0 < \epsilon_2$, such that $\rho' \in \mathcal{D}$ and $\operatorname{Img}\rho' = \operatorname{Img}\rho$. This is so because for the entire $\Delta\epsilon$, the eigenvectors e_k of ρ for which $\lambda_k = 0$, will remain eigenvectors e'_k of ρ' with $\lambda'_k = 0$, and the positive eigenvalues of ρ will change *continuously* with ϵ into positive eigenvalues of ρ'. We know that the extremal points on the set $\mathcal{D}$ are the pure states, but it is useful to formulate conditions for extremality in the form of perturbations as well. The above shows that ρ is extremal in $\mathcal{D}$ if and only if there exists no traceless Hermitian matrix σ such that $P\sigma P = \sigma$, where P and Q are defined by $\operatorname{Img}\rho$ and $\operatorname{Ker}\rho$ as in (5.13). This is equivalent to the condition that the equation $P\sigma P = \sigma$ has only the trivial solution $\sigma = \rho$, up to proportionality. Another formulation is that ρ is extremal in $\mathcal{D}$ if and only if no state $\rho' \neq \rho$ has the same range as ρ. An attempt to visualize this is made in Figure 5.8

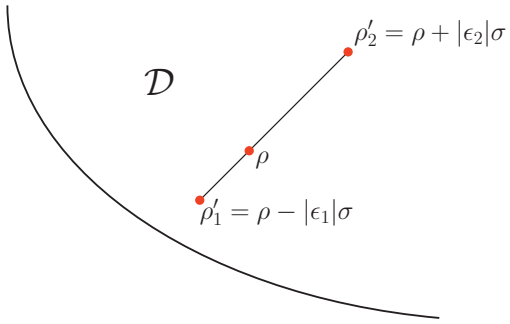


Figure 5.8: Finite perturbations of the density matrix ρ using the Hermitian traceless matrix σ. All states along the line are convex combinations of $\rho'_1 = \rho - |\epsilon_1|\sigma$ and $\rho'_2 = \rho + |\epsilon_2|\sigma$ for every ϵ in the interval $\epsilon_1 < \epsilon < \epsilon_2$, where $\epsilon_1 < 0 < \epsilon_2$. The range $\operatorname{Img}\rho'$ is independent of ϵ in the interval, which means that $\operatorname{Img}\sigma \subset \operatorname{Img}\rho$ or $P\sigma P = \sigma$.

5.5.2 Infinitesimal perturbations

We now assume more generally that $\operatorname{Img}\sigma \not\subseteq \operatorname{Img}\rho$ in the perturbation (5.14). The zero eigenvalues of ρ are then perturbed into eigenvalues of $Q\sigma Q$. Similarly the positive eigenvalues of ρ are perturbed into positive eigenvalues of ρ' determined by $P\sigma P$. Since

the eigenvalues are given by the characteristic polynomial of the matrix, these will carry an ϵ-dependency that includes higher powers. If we restrict to first order in ϵ we get

$$\mathrm{rank}(\rho') = \mathrm{rank}(\rho) + \mathrm{rank}(Q\sigma Q) \tag{5.15}$$

So to ensure that $\mathrm{rank}(\rho) = \mathrm{rank}(\rho')$ and that $\rho' \geq 0$ for ϵ-perturbations in both directions, we thus put $Q\sigma Q = 0$. It is then evident that if we want to increase the rank of ρ by one, we choose σ such that

$$Q\sigma Q = a(zz^\dagger) \tag{5.16}$$

where $z \in \mathrm{Ker}\,\rho$. Since $Q\sigma Q$ is Hermitian we must have $a \in \mathbb{R}$.

5.6 Projection operators on H_N

Using the projections P and Q defined in (5.13) we can define projection operators $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ on H_N, the real Hilbert space of Hermitian $N \times N$ matrices, as follows,

$$\begin{aligned} \mathbf{P}X &= PXP \\ \mathbf{Q}X &= QXQ \\ \mathbf{R}X &= (\mathbf{I} - \mathbf{P} - \mathbf{Q})X \end{aligned} \tag{5.17}$$

Here $\mathbf{I}$ is the identity operator on H_N. It is straightforward to verify that these are complementary projections, *i.e.* $[\mathbf{P}, \mathbf{Q}] = 0$, with $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{Q}^2 = \mathbf{Q}$.

It is possible to see the effects of these operators in a very nice and clear way by using a representation that results in a block form for the matrices. Relative to an orthonormal basis for $\mathbb{C}^N$ with the first basis vectors in $\mathrm{Img}\,\rho$ and the last basis vectors in $\mathrm{Ker}\,\rho$, an Hermitian matrix $X \in H_N$ with rank m and nullity $N - m$ takes the block form

$$X = \begin{pmatrix} A_m & C \\ C^\dagger & B_{N-m} \end{pmatrix} \tag{5.18}$$

where $A_m = A_m^\dagger$ is an $m \times m$ matrix, $B_{N-m} = B_{N-m}^\dagger$ is an $(N-m) \times (N-m)$ matrix and the matrix C has dimensions $m \times (N-m)$. In this basis we also have

$$P = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix} \qquad Q = \begin{pmatrix} 0 & 0 \\ 0 & I_{N-m} \end{pmatrix} \tag{5.19}$$

with I_m and I_{N-m} as the $m \times m$ and $(N-m) \times (N-m)$ identity matrices respectively. The operations $\mathbf{P}X$, $\mathbf{Q}X$ and $\mathbf{R}X$ then becomes

$$\mathbf{P}X = \begin{pmatrix} A_m & 0 \\ 0 & 0 \end{pmatrix} \qquad \mathbf{Q}X = \begin{pmatrix} 0 & 0 \\ 0 & B_{N-m} \end{pmatrix} \qquad \mathbf{R}X = \begin{pmatrix} 0 & C \\ C^\dagger & 0 \end{pmatrix} \tag{5.20}$$

In order to make investigations into the set of PPT states we would like to conduct perturbations that change the ranks of ρ and ρ^P in a controlled manner. In a way which is completely similar to the case for $\mathrm{Img}\,\rho$ and $\mathrm{Ker}\,\rho$ in (5.17), we introduce projection operators $\widetilde{\mathbf{P}}$, $\widetilde{\mathbf{Q}}$ and $\widetilde{\mathbf{R}}$ onto $\mathrm{Img}\,\rho^P$ and $\mathrm{Ker}\,\rho^P$. We first define the matrices $\widetilde{P}$ and $\widetilde{Q} = I - \widetilde{P}$ as orthogonal projections onto $\mathrm{Img}\,\rho^P$ and $\mathrm{Ker}\,\rho^P$ similar to (5.13), and then define

$$\begin{aligned} \widetilde{\mathbf{P}}X &= (\widetilde{P}X^P\widetilde{P})^P \\ \widetilde{\mathbf{Q}}X &= (\widetilde{Q}X^P\widetilde{Q})^P \\ \widetilde{\mathbf{R}}X &= (\mathbf{I} - \widetilde{\mathbf{P}} - \widetilde{\mathbf{Q}})X \end{aligned} \tag{5.21}$$

5.7 Restricted perturbations in $\mathcal{P}$

We have seen that for the perturbation $\rho' = \rho + \epsilon\sigma$ the condition $P\sigma P = \sigma$ ensures that $\mathrm{Img}\,\rho = \mathrm{Img}\,\rho'$ for a finite range of values for ϵ. Also to very good approximation the weaker condition that $Q\sigma Q = 0$ will ensure that the the rank is unchanged for infinitesimal perturbations. We may now use the projection operators (5.17) and (5.21) to impose various restrictions on the perturbation matrix σ.

5.7.1 Extremality in $\mathcal{P}$

The extremality condition for $\mathcal{P}$ is derived in a similar way as the extremality condition for $\mathcal{D}$. We see that ρ is extremal in $\mathcal{P}$ if and only if there exists no $\rho' \in \mathcal{P}$, other than ρ itself, with both $\mathrm{Img}\,\rho = \mathrm{Img}\,\rho'$ and $\mathrm{Img}\,\rho^P = \mathrm{Img}\,\rho'^P$. A more useful formulation is again to say that ρ is extremal in $\mathcal{P}$ if and only if the only solution of the equations $P\sigma P = \sigma$ and $\widetilde{P}\sigma^P\widetilde{P} = \sigma^P$ or equivalently $\mathbf{P}\sigma = \sigma$ and $\widetilde{\mathbf{P}}\sigma = \sigma$, is the trivial solution $\sigma = \rho$. These two equations are equivalent to the single eigenvalue equation

$$(\mathbf{P} + \widetilde{\mathbf{P}})\sigma = 2\sigma \tag{5.22}$$

or alternatively

$$\mathbf{P}\widetilde{\mathbf{P}}\mathbf{P}\sigma = \sigma \qquad \widetilde{\mathbf{P}}\mathbf{P}\widetilde{\mathbf{P}}\sigma = \sigma \tag{5.23}$$

Even though (5.22) does not immediately appear to be a genuine eigenvalue equation, it can be solved by representing σ as a vector in $\mathbb{R}^{N^2}$, and P and $\widetilde{P}$ by $N^2 \times N^2$ matrices. This is outlined in more detail in Section 8.1. The operator $\mathbf{P} + \widetilde{\mathbf{P}}$ will always have an eigenvector $\sigma = \rho$ with $\lambda = 2$ as eigenvalue. If it is the only solution of (5.22) then ρ is an extremal state of $\mathcal{P}$.

We recall from Section 3.6.4 that for ρ to be an edge state there cannot exist any product vector $w = u \otimes v \in \mathrm{Img}\,\rho$ such that $\widetilde{w} = u \otimes v^* \in \mathrm{Img}\,\rho^P$. All entangled extremal PPT states are edge states, although the opposite is not necessarily true. To see this, we observe that an extremality criterion for the state ρ on the set of PPT states $\mathcal{P}$ is that there can exist no product vector $\psi = \phi \otimes \chi$, such that

$$\rho' = \rho - \epsilon(\psi\psi^\dagger) \tag{5.24}$$

is a PPT state for all ϵ in an interval $\epsilon_1 < \epsilon < \epsilon_2$ where $\epsilon_1 < 0 < \epsilon_2$. If we take ρ to be a non-edge state, then there exists a $w = u \otimes v \in \mathrm{Img}\,\rho$ such that $\widetilde{w} = u \otimes v^* \in \mathrm{Img}\,\rho^P$, and an arbitrarily small ϵ-fraction of $ww^\dagger$ may then be subtracted from ρ, so that $\rho' = \rho - \epsilon(ww^\dagger) \in \mathcal{D}$. For the partially transposed state we then have from (3.60)

$$\rho'^P = \rho^P - \epsilon(ww^\dagger)^P = \rho^P - \epsilon(\widetilde{w}\widetilde{w}^\dagger) \tag{5.25}$$

Since, by assumption $\widetilde{w} = u \otimes v^* \in \mathrm{Img}\,\rho^P$, we have that $\rho'^P \in \mathcal{D}$ for the mentioned ϵ-interval. The state ρ' in (5.24) is then a PPT state, and we deduce that ρ cannot be extremal on $\mathcal{P}$. So the non-edge property implies non-extremality, and in effect extremality entails the edge property.

5.7.2 Perturbations preserving PPT property and ranks

The rank and positivity of ρ is preserved by the perturbation (5.14) to first order in ϵ in both directions, if and only if $Q\sigma Q = \mathbf{Q}\sigma = 0$. Similarly, the rank and positivity of ρ^P is preserved if and only if $(\widetilde{Q}\sigma^P\widetilde{Q})^P = \widetilde{\mathbf{Q}}\sigma = 0$. These two equations together can be written as

$$(\mathbf{Q} + \widetilde{\mathbf{Q}})\sigma = 0 \tag{5.26}$$

So the perturbations that preserve the PPT property and the ranks of both ρ and ρ^P to first order in ϵ, are the non-trivial solutions of (5.26). We may want to perturb in different ways, for example such that $\mathrm{Img}\,\rho = \mathrm{Img}\,\rho'$, but not necessarily $\mathrm{Img}\,\rho^P = \mathrm{Img}\,\rho'^P$, *i.e.* we only require $\mathrm{rank}(\rho^P) = \mathrm{rank}(\rho'^P)$. The conditions on σ are then

$$\mathbf{P}\sigma = \sigma \qquad \widetilde{\mathbf{Q}}\sigma = 0 \qquad \Rightarrow \qquad (\mathbf{I} - \mathbf{P} + \widetilde{\mathbf{Q}})\sigma = 0 \tag{5.27}$$

5.7.3 Surfaces of PPT states with specified ranks

It is reasonable to assume that the surfaces of density matrices with specific rank (m, n), are curved in the embedding space. What we do is study the *tangent space* of a surface at given points. This will in particular allow us to calculate the local dimension of the surface. If the point on the surface is sufficiently generic, the local dimension should correspond with the dimension of the whole surface.

We have seen that perturbation matrices σ that preserve the rank of both ρ and ρ^P, are the non-trivial solutions of (5.26). This means that in the block representation (5.18) the matrix σ can expressed on the form

$$X = \begin{pmatrix} A_m & C \\ C^\dagger & 0 \end{pmatrix} \tag{5.28}$$

The number of independent columns in this matrix is the same as for ρ, which means that the rank of ρ and ρ' are equal for small perturbations. If s is the number of linearly independent solutions of $(\mathbf{Q} + \widetilde{\mathbf{Q}})\sigma = 0$, then the dimension of the local tangent space at ρ, *i.e.* the dimension of the surface of density matrices with rank (m, n) at ρ is $s - 1$.

By themselves $\mathbf{Q}\sigma = 0$ and $\widetilde{\mathbf{Q}}\sigma = 0$ define the tangent space to the surface of rank m density matrices at ρ and rank n density matrices at ρ^P respectively. The solutions of $(\mathbf{Q} + \widetilde{\mathbf{Q}})\sigma = 0$ represent the local intersection of these two spaces at ρ, which is the tangent space to the surface of PPT states with rank (m, n) at ρ.

To compute a lower bound on the dimension of this surface we assume that constraints on σ from $\mathbf{Q}\sigma = 0$ and $\widetilde{\mathbf{Q}}\sigma = 0$ are independent. Since Q is the orthogonal projection onto the $(N - m)$-dimensional subspace $\operatorname{Ker}\rho$, the equation $\mathbf{Q}\sigma = 0$ represents $(N - m)^2$ real constraints. Likewise since $|\operatorname{Ker}\rho^P| = N - n$, the equation $\widetilde{\mathbf{Q}}\sigma = 0$ represents $(N - n)^2$ real constraints. The lower bound is then

$$d \geq N^2 - (N - m)^2 - (N - n)^2 - 1 \tag{5.29}$$

5.8 Faces and dual sets of $\mathcal{D}$, $\mathcal{P}$ and $\mathcal{S}$

We can understand extremality and several other geometrical properties of the sets $\mathcal{D}$, $\mathcal{P}$ and $\mathcal{S}$ by looking at the faces of these sets, and of the dual sets $\mathcal{D}^\circ$, $\mathcal{P}^\circ$ and $\mathcal{S}^\circ$.

Faces of $\mathcal{D}$

Given any density matrix ρ, we define a face $\mathcal{F}_\mathcal{D}(\rho)$ on $\mathcal{D}$ as the set of density matrices with a range contained in $\operatorname{Img}\rho$. The face $\mathcal{F}_\mathcal{D}(\rho)$ will in particular contain ρ as an interior point. Note that the boundary of $\mathcal{D}$ consists of both flat and curved segments, where curved entails that the segments are entirely made up of zero-dimensional faces, *i.e.* extremal points, and flat means that the interior of the faces are not extremal points. The projection operator $\mathbf{P}$ in (5.17) thus projects states onto the face $\mathcal{F}_\mathcal{D}(\rho)$. If for any perturbation $\rho' = \rho + \epsilon\sigma$ the traceless Hermitean matrix σ is chosen such that $\operatorname{Img}\rho = \operatorname{Img}\rho'$, or equivalently so that $\mathbf{P}\sigma = \sigma$, we have $\rho, \rho' \in \mathcal{F}_\mathcal{D}(\rho)$. If ρ has rank m, then the rank of the projection operator $\mathbf{P}$ is m^2 and the equation $\mathbf{P}\sigma = \sigma$ has m^2 linearly independent solutions σ_k. The face $\mathcal{F}_\mathcal{D}(\rho)$ together with the positivity conditions on density matrices, form a vector space which is spanned by these σ_k.

The boundary of the face $\mathcal{F}_\mathcal{D}(\rho)$ is found by using the block representation (5.18) for a state $\rho \in \mathcal{F}_\mathcal{D}(\rho)$

$$\rho = \begin{pmatrix} A_m & C \\ C^\dagger & B_{N-m} \end{pmatrix} \tag{5.30}$$

Since the matrix A_m defines the m-dimensional range of ρ, the boundary of $\mathcal{F}_\mathcal{D}(\rho)$ is defined by $\det(A_m) = 0$. These states will then have lesser rank than the states in the interior of $\mathcal{F}_\mathcal{D}(\rho)$. A normalized density matrix on $\partial\mathcal{D}$ with rank m defines a face of dimension $m^2 - 1$. So for most N, the largest faces on $\mathcal{D}_N$, will have a dimension considerably smaller than $N^2 - 2$, which would be the dimension of a facet and of the boundary of a $(N^2 - 1)$-dimensional convex set. In fact, any face of $\mathcal{D}$ with interior states of rank m will be a complete representation of the set $\mathcal{D}_m$ of density matrices on an m-dimensional Hilbert space.

We can perform unitary $U(N)$-transformations on all the density matrices in the face to rotate the face in the full space H_N. However, the subgroups of rotations

$U(m)$ that act purely on the image, and $U(N-m)$ that act purely on the kernel, do not change the *orientation* of the face, but merely rotate it locally. So the remaining elements of the group $U(N)$ that transform the faces in the full space, are the elements of $U(N)/U(m)\times U(N-m)$, which has dimension

$$N^2 - m^2 - (N-m)^2 = 2Nm - 2m^2 \tag{5.31}$$

Combining the dimension of the face with the number of such unitary transformations that change the orientation of the face, we get in total

$$(m^2-1) + (2Nm - 2m^2) = 2Nm - m^2 - 1 \tag{5.32}$$

which is then the dimension of the surface of density matrices of rank m, or equivalently to the dimension of the surface of density matrices of rank $N-m$. If we put $m = N-1$ into (5.32), we obtain the dimension of the set of faces with rank $N-1$, *i.e.* of rank one less than the interior states

$$2N(N-1) - (N-1)^2 - 1 = N^2 - 2 \tag{5.33}$$

which is exactly the dimension of $\partial\mathcal{D}$. We also notice that for $m=1$, which are the pure states, we get a surface of dimension $2N-2$, which is in agreement with results in Section 5.2.

Self duality of $\mathcal{D}$

An essential characteristic of $\mathcal{D}$ is that it is self dual, or $\mathcal{D}^\circ = \mathcal{D}$. By Definition 4.4, the dual set of $\mathcal{D}$ will contain all Hermitian matrices A such that $\mathrm{tr}(A\rho)\geq 0$ for all $\rho\in\mathcal{D}$. Now, if $\rho\in H_N$ is a density matrix we will have $\psi^\dagger\rho\psi\geq 0$ for all $\psi\in\mathcal{H}$. This is equivalent to $\mathrm{tr}[(\psi\psi^\dagger)\rho]\geq 0$ for all $\psi\in\mathcal{H}$. The set of states $\psi\psi^\dagger$ is the set of pure states in $\mathcal{D}$, and since these states constitute the set of extremal points for $\mathcal{D}$, we see that the only set dual to $\mathcal{D}$, is $\mathcal{D}$ itself.

Another important feature is that any face $\mathcal{F}_\mathcal{D}$, is an exposed face. This is related to the fact that any face $\mathcal{F}_\mathcal{D}$ is defined by an m-dimensional subspace $\mathcal{U} = \mathrm{Img}\,\rho$ of a state $\rho\in\mathrm{int}(\mathcal{F}_\mathcal{D})$. Let us choose a density matrix τ

$$\tau = \sum_{i=1}^{N-m}\lambda_i\,(\psi_i\psi_i^\dagger) \tag{5.34}$$

where all $\psi_i\in\mathcal{U}^\perp$. The face $\mathcal{F}_\mathcal{D}$ can then be defined as

$$\mathcal{F}_\mathcal{D}(\mathcal{U}) = \{\rho\in\mathcal{D}_N \,|\, \mathrm{tr}(\tau\rho) = 0\} \tag{5.35}$$

Since $\mathrm{tr}(\tau\rho) = \sum_{i=1}^{N-m}\lambda_i\,(\psi_i^\dagger\rho\psi_i)$ we see that $\psi_i^\dagger\rho\psi_i = 0$. Since ρ is a positive matrix, this means that $\rho\psi_i = 0$ for all $\rho\in\mathcal{F}_\mathcal{D}(\mathcal{U})$, and only those ρ. So the hyperplane

$\mathrm{tr}(\tau\rho) = 0$ that defines the face $\mathcal{F}_{\mathcal{D}}(\mathcal{U})$, intersects $\partial\mathcal{D}$ only in $\mathcal{F}_{\mathcal{D}}(\mathcal{U})$, *i.e.* the face is exposed.

We know that any point x in a convex set X may be written as a convex combination of the extremal points of the set. To establish the minimum number of extremal points which is sufficient, the facial structure of the set is relevant. Theorem 4.1 by Carathéodory, gives the sufficient number of extremal points as $n+1$, where n is the dimension of the set. However, the set $\mathcal{D}$ is an example where we can do better. The dimension of the normalized set is $n = N^2 - 1$, where N is the dimension of the Hilbert space. But as can be seen in (3.3), the spectral representation of a density matrix in terms of its eigenvalues and eigenvectors is a decomposition using only N extremal points.

Any point x in a compact convex set X is either extremal or an interior point of a unique face $\mathcal{F}_1$, which might be the whole set X. If x is not extremal it may be written as a convex combination

$$x = (1 - p_1)s_1 + p_1 t_1 \tag{5.36}$$

where s_1 is an extremal point of $\mathcal{F}_1$ and t_1 is another boundary point of $\mathcal{F}_1$. If t_1 is also extremal on $\mathcal{F}_1$ we simply define $s_2 = t_1$, and we have x as a convex sum $x = (1 - p_1)s_1 + p_1 s_2$ of points s_i which are extremal on X. Otherwise t_1 is an interior point of another proper face $\mathcal{F}_2$ of $\mathcal{F}_1$, and we write

$$t_1 = (1 - p_2)s_2 + p_2 t_2 \tag{5.37}$$

where s_2 is an extremal point of $\mathcal{F}_2$ and t_2 is a boundary point of $\mathcal{F}_2$. Continuing this process we obtain a decomposition of x as a convex combination of extremal points $s_1, s_2, \ldots, s_k \in X$, and a sequence $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \ldots \supset \mathcal{F}_k$ of faces of decreasing dimensions $n \geq n_1 > n_2 > \ldots > n_k = 0$. The length of the sequence is k, and the obvious inequality $k \leq n + 1$ is Carathéodory's theorem. For the set of normalized density matrices the longest possible sequence of face dimensions has length N, it is

$$N^2 - 1 > (N-1)^2 - 1 > \ldots > (N-j)^2 - 1 > \ldots > 8 > 3 > 0 \tag{5.38}$$

So in conclusion, we will always be able to decompose an arbitrary density matrix ρ of rank r as a convex combination of r pure states. These may be built from the eigenvectors of ρ, *i.e.* the spectral decomposition. But as we have seen, there is a considerable ambiguity in the construction of density matrices, so generally there are other possibilities.

Faces of $\mathcal{P}$ and $\mathcal{S}$

A given state ρ in $\mathcal{D}$ defines a face $\mathcal{F}_{\mathcal{D}}(\rho)$, but if ρ is a PPT state then its partial transpose ρ^P also defines (a different) face $\mathcal{F}_{\mathcal{D}}(\rho^P)$ on the set of density matrices. An attempt to illustrate this is done in Figure 5.9.

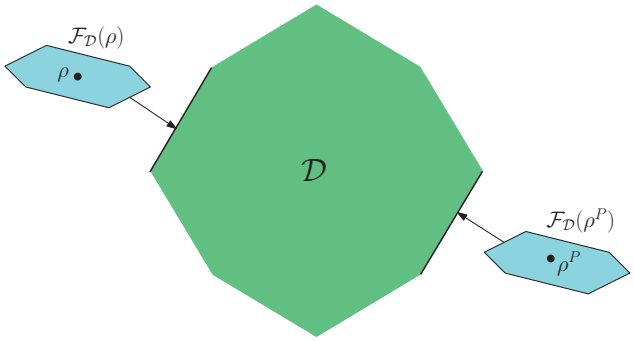


Figure 5.9: Faces $\mathcal{F}_{\mathcal{D}}(\rho)$ and $\mathcal{F}_{\mathcal{D}}(\rho^P)$ on $\mathcal{D}$ containing the PPT state ρ and its partial transpose ρ^P. The faces are defined by the set of density matrices with a range completely contained in $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$ respectively.

The set obtained by taking the partial transpose of all the points in the face $\mathcal{F}_{\mathcal{D}}(\rho^P)$ is itself a face $\mathcal{F}_{\mathcal{D}^P}(\rho)$ on the set $\mathcal{D}^P$. We know that $\rho^P \in \mathrm{int}[\mathcal{F}_{\mathcal{D}}(\rho^P)]$, and then it follows that $\rho \in \mathrm{int}[\mathcal{F}_{\mathcal{D}^P}(\rho)]$. We then realize that $\mathcal{F}_{\mathcal{D}}(\rho)$ and $\mathcal{F}_{\mathcal{D}^P}(\rho)$ must share at least one point, namely ρ. The intersection of the two faces defines a new face on $\mathcal{P}$

$$\mathcal{F}_{\mathcal{P}}(\rho) = \mathcal{F}_{\mathcal{D}}(\rho) \cap \mathcal{F}_{\mathcal{D}^P}(\rho) \tag{5.39}$$

containing ρ. An attempt to illustrate this is made in Figure 5.10.

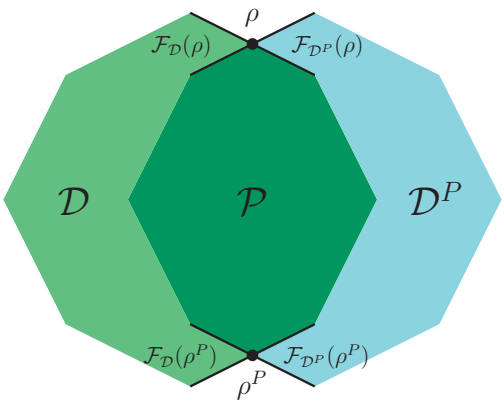


Figure 5.10: A PPT state ρ and its partial transpose ρ^P define faces $\mathcal{F}_{\mathcal{D}}(\rho)$ and $\mathcal{F}_{\mathcal{D}}(\rho^P)$ on $\mathcal{D}$. Taking the partial transpose of all points in $\mathcal{F}_{\mathcal{D}}(\rho^P)$, we get a face $\mathcal{F}_{\mathcal{D}^P}(\rho)$ on $\mathcal{D}^P$ that intersects $\mathcal{F}_{\mathcal{D}}(\rho)$. The intersection defines a face $\mathcal{F}_{\mathcal{P}}(\rho) = \mathcal{F}_{\mathcal{D}}(\rho) \cap \mathcal{F}_{\mathcal{D}^P}(\rho)$ on $\mathcal{P}$, that contains ρ. Likewise, the intersection of $\mathcal{F}_{\mathcal{D}}(\rho^P)$ and $\mathcal{F}_{\mathcal{D}^P}(\rho^P)$ define a face on $\mathcal{P}$ containing ρ^P.

The dimension of the face $\mathcal{F}_{\mathcal{P}}(\rho)$ can then be found by counting the number d of linearly independent solutions of

$$(\mathbf{P} + \widetilde{\mathbf{P}})\sigma = 2\sigma \tag{5.40}$$

Obviously, if and only if $d = 1$ which corresponds to the trivial solution $\sigma = \rho$ itself, we have an extremal point, or a zero-dimensional face. These linearly independent solutions then represent possible directions that preserve both the range of ρ and ρ^P in the perturbation $\rho' = \rho + \epsilon\sigma$.

It is perhaps not easy to see directly that all the faces $\mathcal{F}_\mathcal{P}$ of $\mathcal{P}$ are exposed, as is the case with $\mathcal{D}$. But it follows immediately from the hyperplane construction (5.39), and formal proofs have been given, for instance by Ha and Kye [50]. As for the case of the separable states $\mathcal{S}$, it is not yet known if every face $\mathcal{F}_\mathcal{S}$ is exposed. This is perhaps surprising since the extremal states of $\mathcal{S}$, and therefore of any face on $\mathcal{F}_\mathcal{S}$, is well known by definition. Several works have however been published on the matter including [51, 52], but a clear understanding of the structure of faces on $\mathcal{S}$ has yet to be reached. This has consequences to which we shall return briefly in Section 7.5.4.

A nested sequence of convex sets

We have already observed in Section 3.5.2, that for the set of density matrices $\mathcal{D}$, the set of PPT states $\mathcal{P}$ and the set of separable states $\mathcal{S}$ we have the inclusions

$$\mathcal{S} \subseteq \mathcal{P} \subset \mathcal{D} \tag{5.41}$$

with equality $\mathcal{S} = \mathcal{P}$ for the 2×2 and 2×3 systems only. To illustrate this, let us for a convex body $X \subset \mathbb{R}^N$ define the dual X° by

$$X^\circ = \{y \in \mathbb{R}^N \,|\, 1 + \langle x, y\rangle \geq 0 \;\; \forall\, x \in X\} \tag{5.42}$$

which differs slightly from Definition 4.4 since we are no longer using convex coordinates. If we take the convex body to be a three-dimensional sphere

$$B_r = \{x \in \mathbb{R}^3 \,|\, |x| \leq r\} \tag{5.43}$$

we can use the above description to define B_r°. We quickly realize that the dual is also a sphere $B_r^\circ = B_{1/r} = \{x \in \mathbb{R}^3 \,|\, |x| \leq 1/r\}$. Clearly the sphere B_1 is self dual, while for $r > 1$ the dual is a smaller sphere and for $r < 1$ it is a larger sphere. It is a general feature that if we enlarge a convex body the conditions on the dual becomes more stringent, and hence the dual shrinks. Clearly for the spheres $B_{1/4}$, $B_{1/2}$ and B_1 we will have

$$B_{1/4} \subset B_{1/2} \subset B_1 = B_1^\circ \subset B_{1/2}^\circ \subset B_{1/4}^\circ \tag{5.44}$$

We now use this understanding, along with the self dual property of $\mathcal{D}$ and (5.41) to claim that for a system of dimensions $N_A \times N_B$ with $N = N_A N_B \geq 8$ we have

$$\mathcal{S} \subset \mathcal{P} \subset \mathcal{D} = \mathcal{D}^\circ \subset \mathcal{P}^\circ \subset \mathcal{S}^\circ \tag{5.45}$$

The two sets $\mathcal{P}^\circ$ and $\mathcal{S}^\circ$ are well defined and important convex sets, with which we will make a closer acquaintance later.

Chapter 6

PPT states in the 3×3 system

Presented here is a review of low rank PPT states in the 3×3 system. As first pointed out by Horodecki *et al.* [53], the entangled PPT states of lowest rank in the $N_A \times N_B$ system are states of rank $N_A + N_B - 2$. The only PPT states of lower rank are the pure product states, which are not entangled. For the 3×3 system an entangled state must have a minimum rank of four in order to be extremal on $\mathcal{P}$. While the entangled PPT states of rank $(4, 4)$ can be completely described in a relatively simple way by making use of the UPB construction, this is not the case for the entangled PPT states of rank $(5, 5)$, which show a considerably more complicated structure. We describe PPT states of rank $(4, 4)$ and rank $(5, 5)$, both extremal and separable, and use the results from Chapter 5 to describe methods for tracing surfaces of PPT states of rank $(5, 5)$ in $\mathcal{P}$. We finally make enquiries into certain classes of non-generic PPT states of rank $(5, 5)$.

6.1 PPT states of rank $(4, 4)$

A very important feature in the understanding and construction of the entangled PPT states of rank $(4, 4)$, initially developed by Bennett *et al.* [54], is that of an unextendable product basis (UPB). We first make a formal definition of an UPB for the $N_A \times N_B$ system, and then specialize to the 3×3 system.

Definition 6.1 (UPB)**.** *An unextendible product basis in the $N_A \times N_B$ system is a set of r orthogonal product vectors $w_i = u_i \otimes v_i \in \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ with the property that there exists no product vector in $\mathbb{C}^N = \mathbb{C}^{N_A N_B}$ orthogonal to all of them, and such that any number N_A of the vectors u_i and any number N_B of the vectors v_i are linearly independent.*

The r orthogonal product vectors $w_i = u_i \otimes v_i$ then span an r-dimensional subspace $\mathcal{U} \subset \mathbb{C}^N$, and the orthogonal supplement $\mathcal{V} = \mathcal{U}^\perp$ will then contain no product vectors. There may be more product vectors in $\mathcal{U}$, but they will be linear combinations of $w_i = u_i \otimes v_i$. It is thus clear that for any density matrix ρ where we can construct an UPB in the kernel of ρ, it follows from the range criterion in Section 3.6.4 that the state ρ will be entangled.

For the 3×3 system we know that the entangled states of lowest rank that are extremal on $\mathcal{P}$, have rank $(4, 4)$, and for such states $|\mathrm{Ker}\, \rho| = 5$. According to (3.55)

this is exactly the dimension at which a generic subspace in $\mathbb{C}^9$ will contain a finite number of product vectors. Assuming that $w \in \mathbb{C}^9$ is a product vector

$$w = f \otimes g = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} \otimes \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \tag{6.1}$$

where we have $f, g \in \mathbb{C}^3$. We observe that the factors f_i and g_j appear in several of the components of w, and then express the following requirements on these components so that w is a product vector, essentially a restatement of (3.52) for dimension 3×3

$$\begin{array}{lll} w_1 w_5 = w_2 w_4 & w_1 w_9 = w_3 w_7 & w_4 w_8 = w_5 w_7 \\ w_1 w_6 = w_3 w_4 & w_2 w_6 = w_3 w_5 & w_4 w_9 = w_6 w_7 \\ w_1 w_8 = w_2 w_7 & w_2 w_9 = w_3 w_8 & w_5 w_9 = w_6 w_8 \end{array} \tag{6.2}$$

The factors w_k appear in a non-linear and mixed way in this set of equations, which makes it difficult to analyze in general detail. For the 4×4 system the limit dimension from (3.55) is ten, so here it will be possible to construct an UPB in the kernel of a generic $(6, 6)$ PPT state.

Although a generic five-dimensional subspace for the 3×3 system will contain six product vectors, it is not a generic property that five of them are orthogonal. For the case of the 3×3 system we have $r = 5$ and the orthogonality relations for the u_i and v_i can be visualized as in Figure 6.1.

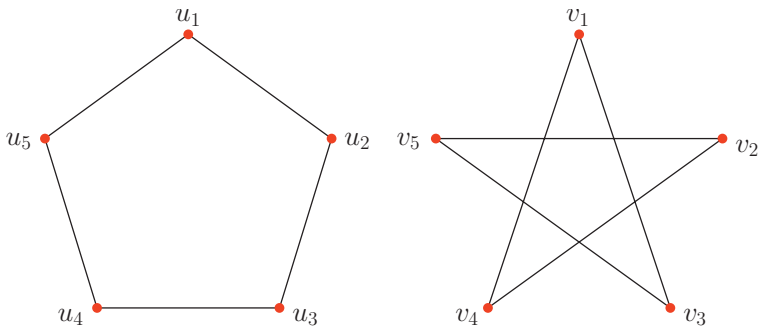


Figure 6.1: UPB construction for a five-dimensional subspace in the 3×3 system. In order to ensure the orthogonality relations $w_i^\dagger w_i = 0$ for the five product vectors $w_i = u_i \otimes v_i$, the orthogonality relations above must be satisfied.

The five-dimensional subspace spanned by the product vectors $w_i = u_i \otimes v_i$ in a UPB is the kernel of the density matrix

$$\rho = \frac{1}{4} \left(I - \sum_{i=1}^{5} w_i w_i^\dagger \right) \tag{6.3}$$

which is proportional to a projection operator. The partial transpose of ρ is

$$\rho^P = \frac{1}{4}\left(I - \sum_{i=1}^{5} \widetilde{w}_i \widetilde{w}_i^\dagger\right) \tag{6.4}$$

with $\widetilde{w}_i = u_i \otimes v_i^*$. Thus both $\rho, \rho^P \geq 0$ by construction. We also see that if the vectors v_i are real we have $\rho = \rho^P$. We can transform the above orthogonal UPB by a product transformation $V = V_A \otimes V_B$ into the unnormalized standard form

$$u = \begin{pmatrix} 1 & 0 & a & b & 0 \\ 0 & 1 & 0 & 1 & a \\ 0 & 0 & b & -a & 1 \end{pmatrix} \qquad v = \begin{pmatrix} 1 & d & 0 & 0 & c \\ 0 & 1 & 1 & c & 0 \\ 0 & -c & 0 & 1 & d \end{pmatrix} \tag{6.5}$$

with a, b, c, d as positive real parameters [55]. Since the UPB vectors w_i in any state of the form (6.3) can be transformed by SL $\times$ SL-transformations into the real standard form (6.5), we realize that all such states are SL-symmetric. The parameters a, b, c, d are determined by the quantities

$$\begin{aligned} s_1 &= -\frac{\det(u_1 u_2 u_4)\ \det(u_1 u_3 u_5)}{\det(u_1 u_2 u_5)\ \det(u_1 u_3 u_4)} = a^2 \\ s_2 &= -\frac{\det(u_1 u_2 u_3)\ \det(u_2 u_4 u_5)}{\det(u_1 u_2 u_4)\ \det(u_2 u_3 u_5)} = \frac{b^2}{a^2} \end{aligned} \tag{6.6}$$

and

$$\begin{aligned} s_3 &= \frac{\det(v_1 v_2 v_3)\ \det(v_1 v_4 v_5)}{\det(v_1 v_2 v_5)\ \det(v_1 v_3 v_4)} = c^2 \\ s_4 &= \frac{\det(v_1 v_3 v_5)\ \det(v_2 v_3 v_4)}{\det(v_1 v_2 v_3)\ \det(v_3 v_4 v_5)} = \frac{d^2}{c^2} \end{aligned} \tag{6.7}$$

These ratios of determinants are invariant under SL $\times$ SL-transformations of the form (3.63). There exist both numerical [55] and analytical evidence [56], that *every* entangled PPT state of rank $(4, 4)$ is SL-equivalent to a state of the form (6.3), with product vectors given by (6.5). The parameters a, b, c, d then define the equivalence classes.

Since the set of $\mathrm{SL}(3, \mathbb{C}) \times \mathrm{SL}(3, \mathbb{C})$-transformations has $16 + 16 = 32$ degrees of freedom, and the four real invariants a, b, c, d add another four, we deduce that the surface of entangled PPT states of rank $(4, 4)$ has dimension 36. In (5.31) we obtained the dimension of the elements of the group $U(N)$ that perform effective rotations of faces (or subspaces) in the full space H_N. The same argument applies of course to general subspaces of dimension r and s such that $N = r + s$. The real dimension of the set of r-dimensional subspaces of an N-dimensional Hilbert space is then

$$d = N^2 - r^2 - s^2 = 2rs \tag{6.8}$$

Thus the set of four-dimensional subspaces of $\mathbb{C}^9$ has dimension 40, and since the surface of entangled PPT states of rank $(4, 4)$ has dimension 36, there cannot exist PPT states

of rank $(4,4)$ in every four-dimensional subspace. The four additional constraints can be seen as the conditions that the parameters a, b, c, d are real and positive.

Separable PPT states of rank $(4,4)$

A separable state of rank $(4,4)$ has the form

$$\rho = \sum_{i=1}^{4} \lambda_i \, (w_i w_i^\dagger) \tag{6.9}$$

with $\lambda_i > 0$, $\sum_{i=1}^{4} \lambda_i = 1$, $w_i^\dagger w_i = 1$ and $w_i = u_i \otimes v_i$. In general any three vectors u_i and any three vectors v_i are linearly independent, and we may perform an SL $\times$ SL-transformation and obtain the standard form

$$u = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \qquad v = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \tag{6.10}$$

In this standard form the v_i are real, so $\rho = \rho^P$. The kernel of a state ρ with this range are all the vectors $z_i = x_i \otimes y_i$ that are orthogonal to all four product vectors w_i in (6.9). Since $\mathrm{Ker}\,\rho$ will contain six product vectors, there will be subsets among the x_i and y_i of three linearly dependent vectors. So a generic five-dimensional subspace will not be $\mathrm{Ker}\,\rho$ for a separable state ρ of rank $(4,4)$. The dimension of the group $\mathrm{SL}(3,\mathbb{C}) \times \mathrm{SL}(3,\mathbb{C})$ is 32, and with three independent coefficients λ_i we see that the surface of separable states of rank $(4,4)$ has dimension 35, one less than the set of entangled PPT states of rank $(4,4)$. So a typical PPT state of rank $(4,4)$ will be entangled (and extremal).

6.2 PPT states of rank $(5,5)$

We have seen that the entangled PPT states of rank $(4,4)$ have a range which is non-generic on the set of four-dimensional subspaces in $\mathbb{C}^9$, and that *all* such states are SL-equivalent to the standard form (6.3). For the case of the entangled PPT states of rank $(5,5)$, *any* generic five-dimensional subspace of $\mathbb{C}^9$ will be the range of such a state, but in addition to the generic cases there also exist a whole range of non-generic cases, which also includes non-extremal states. An attempt to describe how some of these non-generic cases may be understood, is made in Section 6.5.

To describe the generic cases we observe that any given set of five product vectors $w_i = u_i \otimes v_i$ in a generic five-dimensional space may be transformed by an SL $\times$ SL transformation to the standard unnormalized form [57],

$$u = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & p \\ 0 & 0 & 1 & 1 & q \end{pmatrix} \qquad v = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & r \\ 0 & 0 & 1 & 1 & s \end{pmatrix} \tag{6.11}$$

with p, q, r, s generically as complex parameters. By generic we here mean that any three vectors u_i and v_i are linearly independent. There will also be a sixth product vector which is a linear combination of the above five. The parameters p, q, r, s are again determined by ratios of determinants (6.6) and (6.7)

$$s_1 = -\frac{p}{q} \qquad s_2 = q - 1 \qquad s_3 = \frac{r-s}{s} \qquad s_4 = \frac{r}{1-r} \tag{6.12}$$

All the parameters s_i, and therefore also p, q, r, s are invariants in the sense that they cannot be changed by SL $\times$ SL transformations.

Separable rank $(5,5)$ states

In a generic five dimensional subspace of $\mathbb{C}^9$ containing six normalized product vectors $w_i = u_i \otimes v_i$, we may construct a five dimensional set of separable states as convex combinations

$$\rho = \sum_{i=1}^{6} c_i \, (w_i w_i^\dagger) \tag{6.13}$$

with $c_i \geq 0$ and $\sum_i c_i = 1$. Hence all the separable states in the subspace are contained in a simplex with the six pure product states as vertices. The partial transpose of ρ is

$$\rho^P = \sum_{i=1}^{6} c_i \, (\widetilde{w}_i \widetilde{w}_i^\dagger) \tag{6.14}$$

where $\widetilde{w}_i = u_i \otimes v_i^*$ is the partial conjugate of w_i. The six partially conjugated product vectors will be linearly independent in the generic case, hence the separable states in the interior of the simplex will have rank $(5,6)$.

On the facets of the simplex where one coefficient c_i vanishes, ρ will be a rank $(5,5)$ PPT state. In this case five of the product vectors in $\mathrm{Img}\,\rho$ and in $\mathrm{Img}\,\rho^P$ are partial conjugates of each other, whereas the sixth product vectors in the two spaces are related in a more complicated way, unless all the vectors v_i are real.

6.3 The surface of generic PPT states of rank $(5,5)$

A generic five-dimensional subspace of $\mathbb{C}^9$ contains exactly six product vectors $w_i = u_i \otimes v_i$, which can be transformed by SL $\times$ SL-transformations into the standard form (6.11). Thus each such subspace belongs to an equivalence class under SL $\times$ SL-transformations, and each equivalence class corresponds to unique values of the complex invariants p, q, r, s. In total eight real parameters.

We choose two generic five-dimensional subspaces $\mathcal{U}, \widetilde{\mathcal{U}} \subset \mathbb{C}^9$ and then we choose $\mathcal{U} = \mathrm{Img}\,\rho$ and $\widetilde{\mathcal{U}} = \mathrm{Img}\,\rho^P$. Then from (5.13) this defines the Hermitian matrices P and $\widetilde{P}$, and further the operators $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ from (5.17) and (5.21). If we then finally seek solutions for σ of

$$(\mathbf{P} + \widetilde{\mathbf{P}})\sigma = 2\sigma \tag{6.15}$$

where σ is the traceless representation of ρ, then there is typically no solution. Solutions for ρ with $\mathrm{Img}\,\rho = \mathcal{U}$ and $\mathrm{Img}\,\rho^P = \widetilde{\mathcal{U}}$ exist only for special pairs of subspaces $\mathcal{U}$ and $\widetilde{\mathcal{U}}$. If these subspaces admit solutions of (6.15) for ρ then it will generically be one unique (trivial) solution, *i.e.* ρ will be extremal on $\mathcal{P}$. We may instead fix $\mathcal{U} = \mathrm{Img}\,\rho$ and let only $|\widetilde{\mathcal{U}}| = \mathrm{rank}(\rho^P)$ be fixed. There is then always a set of solutions for ρ described by eight real degrees of freedom. The likely role of these eight parameters is to specify the SL-equivalence class to which the five-dimensional subspace $\widetilde{\mathcal{U}} = \mathrm{Img}\,\rho^P$ belongs. The six product vectors in $\mathcal{U}$ are discrete, so they cannot be continuously transformed within this subspace, while it remains fixed. So the only way to vary the subspace $\widetilde{\mathcal{U}} = \mathrm{Img}\,\rho^P$ without changing $\mathcal{U} = \mathrm{Img}\,\rho$ is inevitably to vary the equivalence class of the subspace $\widetilde{\mathcal{U}} = \mathrm{Img}\,\rho^P$.

We may want to fix both $\mathrm{rank}(\rho) = |\mathcal{U}|$ and $\mathrm{rank}(\rho^P) = |\widetilde{\mathcal{U}}|$ to five, but let the subspaces themselves vary. The relevant equation is then (5.26)

$$(\mathbf{Q} + \widetilde{\mathbf{Q}})\sigma = 0 \tag{6.16}$$

In this case the number of linearly independent solutions for σ is 48, and this is then the dimension of the surface of all generic PPT states of rank $(5, 5)$. There are $8 + 8 = 16$ real parameters for SL-equivalence classes of the subspaces $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$, and in addition 32 parameters for the $\mathrm{SL}(3, \mathbb{C}) \times \mathrm{SL}(3, \mathbb{C})$ transformations that determine the specific subspace in each equivalence class.

By (6.8) the set of five-dimensional subspaces has dimension 40, hence one would expect to find an eight-dimensional surface of PPT states of rank $(5, 5)$ in every generic five-dimensional subspace, and this is what we find numerically [57].

In a generic five-dimensional subspace a set of separable states of rank $(5, 5)$ can be constructed. This is a simplex Δ^6 of convex combinations of pure product states $w_i w_i^\dagger$ made up by the six product vectors. The dimension of the surface of PPT states of rank $(5, 5)$ with a fixed range, has however dimension eight. This discrepancy is due to the existence of entangled PPT states of rank $(5, 5)$.

6.4 Tracing the surface

There are several ways of producing walks or tracks on the surface of generic PPT states of rank $(5, 5)$. One possibility is to start at an entangled PPT state ρ of rank $(4, 4)$, produced by using the UPB construction outlined in Section 6.1, and then choose proper directions σ in the perturbation $\rho' = \rho + \epsilon\sigma$, so that ρ' becomes a PPT state of rank $(5, 5)$. Another possibility is to formulate a perturbation $\rho' = \rho + \epsilon\sigma$ as an expansion $\rho(t+\epsilon) = \rho(t) + \epsilon\sigma$ for $\rho = \rho(t)$, and solve the differential equation numerically applying conditions for σ. The starting point $\rho(0)$ is usually chosen as a PPT state of rank $(5, 5)$. A third method is to calculate the eight dimensional tangent space of the surface at ρ, and then make a perturbation $\rho' = \rho + \epsilon\sigma$, where σ is in that tangent space. Then $\mathrm{Img}\,\rho = \mathrm{Img}\,\rho'$, but since in general $\mathrm{Img}\,\rho^P \neq \mathrm{Img}\,\rho'^P$ we have to project ρ' back onto the surface to regain a new PPT state of rank $(5, 5)$.

6.4.1 Perturbing from a rank $(4,4)$ to a rank $(5,5)$ state

We consider the perturbation $\rho' = \rho + \epsilon\sigma$, with ρ as an entangled PPT state of rank $(4,4)$. To perturb these rank $(4,4)$ states is interesting since these states are on the boundary of the surface of PPT states with rank $(5,5)$. If we use the standard form (6.3) with the real product vectors (6.5) we have $\rho = \rho^P$. We again use the projections P and Q defined in (5.13), and the corresponding projections $\widetilde{P}$ and $\widetilde{Q}$ from (5.21). For the special case $\rho = \rho^P$ we have $P = \widetilde{P}$ and $Q = \widetilde{Q}$. From (5.16) the condition for ρ' and ρ'^P to have rank five is

$$Q\sigma Q = \alpha(zz^\dagger) \qquad \widetilde{Q}\sigma^P\widetilde{Q} = \beta(\widetilde{z}\widetilde{z}^\dagger) \tag{6.17}$$

where $\alpha, \beta \neq 0$ are real numbers and

$$z = \sum_{i=1}^{5} c_i w_i \qquad \widetilde{z} = \sum_{i=1}^{5} d_i w_i \qquad c_i, d_i \in \mathbb{C} \tag{6.18}$$

with $w_i \in \operatorname{Ker}\rho$ such that $z^\dagger z = \sum_{i=1}^5 |c_i|^2 = 1$ and $\widetilde{z}^\dagger\widetilde{z} = \sum_{i=1}^5 |d_i|^2 = 1$. Note that if either α or β is allowed to be zero, then ρ' may have rank $(4,5)$ or rank $(5,4)$.

Generically there is one extra sixth product vector $w_6 = \sum_{i=1}^5 a_i w_i \in \operatorname{Ker}\rho$, where $a_i \in \mathbb{R}$. Since the product vectors $w_i = u_i \otimes v_i$ with $u_i, v_i \in \mathbb{R}^3$ for all $i = 1, \ldots, 6$, we have for any Hermitian σ that $w_i^\dagger \sigma w_i = w_i^\dagger \sigma^P w_i$. Using this and (6.17) we get

$$\alpha|c_i|^2 = \beta|d_i|^2 \tag{6.19}$$

for $i = 1, \ldots, 5$, and for the sixth equation

$$\alpha\left|\sum_{i=1}^{5} a_i c_i\right|^2 = \beta\left|\sum_{i=1}^{5} a_i d_i\right|^2 \tag{6.20}$$

From (6.19) and (6.20) it follows that

$$\alpha = \sum_{i=1}^{5} \alpha|c_i|^2 = \sum_{i=1}^{5} \beta|d_i|^2 = \beta \tag{6.21}$$

so in total we have $\alpha = \beta$ and $|c_i| = |d_i|$ for $i = 1, \ldots, 5$. So the coefficients c_i and d_i can differ only by a phase factor for each $i = 1, \ldots, 5$. In total we will have four independent phase factors, since from (6.20) it now follows that $\left|\sum_{i=1}^5 a_i c_i\right| = \left|\sum_{i=1}^5 a_i d_i\right|$.

The vector z which determines the direction $zz^\dagger$ is an arbitrary vector in the five-dimensional subspace $\operatorname{Ker}\rho$, hence after normalization it contains four complex parameters, or eight real parameters. Once z is fixed we can choose from a four parameter family of vectors $\widetilde{z}$, determined by the relative phases between c_i and d_i. Together this gives $8 + 4 = 12$ degrees of freedom. We know that the dimension of the surface of PPT states of rank $(4,4)$ is 36, and this would mean that the subspace of directions that

leads from this surface onto the 48-dimensional surface of PPT states of rank $(5, 5)$, indeed has dimension 12.

For infinitesimal values of ϵ, both ρ' and ρ'^P will have four eigenvalues infinitesimally close to $1/4$ and one eigenvalue $\epsilon\alpha = \epsilon\beta$ very close to zero. With $\alpha > 0$ this means that $\rho', \rho'^P \geq 0$ for $\epsilon > 0$, but not for $\epsilon < 0$. So in this way we never get a PPT state of rank $(4, 5)$ or rank $(5, 4)$, but always rank $(5, 5)$, or for 36 directions a new PPT state of rank $(4, 4)$. This is in complete harmony with numerical investigations [57]. For the more general PPT state of rank $(4, 4)$ obtained by SL $\times$ SL-transformations from the standard form (6.3), the smallest positive eigenvalues of ρ' and ρ'^P are no longer equal, but still tied together so that they approach zero simultaneously.

As described in Section 5.7 we may also restrict the class of perturbations, so that we fix the five-dimensional subspace $\mathrm{Img}\,\rho'$ to be the direct sum of the four-dimensional subspace $\mathrm{Img}\,\rho$ and the one-dimensional subspace of the vector z. Apart from the trivial solution $\sigma = \rho$, we then find five linearly independent possibilities, out of which four give ρ' as a rank $(4, 4)$ PPT state. The fifth is then a unique solution giving ρ' as a rank $(5, 5)$ PPT state. So the dimension of the surface of $(4, 4)$ PPT states with range within the fixed five-dimensional subspace $\mathrm{Img}\,\rho$, is four. When we fix $\mathrm{Img}\,\rho'$ and look for PPT states of rank $(4, 4)$ with range within this space, we eliminate all degrees of freedom corresponding to SL $\times$ SL-transformations. But we still allow variations of the four real invariant parameters a, b, c, d needed to define a rank $(4, 4)$ PPT state.

It is still intriguing that the rank $(5, 5)$ PPT states with a fixed range $\mathrm{Img}\,\rho$, which is an eight-dimensional set, has a set of boundary states that is four-dimensional.

6.4.2 Curves on the surface by integration

We will here look at a method for tracing curves on a surface of PPT states of fixed rank (m, n). The idea is to formulate the perturbation $\rho' = \rho + \epsilon\sigma$ as an expansion $\rho(t + \epsilon) = \rho(t) + \epsilon\sigma$ for $\rho = \rho(t)$, and solve the differential equation that follows from this. The perturbation expansion $\rho(t + \epsilon) = \rho(t) + \epsilon\sigma$ for $\rho = \rho(t)$ is equivalent to the differential equation

$$\frac{\mathrm{d}\rho}{\mathrm{d}t} = \dot{\rho} = \sigma \tag{6.22}$$

Since a matrix $\rho = \sum_{i,\lambda_i \neq 0} \lambda_i \left(\psi_i \psi_i^\dagger\right)$ of less than full rank has no well defined inverse ρ^{-1} it is useful to introduce the *pseudoinverse* of ρ as

$$\rho^+ = \sum_{i,\lambda_i \neq 0} \lambda_i^{-1} \psi_i \psi_i^\dagger \tag{6.23}$$

If ρ is of full rank then $\rho^+ = \rho^{-1}$. The projection P onto $\mathrm{Img}\,\rho$ is $P = \rho^+\rho = \rho\rho^+$. There are similar relations for the projection $\widetilde{P}$ onto $\mathrm{Img}\,\rho^P$ and the pseudoinverse of ρ^P.

If we have $X \in H_N$ with $\mathrm{Img}\,X \subset \mathrm{Img}\,\rho$ then $X = PX = XP$ and $QX = XQ = 0$. We assume that this holds for all t so that

$$\dot{X} = \dot{P}X + P\dot{X} = \dot{X}P + X\dot{P} \tag{6.24}$$

or more useful

$$Q\dot{X} = \dot{P}X \qquad \dot{X}Q = X\dot{P} \tag{6.25}$$

which gives

$$\dot{X} = (P+Q)\dot{X}(P+Q) = P\dot{X}P + X\dot{P} + \dot{P}X \tag{6.26}$$

Now setting $X = \rho$ in (6.25) and multiplying from right and left by the pseudoinverse gives $Q\sigma\rho^+ = \dot{P}P$ and $\rho^+\sigma Q = P\dot{P}$. Since differentiation of $P = P^2$ gives $\dot{P} = \dot{P}P + P\dot{P}$ we get

$$\dot{P} = Q\sigma\rho^+ + \rho^+\sigma Q \tag{6.27}$$

and likewise we use differentiation of $\rho^+ = \rho^+\rho\rho^+$ to get

$$\dot{\rho}^+ = \dot{\rho}^+P + \rho^+\sigma\rho^+ + P\dot{\rho}^+ \tag{6.28}$$

which when multiplied with P on both sides gives $P\dot{\rho}^+P = -\rho^+\sigma\rho^+$. If we use $X = \rho^+$ in (6.26) together with (6.27) we get

$$\dot{\rho}^+ = Q\sigma(\rho^+)^2 + (\rho^+)^2\sigma Q - \rho^+\sigma\rho^+ \tag{6.29}$$

The differential equation $\dot{\rho} = \sigma$ may be integrated together with (6.27) and (6.29) as long as we accompany this with proper specifications for how to calculate the direction σ for each ρ. For the partial transpose ρ^P there are completely similar equations relating the projections $\widetilde{P}$ and $\widetilde{Q}$ with ρ^P and $(\rho^P)^+$ that must be satisfied.

We can now generate a curve $\rho(t)$ on the 48-dimensional surface of PPT states of rank $(5,5)$. Since this essentially is a differential equation we will find a curve containing the state $\rho(0)$ as a starting point. The constraints are then again (5.26)

$$(\mathbf{Q} + \widetilde{\mathbf{Q}})\sigma = 0 \tag{6.30}$$

We may also want to generate a curve that lies on the eight-dimensional surface of PPT states of rank $(5,5)$, but with $\mathrm{Img}\,\rho$ fixed, while $\mathrm{Img}\,\rho^P$ is allowed to change. We then have the condition

$$(\mathbf{I} - \mathbf{P} + \widetilde{\mathbf{Q}})\sigma = 0 \tag{6.31}$$

again. In this case P is constant, but $\widetilde{Q} = \widetilde{Q}(t)$ would then change.

We have done numerical integrations to generate such curves. The eigenvalues of ρ and ρ^P stay remarkably similar along the curve, yet they are not identical. The condition that at least one eigenvalue of ρ (and ρ^P) goes to zero, defines the boundary of the surface. When ρ (and ρ^P) get one dominant eigenvalue, we interpret this as an indication that ρ (and ρ^P) approach a pure product state.

6.4.3 Tracing the surface by repeated projections

If we choose any generic rank $(5,5)$ PPT state ρ, which then fixes the subspace $\mathrm{Img}\,\rho$, we can calculate the eight linearly independent directions for the perturbation $\rho' = \rho + \epsilon\sigma$ that solve (6.31). Equation (6.31) guarantees that $\mathrm{Img}\,\rho' = \mathrm{Img}\,\rho$ for small finite values

of ϵ, and that $\text{rank}(\rho'^P) = \text{rank}(\rho^P)$ for infinitesimal ϵ. Thus the eight solutions of (6.31) define (at ρ) a tangent plane to the eight-dimensional surface of PPT states we want to track. We choose σ to lie in this plane, and then a finite value of ϵ and compute $\rho' = \rho + \epsilon\sigma$. Then ρ' will have the same range as ρ, but ρ'^P will only have approximately the same rank as ρ^P, in fact ρ'^P will in general have full rank.

To project ρ' back onto the surface of rank $(5, 5)$ PPT states, whilst keeping the range of ρ, we use an iterative procedure. When we compute the nine eigenvalues $\lambda_1, \ldots, \lambda_9$ of ρ'^P we use the eigenvectors $e_1, \ldots, e_5$ corresponding to the five dominant eigenvalues $\lambda_1, \ldots, \lambda_5$ to construct a projection operator onto the subspace spanned by these e_n

$$\widetilde{P}_5 = \sum_{n=1}^{5} e_n e_n^\dagger \tag{6.32}$$

For sufficiently small steps ϵ the corresponding change $\Delta\lambda_n(\epsilon)$ in the eigenvalues will also be small, so that the eigenvectors e_n in (6.32) are still the dominant eigenvalues.

The projection P is onto $\text{Img}\,\rho$ throughout the whole procedure, thus ensuring that the ranges of ρ, ρ' and the new state ρ'' are the same. Projecting ρ'^P onto the range defined by its five dominant eigenvalues with the operator $\widetilde{P}_5$ insures that $\text{rank}(\rho'^P)$ becomes five again. Unfortunately these operations do not work at the same time since adjusting ρ'^P changes ρ' and vice versa. Repeating the procedure

$$\rho'_{n+1} = \gamma P \left[\rho'_n + \{ \widetilde{P}_5 \rho'^P_n \widetilde{P}_5 \}^P \right] P \tag{6.33}$$

a sufficient number of times, projects the state ρ' with good precision to a new state ρ'' on the surface of rank $(5, 5)$ PPT states. The projection P is thus kept throughout the iterative procedure, but $\widetilde{P}_5$ is updated for every iteration. What effectively happens is that since $\{ \widetilde{P}_5 \rho'^P_n \widetilde{P}_5 \}^P \neq \rho'_n$, an average of these two states is projected by P so that $\text{Img}\,\rho'_n = \text{Img}\,\rho'_{n+1}$ for each iteration. The factor γ is a normalization factor, since the projections generally do not preserve the norm. This method to track the surface is in some ways preferable to ordinary perturbation methods, where the step size in general is much smaller, and progress therefore much slower.

For many generic five-dimensional subspaces, we have numerically constructed a large number of extremal PPT states $\rho_1, \ldots, \rho_k$ of all ranks ≤ 5 in the same subspace. We have then utilized the above scheme to walk from a state ρ_1 to all the other states $\rho_2, \ldots, \rho_k$. It is reasonable to assume that this tracking between states is transitive, *i.e.* if ρ_1 is connected to ρ_2 and ρ_2 connected to ρ_3, then ρ_1 is connected to ρ_3. Based on this we conclude that for generic cases all the states in the same subspace lie on one connected surface.

We have also used this method to track the curvature of the eight-dimensional surface of rank $(5, 5)$ PPT states with a fixed range. As seen in Figure 6.2, each step generated along the curve by our tracking scheme is inherently two-dimensional.

For each step we know the state ρ, ρ' and ρ'', so the distance between each state can be calculated using the Hilbert-Schmidt metric

$$|\rho_i - \rho_j|^2 = \text{tr}\left[(\rho_i - \rho_j)^2 \right] \tag{6.34}$$

The curvature is given by $\kappa = |d\theta/ds|$. Since all sides of the triangle are known we can calculate the angle $d\theta$ for each step, and if the curvature is sufficiently small we can assume to a very good approximation that $ds \cong |\rho - \rho''|$.

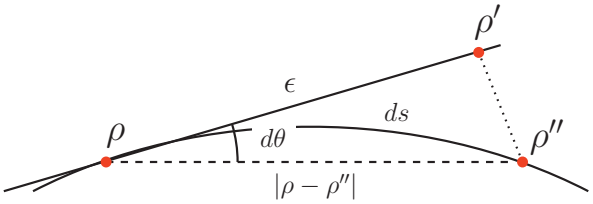


Figure 6.2: The eight-dimensional surface of rank $(5, 5)$ PPT states with a fixed range, as a two-dimensional projection. A finite perturbation $\rho \mapsto \rho'$ is made along the tangent plane to the surface at ρ. The projection scheme (6.33) is then made so that ρ'' is a rank $(5, 5)$ PPT state with the same range as ρ. The changes $d\theta$ and ds are then used to calculate the local curvature. The curvature portrayed here is extensively exaggerated.

In general, for sufficiently small steps ϵ we have $|\rho - \rho'| \simeq |\rho - \rho''| \gg |\rho' - \rho''|$, so the triangle is to very good approximation a *skinny* triangle. This method will not return good values when the curvature becomes to large because then $ds \not\simeq |\rho - \rho''|$, or when the steps become too small because then the calculation of the angle $d\theta$ is compromised.

We can use this method to calculate the curvature along a curve between a pure product state and an extremal rank $(5, 5)$ PPT state on the surface, and this is pictured in Figure 6.3.

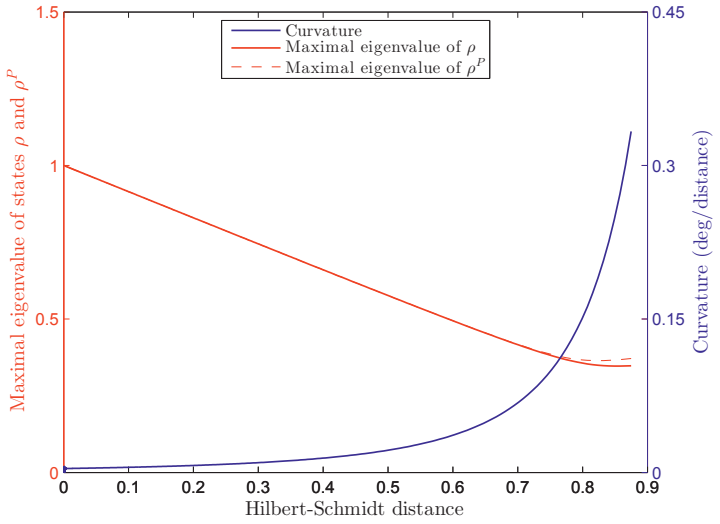


Figure 6.3: We see the curvature $\kappa = |d\theta/ds|$ for the curve generated by tracking the surface of rank $(5, 5)$ PPT states within a fixed range. The tracking is between one of the six pure product states in the subspace and an extremal rank $(5, 5)$ PPT state on the interior of the surface. Also plotted are the maximum eigenvalues of ρ and ρ^P.

We note that the curvature is quite small along the entire length, and this appears to be a general feature for these cases. The estimates of the curvature very close to the pure product state are very uncertain, and it is quite possible that the derivative $d\theta/ds$ has a discontinuity at the product state.

We can also walk from an extremal rank $(5, 5)$ PPT state towards an extremal rank $(4, 4)$ PPT state on the boundary of the eight-dimensional surface. This is seen in Figure 6.4. In this case, the curvature of the surface increases towards the boundary.

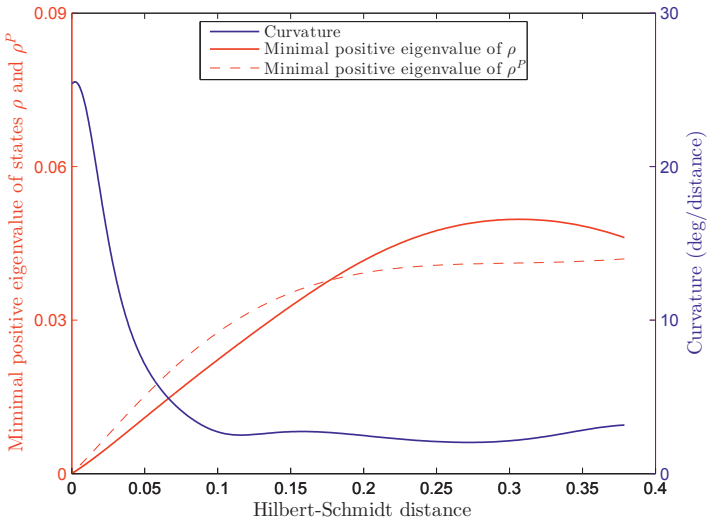


Figure 6.4: The curvature $\kappa = |d\theta/ds|$ for the curve generated by tracking the surface of rank $(5, 5)$ PPT states within a fixed range space is shown. The tracking connects an extremal rank $(5, 5)$ PPT state on the interior of the surface, and an extremal rank $(4, 4)$ PPT state on the boundary of the surface. Also plotted are the minimal positive eigenvalues of ρ and ρ^P.

6.5 Non-generic PPT states of rank $(5, 5)$

A very important characteristic of PPT states is the number of product vectors in the range and the kernel of the state. We have seen that the dimension of the set of separable states of rank $(5, 5)$ is smaller than the dimension of PPT states of same rank, so a generic PPT state of rank $(5, 5)$ will be entangled. It is also a fact that the range of such states contains six product vectors, while the kernel contains none.

6.5.1 Non-generic orthogonal subspaces

From Section 3.6 we recollect the use of the notation $\{n_{\text{img}}, \widetilde{n}_{\text{img}}; n_{\text{ker}}\}$ for a density matrix ρ. Here n_{img} is the number of product vectors in $\operatorname{Img}\rho$, $\widetilde{n}_{\text{img}}$ is the number of product vectors in $\operatorname{Img}\rho^P$ and n_{ker} is number of product vectors in $\operatorname{Ker}\rho$. Like in Section 3.6.3, for a density matrix ρ we let the product vectors in $\operatorname{Img}\rho$ be

$$w_i = u_i \otimes v_i \qquad i = 1, \ldots, n_{\text{img}} \tag{6.35}$$

and likewise for $\text{Ker}\,\rho$:

$$z_j = x_j \otimes y_j \qquad j = 1, \ldots, n_{\text{ker}} \tag{6.36}$$

We know that $z_i = x_i \otimes y_i \in \text{Ker}\,\rho$ if and only if $\widetilde{z}_i = x_i \otimes y_i^* \in \text{Ker}\,\rho^P$, so the number of product vectors in $\text{Ker}\,\rho$ is always equal to the number in $\text{Ker}\,\rho^P$, or more formally $n_{\text{ker}} = \widetilde{n}_{\text{ker}}$.

By definition, for a generic set of vectors in $\mathbb{C}^3$ any subset of three vectors will be linearly independent. A non-zero vector $x \in \mathbb{C}^3$ can at most be orthogonal to two vectors u_i, and a non-zero vector $y \in \mathbb{C}^3$ can at most be orthogonal to two v_i, hence the product vector $z = x \otimes y$ can at most be orthogonal to four w_i. Since $\text{Ker}\,\rho = (\text{Img}\,\rho)^\perp$ and since the range of a generic rank $(5, 5)$ PPT state ρ contains six product vectors $w_i = u_i \otimes v_i$, it is clear that it is not possible to have a product vector in $\text{Ker}\,\rho$ in this case. Thus generic states must have $n_{\text{ker}} = 0$.

In order to construct pairs of orthogonal subspaces $\mathcal{U} \subset \mathbb{C}^9$ and $\mathcal{V} = \mathcal{U}^\perp \subset \mathbb{C}^9$ with $|\mathcal{U}| = 5$ and $|\mathcal{V}| = 4$, such that $\mathcal{V}$ contains one or more product vectors, we must alter the generic linear dependencies of the vectors u_i and v_i. Instead of the generic condition that any subset with *three* vectors $u_i \in \mathbb{C}^3$ and $v_i \in \mathbb{C}^3$ must be linearly independent, we introduce the conditions that any subset with *four* vectors must be linearly independent. Only by appealing to these non-generic characteristics is it possible to construct the necessary orthogonality relations, as described in Section 3.6.3.

In the above sense it is to a certain degree the *subspaces* $\mathcal{U}$ and $\mathcal{V}$ that are relevant here, and to a lesser degree the states themselves. A general characterization of two orthogonal subspaces $\mathcal{U}$ and $\mathcal{V}$ with regard to the number of product vectors they contain, can be written as $\{n_u; n_v\}$.

Since $z_j = x_j \otimes y_j \in \text{Ker}\,\rho$ if and only if $\widetilde{z}_j = x_j \otimes y_j^* \in \text{Ker}\,\rho^P$ it is clear that the kernels of ρ and ρ^P are related when they contain product vectors. In particular, if y_j is real then $z_j = \widetilde{z}_j$. As long as $n_{\text{ker}} \leq 4$ we can always choose a standard form where all the vectors x_j and y_j are real.

By making use of the non-generic structures discussed here, it is possible to define standard forms for several combinations of orthogonal non-generic subspaces $\mathcal{U}$ and $\mathcal{V}$, *i.e.* with various $\{n_u; n_v\}$. In Paper 4 we have constructed a number of such standard forms, and then produced PPT states of rank $(5, 5)$ with $\text{Img}\,\rho = \mathcal{U}$ and $\text{Ker}\,\rho = \mathcal{V}$ for $0 < n_{\text{ker}} \leq 4$.

6.5.2 The case $n_{\text{ker}} = 4$

As one example, we take the case for $n_{\text{ker}} = 4$. Given a PPT state ρ and four product vectors $z_j = x_j \otimes y_j$ in the kernel of ρ, in some definite but arbitrary order. We assume that any three vectors $x_j \in \mathbb{C}^3$ and any three vectors $y_j \in \mathbb{C}^3$ are linearly independent. Then we may perform a product transformation, and subsequent normalizations, so that the vectors take the form

$$x = y = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \tag{6.37}$$

and for the product vectors z_j

$$z = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \tag{6.38}$$

The transformation is unique. In the space spanned by the z_j there exist no other product vectors. The real form of the $z_j \in \mathrm{Ker}\,\rho$ implies that $z_j \in \mathrm{Ker}\,\rho^P$.

It is equally easy to see that there exist exactly six product vectors $w_i = u_i \otimes v_i$ in the orthogonal subspace. In fact, in order to have $(u_i \otimes v_i) \perp (x_j \otimes y_j)$ for all $i = 1, \ldots, 6$ and $j = 1, \ldots, 4$, we must have for each pair i, j that either $u_i \perp x_j$ or $v_i \perp y_j$. Since any three vectors x_j and any three vectors y_j are linearly independent, a vector u_i can be orthogonal to at most two vectors x_j, and a vector v_i can be orthogonal to at most two vectors y_j. This gives the six possibilities for orthogonality listed in the table.

$u_i \otimes v_i$	$u_i \perp x_k, x_l$	$v_i \perp y_m, y_n$
i	k, l	m, n
1	2, 3	1, 4
2	1, 3	2, 4
3	1, 2	3, 4
4	1, 4	2, 3
5	2, 4	1, 3
6	3, 4	1, 2

Table 6.1: Possibilities for a product vector $u_i \otimes v_i$ to be orthogonal to all four product vectors $x_j \otimes y_j$.

The unique solution is the following list of vectors,

$$u = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & -1 & 0 \end{pmatrix} \qquad v = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{6.39}$$

and for the product vectors w_i

$$w = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.40}$$

The five-dimensional subspace $\operatorname{Img}\rho$ spanned by the w_i given by (6.40) defines a face $\mathcal{F} \subset \mathcal{D}$ of dimension $5^2 - 1 = 24$. Recall that

$$\rho z_j = \rho^P z_j = 0 \qquad j = 1, \ldots, 4 \tag{6.41}$$

In terms of the face $\mathcal{F}$, the above means that $\rho \in \mathcal{F}$ and $\rho^P \in \mathcal{F}$, or equivalently, that $\rho \in \mathcal{F} \cap \mathcal{F}^P$. Effectively (6.41) restricts ρ to have the form (6.43), with $c_i \in \mathbb{R}$.

Note that $\rho = \rho^P$. Since $\mathcal{F}$ is a face on $\mathcal{D}$, we understand that $\mathcal{F}^P$ is a face on $\mathcal{D}^P$, and the intersection $\mathcal{G} = \mathcal{F} \cap \mathcal{F}^P$ is a face on $\mathcal{P} = \mathcal{D} \cap \mathcal{D}^P$. Equation (6.43) shows that the face $\mathcal{G}$ has dimension five.

The state ρ defined in (6.43) is a linear combination, but not necessarily a convex combination, of the six pure product states,

$$\rho = \frac{1}{2} \sum_{i=1}^{6} c_i \, (w_i w_i^\dagger) \tag{6.42}$$

$$\rho = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_1 + c_5 & -c_1 & 0 & 0 & 0 & 0 & -c_5 & 0 \\ 0 & -c_1 & c_1 + c_4 & 0 & 0 & -c_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_2 + c_6 & 0 & -c_2 & -c_6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -c_4 & -c_2 & 0 & c_2 + c_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & -c_6 & 0 & 0 & c_3 + c_6 & -c_3 & 0 \\ 0 & -c_5 & 0 & 0 & 0 & 0 & -c_3 & c_3 + c_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{6.43}$$

From (6.43) we observe that the normalization condition for ρ becomes

$$\operatorname{Tr}\rho = \sum_{i=1}^{6} c_i = 1 \tag{6.44}$$

Since the matrices $w_i w_i^\dagger$ are linearly independent, and there are no other pure product states in $\operatorname{Img}\rho$, the state ρ is separable if and only if all the c_i are non-negative. However,

we will now see that it is possible for ρ to be an entangled PPT state even if one of the coefficients is negative, and this observation will lead us to a scheme for constructing the extremal PPT states of rank $(4, 4)$, a scheme which does not rely on the UPB method discussed briefly in Section 6.1.

An eigenvalue of ρ, and of $\rho^P = \rho$, is a root of the characteristic polynomial

$$\det(\rho - \lambda I) = -\lambda^4 f(\lambda) \tag{6.45}$$

with

$$f(\lambda) = (\lambda^5 - d_4\lambda^4 + d_3\lambda^3 - d_2\lambda^2 + d_1\lambda - d_0) \tag{6.46}$$

The constant term in $f(\lambda)$ is

$$d_0 = \frac{3}{16} \sum_{i=1}^{6} \prod_{j \neq i} c_j = \frac{3}{16} \Big(\prod_{j=1}^{6} c_j \Big) \sum_{i=1}^{6} \frac{1}{c_i} \tag{6.47}$$

If we start with $c_i > 0$ for $i = 1, \ldots, 6$, then ρ is a rank $(5, 5)$ separable state. If we then change the coefficients continuously, ρ will continue to have five positive eigenvalues until we get $d_0 = 0$. Hence $d_0 = 0$ defines the boundary of the set of density matrices, and also of the set of PPT states since $\rho = \rho^P$.

We know that the boundary is not reached before at least one coefficient becomes zero or negative. If two coefficients become zero simultaneously, then $d_0 = 0$ and we have reached a boundary state which is separable. To get negative coefficients while ρ is a rank $(5, 5)$ PPT state we have to make one coefficient negative before the others. Let us say, for example, that $c_1 < 0$, and that we want to make also c_2 negative, while $c_i > 0$ for $i = 3, \ldots, 6$. Then we first have to make $c_2 = 0$, in which case $d_0 = 3\, c_1 c_3 c_4 c_5 c_6/16 < 0$ and we have already crossed the boundary $d_0 = 0$.

In conclusion, the entangled boundary states have $c_i \neq 0$ for $i = 1, \ldots, 6$, and they have one negative and five positive coefficients c_i satisfying

$$\sum_{i=1}^{6} \frac{1}{c_i} = 0 \tag{6.48}$$

Thus the boundary $d_0 = 0$ consists of two types of states.

1. Separable states that are convex combinations of up to four of the pure product states $w_i w_i^\dagger$.

2. Rank $(4, 4)$ entangled PPT states that are linear combinations of all the six pure product states $w_i w_i^\dagger$ with one negative coefficient.

Figure 6.5 shows a two dimensional section through the five dimensional face of $\mathcal{P}$ defined by (6.42) and (6.44).

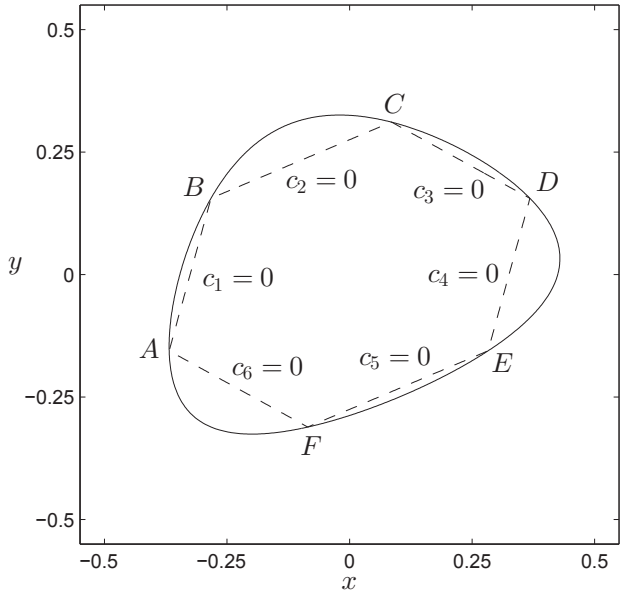


Figure 6.5: A two dimensional cut through the five dimensional face on $\mathcal{P}$ of states given by (6.43). The outer curve is the common boundary of $\mathcal{D}$ and $\mathcal{P}$, which consists of rank $(4,4)$ PPT states. The cut through the simplex in $\mathcal{S}$ is the hexagon with corners A to F. The region between the two curves consists of non-extremal entangled PPT states of rank $(5,5)$. The coordinates x, y are dimensionless.

The section through the separable states $\mathcal{S}$, is seen as the hexagon with corners A to F. On the boundary of the hexagon (dashed), exactly one of the coefficients c_i is zero. The region between the two curves consists of non-extremal entangled PPT states of rank $(5,5)$ with exactly one coefficient negative. The hexagon is reflection symmetric about two axes. In Table 6.2 are listed the coefficients c_i that define the states A to F by (6.42), multiplied by 12.

i	A	B	C	D	E	F
1	0	0	3	6	6	3
2	1	0	0	1	2	2
3	6	3	0	0	3	6
4	2	2	1	0	0	1
5	3	6	6	3	0	0
6	0	1	2	2	1	0

Table 6.2: The coefficients c_i, multiplied by 12, for the states A to F in Figure 6.5.

The most general rank $(4,4)$ entangled PPT states

Consider a general rank $(4,4)$ entangled PPT state ρ in 3×3 dimensions. It is known that any such state is extremal, and has exactly six product vectors in its kernel. We

can now see that it is SL-equivalent, in no less than 360 different ways, to such states on the boundary of the five dimensional face of $\mathcal{P}$ that we have described here. The 360 transformations are found in the following way.

Pick any four of the six product vectors in $\mathrm{Ker}\,\rho$, this can be done in 15 different ways. Order them next in one of the 24 possible ways. Altogether there are $24 \cdot 15 = 360$ possibilities. There is then a unique product transformation that transforms the four product vectors to the form given in (6.37). We know that it must transform the state ρ into one of the rank $(4, 4)$ states described by Equations (6.42), (6.44), and (6.48), since these are the only rank $(4, 4)$ entangled PPT states of this form.

In conclusion, the scheme derived in this section represents an alternative method for producing the extremal PPT states of rank $(4, 4)$. It is alternative in the sense that it does not rely on the UPB construction.

6.5.3 Other values of $n_{\mathbf{ker}}$

While we have seen that PPT states constructed from (6.37) and (6.39) are exclusively non-extremal, there are standard forms that generically produces extremal states.

We can for example construct a standard form for which $n_{\mathrm{ker}} = 1$. This standard form also has $n_{\mathrm{img}} = 6$, and generically produces extremal $\{6, 6; 1\}$ PPT states. For $z_1 = x_1 \otimes y_1 \in \mathrm{Ker}\,\rho$ and $w_i = u_i \otimes v_i \in \mathrm{Img}\,\rho$ with $i = 1, \ldots, 6$, we now have a more limited set of orthogonality relations

$$x_1 \perp u_i \qquad i = 1, 2, 3 \tag{6.49}$$

$$y_1 \perp v_i \qquad i = 4, 5, 6 \tag{6.50}$$

Now $z_1 = x_1 \otimes y_1$ is the only product vector allowed in $\mathrm{Ker}\,\rho$. It is also possible to produce standard forms which allow an infinite number of product vectors in the range. This is possible if there for a subset of the $w_i \in \mathrm{Img}\,\rho$ exist linear combinations $\sum_i c_i w_i$ which are again product vectors, and where these linear combinations represent an underdetermined set of equations. In Paper 4 we have constructed two standard forms with $n_{\mathrm{ker}} = 2$, which generically produce $\{6, 6; 2\}$ and $\{\infty, \infty; 2\}$ states. The former are exclusively extremal states, while the latter states are both extremal and non-extremal. Also, the form presented in Paper 4 which has $n_{\mathrm{ker}} = 3$ produces only non-extremal states.

There is however no reason to limit the possible values of n_{img} to six or infinity. In our random searches for SL-symmetric PPT states of rank $(5, 5)$ we have found, by complete accident, several non-extremal $\{2, 2; 1\}$-states. It is unclear why this combination has appeared several times. Using the subspaces from these states we have produced extremal states with the same range, and these are $\{2, 6; 1\}$-states.

It appears that for any five dimensional subspace of $\mathbb{C}^9$ to contain less than six product vectors, the set of equations (3.52) must have degenerate solutions. In this sense the only possibilities are $1 < n_{\mathrm{img}} \leq 6$ and $n_{\mathrm{img}} = \infty$. An analysis regarding these matters is given in [56].

Chapter 7

Positive maps and entanglement witnesses

We start by defining entanglement witnesses, some of their properties and why they are useful and important. We then move on to a different type of mathematical objects, namely maps and more specifically positive maps. The one-to-one correspondence between witnesses and positive maps, the Choi-Jamiolkowski isomorphism, is then described. Completely positive maps is defined, and the important distinction between decomposable and non-decomposable maps, or equivalently decomposable and non-decomposable witnesses, is made. The zeros of a witness W are defined as the product vectors $\phi \otimes \chi$ satisfying $(\phi \otimes \chi)^\dagger W(\phi \otimes \chi) = 0$. These zeros are used to define boundary witnesses and extremality for witnesses. An important characteristic here is whether a zero is quadratic or quartic. A method to search for generic extremal witnesses with quadratic zeros is outlined, and some of the results are discussed. Positive maps with the unital and trace preserving property is discussed, and a method for transforming a positive map to this form is constructed. Visualizations of how extremal positive maps act on various two-dimensional cross sections in $\mathcal{D}_3$, are performed and discussed. The facial structure of the set of separable states $\mathcal{S}$ is closely related to the facial structure of the set of entanglement witnesses $\mathcal{S}^\circ$. This relation is used to develop an understanding of the faces of both these sets. Optimal entanglement witnesses is then defined and discussed, and an attempt to understand the link between extremality and optimality for entanglement witnesses is then made. Finally, a description of the so-called SPA separability conjecture is included.

7.1 Entanglement witnesses

We know that a density matrix $\rho \in H_N = H_{N_A} \otimes H_{N_B}$ on a Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ is separable if and only if it is of the form

$$\rho = \sum_k p_k \left(\mu_k \otimes \tau_k\right) \tag{7.1}$$

where $\mu_k \in \mathcal{D}_{N_A}$, $\tau_k \in \mathcal{D}_{N_B}$ with $p_k > 0$ and $\sum_k p_k = 1$. The partial transpose

$$\rho^P = \sum_k p_k \, (\mu_k \otimes \tau_k^T) \tag{7.2}$$

is then always positive, *i.e.* $\mathcal{S} \subseteq \mathcal{P}$. This gives rise to the Peres-Hordecki separability criterion discussed in Section 3.5.2. Since for $N \geq 8$ there exist entangled PPT states, an alternative criterion involving entanglement witnesses is more useful.

We start by defining the dual set of the separable states $\mathcal{S}$

$$\mathcal{S}^\circ = \{W \in H_N \,|\, \mathrm{tr}(W\rho) \geq 0 \;\; \forall \, \rho \in \mathcal{S}\} \tag{7.3}$$

The dual of $\mathcal{S}^\circ$ is $\mathcal{S}$, or $\mathcal{S}^{\circ\circ} = \mathcal{S}$, and thus a state $\rho \in \mathcal{D}$ is separable if and only if $\mathrm{tr}(W\rho) \geq 0$ for all $W \in \mathcal{S}^\circ$. Thus, if any $W \in \mathcal{S}^\circ$ exists such that $\mathrm{tr}(W\rho) < 0$, we can deduce that the state ρ must be entangled. The Hermitian operator W is then an *entanglement witness* for ρ. We state the Horodecki criterion, originally given by Michal, Pawel and Ryszard Horodecki [33], in the form

Theorem 7.1 (Horodecki criterion)**.** *For any entangled state $\rho \in H_N$ on the Hilbert space $\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B$ there exists an Hermitian operator W such that*

$$\mathrm{tr}(W\rho) < 0 \quad \textit{and} \quad \mathrm{tr}(W\tau) \geq 0 \tag{7.4}$$

for all states $\tau \in \mathcal{S}$.

This two-way implication makes entanglement witnesses powerful tools for detecting entanglement experimentally, and such schemes have been performed [58, 59]. Note that this testimony of an entanglement witness is of a statistical nature, in that it is the expectation value $\mathrm{tr}(W\rho)$ which is used. It should also be emphasized that there is no universal entanglement witness. Each entangled state is revealed by its own subset of $\mathcal{S}^\circ$, and this fact leads to the notion of *optimal entanglement witnesses*, to which we shall return briefly in Section 7.6.

Note that since the set of entanglement witnesses $\mathcal{S}^\circ$ is here defined as the dual cone of $\mathcal{S}$, this implies that it is a closed and convex set. It is therefore completely characterized by its extremal points, and the objective must be to understand the structure of these extremal points.

The extremal points of $\mathcal{S}$ are known to be the pure product states $\rho = \psi\psi^\dagger = \phi\phi^\dagger \otimes \chi\chi^\dagger$ with $\phi^\dagger\phi = \chi^\dagger\chi = 1$ for normalization. A matrix $M \in H_N$ can then be an entanglement witness $W \in \mathcal{S}^\circ$ if and only if its expectation value in every such pure product state

$$\mathrm{tr}(W\rho) = \psi^\dagger W \psi = (\phi \otimes \chi)^\dagger W (\phi \otimes \chi) = \sum_{ij;kl} \phi_i^* \chi_j^* \, W_{ij;kl} \, \phi_k \chi_l \tag{7.5}$$

is non-negative. Here ϕ_n, χ_n are the components of $\phi \in \mathcal{H}_A$ and $\chi \in \mathcal{H}_B$ respectively. Thus the condition that $W \in \mathcal{S}^\circ$ is that the biquadratic form (7.5) is non-negative for all $\phi \in \mathcal{H}_A$ and $\chi \in \mathcal{H}_B$

$$f_W(\phi, \chi) = \sum_{ij;kl} \phi_i^* \chi_j^* \, W_{ij;kl} \, \phi_k \chi_l \geq 0 \tag{7.6}$$

It should be added that since $f_{W^P}(\phi,\chi) = f_W(\phi,\chi^*)$ we have that $W^P \in \mathcal{S}^\circ$ if and only if $W \in \mathcal{S}^\circ$.

Non-negative polynomials

The problem (7.6) is related to a set of larger fundamental problems in mathematics. On Hilbert's list of 23 unsolved problems published in 1900 [60], the 17th problem is whether every non-negative real polynomial $f(x_1,\ldots,x_n)$ is a sum of squares of rational functions

$$f(x_1,\ldots,x_n) = \sum_k \left[g_k(x_1,\ldots,x_n)\right]^2 \tag{7.7}$$

In 1888 Hilbert had himself disproved a weaker version of the 17th problem, namely that there are non-negative polynomials that are not sums of squares of polynomials. Artin finally proved in 1927 that the answer to the 17th problem is yes [61], but it took another 39 years before Motzkin [62] gave the first explicit example of such a polynomial in 1966

$$M(x_1,x_2,x_3) = x_3^6 + x_1^4x_2^2 + x_1^2x_2^4 - 3x_1^2x_2^2x_3^2 \tag{7.8}$$

Such a function as a biquadratic form was discussed by Choi in 1975 [63], and an example of such a function as a biquadratic form, *i.e.* of the type (7.6), was given in 1977 by Choi and Lam [64]

$$\begin{aligned} f(x,y) = {} & (x_1^2+x_2^2+x_3^2)(y_1^2+y_2^2+y_3^2) - x_1^2y_3^2 - x_2^2y_1^2 - x_3^2y_2^2 \\ & - 2(x_1x_2y_1y_2 + x_2x_3y_2y_3 + x_3x_1y_3y_1) \end{aligned} \tag{7.9}$$

which we shall encounter in some detail later.

7.2 Positive maps

It is possible to regard a Hermitian matrix $M \in H_N = H_{N_A} \otimes H_{N_B}$ as a linear transformation $\mathbf{L}_M$ from H_{N_A} to H_{N_B}. We define the real linear map $\mathbf{L}_M : H_{N_A} \mapsto H_{N_B}$ written as $Y = \mathbf{L}_M X$ with

$$Y_{jl} = \sum_{i,k} M_{ij;kl} X_{ki} \tag{7.10}$$

where $X \in H_{N_A}$ and $Y \in H_{N_B}$. The correspondence $M \leftrightarrow \mathbf{L}_M$ is a vector space isomorphism between the set of Hermitian matrices H_N and the space of real linear maps $\mathbf{L} : H_{N_A} \mapsto H_{N_B}$. A slightly different version of this isomorphism is the Jamiolkowski isomorphism $M \leftrightarrow \mathbf{J}_M$ by which $\mathbf{J}_M X = \mathbf{L}_M(X^T)$ [65]. The correspondences $M \leftrightarrow \mathbf{L}_M$ and $M \leftrightarrow \mathbf{J}_M$ can be linked by using the Hermitean matrix $\hat{M} \in H_N$ to define the alternative map $\mathbf{L}_{\hat{M}} : H_{N_A} \mapsto H_{N_B}$

$$Y_{jl} = \sum_{i,k} \hat{M}_{jl;ik} X_{ik} \tag{7.11}$$

The relation $M_{ij;kl} = \hat{M}_{jl;ki}$ then defines the link between these two equivalent descriptions, which means that $Y = \mathbf{L}_{\hat{M}} X = \mathbf{L}_M(X^T)$, so that effectively $\mathbf{L}_{\hat{M}} = \mathbf{J}_M$.

The transposed real linear map $\mathbf{L}_M^T : H_{N_B} \mapsto H_{N_A}$ is defined such that $X = \mathbf{L}_M^T Y$ when

$$X_{ik} = \sum_{j,l} M_{ij;kl} Y_{lj} \tag{7.12}$$

The transpose of $\mathbf{L}_M$ is with respect to the natural scalar product (3.8) with $\langle X, Y\rangle = \mathrm{tr}(XY)$ for Hermitian matrices. In fact, for any $X \in H_{N_A}$ and $Y \in H_{N_B}$, we have

$$\langle \mathbf{L}_M X, Y\rangle = \mathrm{tr}(\mathbf{L}_M X, Y) = \sum_{ij;kl} M_{ij;kl} X_{ki} Y_{lj} = \mathrm{tr}[X(\mathbf{L}_M^T Y)] = \langle X, \mathbf{L}_M^T Y\rangle \tag{7.13}$$

We now define the essential type of linear maps connected to the description of entanglement, namely the positive maps

Definition 7.1 (Positive map)**.** *A linear map $\mathbf{L}_M : H_{N_A} \mapsto H_{N_B}$ is a positive map if it maps any positive matrix $\mu \in H_{N_A}$ into a positive matrix $\tau \in H_{N_B}$.*

A short way of writing that $\mathbf{L}_M$ is a positive map is $\mathbf{L}_M > 0$. It is also evidently clear from the convex cone structure of positive operators described in Figure 5.1, that positive here means positive semidefinite. We see from

$$\mathbf{L}_M(X^\dagger) = [\mathbf{L}_M(X)]^\dagger \tag{7.14}$$

that positivity and Hermitian preserving are equivalent characteristics of linear maps.

7.2.1 Choi-Jamiolkowski isomorphism

The maps $\mathbf{L}_M$ and $\mathbf{L}_M^T$ act on rank one projection operators $\phi\phi^\dagger \in H_{N_A}$ and $\chi\chi^\dagger \in H_{N_B}$ according to

$$\begin{aligned} \mathbf{L}_M(\phi\phi^\dagger) &= (\phi \otimes I)^\dagger M(\phi \otimes I) \\ \mathbf{L}_M^T(\chi\chi^\dagger) &= (I \otimes \chi)^\dagger M(I \otimes \chi) \end{aligned} \tag{7.15}$$

This means that $\phi \otimes I$ is an $N \times N_B$ matrix such that $(\phi \otimes I)\chi = \phi \otimes \chi$, and likewise $I \otimes \chi$ is an $N \times N_A$ matrix such that $(I \otimes \chi)\phi = \phi \otimes \chi$. It then follows that

$$\chi^\dagger[\mathbf{L}_M(\phi\phi^\dagger)]\chi = \phi^\dagger[\mathbf{L}_M^T(\chi\chi^\dagger)]\phi = (\phi \otimes \chi)^\dagger M(\phi \otimes \chi) \tag{7.16}$$

The condition $f_M(\phi, \chi) = (\phi \otimes \chi)^\dagger M(\phi \otimes \chi) \geq 0$ for all $\phi \in \mathbb{C}^{N_A}$ and $\chi \in \mathbb{C}^{N_B}$ means that $\mathbf{L}_M(\phi\phi^\dagger)$ is a positive matrix for every $\phi \in \mathbb{C}^{N_A}$. So, if M is an entanglement witness W then (7.16) implies that $\mathbf{L}_M$ is a positive linear map $\mathbf{L}_M : H_{N_A} \mapsto H_{N_B}$, mapping positive semidefinite matrices $\sum_i p_i(\phi_i\phi_i^\dagger) \in H_{N_A}$ to positive semidefinite matrices $\sum_i p_i \mathbf{L}_M(\phi_i\phi_i^\dagger) \in H_{N_B}$. In a completely similar way $\mathbf{L}_M^T$ must be a positive linear map $\mathbf{L}_M^T : H_{N_B} \mapsto H_{N_A}$.

So the condition defining M to belong to the subclass of Hermitian operators known as entanglement witnesses or $M \in \mathcal{S}^\circ \subset H_N$, is precisely the condition defining $\mathbf{L}_M$ and $\mathbf{L}_M^T$ as positive maps. The correspondence $W \leftrightarrow \mathbf{L}_W$ is a vector space isomorphism between the set of entanglement witnesses and the set of positive maps. The slightly

alternative correspondence $W \leftrightarrow \mathbf{J}_W$ is the celebrated Choi-Jamiolkowski isomorphism. Sometimes the Hermitian matrix W related to the positive map $\mathbf{J}_W$ is called the *Choi-matrix* C_J of $\mathbf{J}_W$. These considerations relate the theory of entanglement to the theory of positive maps, developed originally to a large part by Størmer [66].

7.2.2 Completely positive maps

An obvious condition for a map to be a *physical map* $\mathbf{L}_M : H_{N_A} \mapsto H_{N_B}$ transforming physical states into physical states, is that it should be positive. A less obvious condition is that it should be *completely positive*

Definition 7.2 (Completely positive)**.** *A positive map $\mathbf{L}_M : H_{N_A} \mapsto H_{N_B}$ is completely positive if the extended map $\mathbf{I}_k \otimes \mathbf{L}_M : H_k \otimes H_{N_A} \mapsto H_k \otimes H_{N_B}$ is positive, where the auxiliary map $\mathbf{I}_k : H_k \mapsto H_k$ with any $k > 0$, is the identity map.*

The reason why complete positivity is important with regards to quantum operations, is that a physical (positive) map that transforms any local part of a quantum state into a new quantum state should also keep the total state physical (positive), regardless of how these other parts look, and what their dimension k is.

The most general form of a quantum operation, or completely positive map is

$$\mathbf{L}_M X = \sum_r V_r X V_r^\dagger \tag{7.17}$$

where V_r are $N_B \times N_A$ matrices known as Kraus operators. With our definition (7.10) of the Hermitian matrix M corresponding to the map $\mathbf{L}_M$ we write the elements $M_{ij;kl}$ corresponding to a completely positive map as

$$M_{ij;kl} = (M^P)_{il;kj} = \sum_r (V_r)_{jk} (V_r)^*_{li} \tag{7.18}$$

So, we see that with our version of the Choi-Jamiolkowski isomorphism the positive map $\mathbf{L}_M$ is completely positive if and only if the partial transpose M^P is a positive matrix. The more commonly used map $\mathbf{J}_M$ is completely positive if and only if M is a positive matrix.

Any separable state ρ, as defined in (3.28), remains positive under the application of positive maps to any of its local parts. Thus any positive map $\mathbf{L}_M : H_{N_A} \mapsto H_{N_B}$ is completely positive on the separable states

$$(\mathbf{L}_M \otimes \mathbf{I})\rho = (\mathbf{L}_M \otimes \mathbf{I}) \left[\sum_r p_r \left(\mu_r \otimes \tau_r \right) \right] = \sum_r p_r \left(\mathbf{L}_M \mu_r \otimes \tau_r \right) \tag{7.19}$$

which is clearly a separable state if $\mathbf{L}_M$ is a positive map. Only separable states have the property that they stay positive under *all* local quantum operations. This means that if ρ is entangled there always exist a positive map $\mathbf{L}_M$ which when applied to a local part of the system, transforms ρ into a non-state. This is essentially just a restatement of the Horodecki-criterion.

The partial transposition map

The transposition map $\mathbf{T} : X_{ij} \mapsto X_{ji}$ is a positive map, but this is not the case for the partial transposition map $\mathbf{I} \otimes \mathbf{T}$, where $\mathbf{T} : H_{N_B} \mapsto H_{N_B}$. This is clear by the existence of NPT states. So the transposition map is an example of a map which is positive, but not completely positive.

The positivity condition that comes from applying a tensor product of positive maps to different local parts of a quantum state, is expressed as positivity conditions on the eigenvalues of the resulting matrix. In general this will be a characteristic polynomial of order at least equal to the rank of the state, and this is highly non-linear in ρ, and therefore much more difficult to analyze. An entanglement witness W however, effectively defines a hyperplane $\mathrm{tr}(W\rho) = 0$ in H_N, naturally dividing $\mathcal{D}_N$ into two regions, *i.e.* in a linear sense. In total, this means that a single positive map can in principle reveal the entanglement of many different states for which we would need many different entanglement witnesses. The Peres (or PPT) condition associated with the transposition map is a striking example, since it can reveal entanglement in all NPT states, something that no single entanglement witness can do.

From a physical point of view an entanglement witness W is more directly useful than a positive map, because it is an observable, and its expectation value $\mathrm{tr}(W\rho)$ in a state ρ can be measured.

7.2.3 Decomposable maps

A certain subclass of linear maps are the *decomposable maps*. The common way of defining a decomposable map is taken from Størmer [67]

Definition 7.3 (Decomposable map)**.** *A decomposable map* $\mathbf{L}_D : H_{N_A} \mapsto H_{N_B}$ *is a map that can be written*

$$\mathbf{L}_D = \mathbf{C}_1 + \mathbf{C}_2\mathbf{T}_A \tag{7.20}$$

where $\mathbf{C}_1$ *and* $\mathbf{C}_2$ *are completely positive, and* $\mathbf{T}_A$ *is transposition on system A.*

Any map is either a decomposable map or a non-decomposable map. We have seen that the map $\mathbf{L}_M : H_{N_A} \mapsto H_{N_B}$ corresponding to $M \in H_N$ as defined in (7.10), is completely positive if and only if the matrix M^P is positive. Hence $\mathbf{L}_M\mathbf{T}_A = \mathbf{L}_{M^P}$ is completely positive if and only if M is a positive matrix. Further from (7.20), any linear map $\mathbf{L}_M$ is decomposable if $M = \mu + \tau^P$, with μ and τ as positive matrices. We thus define any entanglement witness of the form

$$D = \mu + \tau^P \qquad \mu, \tau \geq 0 \tag{7.21}$$

to be a decomposable witness. The expectation value of the operator D in a PPT state ρ is

$$\mathrm{tr}(D\rho) = \mathrm{tr}(\mu\rho) + \mathrm{tr}(\tau^P\rho) = \mathrm{tr}(\mu\rho) + \mathrm{tr}(\tau\rho^P) \geq 0 \tag{7.22}$$

It is thus evident that any witness with the canonical form (7.21) has non-negative expectation value $\mathrm{tr}(D\rho)$ in any PPT state, and therefore cannot be used to reveal any

entangled PPT state. We observe that since $\mathrm{tr}(D\rho) \geq 0$ for any PPT state, this also ensures that D is indeed an entanglement witness, since $\mathcal{S} \subseteq \mathcal{P}$.

We have seen that the total set of entanglement witnesses $\mathcal{S}^\circ$ is the dual of the set of separable states $\mathcal{S}$. From (7.22) and Definition 4.4 of dual cones, we realize that in much the same way, the set of PPT states $\mathcal{P}$ are dual to the set of decomposable witnesses $\mathcal{P}^\circ$. In fact from $\mathcal{P} = \mathcal{D} \cap \mathcal{D}^P$ and (7.21), we can deduce that the set $\mathcal{P}^\circ$ is the convex hull of $\mathcal{D}$ and $\mathcal{D}^P$. The term convex hull of a set of points X is defined and outlined in Section 4.2 to be the smallest convex set that contains X. We realize that an extremal point of $\mathcal{P}^\circ$ is an extremal point of either $\mathcal{D}$ or $\mathcal{D}^P$. That is, it must be either a pure state $\rho_1 = \psi\psi^\dagger$ or a partially transposed pure state $\rho_2 = (\eta\eta^\dagger)^P$.

We can visualize this in a geometric manner, and this is done in Figure 7.1 with D as a decomposable witness and W as a non-decomposable witness.

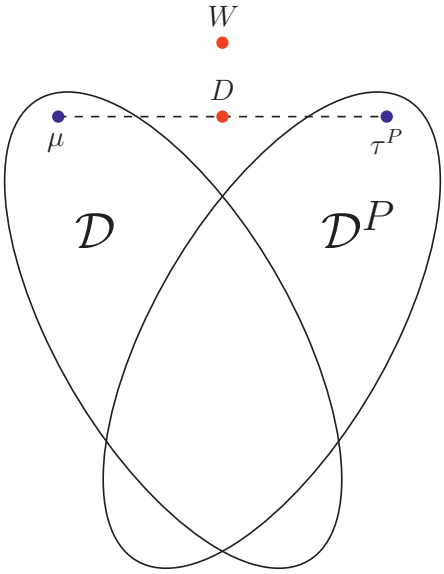


Figure 7.1: We see the set of density matrices $\mathcal{D}$ and partially transposed density matrices $\mathcal{D}^P$. The convex hull of these sets constitutes the set of decomposable witnesses $\mathcal{P}^\circ$. It is obviously possible to express the witness D in the form (7.21), but not the witness W, which is then a non-decomposable witness.

The two types of extremal states ρ_1 and ρ_2 of $\mathcal{P}^\circ$, can be associated with the bi-quadratic forms

$$\begin{aligned} f_{\rho_1}(\phi,\chi) &= \left|\psi^\dagger(\phi\otimes\chi)\right|^2 = \left|\sum_{k,l}\psi^*_{kl}\,\phi_k\chi_l\right|^2 \geq 0 \\ f_{\rho_2}(\phi,\chi) &= \left|\eta^\dagger(\phi\otimes\chi^*)\right|^2 = \left|\sum_{k,l}\eta^*_{kl}\,\phi_k\chi^*_l\right|^2 \geq 0 \end{aligned} \tag{7.23}$$

where ψ_{kl}, η_{kl}, ϕ_k and χ_l are components of $\psi, \eta \in \mathcal{H}_{AB}$, $\phi \in \mathcal{H}_A$ and $\chi \in \mathcal{H}_B$ respectively. This proves that $\rho_1 = \psi\psi^\dagger$ and $\rho_2 = (\eta\eta^\dagger)^P$ are indeed entanglement witnesses.

It was shown by [66, 68, 69] that all witnesses of the form $\rho_1 = \psi\psi^\dagger$ and $\rho_2 = (\eta\eta^\dagger)^P$ are extremal on $\mathcal{S}^\circ$, and that the corresponding positive maps are extremal on the total set of positive maps. Since $\mathcal{P}^\circ$ is a convex subset of $\mathcal{S}^\circ$, we deduce that all witnesses of type ρ_1 and ρ_2 are extremal on $\mathcal{P}^\circ$, and as we have argued, they are the only extremal points.

It was shown by Woronowicz [70] that all positive maps $\mathbf{L} : H_{N_A} \mapsto H_{N_B}$ with $N = N_A N_B < 8$ are decomposable. We know that for every entangled state ρ there exists at least one entanglement witness W [33], *i.e.* at least one W such that $\mathrm{tr}(W\rho) < 0$, and that if the state ρ is an entangled PPT state this witness W must be non-decomposable. It is thus easy to see that for the dimensions $N = N_A N_B < 8$ entangled PPT states cannot exist, and the Peres criterion is both necessary and sufficient for separability. It is the same logic that enforces the existence of entangled PPT states for dimensions $N = N_A N_B \geq 8$.

It is thus evident that in higher dimensions the development of practical and useful separability criteria is a problem closely related to the understanding of the non-decomposable positive maps and entanglement witnesses.

7.3 Zeros of entanglement witnesses

We here investigate the zero points of entanglement witnesses. The zero points on the set of product vectors $\phi \otimes \chi$ is of fundamental importance to the understanding of entanglement witnesses, especially the geometrical aspects.

7.3.1 Primary constraints

We have seen that, with $f_W(\phi, \chi)$ defined as in (7.6), an entanglement witness W satisfies the infinite set of inequalities

$$f_W(\phi, \chi) \geq 0 \qquad \phi \in \mathcal{H}_A,\, \chi \in \mathcal{H}_B \tag{7.24}$$

These constraints, which are linear in W, are in (7.6) represented through a biquadratic form. Observe that the normalization conditions $\phi^\dagger\phi = \chi^\dagger\chi = 1$ are independent of the constraints (7.24). If W is situated on the boundary $\partial\mathcal{S}^\circ$ it means that at least one of these inequalities is an equality. We call the pair (ϕ_0, χ_0) a zero of the witness W if $f_W(\phi_0, \chi_0) = 0$, where the point $(a\phi_0, b\chi_0)$ is counted as the same zero for all $a, b \in \mathbb{C}$. For a witness W we formally label the set of zeros as

$$\Pi_W = \{(\phi, \chi),\, \phi \in \mathcal{H}_A,\, \chi \in \mathcal{H}_B \,|\, f_W(\phi, \chi) = 0\} \tag{7.25}$$

7.3.2 Secondary constraints

From the primary constraints (7.24) one can deduce more stringent constraints on W when (ϕ, χ) is close to a zero (ϕ_0, χ_0). These secondary constraints are both equalities and inequalities, and like the primary constraints they are linear in W.

The secondary constraints on $f_W(\phi,\chi)$ near a zero (ϕ_0,χ_0) are along the lines of those for a real polynomial $f(t)$ of degree four, which satisfies $f(t)\geq 0$ for all t. If we choose a zero point $f(t_0)=0$ we also require that $f'(t_0)=0$ and that $f''(t_0)\geq 0$. In the case that $f''(t_0)=0$ we also get $f^{(3)}(t_0)=0$. These requirements must hold for all zero points $t_0,\ldots,t_n$. Since the set of zeros of a witness W are roots of a polynomial equation in several variables, we deduce that the set of zeros Π_W of a witness W consists of at most a finite number of components, where each component is either an isolated point or a continuous connected surface.

To investigate the behaviour of $f_W(\phi,\chi)$ around the zero (ϕ_0,χ_0) we first introduce two directions $J_0\in\mathcal{H}_A$ and $K_0\in\mathcal{H}_B$ such that $\phi_0^\dagger J_0=\chi_0^\dagger K_0=0$. We also introduce two variables $x\in\mathbb{R}^{2N_A-2}$ and $y\in\mathbb{R}^{2N_B-2}$. The vectors $\xi=J_0x$ and $\zeta=K_0y$ are then well defined since the orthogonality conditions remove two real degrees of freedom from J_0 and K_0. Our biquadratic form around the fixed zero point (ϕ_0,χ_0) is then a real inhomogeneous polynomial which is quadratic in both x and y

$$f(x,y)=\{(\phi_0+\xi)\otimes(\chi_0+\zeta)\}^\dagger W\{(\phi_0+\xi)\otimes(\chi_0+\zeta)\} \tag{7.26}$$

The linear term of the polynomial is

$$\begin{aligned} f_1(x,y) &= 2\,\mathrm{Re}\{(\xi\otimes\chi_0)^\dagger W(\phi_0\otimes\chi_0)+(\phi_0\otimes\zeta)^\dagger W(\phi_0\otimes\chi_0)\} \\ &= x^T\mathrm{D}_x f(0,0)+y^T\mathrm{D}_y f(0,0) \end{aligned} \tag{7.27}$$

in terms of the gradient

$$\begin{aligned} \mathrm{D}_x f(0,0) &= 2\,\mathrm{Re}\{(J_0\otimes\chi_0)^\dagger W(\phi_0\otimes\chi_0)\} \\ \mathrm{D}_y f(0,0) &= 2\,\mathrm{Re}\{(\phi_0\otimes K_0)^\dagger W(\phi_0\otimes\chi_0)\} \end{aligned} \tag{7.28}$$

The quadratic term of (7.26) is

$$\begin{aligned} f_2(x,y) &= (\xi\otimes\chi_0)^\dagger W(\xi\otimes\chi_0)+(\phi_0\otimes\zeta)^\dagger W(\phi_0\otimes\zeta) \\ &\quad +2\,\mathrm{Re}\{(\phi_0\otimes\zeta)^\dagger W(\xi\otimes\chi_0)+(\phi_0\otimes\chi_0)^\dagger W(\xi\otimes\zeta)\} \\ &= z^T G_W z \end{aligned} \tag{7.29}$$

where $z^T=(x^T,y^T)$ and $2\,G_W=\mathrm{D}^2 f(0,0)$ is the second derivative, or Hessian matrix, which is always real and symmetric

$$G_W=\mathrm{Re}\begin{bmatrix} g_{xx} & g_{yx}^T \\ g_{yx} & g_{yy}\end{bmatrix} \tag{7.30}$$

with the coefficients

$$\begin{aligned} g_{xx} &= (J_0\otimes\chi_0)^\dagger W(J_0\otimes\chi_0) \\ g_{xy} &= (\phi_0\otimes K_0)^\dagger W(J_0\otimes\chi_0)+(\phi_0\otimes K_0^*)^\dagger W^P(J_0\otimes\chi_0^*) \\ g_{yy} &= (\phi_0\otimes K_0)^\dagger W(\phi_0\otimes K_0) \end{aligned} \tag{7.31}$$

The cubic term is like the linear term but with $\phi_0 \leftrightarrow \xi$ and $\chi_0 \leftrightarrow \zeta$

$$f_3(x,y) = 2\,\mathrm{Re}\{(\phi_0 \otimes \zeta)^\dagger W(\xi \otimes \zeta) + (\xi \otimes \chi_0)^\dagger W(\xi \otimes \zeta)\} \tag{7.32}$$

and the quartic term is

$$f_4(x,y) = f_W(\xi,\zeta) = (\xi \otimes \zeta)^\dagger W(\xi \otimes \zeta) \tag{7.33}$$

Constraint operators $\mathbf{T}_k$

From the fact that (ϕ_0, χ_0) is a zero point for the witness W, we need $f(0,0) = 0$ or in the form of a constraint operator $\mathbf{T}_0 : H_N \mapsto \mathbb{R}$

$$\mathbf{T}_0 W = (\phi_0 \otimes \chi_0)^\dagger W(\phi_0 \otimes \chi_0) = 0 \tag{7.34}$$

which is one linear constraint on W. Because $(x,y) = (0,0)$ is a minimum of the polynomial, the linear term must also vanish, and we introduce another operator $\mathbf{T}_1 : H_N \mapsto \mathbb{R}^{2(N_A+N_B-2)}$ which gives another linear system of constraints

$$\mathbf{T}_1 W = \begin{bmatrix} \mathrm{D}_x f(0,0) \\ \mathrm{D}_y f(0,0) \end{bmatrix} = 0 \tag{7.35}$$

It should be noted that the constraints $\mathbf{T}_0$ and $\mathbf{T}_1$ are the same for any witness W with a zero at (ϕ_0, χ_0). All of these constraints are linearly independent. We note that the vanishing of the constant and linear terms of $f(x,y)$ is equivalent to the conditions

$$\begin{aligned} (\phi \otimes \chi_0)^\dagger W(\phi_0 \otimes \chi_0) &= 0 \qquad \forall\, \phi \in \mathcal{H}_A \\ (\phi_0 \otimes \chi)^\dagger W(\phi_0 \otimes \chi_0) &= 0 \qquad \forall\, \chi \in \mathcal{H}_B \end{aligned} \tag{7.36}$$

which in dimensions $N = N_A N_B$ produce $2(N_A + N_B) - 3$ real constraints. So, the total number of constraints given by $\mathbf{T}_0$ and $\mathbf{T}_1$ is

$$c_{01}(\phi_0, \chi_0) = 2(N_A + N_B) - 3 \tag{7.37}$$

Quadratic zeros

The quadratic term (7.29) of the polynomial $f_W(x,y)$ has to be non-negative. The inequalities

$$z^T G_W z \geq 0 \qquad \forall\, z \in \mathbb{R}^{2(N_A+N_B-2)} \tag{7.38}$$

are secondary inequality constraints. They are linear in W, and equivalent to the non-linear constraints that all eigenvalues of the Hessian matrix G_W must be non-negative. We then face two different alternatives

Definition 7.4 (Quadratic zeros)**.** *In the case that the inequalities $z^T G_W z > 0$ holds for all $z \neq 0$, or equivalently all eigenvalues of G_W are positive, the zero (ϕ_0, χ_0) is a quadratic zero of the entanglement witness W.*

In the case of a quadratic zero, $\mathbf{T}_0$ and $\mathbf{T}_1$ are the only equality constraints placed on W by the existence of the zero (ϕ_0, χ_0). We then state the following theorem

Theorem 7.2 (Isolated zeros)**.** *A quadratic zero is always isolated. Hence, an entanglement witness can only have a finite number of quadratic zeros.*

Since the constraints from $\mathbf{T}_0$ and $\mathbf{T}_1$ for the constant and linear term in $f_W(x, y)$, and their relations to quadratic zeros are relatively well understood, we will move on to the constraints from $\mathbf{T}_2$ and $\mathbf{T}_3$.

Quartic zeros

Let us assume that the Hessian matrix G_W has $K \geq 1$ eigenvalues equal to zero. Then $G_W z_i = 0$ for $i = 1, \ldots, K$, and the directions z_i are quartic directions, or *Hessian zeros* at (ϕ_0, χ_0). Any quartic direction z is a real linear combination

$$z = \sum_{i=1}^{K} d_i z_i \tag{7.39}$$

of basis vectors $z_i \in \operatorname{Ker} G_W$. The K linearly independent eigenvectors z_i define a system of linear constraints on W, which with $\mathbf{T}_2 : H_N \mapsto \mathbb{R}^{2K(N_A+N_B-2)}$ is

$$(\mathbf{T}_2 W)_i = G_W z_i = 0 \qquad i = 1, \ldots, K \tag{7.40}$$

These constraints ensure that $G_W z = 0$ and hence $f_2(x, y) = z^T G_W z = 0$. If we are to make the qubic term (7.32) vanish, this leads to a linear system $\mathbf{T}_3 W = 0$ containing $\binom{K+2}{3}$ equations [71]. So the total number of constraints from $\mathbf{T}_2$ and $\mathbf{T}_3$ is therefore

$$c_{23} = 2K(N_A + N_B - 2) + \binom{K+2}{3} \tag{7.41}$$

Note that the constraints from $\mathbf{T}_0$ and $\mathbf{T}_1$ address other terms in the polynomial than those of $\mathbf{T}_2$ and $\mathbf{T}_3$, so the two sets of constraints should be independent. Furthermore for the case of only one quartic direction $K = 1$, all c_{23} equations in the total system given by $\mathbf{T}_2$ and $\mathbf{T}_3$ should be linearly independent. The case for $K > 1$ is far less clear, but it is likely that the constraints are not independent.

7.3.3 Summary of constraints from zeros

Each zero in Π_W associated with an entanglement witness W defines zeroth and first order equality constraints from $\mathbf{T}_0$ and $\mathbf{T}_1$. If we define the set of quartic zeros of W as $\widetilde{\Pi}_W \subseteq \Pi_W$, then each zero in $\widetilde{\Pi}_W$ introduces additional second and third order constraints from $\mathbf{T}_2$ and $\mathbf{T}_3$. For each zero (ϕ_0, χ_0), the constraint operator $\mathbf{T}_2$ defines the following important distinction

$$\underbrace{\mathbf{T}_2 W > 0}_{\text{Quadratic zero}} \qquad\qquad \underbrace{\mathbf{T}_2 W = 0}_{\text{Quartic zero}} \tag{7.42}$$

Let us define $\mathbf{U}_i$ to be the direct sum of constraints from $\mathbf{T}_i$ over all the zeros in Π_W. We then write the combinations of constraints as

$$\mathbf{U}_{01} = \mathbf{U}_0 \oplus \mathbf{U}_1 \qquad \mathbf{U}_{23} = \mathbf{U}_2 \oplus \mathbf{U}_3 \tag{7.43}$$

The total system of constraints is then

$$\mathbf{U}_W = \mathbf{U}_{01} \oplus \mathbf{U}_{23} \tag{7.44}$$

We observe that the constraints $\mathbf{U}_{01}$ are completely determined by the zeros Π_W, whereas $\mathbf{U}_{23}$ depends only on $\widetilde{\Pi}_W$ and the kernel of the Hessian at each quartic zero.

We here define any witness with at least one zero (ϕ_0, χ_0), *i.e.* a boundary witness, to be a *quadratic entanglement witness* if all the zeros Π_W are quadratic zeros. Similarly, a boundary witness with at least one quartic zero, will be a *quartic entanglement witness*. Note that if W is a quadratic witness, then $\widetilde{\Pi}_W$ is empty, so $\mathbf{U}_W = \mathbf{U}_{01}$. This classification of boundary witnesses is fundamentally important, since the two classes have very different properties.

7.4 Extremal entanglement witnesses

We will here describe two conditions for W to be an extremal witness. The first is based on the set of zeros Π_W of the witness W, which define faces on $\mathcal{S}^\circ$. The other is related to constraints on W. The two conditions are closely related. These extremality conditions, and their relation to the facial structure of $\mathcal{S}^\circ$ can be used to develop a method to search for extremal witnesses. We then outline some results about quadratic extremal witnesses, and briefly describe extremal quartic witnesses, which appear to be more difficult to understand.

7.4.1 Extremaliy from zeros

If a witness $W \in \mathcal{S}^\circ$ is not an extremal witness then it can be written as a genuine convex combination of two other witnesses $\Lambda, \Sigma \in \mathcal{S}^\circ$ and

$$W = (1-p)\Lambda + p\Sigma \qquad 0 < p < 1 \tag{7.45}$$

We can then also write the biquadrtic form as a convex combination

$$f_W(\phi, \chi) = (1-p) f_\Lambda(\phi, \chi) + p f_\Sigma(\phi, \chi) \tag{7.46}$$

and also the Hessian matrix defined at a zero (ϕ_0, χ_0) of W as

$$G_W = (1-p) G_\Lambda + p G_\Sigma \tag{7.47}$$

Since Λ and Σ are witnesses we know that $f_\Lambda, f_\Sigma \geq 0$ for all (ϕ, χ), and likewise that $G_\Lambda, G_\Sigma \geq 0$ for all (ϕ, χ). This means that (ϕ_0, χ_0) is a zero of the witness W in (7.45), if and only if it is also a zero of both witnesses Λ and Σ. Please note that it is possible

for Λ to have a zero which is not a zero of W, as long as the zero is not shared by Σ. Because of (7.47) we also understand that z is a Hessian zero of W at (ϕ_0, χ_0) if and only if it is a Hessian zero at (ϕ_0, χ_0) of both Λ and Σ. We can see that all witnesses W in the line segment (7.45) have exactly the same zeros and Hessian zeros.

If the witnesses Λ and Σ described above are the extremal points of a line segment in $\mathcal{S}^\circ$, then the line segment cannot be prolonged within $\mathcal{S}^\circ$ in either direction. Now, both Λ and Σ must have at least one additional zero, or alternatively one additional Hessian zero in addition to the zeros and Hessian zeros of witnesses in the interior of the line segment [71]. We can summarize the above in an important theorem regarding extremal entanglement witnesses

Theorem 7.3 (Extremality condition I)**.** *A witness W is extremal on $\mathcal{S}^\circ$ if and only if no witness $\Theta \neq W$ has a set of zeros Π_Θ and Hessian zeros which includes the zeros Π_W and Hessian zeros of W.*

An equivalent condition in terms of the constraints operator $\mathbf{U}_W = \oplus_{i=1}^4 \mathbf{U}_i$ for the zeros and Hessian zeros, is that there can be no witness $\Theta \neq W$ that satisfies the constraints $\mathbf{U}_W \Theta = 0$. From this we can deduce that if none of the witnesses on the line segment (7.45) are quartic witnesses, then the witnesses Λ and Σ which are extremal to the line segment, will have at least one additional zero (ϕ_0, χ_0) not shared by other witnesses in the segment.

7.4.2 Faces of $\mathcal{S}^\circ$

If we extend the above discussion from one-dimensional line segments to faces of larger dimensions, we can see that every face $\mathcal{F} \in \mathcal{S}^\circ$ is uniquely characterized by a set of zeros and Hessian zeros that is the complete set of zeros and Hessian zeros of every witness in $\text{int}(\mathcal{F})$. Every witness on $\partial\mathcal{F}$ is a witness having the zeros and Hessian zeros characteristic of witnesses in $\text{int}(\mathcal{F})$, plus at least one more zero or Hessian zero. Roundly formulated we can say that the boundary of $\mathcal{S}^\circ$ is a hierarchy of faces, and that the number of zeros and Hessian zeros that define the faces, increase for each step we descend through the dimensions of these faces.

Once the zeros Π_W and accompanying Hessian zeros of a witness W are known, it is possible to find a finite perturbation of W within the unique face of $\mathcal{S}^\circ$ to which W is an interior point. The most general direction for such a perturbation is a traceless matrix $\Gamma \in \text{Ker}\, \mathbf{U}_W$. Note that as the number of zeros and Hessian zeros increase, so does the number of constraints, so $\text{Ker}\, \mathbf{U}_W$ and thereby the possible choices for Γ decreases. Computationally, if we have a general $\Gamma' \in \text{Ker}\, \mathbf{U}_W$, it can be made traceless by the transformation $\Gamma = \Gamma' - \text{tr}(\Gamma')W$. A result which supports the description of the hierarchy of faces $\mathcal{F}_i$ on the set $\mathcal{S}^\circ$ is [71]

Theorem 7.4 (Facial structure of $\mathcal{S}^\circ$)**.** *Let W be a witness, and let $\Gamma \neq 0$ be a traceless Hermitian operator. Then*

$$\Theta = W + t\Gamma \tag{7.48}$$

is a witness for t in some interval $[t_1, t_2]$, with $t_1 < 0 < t_2$, if and only if $\Gamma \in \text{Ker}\, \mathbf{U}_W$.

The maximal value of t_2 is the value of t where Θ acquires a new zero or Hessian zero, and the minimal value t_1 is determined in the same way. So the directions $\Gamma \in \operatorname{Ker} \mathbf{U}_W$ are the only directions which allow us to descend through the hierarchy of faces $\mathcal{F}_i$ in this way. A generic direction $\Gamma \notin \operatorname{Ker} \mathbf{U}_W$ either leads out of $\partial\mathcal{S}^\circ$ into $\mathcal{S}^\circ$, or in the opposite direction, out of $\mathcal{S}^\circ$. Only a $\Gamma \in \operatorname{Ker} \mathbf{U}_W$ ensures that we stay on the face defined by $\operatorname{Ker} \mathbf{U}_W$. A geometrical way of understanding Theorem 7.4 is to look at the face

$$\mathcal{F}_W = (\operatorname{Ker} \mathbf{U}_W) \cap \mathcal{S}^\circ \tag{7.49}$$

Evidently, as $\operatorname{Ker} \mathbf{U}_W$ decreases with the addition of new zeros and Hessian zeros, the face $\mathcal{F}_W$ also shrinks in dimension. If the face $\mathcal{F}_W$ becomes a single point W, or a zero-dimensional face, then W is an extremal point on the set of entanglement witnesses $\mathcal{S}^\circ$. Thus an algorithm which produces a decreasing sequence of faces $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \ldots \supset \mathcal{F}_n$ of $\mathcal{S}^\circ$, where every face $\mathcal{F}_j$ is a face of every $\mathcal{F}_i$ with $i < j$, will eventually converge to an extremal point of all these faces. All this can also be formulated into a another quite useful extremality condition

Theorem 7.5 (Extremality condition II)**.** *Let* $\mathbf{U}_W$ *be the constraint operator as defined in* (7.44)*. An entanglement witness* W *is extremal if and only if* $\operatorname{Ker} \mathbf{U}_W$ *is one-dimensional, i.e. spanned by* W*.*

Thus once the zeros and Hessian zeros of the witness W are known, we can test for extremality by computing the dimension of $\operatorname{Ker} \mathbf{U}_W$. This allows for a numerical implementation of the extremality criterion.

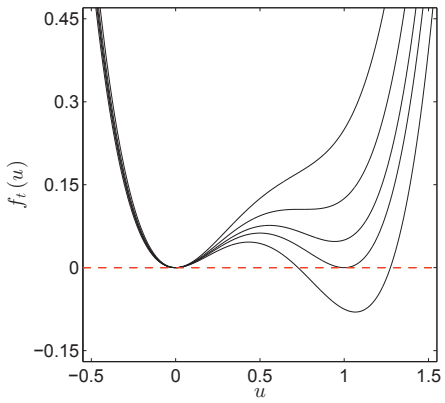


Figure 7.2: Model for how a new local minimum of $f_t(u)$ appears as t varies, turning into a zero and then a negative minimum as the witness leaves $\mathcal{S}^\circ$. The model function is $f_t(u) = u^2[(u-1)^2 + 1 - t]$.

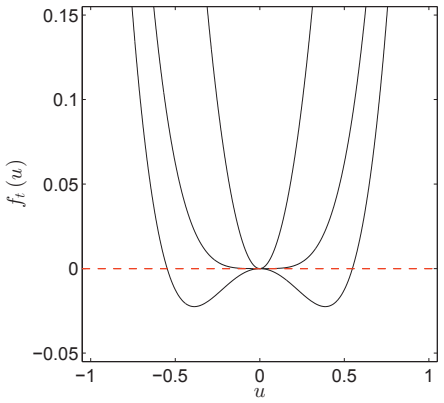


Figure 7.3: A quadratic zero turns into a quartic zero. And then new negative minima, which indicates that $\Theta(t)$ no longer is in the set $\mathcal{S}^\circ$, appears. The model function is $f_t(u) = u^2(u^2 + 1 - t)$.

We see depicted in Figure 7.2 a model of how a new isolated quadratic zero of the witness $\Theta(t)$ appears. Each curve is for different values of t in the perturbation (7.48),

and represents the function

$$f_t(u) = f_{W+t\Gamma}(\phi + u\phi', \chi + u\chi') \tag{7.50}$$

for different witnesses $\Theta(t)$. An initial zero point exists at $u = 0$, and the perturbation progresses until a new witness with a zero at $u = 1$ is reached. It is quite clear that for the lowest curve, $\Theta(t)$ is no longer an entanglement witness. Likewise in Figure 7.3 we see how an existing quadratic zero turns into a quartic zero.

This discussion can be formulated into an algorithm that can be used to search numerically for extremal entanglement witnesses. A somewhat more formal description of this algorithm is outlined in Section 8.6.

7.4.3 Quadratic extremal witnesses

It is interesting to try and understand the minimum number of zeros n_c a witness must have in order to be extremal on $\mathcal{S}^\circ$. The case for quadratic witnesses is far clearer than for the quartic witnesses, largely because the number of constraints for a quadratic zero is easier to handle. We remember that if (ϕ_0, χ_0) is a zero of the witness W, then the partially conjugated zero (ϕ_0, χ_0^*) is a zero of W^P. We now make the useful definition

Definition 7.5 (Spanning property)**.** *If the collection of zeros Π_W of an entanglement witness $W \in H_N$ span the Hilbert space $\mathbb{C}^N$, then W has spanning property. In the case that also the partially conjugated zeros Π_W^* span $\mathbb{C}^N$, the witness W has double spanning property.*

A result that effectively determines the minimum number of zeros $n_c = |\Pi_W|$ for a quadratic extremal witness is the following theorem

Theorem 7.6 (Double spanning)**.** *A quadratic extremal witness W has the double spanning property.*

This is easy to see by appealing to a projection $P = \psi\psi^\dagger$. If the set of zeros of W do *not* span the Hilbert space, then such a projection exists, with $\psi \perp \Pi_W$. The projection P will then be witness with a set of zeros including Π_W, and hence W cannot be extremal.

We return briefly to the case of decomposable witnesses, and observe that the bi-quadratic form for the decomposable witness $D = \mu + \tau^P$, with $\mu, \tau \in \mathcal{D}$ is

$$f_D(\phi, \chi) = f_\mu(\phi, \chi) + f_\tau(\phi, \chi^*) \tag{7.51}$$

If (ϕ_0, χ_0) is zero of D then (7.51) shows that (ϕ_0, χ_0) must be a zero of μ and (ϕ_0, χ_0^*) must be a zero of τ. Since both $\mu, \tau \in \mathcal{D}$ we can deduce that the zeros of $D = \mu + \tau^P$ span the Hilbert space only if $\mu = 0$, and the partially conjugated zeros of D span the Hilbert space only if $\tau = 0$. We then easily see from Theorem 7.6 that a quadratic extremal witness is always non-decomposable.

Minimum number of zeros from constraints

If we count the number of constraints c_{01} from one quadratic zero, we collect from (7.37) that $c_{01} = 2\,(N_A + N_B) - 3$. With n quadratic zeros we will have nc_{01} constraints, and since $|\mathcal{S}^\circ| = N^2 - 1$ these will for the generic case be linearly independent as long as $nc_{01} \leq N^2 - 1$. A lower bound on the number of zeros of quadratic witness is then

$$n_c = \left\lceil \frac{(N_A N_B)^2 - 1}{2(N_A + N_B) - 3} \right\rceil \tag{7.52}$$

Please note that for $N_A = 2$ we get $n_c = N - 1$ which is weaker than the lower bound given by Theorem 7.6. For all $N_A, N_B \geq 3$ we get $n_c \geq N$, which is consistent with Theorem 7.6. In Table 7.1 we see the different relevant numbers for a selection of dimensions $N_A \times N_B$. The number m_{ind} is the numerically calculated number of linearly independent constraints arising from n_c randomly chosen (generic) pairs (ϕ, χ). For each of these pairs the constraint operator $\mathbf{U}_{01}$ is built, and m_{ind} is then calculated.

$N_A \times N_B$	N^2	c_{01}	n_c	m_{ind}
2×2	16	5	3	14
2×3	36	7	5	34
2×4	64	9	7	62
2×5	100	11	9	98
3×3	81	9	9	81
3×4	144	11	13	143
3×5	225	13	18	225
4×4	256	13	20	256
4×5	400	15	27	400
5×5	625	17	37	625

Table 7.1: Numbers related to a generic quadratic witness W in dimension $N_A \times N_B$. c_{01} is the number of linearly independent constraints from each zero, n_c is the estimated minimum number of zeros required for W to be extremal and m_{ind} is the actual number of independent constraints from n_c random product vectors in $\mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$.

In Theorem 7.6 we established that a quadratic extremal witness W always has the double spanning property. But is the converse true? From Table 7.1 we see that in dimensions higher than $2 \times N_B$ and 3×3, the minimum number n_c of zeros of a quadratic extremal witness must be larger than the Hilbert space dimension N. Hence non-extremal witnesses with the double spanning property must be very common. So we conclude that the converse of Theorem 7.6 is true generically only for the $2 \times N_B$ and 3×3 systems. These observations give a picture of increasingly complicated geometry of the quadratic witnesses as dimensions increase. In dimensions 2×2 and 2×3 all witnesses are decomposable, while for the $2 \times N_B$ and 3×3 systems all quadratic witnesses have the double spanning property, and for larger dimensions there exist an

abundance of non-extremal quadratic witnesses with the double spanning property. We make further observations from Table 7.1 regarding quadratic witnesses:

- In order to produce an extremal witness we need to use all constraints contained in the zeros, and because of the extremality condition in Theorem 7.5 the number of these must be $N^2 - 1$, which is equal to the dimension of the real vector space H_N of Hermitian matrices, minus one. The demand that $|\mathrm{Ker}\,\mathbf{U}_W| = 1$ for the final step, takes care of the normalization of the witness W. For many of the dimensions $N_A \times N_B$ we observe that the number of independent constraints from a *random* set of product vectors m_{ind} is equal to the dimension of the real vector space H_N of Hermitian matrices. In these cases there does not exist an entanglement witness with these product vectors as zeros. So we conclude that the zeros of extremal witnesses in these systems must satisfy some relations that reduce the number of independent constraints m_{ind} by one.

- In the case $N_A = 2$ we see that $n_c = N - 1$, so if n_c is the correct minimum number of zeros, then Theorem 7.6 appears to fail. It looks however like there is an intrinsic degeneracy that reduces the number of constraints by one. For example in the 2×4 system the actual number of constraints from $n_c = 7$ random product vectors is 62, instead of $n_c c_{01} = 63$. This degeneracy implies that one extra zero is needed. With $N - 1$ zeros there exists a vector ψ orthogonal to all the zeros, and another vector η orthogonal to all the partially conjugated zeros. And since both $\psi\psi^\dagger$ and $(\eta\eta^\dagger)^P$ lie in $\mathrm{Ker}\,\mathbf{U}_W$, we must have $|\mathrm{Ker}\,\mathbf{U}_W| \geq 2$.

- The 3×4 system is special in that the number of constraints from a generic random set of $n_c = 13$ product vectors, add up to exactly the number needed to define a unique $\Lambda \in H_N$ with $\mathbf{U}_\Lambda \Lambda = 0$. This does not however imply that Λ defined in this way from a *random* set of product vectors will be a witness, since we generally will not have $(\phi \otimes \chi)^\dagger \Lambda (\phi \otimes \chi) \geq 0$ for all the product vectors $\phi \otimes \chi$.

Continuing in the vein of the last point above, we may ask a very interesting and important question: How do we choose a set of product vectors $\phi_i \otimes \chi_i$ such that they serve as the zeros of an extremal witness W? It appears that the sets of product vectors $\phi_i \otimes \chi_i \in \mathcal{S}$ that act as the zeros Π_W for a witness W, are quite special and non-generic sets of product vectors. We shall return to these matters later in Section 7.5.3.

Generic quadratic extremal witnesses in the 3×3 system

We have implemented Algorithm 1, and used it to locate numerical examples of extremal witnesses in the 2×4, 3×3 and 3×4 systems. In the process we produce extremal and non-extremal witnesses situated on a hierarchy of successively lower dimensional faces of $\mathcal{S}^\circ$. We define a class of extremal witnesses to be generic if such witnesses can be found with non-zero probability when the search direction Γ in (7.48) is chosen at random within $\mathrm{Ker}\,\mathbf{U}_W$, in every iteration. An overwhelming majority of the extremal witnesses found in the numerical random searches are quadratic. We have

also investigated quadratic witnesses with a number of zeros exceeding the minimum number n_c given in Table 7.1, and these appear to be non-generic.

- Generally, a single new quadratic zero appears in every iteration k when the boundary of the current face is reached. As long as the number of zeros k in the chain of iterations is small, there is no redundancy between constraints from existing zeros and the new zero. A redundancy appears typically with the seventh zero in 2×4, with the ninth in 3×3 and never in 3×4. Assuming non-redundancy, this gives a hierarchy of faces of $\mathcal{S}^\circ$ of dimension N^2-1-kc_{01} where $c_{01}=2(N_A+N_B)-3$.

- The extremal witnesses have the expected number of zeros as listed in Table 7.1. The zeros (ϕ_0,χ_0) and the partially conjugated zeros (ϕ_0,χ_0^*), span the Hilbert space, in accordance with Theorem 7.6.

- Every witness W and its partial transpose W^P have full rank, and a generic quadratic extremal witness W and its partial transpose W^P has at least one negative eigenvalue.

In Figure 7.4 we see some data for the eigenvalues of 171 quadratic extremal witnesses in the 3×3 system, found in numerical searches using Algorithm 1.

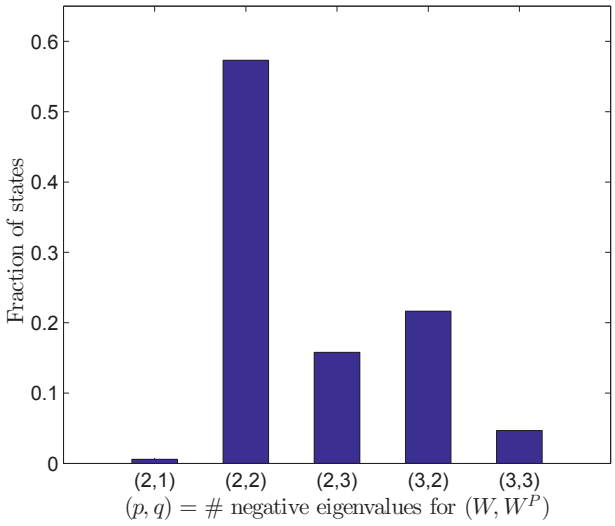


Figure 7.4: Classification of generic quadratic extremal witnesses in the 3×3 system found numerically by using the algorithm based on Theorems 7.4 and 7.5, along with (7.49), which is found in Section 8.6. A witness W of type (p,q) has p negative eigenvalues and its partial transpose W^P has q negative eigenvalues.

Since it is the negative part of the eigenvalue spectrum which makes an entanglement witness W useful, it seems obvious that the number of negative eigenvalues p and their

magnitude $|\lambda_k|$ is important. It could be argued that p, and the magnitude of these negative eigenvalues, is related to the size of the set $\{\rho \in \mathcal{D} \,|\, \mathrm{tr}(W\rho) < 0\}$ which W reveals.

Non-generic quadratic extremal witnesses in the 3×3 system

In order to demonstrate that a quadratic witness may have more than the minimum number of zeros, we have constructed quadratic extremal witnesses in the 3×3 system with ten zeros rather than the generic nine zeros. In order to produce a witness with a larger number of zeros than the minimum n_c required for extremality, we first produce a set of $k < n_c$ zeros using less than the generic number of constraints. This means that the set of k product vectors must have a degree of linear dependence. In this way, when the generic number of zeros n_c is reached, there are still enough degrees of freedom to create one extra zero, and make W extremal.

For the 3×3 system it appears that if we choose $k = 8$ and $k = 9$, the constraints leave only the decomposable witnesses [71], none of which are quadratic and extremal. For $k = 7$ however, there is room for additional zeros so that we can create a non-decomposable quadratic extremal witness. Denote $\mathbf{U}_{01}^{(k)}$ as the linear system of constraints for k zeros. We choose seven zeros (ϕ_i, χ_i) such that $|\mathrm{Ker}\, \mathbf{U}_{01}^{(7)}| = 22$, which is non-generic. A decomposable witness with all these seven (ϕ_i, χ_i) as zeros has the form $D_7 = \mu + \tau^P$ where $\mu, \tau \geq 0$ has rank three. Hence the set of such decomposable witnesses is 18-dimensional, so that there are four dimensions in $\mathrm{Ker}\, \mathbf{U}_{01}^{(7)}$ orthogonal to the face of decomposable witnesses. Defining Γ_7 to lie in these four dimensions, one can walk towards the boundary of the face

$$\mathcal{F}_7 = \mathrm{Ker}\, \mathbf{U}_{01}^{(7)} \cap \mathcal{S}^\circ \tag{7.53}$$

and find $\Theta_8 = D_7 + t_c\Gamma_7$ with eight zeros. Θ_8 is now guaranteed to be non-decomposable. Let now $\mathbf{U}_{01}^{(8)}$ be the constraint operator defined by these eight zeros, and define the face

$$\mathcal{F}_8 = \mathrm{Ker}\, \mathbf{U}_{01}^{(8)} \cap \mathcal{S}^\circ \tag{7.54}$$

We find that $|\mathrm{Ker}\, \mathbf{U}_{01}^{(7)}| - |\mathrm{Ker}\, \mathbf{U}_{01}^{(8)}| = 9$. Defining $\Gamma_8 \in \mathrm{Ker}\, \mathbf{U}_{01}^{(8)}$, we locate a $\Theta_9 \in \mathcal{F}_9$ on the boundary of $\mathcal{F}_8$, with nine zeros. The essential thing is that we now have $|\mathrm{Ker}\, \mathbf{U}_{01}^{(9)}| = 4$, hence there is still freedom to move along $\mathcal{F}_9$. Doing so produces a quadratic extremal witness in the 3×3 system, with ten zeros.

The above method involved the construction of non-generic structures in order to produce non-generic properties of extremal quadratic witnesses. We can also find generic faces on $\mathcal{S}^\circ$ in the 2×4 and 3×3 systems, that contain extremal quadratic witnesses with nine and ten zeros respectively. The next-to-extremal faces of $\mathcal{S}^\circ$ in these systems reveal a special geometry, which is related to the presence of decomposable witnesses on the boundary of these faces.

What we typically find in the kth iteration when using Algorithm 1, is a face $\mathcal{F}_k$ of $\mathcal{S}^\circ$ with interior points that are quadratic witnesses with k zeros. In the 3×3 system, $\mathcal{F}_8$ is the last, *i.e.* next-to-extremal, face found before a generic extremal quadratic

witness is reached. These particular faces have a special geometry, because the number of zeros is one less than the dimension of the Hilbert space, and as a result of this, a part of the boundary is a line segment of decomposable witnesses

$$M = (1-p)\psi\psi^\dagger + p(\eta\eta^\dagger)^P \qquad 0 \le p \le 1 \tag{7.55}$$

where ψ is orthogonal to the $N-1$ product vectors $\phi_i \otimes \chi_i$ which are zeros of all the witnesses in the interior of $\mathcal{F}_{N-1}$, and η to all $\phi_i \otimes \chi_i^*$. The line segment (7.55) of decomposable witnesses must be on the boundary of $\mathcal{F}_{N-1}$ because the interior consists of quadratic witnesses with a fixed set of $N-1$ quadratic zeros, whereas the decomposable witnesses have additional zeros, in fact an infinity of quartic zeros.

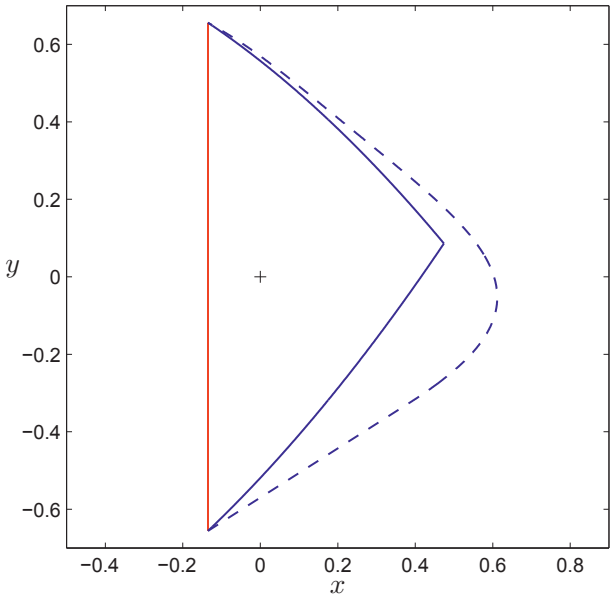


Figure 7.5: A special two-dimensional section of an eight-dimensional face $\mathcal{F}_8 \subset \mathcal{S}^\circ$ obtained by using Algorithm 1 for the 3×3 system. The solid lines represent the boundary $\partial\mathcal{F}_8$. The red line is a segment of decomposable witnesses, with the upper endpoint being a pure state $\psi\psi^\dagger$ and the lower end point a partial transpose of a pure state $(\eta\eta^\dagger)^P$. The "+" represents a witness in the interior of the face, with eight zeros. This is the starting point for perturbations of the type (7.48) along the face, with two orthogonal traceless directions Γ_1, Γ_2 as axes. The perturbations identify different boundaries: The curved solid boundary in blue, which consists of extremal witnesses with nine zeros, except at the kink where the witness has ten zeros. The dashed blue line is where at least one zero becomes quartic, and lies outside the set of entanglement witnesses, except at the two endpoints $\psi\psi^\dagger$ and $(\eta\eta^\dagger)^P$.

For the face $\mathcal{F}_8$ in the 3×3 system, the remaining boundary, with the line segment (7.55) taken away, will consist of quadratic extremal witnesses with nine or ten zeros. Since the face $\mathcal{F}_8$ is eight-dimensional, this curved boundary will be seven-dimensional. Any two-dimensional section passing through $\mathcal{F}_8$ that includes the line segment (7.55) is shaped like a D. In Figure 7.5 we see such a D-shaped cross section, constructed by

using Algorithm 1, and then making a large number of paths Γ on $\mathcal{F}_8$, locating the boundary $\partial\mathcal{F}_8$.

All along the blue curve consisting of quadratic witnesses with nine zeros, eight of those zeros are shared by all witnesses on the face, they in fact define the face $\mathcal{F}_8$. But the ninth zero changes continuously along the curve, and since these two continuous curves meet in one point, this point is a witness with two quadratic zeros in addition to the eight zeros defining the face.

Even though these so-called D-faces are found in completely generic searches, please observe that there is no guarantee that any choice of eight random product vectors define such a D-shaped face. In other, and generally higher dimensions $N_A \times N_B$, line segments of the type (7.55) constructed from $N-1$ zeros, also define D-shaped faces, but the curved part of these faces will in general not consist of extremal witnesses. This is related to the fact that for most of these systems the minimum number of zeros for extremailty n_c, as given by (7.52), is larger than the dimension $N = N_A N_B$ of the Hilbert space. As we have seen in Table 7.1 this happens for systems with dimensions 3×4 and larger.

7.4.4 Quartic extremal witnesses

In each iteration of Algorithm 1 the critical parameter t_c is reached when either a new quadratic zero appears, or a new zero eigenvalue of the Hessian matrix appears at one of the existing zeros. This is depicted in Figures 7.2 and 7.3. Our experience is that a random search most often produces a quadratic extremal witness, and only rarely a quartic extremal witness.

In order to quantify this somewhat, we have made random searches for quartic witnesses in dimension 3×3 in the following manner: We took 58 hierarchies of faces $\mathcal{F}_i^{(k)}$ of quadratic witnesses which we generated by Algorithm 1, and generated 100 random perturbations Γ away from the quadratic witness $W \in \mathrm{int}(\mathcal{F}_k)$ found on each face k. Since we have $k = 1, \ldots, 8$ faces and $i = 1, \ldots, 58$ extremal witnesses we get totally $8 \cdot 58 \cdot 100 = 46400$ trials. For each trial we computed t_c' as the smallest t resulting in a zero eigenvalue of the Hessian at a zero. The value t_c' can then be considered an upper bound for t_c. At $t = t_c'$ we test whether $W + t_c'\Gamma$ is still an entanglement witness, in which case it is a quartic witness. If it is not, this will mean that there has appeared a new quadratic zero for some $t < t_c'$. We found with a tolerance of $\pm 10^{-14}$ that only approximately 0.2% of the trials returned a quartic witness. We conclude that the fraction of quartic extremal witnesses among the generic extremal witnesses is small but non-zero.

We have seen that the pure states $\rho_1 = \psi\psi^\dagger$ and partially transposed pure states $\rho_2 = (\eta\eta^\dagger)^P$ are witnesses that are extremal on both the set of decomposable witnesses $\mathcal{P}^\circ$ and the total set of witnesses $\mathcal{S}^\circ$. An easy way to see that the pure state $\rho_1 = \psi\psi^\dagger$ is a quartic witness, is to invoke the Schmidt decomposition scheme (3.21) on ψ. The zeros of ρ_1 are the product vectors $\phi \otimes \chi$ orthogonal to ψ. In a Schmidt decomposition of ψ we obtain the singular values $c_i > 0$ and orthonormal bases $u_i \in \mathcal{H}_A$ and $v_i \in \mathcal{H}_B$ such that $\psi = \sum_{i=1}^r c_i(u_i \otimes v_i)$, with r as the Schmidt rank of ψ and where $1 \leq r \leq \min(N_A, N_B)$.

A general product vector in the bases u_i and v_i is then

$$\phi \otimes \chi = \sum_{i,j}^{N_A N_B} a_i b_j (u_i \otimes v_j) \tag{7.56}$$

and the condition for $\psi \perp \phi \otimes \chi$ is then $\sum_{i=1}^{r} c_i a_i b_i = 0$, which for all $N_A, N_B \geq 0$ and any r gives a connected and continuous set of zeros. Since all quadratic witnesses have isolated zeros, the witness $\rho_1 = \psi\psi^\dagger$ is a quartic witness. An identical argument can be used for witnesses of the form $\rho_2 = (\eta\eta^\dagger)^P$.

Indeed the pure state $P = \psi\psi^\dagger$ is an example where the zeros Π_P do not span the Hilbert space, but where the witness P is extremal nonetheless. The partially conjugated zeros Π_P^* may however span the Hilbert space. The witness P exemplifies that Theorem 7.6 cannot be extended to quartic witnesses, but as we shall see below there are more illustrious examples of this. Again this is a consequence of the fact that a quartic zero usually carry more constraints than a quadratic.

The Choi-Lam witness

Define, like Ha and Kye [72]

$$\Omega(a,b,c;\theta) = \begin{pmatrix} a & . & . & . & -\mathrm{e}^{\mathrm{i}\theta} & . & . & . & -\mathrm{e}^{-\mathrm{i}\theta} \\ . & c & . & . & . & . & . & . & . \\ . & . & b & . & . & . & . & . & . \\ . & . & . & b & . & . & . & . & . \\ -\mathrm{e}^{-\mathrm{i}\theta} & . & . & . & a & . & . & . & -\mathrm{e}^{\mathrm{i}\theta} \\ . & . & . & . & . & c & . & . & . \\ . & . & . & . & . & . & c & . & . \\ . & . & . & . & . & . & . & b & . \\ -\mathrm{e}^{\mathrm{i}\theta} & . & . & . & -\mathrm{e}^{-\mathrm{i}\theta} & . & . & . & a \end{pmatrix} \tag{7.57}$$

We write dots instead of zeros to make it more readable. The special case $W_C = \Omega(1,0,1;0)$ is the Choi-Lam witness, one of the first examples of a non-decomposable witness, developed by Choi and Lam [64]. The set of zeros of W_C consists of three isolated quartic zeros

$$e_{13} = e_1 \otimes e_3 \qquad e_{21} = e_2 \otimes e_1 \qquad e_{32} = e_3 \otimes e_2 \tag{7.58}$$

where e_1, e_2, e_3 are the natural basis vectors in $\mathbb{C}^3$, in addition to a continuum of zeros $\phi \otimes \chi$ where

$$\phi = e_1 + \mathrm{e}^{\mathrm{i}\alpha} e_2 + \mathrm{e}^{\mathrm{i}\beta} e_3 \qquad \chi = \phi^* \tag{7.59}$$

with the coefficients $\alpha, \beta \in \mathbb{R}$. The product vectors defined in (7.59) span a seven-dimensional subspace V consisting of all vectors $\psi \in \mathbb{C}^9$ with components $\psi_1 = \psi_5 = \psi_9$. The three product vectors defined in (7.58) have $\psi_1 = \psi_5 = \psi_9 = 0$ and lie in the same subspace V. The Choi-Lam witness provides a good example of a witness which is

extremal but does not have spanning property, and thus underlines the restriction of Theorem 7.6 to quadratic witnesses. It should also be emphasized that there do exist quartic witnesses with a set of zeros that span the Hilbert space. So the spanning property is in no way exclusive to the quadratic witnesses, but only guaranteed for this type by Theorem 7.6.

It can be proven numerically that the Choi-Lam witness is extremal on $\mathcal{S}^\circ$. Using all the quadratic and quartic constraints $\mathbf{U}_W$ from all the three isolated zeros e_1, e_2, e_3 and any single zero $\phi(\alpha, \beta)$ from the continuum, one can show that $|\mathrm{Ker}\, \mathbf{U}_{W_C}| = 1$, and this uniquely defines the Choi-Lam witness. The extended class of Choi-type witnesses (7.57), has many different properties for varying values of the parameters a, b, c, θ. For example, it is known that the Choi-Lam witness $W_C = \Omega(1, 0, 1; 0)$ is extremal but not exposed, but the extended witness given in (7.57) is also exposed for some combinations of a, b, c, θ.

7.5 Faces of $\mathcal{S}$ and $\mathcal{S}^\circ$

As we have seen, an understanding of the geometry of the set of entanglement witnesses $\mathcal{S}^\circ$ can be made by studying the faces of this set, and their related zeros. These zeros are pure product vectors $\psi = \phi \otimes \chi \in \mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ which correspond to states $\psi\psi^\dagger \in \mathcal{S}$. So our understanding of $\mathcal{S}^\circ$ is then related to an understanding of the faces $\mathcal{F}_\mathcal{S}$ of the separable states $\mathcal{S}$. Some studies have been made into the structure of $\mathcal{F}_\mathcal{S}$. Alfsen and Shultz [51] describe two special types of faces in $\mathcal{S}$, as either a special class of simplexes or direct convex sums of faces isomorphic to matrix algebras. It is our understanding that these two categories correspond to certain types of quadratic and quartic witnesses respectively, where our understanding of the quadratic witnesses is good. It is known that the set of entanglement witnesses $\mathcal{S}^\circ$ has unexposed faces, but it is unknown as to whether this is the case for $\mathcal{S}$, though several works have been published on the matter, notably by Chruściński and Sarbicki [76, 77].

7.5.1 Duality of exposed faces of $\mathcal{S}$ and $\mathcal{S}^\circ$

It is possible to define the dual of a subset $\mathcal{X} \subset \mathcal{S}$ by

$$\mathcal{X}^\circ = \{W \in \mathcal{S}^\circ \,|\, \mathrm{tr}(W\rho) = 0 \quad \forall\, \rho \in \mathcal{X}\} \tag{7.60}$$

and similarly from any subset $\mathcal{Y} \subset \mathcal{S}^\circ$, the dual

$$\mathcal{Y}^\circ = \{\rho \in \mathcal{S} \,|\, \mathrm{tr}(W\rho) = 0 \quad \forall\, W \in \mathcal{Y}\} \tag{7.61}$$

Since no witnesses in the interior of $\mathcal{S}^\circ$ have zeros, we must assume that $\mathcal{X} \subset \partial\mathcal{S}$ and $\mathcal{Y} \subset \partial\mathcal{S}^\circ$, and then follows that $\mathcal{X}^\circ \subset \partial\mathcal{S}^\circ$ and $\mathcal{Y}^\circ \subset \partial\mathcal{S}$.

It is clear that in any system of finite dimension, we would need a finite number k of separable states $\rho_1, \ldots, \rho_k \in \mathcal{X}$ to define $\mathcal{X}^\circ = \{W \in \mathcal{S}^\circ \,|\, \mathrm{tr}(W\rho_i) = 0 \text{ for } i = 1, \ldots, k\}$.

If we then construct for instance the evenly weighted separable state $\rho_0 = \frac{1}{k}\sum_{i=1}^{k}\rho_i \in \mathcal{X}$ we see that the set $\mathcal{X}^\circ$ can be defined as

$$\mathcal{X}^\circ = \{\rho_0\}^\circ = \{W \in \mathcal{S}^\circ \,|\, \mathrm{tr}(W\rho_0) = 0\} \tag{7.62}$$

This is possible because $\mathrm{tr}(W\rho_i) \geq 0$ for $i = 1, \ldots, k$, which means that $\mathrm{tr}(W\rho_0) = 0 \Rightarrow \mathrm{tr}(W\rho_i) = 0$ for $i = 1, \ldots, k$. By the same argument we can show that it is always possible to find a witness W_0 such that

$$\mathcal{Y}^\circ = \{W_0\}^\circ = \{\rho \in \mathcal{S} \,|\, \mathrm{tr}(W_0\rho) = 0\} \tag{7.63}$$

We now assume that the witness $W \in \mathcal{X}^\circ$ can be written as

$$W = (1-p)W_1 + pW_2 \qquad\qquad 0 \leq p \leq 1 \tag{7.64}$$

where $W_1, W_2 \in \mathcal{S}^\circ$. Since W_1 and W_2 are witnesses we must have $\mathrm{tr}(W_1\rho) \geq 0$ and $\mathrm{tr}(W_2\rho) \geq 0$ for all $\rho \in \mathcal{S}$. From (7.64) we get that

$$\mathrm{tr}(W\rho) = (1-p)\,\mathrm{tr}(W_1\rho) + p\,\mathrm{tr}(W_2\rho) = 0 \tag{7.65}$$

which again means that $\mathrm{tr}(W_1\rho) = \mathrm{tr}(W_2\rho) = 0$. So W_1 and W_2 must lie in $\mathcal{X}^\circ$, which means that $\mathcal{X}^\circ$ is a face. Since we see from (7.62) that $\mathcal{X}^\circ$ is dual to one $\rho_0 \in \mathcal{X} \subset \mathcal{S}$, the equation $\mathrm{tr}(\Lambda\rho_0) = 0$ with $\Lambda \in H_N$ defines a hyperplane of dimension $N^2 - 2$ in the (N^2-1)-dimensional space of Hermitean $N \times N$ matrices of unit trace. The $\Lambda \in H_N$ which are also witnesses, then lie exclusively in the face $\mathcal{X}^\circ$, which is then an exposed face of $\mathcal{S}^\circ$. By similar argument we also conclude that $\mathcal{Y}^\circ$ is an exposed face of the set of separable states $\mathcal{S}$.

We have seen that the two sets $\mathcal{X}^\circ$ and $\mathcal{Y}^\circ$ are exposed faces on $\mathcal{S}^\circ$ and $\mathcal{S}$. We have also observed that $\mathcal{X}^\circ$ is the dual of a separable state ρ_0 and that $\mathcal{Y}^\circ$ is the dual of a witness W_0. We can now summarize this into a result that expresses a duality between exposed faces of $\mathcal{S}$ and $\mathcal{S}^\circ$

Theorem 7.7 (Duality of exposed faces in $\mathcal{S}$ and $\mathcal{S}^\circ$)**.** *There is an isomorphic relation between exposed faces of $\mathcal{S}$ and exposed faces of $\mathcal{S}^\circ$. The faces in each pair are orthogonal to each other.*

We know that in any nested hierarchy of faces $\mathcal{F}_1 \supset \mathcal{F}_2 \ldots \supset \mathcal{F}_k$, extremal points are inherited. It thus follows that the extremal points of the face $\mathcal{Y}^\circ \subset \mathcal{S}$ are the zeros which are common to all the witnesses in $\mathcal{Y} \subset \mathcal{S}^\circ$. So an important result is that

Theorem 7.8 (Zeros of W and exposed faces of $\mathcal{S}$)**.** *A set of product vectors $\psi_i = \phi_i \otimes \chi_i \in \mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ is the complete set of zeros of some witness W if and only if all $\rho_i = \psi_i\psi_i^\dagger$ are the extremal points of an exposed face of $\mathcal{S}$.*

It is unknown whether unexposed faces of $\mathcal{S}$ exist. We know that for any compact convex set, an unexposed face must be contained in an exposed face, so any hypothetical unexposed face of $\mathcal{S}$ must lie in a larger exposed face.

7.5.2 Unexposed faces of $\mathcal{S}^\circ$

A version of the Choi-Lam witness in (7.57) can be written $W_C = \Omega(1,0,1;0)$. This witness is extremal, and represents a zero-dimensional face $\mathcal{F}_{\mathcal{S}^\circ}(W_C)$ which is unexposed. The three isolated product vectors from (7.58) define the states $e_1e_1^\dagger$, $e_2e_2^\dagger$, $e_3e_3^\dagger$, and the continuum of product vectors defined in (7.59) define the states $\phi(\alpha,\beta)\phi^\dagger(\alpha,\beta)$. The dual face $W_C^\circ \subset \mathcal{S}$ then has the states $e_1e_1^\dagger$, $e_2e_2^\dagger$, $e_3e_3^\dagger$ and $\phi(\alpha,\beta)\phi^\dagger(\alpha,\beta)$ as extremal points. The face W_C° has dimension 21, and the separable states in the interior of this face have rank seven, since the zeros span a seven-dimensional subspace.

Numerically we find that the constraints $\mathbf{U}_0$ and $\mathbf{U}_1$ associated with the zeros of W_C define a four-dimensional face of $\mathcal{S}^\circ$. The rest of the constraints comes from $\mathbf{U}_2$ and $\mathbf{U}_3$ which are associated with the quartic nature of the zeros. This is an example of how faces of $\mathcal{S}^\circ$ can be unexposed. In general, a witness W having one or more isolated quartic zeros, will be an interior point of an unexposed face on $\mathcal{S}^\circ$. This unexposed face is then a face of a larger exposed face consisting of witnesses having the same zeros as W, but such that for these witnesses all these isolated zeros are quadratic.

7.5.3 Simplex faces and other faces of $\mathcal{S}$

We here distinguish between two different types of faces on $\mathcal{S}$, namely the faces which take the form of a simplex, and the faces that do not. The simplex type of faces are closely related to quadratic witnesses, while other types are connected to quartic witnesses.

Simplex faces of $\mathcal{S}$

From Theorem 7.7 we find the relation between zeros of witnesses and exposed faces on $\mathcal{S}$. The exposed faces of $\mathcal{S}$ defined by extremal quadratic witnesses are simplexes. We discussed simplexes as convex sets more thoroughly in Section 4.5, and we defined an n-simplex Δ^n to be a simplex with $n+1$ vertices and an interior with dimension n. Given the product vectors $\psi_i = \phi_i \otimes \chi_i$ for $i = 1,\ldots,k$ so that $n = k-1$, as zeros of a generic extremal quadratic witness W, the corresponding product states $\rho_i = \psi_i\psi_i^\dagger$ are the vertices of an n-simplex, which is the exposed face W° on $\mathcal{S}$ dual to W.

Let ρ be an interior point of this face

$$\rho = \sum_{i=1}^{k} p_i\rho_i \qquad\qquad \sum_i p_i = 1 \qquad\qquad p_i > 0 \tag{7.66}$$

From Theorem 7.6 we understand that ρ constructed from all the zeros $\psi_i = \phi_i \otimes \chi_i$ and ρ^P constructed from all the zeros $\psi_i = \phi_i \otimes \chi_i^*$, have full rank $N = N_AN_B$. Thus for these states we will have $\rho \in \text{int}(\mathcal{P}) \subset \text{int}(\mathcal{D})$. These full rank states make up the interior of the simplex Δ^{N-1}. For instance for the 3×3 system, these states will have rank nine, and the entire simplex Δ^8 will be a face of $\mathcal{S}$, *i.e.* lying on $\partial\mathcal{S}$. But apart from the extremal states $\rho_i = \psi_i\psi_i^\dagger$ the simplex Δ^8 will lie in the interior of $\mathcal{P}$. The boundary of Δ^8 will consist of other faces of $\mathcal{S}$ defined by less than nine zeros,

accordingly these faces are located inside faces of $\mathcal{D}$, *i.e.* on $\partial\mathcal{D}$. In dimensions that are large enough (Table 7.1), such a witness has more than $N = N_A N_B$ zeros, and define a simplex face of $\mathcal{S}$ for which also the boundary points may lie in the interior of $\mathcal{D}$. In this way the geometry of $\mathcal{S}$ in relation to $\mathcal{D}$, appear to become more and more complicated as the dimensions increase.

One question that could be raised is whether *any* generic simplex $\Delta^{N-1} \in \mathcal{D}_N$, with n_c vertices $\psi_i\psi_i^\dagger$ with $\psi_i = \phi_i \otimes \chi_i$, define an exposed simplex face of $\mathcal{S}$. If $n_c > N$, which is the case when $N_A, N_B > 3$, then the simplex is overdetermined by the fact that the largest possible rank of the set ψ_i is N. If the answer is in the affirmative, then any such generic collection of vertices ψ_i will be zeros of an extremal quadratic witness.

We have made investigations in the 3×3 system, into the geometrical *orientation* of 8-simplexes $\Delta^8 \in \mathcal{D}_9$ relative to the maximally mixed state $\rho_0 = I/9$. On such a simplex there always exists a (usually unique) state $\rho_{\min}$ closest to ρ_0. The minimum distance is given by the Hilbert-Schmidt norm $d_{\min} = |\rho_{\min} - \rho_0|$. In dimension $N = N_A N_B$ the distance from any pure state to the maximally mixed state ρ_0 is $\sqrt{(N-1)/N}$, and in our case this upper limit is $d_{\min} \leq \sqrt{8}/3 \approx 0.94$.

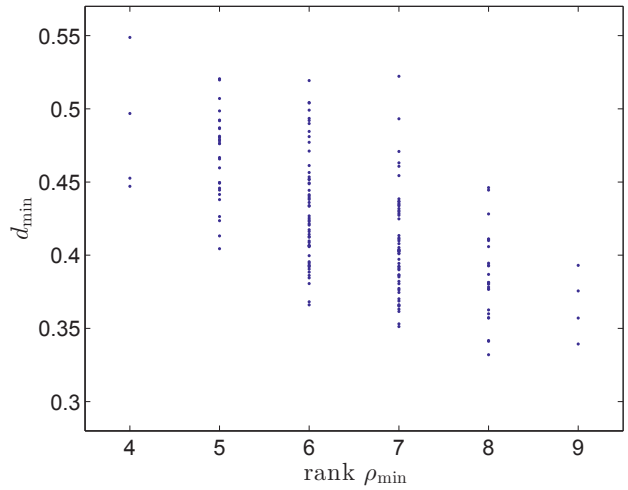


Figure 7.6: The rank of $\rho_{\min}$ and its distance $d_{\min} = |\rho_{\min} - \rho_0|$ from the maximally mixed state ρ_0, for a large number of 8-simplexes Δ^8 in dimension 3×3. The distance $d_{\min}$ is in the Hilbert-Schmidt metric. The 8-simplexes are faces on $\mathcal{S}$, *i.e.*, they are constructed from the zeros of a large number of extremal quadratic witnesses found in random searches using Algorithm 1. A similar scheme for randomly generated 8-simplexes return approximately 90% states of rank nine.

For the 3×3 system we find numerically, if we generate a large number of 8-simplexes defined by a random set of nine pure product states, that for approximately 90% of the cases, the state $\rho_{\min}$ will have full rank, *i.e* it will lie in the interior of Δ^8 and $\mathcal{D}_9$. For the remaining cases we find that $\rho_{\min}$ mostly has rank eight, and very rarely rank seven. Figure 7.6 shows the results for the sample of 171 extremal quadratic

witnesses we have generated using Algorithm 1. For the cases in Figure 7.6 it happens in only about 2% of the cases that $\rho_{\min}$ has full rank. In the remaining cases $\rho_{\min}$ lies in $\partial\Delta^8$ and $\partial\mathcal{D}_9$. We see a tendency that the minimum distance $d_{\min}$ is smaller when $\text{rank}(\rho_{\min})$ is higher. This is because the most regular simplex faces are positioned most symmetrically relative to the maximally mixed state.

Thus Figure 7.6 clearly verifies numerically that the orientations of randomly generated simplexes relative to $\rho_0 = I/9$, has a distribution which is different than for the simplexes which are constructed from zeros of extremal quadratic witnesses. So it is likely that any random simplex Δ^8 will in general not define zeros of an entanglement witness, and that the simplex will then not be a face of $\mathcal{S}$.

Other types of faces of $\mathcal{S}$

Since any entanglement witness $W \in \partial\mathcal{S}^\circ$ lies on an exposed face of $\mathcal{S}^\circ$, it follows effectively from Theorem 7.7 that to every entanglement witness $W \in \partial\mathcal{S}^\circ$ there corresponds an exposed face on $\mathcal{S}$. The large variety of entanglement witnesses implies a similar variety of faces of $\mathcal{S}$. The class of simplex faces have only a finite number of extremal points, while other faces may have only continuous sets of extremal points, and further we may have faces with both discrete and continuous subsets of extremal points.

We have seen that quadratic witnesses can be related to simplex faces of $\mathcal{S}$. Another well known example of an extremal witness is the Choi-Lam witness W_C, which is quartic and extremal. It has three discrete zeros and one continuous set of zeros, and these zeros define extremal points of a face of $\mathcal{S}$ which is quite different from the simplex faces.

If we consider a pure state $W = \psi\psi^\dagger$ which is quartic and extremal on the set of decomposable witnesses $\mathcal{P}^\circ$. The zeros of W are the product vectors which are orthogonal to ψ. Using the standard basis e_i for $\mathbb{C}^3$ we can look at two very simple examples. We write a general product vector as

$$\phi \otimes \chi = \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} \otimes \begin{pmatrix} g_1 \\ g_2 \\ g_3 \end{pmatrix} \tag{7.67}$$

with $f_i, g_i \in \mathbb{C}$

- If we take $\psi = e_1 \otimes e_1$, then all components ψ_i are zero, except $\psi_1 = 1$. So for $W = \psi\psi^\dagger$ we get

$$(\phi \otimes \chi)^\dagger W (\phi \otimes \chi) = 0 \qquad \Rightarrow \qquad f_1 g_1 = 0 \tag{7.68}$$

  with either $f_1 = 0$ or $g_1 = 0$. The product vectors $\phi \otimes \chi$ satisfying this, produce product states belonging to a full matrix algebra in dimension 2×3 or 3×2, depending on whether $f_1 = 0$ or $g_1 = 0$. They are extremal points on a face of the type discussed by Alfsen and Shultz [51], which they call a convex combination of matrix algebras.

- We can choose the Bell type entangled state $\psi = \sqrt{\frac{1}{3}}\left(|11\rangle + |22\rangle + |33\rangle\right)$. Keeping in mind that we write $e_i = |i\rangle$, we get for $W = \psi\psi^\dagger$

$$(\phi \otimes \chi)^\dagger W(\phi \otimes \chi) = 0 \qquad \Rightarrow \qquad f_1 g_1 + f_2 g_2 + f_3 g_3 = 0 \tag{7.69}$$

  This means that for every vector ϕ, the vector χ defines a two-dimensional subspace of vectors satisfying the orthogonality condition, and vice versa. These pure product states are then the extremal points of a face on $\mathcal{S}$ which is neither a simplex nor a convex combination of matrix algebras.

So it is pretty clear that there is a great variety of faces on $\mathcal{S}$, and the structure of these faces is far from understood. As we shall immediately see, this has consequences for the convex decomposition of separable states, where we naturally want to be able to find a decomposition into the minimum number of pure product states.

7.5.4 Minimal convex decomposition of separable states

In Section 5.8 we constructed a procedure for decomposing an arbitrary state $\rho \in \mathcal{D}$ of rank r, as a convex combination of r pure states $\psi\psi^\dagger$

$$\rho = \sum_{i=1}^{r} p_i \left(\psi_i \psi_i^\dagger\right) \qquad p_i > 0 \qquad \sum_i p_i = 1 \tag{7.70}$$

The essential characteristic of $\mathcal{D}$ that ensures the possibility of such a procedure, is the structure of faces $\mathcal{F}_\mathcal{D}$ on $\mathcal{D}$. By the Werner definition [21], a state ρ on $\mathbb{C}^N$ is separable if it can be expressed as

$$\rho = \sum_{i=1}^{s} p_i \left(\mu_i \otimes \tau_i\right) = \sum_{i=1}^{s} p_i \left(\phi_i \phi_i^\dagger \otimes \chi_i \chi_i^\dagger\right) \tag{7.71}$$

It is obvious that if we pick a set of $n \leq N$ linearly independent product vectors $\phi_i \otimes \chi_i$ we can invariably form a separable state of the form (7.71), with rank $n = s$. But if we approach from the other end, given a separable state ρ, what is the minimum number $s_{\min}$ necessary to write ρ on the form (7.71). This question is closely related to the structure of faces of $\mathcal{S}$. If we use the procedure in Section 5.8 in order to decompose $\rho \in \mathcal{S}$ on form (7.71), it might happen, that the face $\mathcal{F}_2$ is the dual of a quadratic witness. It will then contain only a finite number of pure product states, much smaller than the dimension $N^2 - 1$ of the set of normalized separable states. Recall that the pure product states in such a face are the zeros of this quadratic witness, and the number of zeros is largest when the witness is extremal. In the generic case the number of zeros of a quadratic extremal witness is n_c, as given by (7.52).

It therefore seems likely that also in the decomposition of separable states one could do much better than the N^2 pure product states guaranteed by Carathéodory's theorem.

7.6 Optimal witnesses

The fact that there does not exist a universal entanglement witness, gives rise to the notion of *optimal* witnesses. There are many, and equivalent, ways of defining optimality for witnesses, but one possible way is

Definition 7.6 (Optimal witness)**.** *Let the set of entangled states revealed by an entanglement witness $W \in \mathcal{S}^\circ$ be*

$$\Xi_W = \{\rho \in \mathcal{D}_N \,|\, \mathrm{tr}(W\rho) < 0\} \tag{7.72}$$

Then a witness $W_1 \in \mathcal{S}^\circ$ is an optimal witness if and only if there does not exist a witness $W_2 \in \mathcal{S}^\circ$ such that $\Xi_{W_1} \subset \Xi_{W_2}$.

Given a witness W, the expression $\mathrm{tr}(W\rho)$ linearly defines a hyperplane in the set of Hermitean matrices H_N. This hyperplane will split $\mathcal{D}_N$ into two regions. One region with $\mathrm{tr}(W\rho) > 0$, which for W to be a witness must include the complete set of separable states $\mathcal{S}$, and another where $\mathrm{tr}(W\rho) < 0$ which is identical to Ξ_W. An attempt to visualize this geometrically is made in Figure 7.7.

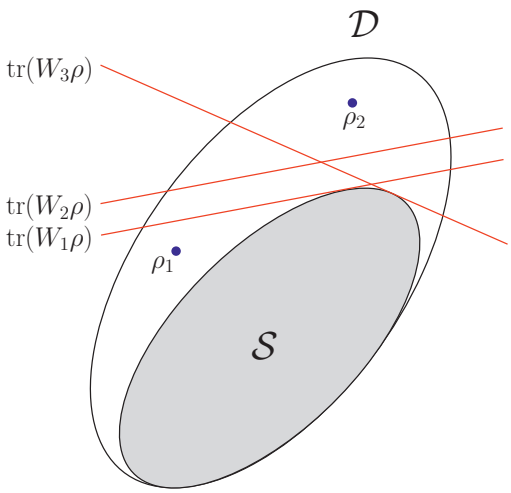


Figure 7.7: We see the sets $\mathcal{S}$ and $\mathcal{D}$, and the red lines identify the three different hyperplanes $\mathrm{tr}(W_k\rho)$ defined by three witnesses W_1, W_2 and W_3. Since $\mathrm{tr}(W_k\rho_1) < 0$ for all W_k, it is revealed by all three witnesses, and since $\mathrm{tr}(W_k\rho_2) > 0$ for all W_k, it is not revealed by any of the three witnesses.

We see red lines that identify the different hyperplanes defined by the three witnesses W_1, W_2 and W_3. We observe that the hyperplane $\mathrm{tr}(W_2\rho)$ is not tangent to $\mathcal{S}$, which in fact means that W_2 cannot be optimal. This is because there is always a witness W_1 with a hyperplane $\mathrm{tr}(W_1\rho)$ parallel to $\mathrm{tr}(W_2\rho)$ which is tangent to $\mathcal{S}$, and thus $\Xi_{W_2} \subset \Xi_{W_1}$. Generally, the hyperplane $\mathrm{tr}(W\rho)$ defined by any witness W, cannot

intersect the interior of $\mathcal{S}$, because then we would have $\mathrm{tr}(W\rho) < 0$ for some separable state ρ. So a necessary condition for W to be optimal is that $\mathrm{tr}(W\rho)$ must be tangent to $\mathcal{S}$, and thus W must have at least one zero, *i.e.* $W \in \partial\mathcal{S}^\circ$. It is important to emphasize that this is a necessary, but not sufficient condition. We see that the entangled state ρ_1 is revealed by all three witnesses, while ρ_2 is revealed by none of them.

There are also more constructive, but equivalent, ways of defining optimality of witnesses, and this was done originally by Lewenstein *et al.* [78, 79]. They introduced two different concepts of optimality, and gave two necessary and sufficient conditions for optimality, rather similar to the condition for extremality. These definitions are made by looking at ways in which a witness W *cannot* be expressed. Assume that for a witness W the genuine convex combination $W = (1-p)M_1 + pM_2$ with M_1, M_2 in a specified subset of the Hermitian matrices, is *not* possible, *i.e.*

$$W \neq (1-p)M_1 + pM_2 \qquad 0 < p < 1 \tag{7.73}$$

Based on (7.73) the following definitions are made

$$\begin{aligned} M_1, M_2 \in \mathcal{S}^\circ \quad &\Rightarrow \quad W \text{ is extremal} \\ M_1 \in \mathcal{S}^\circ, M_2 \in \mathcal{D} \quad &\Rightarrow \quad W \text{ is optimal} \\ M_1 \in \mathcal{S}^\circ, M_2 \in \mathcal{P}^\circ \quad &\Rightarrow \quad W \text{ is non-decomposable optimal} \end{aligned} \tag{7.74}$$

The first, being the well known definition of extremality. Obviously, a witness which is extremal is optimal by both optimality criteria. Note that the criterion for optimality is not invariant under partial transposition, hence a witness W may be optimal while its partial transpose W^P is not optimal. The criterion for non-decomposable optimality, on the other hand, is invariant under partial transposition. The condition of being extremal is stricter than those of being optimal or non-decomposable optimal. Accordingly there should, in general, exist plenty of non-extremal witnesses that are either optimal or non-decomposable optimal. Keeping in mind Definition 7.5, the following sufficient optimality conditions were proved by Lewenstein *et al.* [78]

Theorem 7.9 (Optimality condition)**.** *A witness W is optimal if it has spanning property, and it is non-decomposable optimal if it has double spanning property.*

The opposite is not necessarily true. In the 3×3 system we have the Choi-Lam witness $W_C = \Omega(1, 0, 1; 0)$ from (7.57), which is extremal, and therefore optimal, but Π_{W_C} only spans a seven-dimensional subspace of $\mathbb{C}^9$. For a quadratic witness however these conditions are not only sufficient, but also necessary [71]

Theorem 7.10 (Optimality for quadratic witnesses)**.** *Let W be a quadratic witness. Then W is optimal if and only if it has spanning property, and it is non-decomposable optimal if and only if it has double spanning property.*

It is known that the equivalence between optimality and the spanning property holds also in the case of decomposable witnesses in $2 \times N_B$ dimensions [80], but that for

$N_A, N_B \geq 3$ there exist optimal decomposable witnesses without the spanning property [81]. By Theorem 7.10 a test for optimality of a witness W involves finding the complete set of zeros Π_W. If this set spans the Hilbert space, then W is optimal, if it does not then W is either non-optimal quadratic or quartic. We must then calculate $\mathrm{Ker}\, G_W$ and thus verify whether W is quadratic or not. If it is, then then the test ends, with the result that W is non-optimal. If it turns out that the witness is quartic, we need further analysis.

The definition of optimality given in (7.74), can be slightly reformulated to read: If W is a witness, and there exists a real $\lambda > 0$ and positive matrix $P \geq 0$ such that $W' = W - \lambda P \in \mathcal{S}^\circ$, then W is not optimal. Using this formulation, an optimality test for quartic witnesses has been constructed [78]. Since we do not have a complete understanding of constraints from quartic zeros, the relation between non-decomposable optimality and extremality of quartic witnesses still remains an open problem for future research.

In dimensions 3×3 and $2 \times N_B$, a witness with doubly spanning zeros is generically extremal by Theorem 7.6, and hence both optimal and non-decomposable optimal. Repeating our discussion in Section 7.4.3, we remember that in higher dimensions the number of zeros n_c of a quadratic extremal witness must be larger than the Hilbert space dimension $N = N_A N_B$. Hence non-extremal witnesses with doubly spanning zeros must be very common, so that non-decomposable optimality is a significantly weaker property than extremality. In particular, when we search for generic quadratic extremal witnesses and reach the stage where the witnesses have N or more zeros, then every face of $\mathcal{S}^\circ$ we encounter will consist entirely of non-decomposable optimal witnesses, most of which are not extremal.

So, to conclude some of our discussion here: For the 2×2 and 2×3 systems, every witness is decomposable [33, 66]. In dimensions $2 \times N_B$ with $N_B > 3$ and 3×3, a generic quadratic witness is either non-decomposable optimal and extremal, or it is neither. In higher dimensions there exist an abundance of non-decomposable optimal non-extremal witnesses.

The SPA conjecture

We have seen that even though they are very useful as an operation for detecting entanglement, maps which are positive but not completely positive, are non-physical. Horodecki and Ekert introduced the concept of a "structural physical approximation" (SPA) of a non-physical positive map [82]. The SPA is a physical map and may be implemented in an experimental setup. The transposition map for instance, is non-physical, since it is not completely positive.

The SPA of the corresponding witness W is defined by using the so-called depolarizing channel

$$\Sigma(p) = (1-p)W + \frac{p}{N} I \tag{7.75}$$

where $p \in X = [0, 1]$. The SPA of the witness W is defined to be the positive matrix

$\Sigma(p)$ of the form (7.75) which is closest to W. Put

$$\begin{aligned} p_0 &= \{\min(X) \,|\, \Sigma(p) \in \mathcal{S}\} \\ p_1 &= \{\min(X) \,|\, \Sigma(p) \in \mathcal{D}\} \\ p_2 &= \{\min(X) \,|\, [\Sigma(p)]^P \in \mathcal{D}\} \end{aligned} \tag{7.76}$$

The SPA of W is then defined as $W_{\mathrm{SPA}} = \Sigma(p_1)$. Clearly $p_0 \geq p_1$ and $p_0 \geq p_2$, but we may have either $p_1 < p_2$, $p_1 = p_2$ or $p_1 > p_2$. The SPA of W is a PPT state if and only if $p_1 \geq p_2$. It follows directly from the definitions that the SPA of the witness W^P is $[\Sigma(p_2)]^P$, and this is a PPT state if and only if $p_2 \geq p_1$.

The so called *SPA separability conjecture* states that the SPA of an optimal entanglement witness is always a separable density matrix [83]. Thus for optimal witnesses we should have $p_0 = p_1 \geq p_2$. One reason for the interest in this conjecture is that separability simplifies the physical implementation of the corresponding map.

Counterexamples in the form of the generalized Choi-Lam witness (7.57), have been given by Ha and Kye, and by Størmer [73, 85]. With the parameter values considered by Størmer we find that the witnesses are not extremal, and their SPAs are entangled because they are not PPT states. More interesting are the parameter values considered by Ha and Kye, because they give extremal witnesses with SPAs that are entangled PPT states.

The above counterexamples are non-decomposable witnesses, but the case for decomposable witnesses has also been closed by Chruściński and Sarbicki [84]. They show that the SPA of an optimal but non-extremal decomposable witness, of the form $W = \sigma^P$ with $\sigma \geq 0$ of rank three, may be entangled, although it is automatically a PPT state.

The SPA conjecture applies directly to extremal entanglement witnesses, since they are optimal. If W is an extremal witness, then so is the partial transpose W^P. However, if the SPA of W is a PPT state, then the SPA of W^P is usually not. This means that the original SPA conjecture cannot be true both for W and W^P, simply because separable states are PPT states. Ha and Kye gave essentially this argument in [85]. The extremal witnesses that we have produced in a random search by Algorithm 1, support this argument, in that about half of the witnesses we find, have an SPA which is a PPT state.

An obvious modification of the conjecture would be to redefine the SPA as the nearest PPT matrix instead of the nearest positive matrix. In the case of a pair of extremal witnesses W and W^P, the modified conjecture holds either for both or for none. However, the counterexamples cited above disprove also this modified separability conjecture in its full generality. It is nevertheless the case that these known counterexamples are rather non-generic.

If we regard separability as an essential property of the SPA, it seems that the natural solution would be to define the SPA by the parameter value $p = p_0$, where $\Sigma(p)$ first becomes separable, rather than by the value $p = p_1$, where $\Sigma(p)$ first becomes positive, or by $p = \max(p_1, p_2)$, where $\Sigma(p)$ first becomes a PPT state.

7.7 Unital and trace preserving positive maps

We now pick up the discussion of positive maps from Section 7.2. For a map $\mathbf{L}_M : H_{N_A} \mapsto H_{N_B}$ we will now define two very useful properties

Definition 7.7 (Unital and trace preserving)**.** *The map $\mathbf{L}_M : H_{N_A} \mapsto H_{N_B}$ is unital if it maps the identity $I_A \in H_{N_A}$ to the identity $I_B \in H_{N_B}$, and trace preserving if* $\mathrm{tr}(\mathbf{L}_M X) = \mathrm{tr}(X)$ *for all $X \in H_{N_A}$*

The trace preserving property is quite useful since a physical map should preserve probability, and the unital property naturally defines a fixed point for the map. Using (7.11) we must have

$$\sum_{i,k} \hat{M}_{jl;ik}\delta_{ik} = \sum_{i} \hat{M}_{jl;ii} = \delta_{jl} \tag{7.77}$$

and for the trace preserving property

$$\sum_{j,i,k} \hat{M}_{jj;ik}X_{ik} = \sum_{i} X_{ii} \qquad \Rightarrow \qquad \sum_{j} \hat{M}_{jj;ik} = \delta_{ik} \tag{7.78}$$

Thus the map $\mathbf{L}_{\hat{M}}$ is trace preserving if and only if the map $\mathbf{L}^T_{\hat{M}}$ is unital.

In terms of the corresponding Hermitian matrix $M = M_{ij;kl}$ defined by (7.10) we can write the condition that $\mathbf{L}_M$ be unital as

$$\mathbf{L}_M I_A = \mathrm{tr}_A(M) = I_B \tag{7.79}$$

and to be trace preserving

$$\mathbf{L}^T_M I_B = \mathrm{tr}_B(M) = I_A \tag{7.80}$$

Equation (7.79) implies that $\mathrm{tr}(M) = \mathrm{tr}(I_B) = N_B$, whereas (7.80) implies that $\mathrm{tr}(M) = \mathrm{tr}(I_A) = N_A$. So it appears that the map may be both unital and trace preserving only if $N_A = N_B$. If $N_A \neq N_B$ the proper condition can be adjusted to

$$\mathbf{L}_M E_0 = F_0 \qquad\qquad \mathbf{L}^T_M F_0 = E_0 \tag{7.81}$$

where $E_0 = I_A/\sqrt{N_A}$ and $F_0 = I_B/\sqrt{N_B}$. We can then call the map unital and trace preserving if (7.81) holds, even if $N_A \neq N_B$.

7.7.1 Transforming positive maps

We can argue, very much along the lines of Størmer [73], that every positive map $\mathbf{L}_W$ may be transformed into a unital and trace preserving form through a product transformation on the corresponding witness $W \in \mathcal{S}^\circ$

$$W \mapsto \widetilde{W} = (U \otimes V)W(U \otimes V)^\dagger \tag{7.82}$$

where $U \in \mathrm{GL}(N_A, \mathbb{C})$ and $V \in \mathrm{GL}(N_B, \mathbb{C})$. A product transformation of this kind preserves all the essential characteristics of W. For instance, if W is extremal in $\mathcal{S}^\circ$ and

non-decomposable, then so is $\widetilde{W}$, and a zero $\phi \otimes \chi$ of W corresponds to a zero $\widetilde{\phi} \otimes \widetilde{\chi}$ of $\widetilde{W}$, where

$$\widetilde{\phi} = (U^\dagger)^{-1}\phi \qquad \widetilde{\chi} = (V^\dagger)^{-1}\chi \tag{7.83}$$

It has been proved by Leinaas *et al.* [74] that such a transformation is always possible for a more limited set of matrices, namely the set of witnesses having no zeros, *i.e.* witnesses not on the boundary of $\mathcal{S}^\circ$. As we have seen however, it is the set of witnesses on the boundary $\partial\mathcal{S}^\circ$ which are the most interesting.

In terms of the transformed map $\widetilde{\mathbf{L}}_W = \mathbf{L}_{\widetilde{W}}$, the conditions to be fulfilled are that $\widetilde{\mathbf{L}}_W I_A = I_B$ and $\widetilde{\mathbf{L}}_W^T I_B = I_A$. We will assume here that $N_A = N_B$, otherwise proper adjustments are easily made according to (7.81). In index notation (7.82) reads as follows

$$\widetilde{W}_{ij;kl} = \sum_{a,b,c,d} U_{ia} V_{jb}\, W_{ab;cd}\, U^*_{kc} V^*_{ld} \tag{7.84}$$

It is possible to show that the maps $\mathbf{L}_W : H_{N_A} \mapsto H_{N_B}$ and $\mathbf{L}_W^T : H_{N_B} \mapsto H_{N_A}$ act on $X \in H_{N_A}$ and $Y \in H_{N_B}$ respectively as

$$\mathbf{L}_W X = \mathrm{tr}_A[W(X \otimes I)] \qquad\qquad \mathbf{L}_W^T Y = \mathrm{tr}_B[W(I \otimes Y)] \tag{7.85}$$

and using this, we may write the transformation $Y = \widetilde{\mathbf{L}}_W X$ as

$$Y_{jl} = \sum_{i,k} \widetilde{W}_{ij;kl}\, X_{ki} = \sum_{i,k,a,b,c,d} V_{jb}[W_{ab;cd}(U^*_{kc}\, X_{ki}\, U_{ia})]V^*_{ld} \tag{7.86}$$

or more compactly as

$$Y = \widetilde{\mathbf{L}}_W X = V[\mathbf{L}_W(U^\dagger X U)]V^\dagger \tag{7.87}$$

By similar argument for the transformation $X = \widetilde{\mathbf{L}}_W^T Y$ we get

$$X = \widetilde{\mathbf{L}}_W^T Y = U[\mathbf{L}_W^T(V^\dagger Y V)]U^\dagger \tag{7.88}$$

So the conditions for $\widetilde{\mathbf{L}}_W$ to be unital and trace preserving is that

$$\begin{aligned} \widetilde{\mathbf{L}}_W I_A &= V[\mathbf{L}_W(U^\dagger U)]V^\dagger = I_B \\ \widetilde{\mathbf{L}}_W^T I_B &= U[\mathbf{L}_W^T(V^\dagger V)]U^\dagger = I_A \end{aligned} \tag{7.89}$$

So, in total, the problem to be solved is to find operators U and V such that

$$\mathbf{L}_W(U^\dagger U) = (V^\dagger V)^{-1} \qquad \mathbf{L}_W^T(V^\dagger V) = (U^\dagger U)^{-1} \tag{7.90}$$

We have used an iteration scheme, presented in Section 8.7, to solve (7.90) for a large number of numerically produced extremal entanglement witnesses [71, 75], and also on many non-extremal witnesses. Numerically our attempts, which are in the thousands, always converge to a unique $\widetilde{W}$ for each W.

7.7.2 Extremal positive maps

We have seen in Section 7.4 how entanglement witnesses, and in particular extremal witnesses, are understood in terms of their zeros, and we have to understand what the zeros of the witness W imply for the corresponding map $\mathbf{L}_W$.

If we let (ϕ_0, χ_0) be an isolated zero of W, and if we further define $Y = \mathbf{L}_W(\phi_0\phi_0^\dagger)$ and $X = \mathbf{L}_W^T(\chi_0\chi_0^\dagger)$, we can put

$$\chi_0^\dagger Y \chi_0 = \phi_0^\dagger X \phi_0 = (\phi_0 \otimes \chi_0)^\dagger W(\phi_0 \otimes \chi_0) = 0 \tag{7.91}$$

Since $X, Y \geq 0$ it follows that

$$X\phi_0 = 0 \qquad\qquad Y\chi_0 = 0 \tag{7.92}$$

Since the zero (ϕ_0, χ_0) is an isolated zero, we also know that X and Y have no other zeros in $\mathbb{C}^{N_A}$ and $\mathbb{C}^{N_B}$ respectively. Hence X has rank $N_A - 1$ and Y has rank $N_B - 1$.

When W is an extremal entanglement witness, the corresponding maps $\mathbf{L}_W$ and $\mathbf{L}_W^T$ are extremal positive maps. An isolated zero (ϕ_0, χ_0) of W defines a rank one state $\phi_0\phi_0^\dagger \in \mathcal{D}_{N_A}$ mapped by $\mathbf{L}_W$ into a rank $N_B - 1$ state in $\mathcal{D}_{N_B}$. It also defines a rank one state $\chi_0\chi_0^\dagger \in \mathcal{D}_{N_B}$ mapped by $\mathbf{L}_W^T$ into a rank $N_A - 1$ state in $\mathcal{D}_{N_A}$. So the zero (ϕ_0, χ_0) of W defines a point on $\partial\mathcal{D}_{N_A}$ which is mapped by $\mathbf{L}_W$ to a point on $\partial\mathcal{D}_{N_B}$, and likewise a point on $\partial\mathcal{D}_{N_B}$ which is mapped by $\mathbf{L}_W^T$ to a point on $\partial\mathcal{D}_{N_A}$. We can conclude that the map $\mathbf{L}_W$ is extremal precisely because the image $\mathbf{L}_W\mathcal{D}_{N_A}$ inside $\mathcal{D}_{N_B}$ touches $\partial\mathcal{D}_{N_B}$ in as many points as possible.

The extremal maps $\mathbf{L}_W : H_{N_A} \mapsto H_{N_B}$ that correspond to generic quadratic extremal witnesses, thus map only a finite number of points on $\partial\mathcal{D}_{N_A}$ to $\partial\mathcal{D}_{N_B}$, and they map all other points to the interior of $\mathcal{D}_{N_B}$. So an extremal map of this type is contractive, which means that it maps the set of density matrices onto a subset of itself, which again entails that it reduces volumes.

A very different type of extremal positive maps are those that carry positive rank one operators to positive rank one operators, and these are called *rank one preservers*. Two simple examples of such maps are the identity map $\mathbf{I} : X \mapsto X$ and the transposition map $\mathbf{T} : X \mapsto X^T$. But more generally any map $\mathbf{U} : X \mapsto UXU^\dagger$ will be a rank one preserver, and then obviously also $\mathbf{UT}$. These positive maps are extremal because they map $\partial\mathcal{D}_{N_A}$ completely onto $\partial\mathcal{D}_{N_B}$.

7.7.3 Visualizations of extremal positive maps

Here we consider the qutrit case $N_A = N_B = 3$. The set of normalized density matrices $\mathcal{D}_3$ has dimension $3^2 - 1 = 8$, and as we have seen in Section 7.7.1, in this case a positive linear map can always be transformed into a unital and trace preserving form. Given an unital and trace preserving positive map $\mathbf{L}_W : H_3 \mapsto H_3$ it will map $\mathcal{D}_3$ into $\mathcal{D}_3$, and the maximally mixed state $I/3$ to itself. It is possible to plot two-dimensional planar sections of the density matrices in order to see what $\mathbf{L}_W\mathcal{D}_3$ looks like inside $\mathcal{D}_3$ for various choices of sections. We have seen earlier, in Section 5.4, how to perform such plots by choosing an origin ρ_0 and two traceless and orthogonal directions σ_1 and σ_2.

These together define the plane $\mathcal{Z} \in H_3$ containing the states ρ_0, ρ_1 and ρ_2, with the two latter being the density matrices corresponding to σ_1 and σ_2. See expression (5.9) and Figure 5.1. A matrix in $\mathcal{Z}$ is then specified by a coordinate pair (x, y)

$$X = \rho_0 + x\sigma_1 + y\sigma_2 \tag{7.93}$$

Once the plane $\mathcal{Z}$ is known, the procedure described in Section 5.4 and illustrated in Figure 5.2, can be used to calculate the border of $\mathcal{D}_3 \subset \mathcal{Z}$.

In order to visualize maps in this way, we let the positive map $\mathbf{L}_W$ be extremal, corresponding to an extremal witness W, and then use $\mathbf{L}_W$ to map $\mathcal{D}_3$ into $\mathcal{D}_3$ for different types of sections $\mathcal{Z}$.

The positive map $\mathbf{L}_W$ maps the plane $\mathcal{Z}$ defined by the states $\rho_i \in \mathcal{D}_3$ with $i = 0, 1, 2$ into the plane $\mathcal{Z}' = \mathbf{L}_W\mathcal{Z}$ defined by the states $\rho_i' = \mathbf{L}_W\rho_i \in \mathcal{D}_3$, in such a way that X as given by (7.93) is mapped into

$$X' = \mathbf{L}_W X = \rho_0' + x\sigma_1' + y\sigma_2' \tag{7.94}$$

Since $\mathbf{L}_W$ is unital and trace preserving, $\mathrm{tr}(\sigma_1') = \mathrm{tr}(\sigma_2') = 0$, but we must also require that $\mathrm{tr}({\sigma_1'}^2) = \mathrm{tr}({\sigma_2'}^2) = 1$ and $\mathrm{tr}(\sigma_1'\sigma_2') = 0$. This is to ensure that the plots reproduce the distances in the image plane $\mathcal{Z}'$ of the map $\mathbf{L}_W$ faithfully. It is in general not possible to make both $\mathrm{tr}(\sigma_1\sigma_2) = 0$ and $\mathrm{tr}(\sigma_1'\sigma_2') = 0$, so the plot will give a distorted representation of the plane $\mathcal{Z}$. Note that in order to make $\mathrm{tr}(\sigma_1'\sigma_2') = 0$ we have consistently rotated the σ_2' or y-axis, so that ρ_1' always has coordinates $(x > 0, y = 0)$ whereas ρ_2' always has $y > 0$, but in general $x \neq 0$. We plot $\partial\mathcal{D}_3$ in $\mathcal{Z}'$ in blue, and the image of $\partial\mathcal{D}_3$ in $\mathcal{Z}$ by a red curve. The "+" represents the orthogonal projection of the maximally mixed state $I/3$ on the image plane $\mathcal{Z}' = \mathbf{L}_W\mathcal{Z}$. If $\rho_0 = I/3$, then the "+" has coordinates $(0, 0)$, otherwise it will be shifted from the origin by $\mathbf{L}_W$.

Generic quadratic extremal witness

The witness W is drawn from a sample of extremal witnesses produced numerically using Algorithm 1 [71, 75]. The iteration procedure described in Section 8.7 is used to convert the map $\mathbf{L}_W$ into unital and trace preserving form. Figure 7.8 shows a section with the maximally mixed state at the origin, that is $\rho_0 = \rho_0' = I/3$. We have chosen the plane $\mathcal{Z}$ to also go through the pure states $\rho_1 = \phi_1\phi_1^\dagger$ and $\rho_2 = \phi_2\phi_2^\dagger$ corresponding to two arbitrarily chosen zeros $\phi_1 \otimes \chi_1$ and $\phi_2 \otimes \chi_2$ of the extremal witness W. In the section $\mathcal{Z} \cap \mathcal{D}_3$ enclosed by the red curve, the two pure states ρ_1 and ρ_2 which are joined in the plot by a straight line, are both mapped into rank two states on $\partial\mathcal{D}_3$. States on the line segment, originally rank two states, are mapped into rank three states, but they remain convex combinations of the two states ρ_1' and ρ_2'. The section $\mathcal{Z}' \cap \mathcal{D}_3$, enclosed by the blue curve, are all states of full rank, while the blue boundary are rank two states.

In Figure 7.9 we have also chosen $\rho_1 = \phi_1\phi_1^\dagger$, $\rho_2 = \phi_2\phi_2^\dagger$, but ρ_0 is now an even mix of the three pure states $\rho_1 = \phi_1\phi_1^\dagger$, $\rho_2 = \phi_2\phi_2^\dagger$, and $\rho_3 = \phi_3\phi_3^\dagger$. We also observe that the projection of the maximally mixed state $I/3$ is off center. The triangle is a two-dimensional face of an eight-dimensional simplex defined by all nine zeros of the

witness W. The three vertices ρ_1, ρ_2 and ρ_3 have been mapped into rank two states, and the original lines of rank two states have been mapped into the interior of $\mathcal{D}_3$, *i.e.* into rank three states.

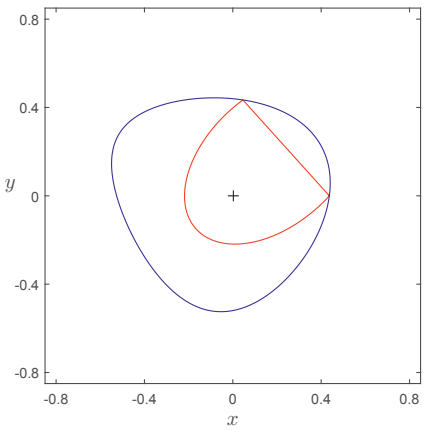


Figure 7.8: $\rho_1 = \phi_1\phi_1^\dagger$ and $\rho_2 = \phi_2\phi_2^\dagger$, and the origin ρ_0 is the maximally mixed state $I/3$. The origin is therefore unchanged under the map $\mathbf{L}_W$.

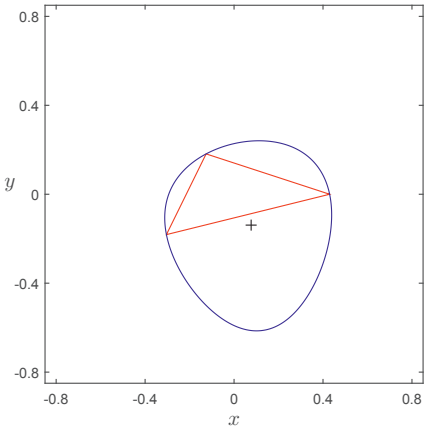


Figure 7.9: $\rho_1 = \phi_1\phi_1^\dagger$ and $\rho_2 = \phi_2\phi_2^\dagger$. The origin ρ_0 is an even mix of the pure states ρ_1, ρ_2, and $\rho_3 = \phi_3\phi_3^\dagger$, and is therefore shifted off center.

In Figure 7.10, the state $\rho_1' = \mathbf{L}_W\rho_1$ has rank two and lies on the boundary $\partial\mathcal{D}_3$, whereas $\rho_2' = \mathbf{L}_W\rho_2$ has full rank and lies in the interior of $\mathcal{D}_3$. In Figure 7.11 both states ρ_1 and ρ_2 are random states of rank three. Since this means that also ρ_1' and ρ_2' are states of rank three, both the blue and red curves portray generic boundaries of $\mathcal{D}_3$.

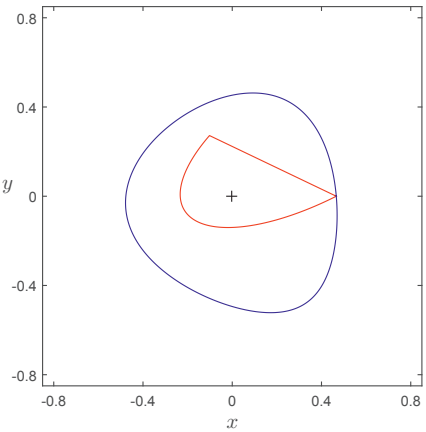


Figure 7.10: $\rho_1 = \phi_4\phi_4^\dagger$, ρ_2 is a random pure state and $\rho_0 = I/3$. $\rho_1' = \mathbf{L}_W\rho_1$ has rank two and lies on $\partial\mathcal{D}_3$, whereas $\rho_2' = \mathbf{L}_W\rho_2$ has full rank and lies in the interior of the set of density matrices $\mathcal{D}_3$.

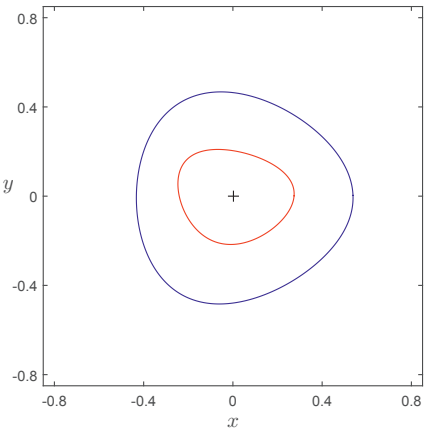


Figure 7.11: Both ρ_1 and ρ_2 are random states of rank three, and $\rho_0 = I/3$. Both curves represent generic boundaries of $\mathcal{D}_3$. We observe that the image of $\mathcal{Z} \cap \mathcal{D}_3$ is completely contracted inside $\mathcal{Z}' \cap \mathcal{D}_3$.

The Choi-Lam map

The unital and trace preserving map

$$\mathbf{L}_C : X \mapsto Y = \frac{1}{2}\begin{pmatrix} X_{11}+X_{33} & -X_{12} & -X_{13} \\ -X_{21} & X_{11}+X_{22} & -X_{23} \\ -X_{31} & -X_{32} & X_{22}+X_{33} \end{pmatrix} \tag{7.95}$$

was introduced by Choi and Lam in 1977 as the first example of an extremal non-decomposable positive map [63, 64]. It has been generalized to a one parameter family of extremal positive maps, but it is still one of the very few known analytical examples of such maps. As mentioned in Section 7.4.4 the corresponding witness W has three isolated zeros

$$e_{13} = e_1 \otimes e_3 \qquad e_{21} = e_2 \otimes e_1 \qquad e_{32} = e_3 \otimes e_2 \tag{7.96}$$

where e_1, e_2, e_3 are the natural basis vectors in $\mathbb{C}^3$, in addition to a continuum of zeros $\phi \otimes \chi$ where

$$\phi = e_1 + \mathrm{e}^{\mathrm{i}\alpha} e_2 + \mathrm{e}^{\mathrm{i}\beta} e_3 \qquad \chi = \phi^* \tag{7.97}$$

and $\alpha, \beta \in \mathbb{R}$. The states defined by the three isolated zeros of the Choi-Lam witness is then

$$\rho_1 = e_1 e_1^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \rho_2 = e_2 e_2^\dagger = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \rho_3 = e_3 e_3^\dagger = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{7.98}$$

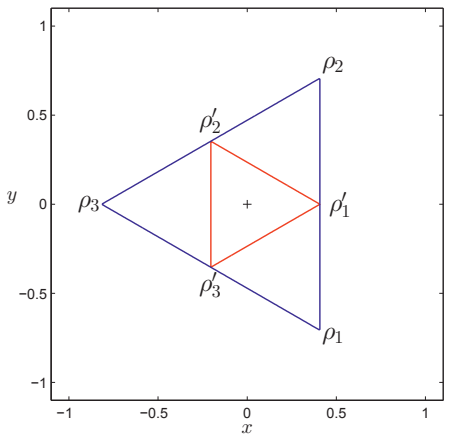


Figure 7.12: The map $\mathbf{L}_C$ in the plane of diagonal matrices. We use $\rho_0 = I/3$, $\rho_1 = e_1 e_1^\dagger$ and $\rho_2 = e_2 e_2^\dagger$. The image of $\mathcal{Z} \cap \mathcal{D}_3$, shown by the red curve, is rotated 60° and contracted by a factor one half.

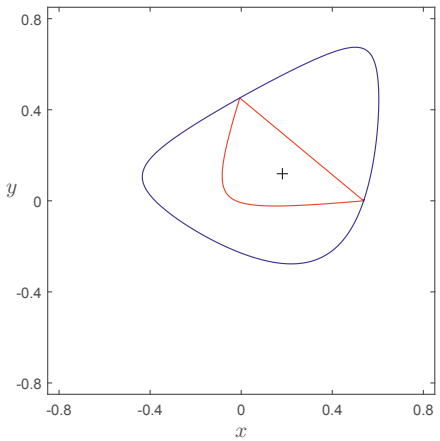


Figure 7.13: Here we see another section with $\rho_1 = e_1 e_1^\dagger$ and $\rho_2 = e_2 e_2^\dagger$. Here, the origin ρ_0 is however chosen as a random state of rank three, so the sector $\mathcal{Z} \cap \mathcal{D}_3$ contains neither $e_3 e_3^\dagger$ nor $I/3$.

If we take $\mathcal{Z}$ to be the plane containing the states ρ_1, ρ_2 and ρ_3, this gives a very special section through $\mathcal{D}_3$, namely a section containing the diagonal matrices. This section is mapped into itself by $\mathbf{L}_C$

$$\rho_1' = \frac{\rho_1 + \rho_2}{2} \qquad \rho_2' = \frac{\rho_2 + \rho_3}{2} \qquad \rho_3' = \frac{\rho_3 + \rho_1}{2} \tag{7.99}$$

We observe from (7.99) that the map $\mathbf{L}_C$ for the diagonal matrices represents a rotation of 60°, followed by a contraction by one half. This is seen in Figure 7.12. Since it is diagonal, the maximally mixed state is contained in the region $\mathcal{Z}$. In Figure 7.13 we see another section defined by $\rho_1 = e_1 e_1^\dagger$ and $\rho_2 = e_2 e_2^\dagger$. In this case however, the origin is chosen as a random state of rank three, so the sector $\mathcal{Z} \cap \mathcal{D}_3$ contains neither $e_3 e_3^\dagger$ nor $I/3$.

If we now try somehow to involve the quartic zeros (7.97), we can start by constructing the state

$$\rho(\alpha, \beta) = \frac{1}{3}\, \phi(\alpha, \beta)\phi^\dagger(\alpha, \beta) = \rho_0 + \sigma(\alpha, \beta) \tag{7.100}$$

These states define a two-dimensional curved surface of pure states in $\mathcal{S}$, which will be mapped by $\mathbf{L}_C$ to the boundary $\partial\mathcal{D}_3$. The matrix $\sigma(\alpha, \beta)$ is a completely off-diagonal matrix. We can see directly from (7.95) that with $\rho_0 = I/3$, the state $\rho(\alpha, \beta)$ is mapped into

$$\mathbf{L}_C \rho(\alpha, \beta) = \rho_0 - \frac{1}{2}\, \sigma(\alpha, \beta) = \frac{1}{2}\, [3\rho_0 - \rho(\alpha, \beta)] \tag{7.101}$$

If we choose *e.g.* the α-direction on this surface, we can define the derivative of $\phi(\alpha, \beta)$ in this direction by

$$\xi(\alpha) = \frac{\partial}{\partial \alpha} \phi(\alpha, \beta) = \mathrm{i} e^{\mathrm{i}\alpha} e_2 \tag{7.102}$$

Then the matrix

$$D(\alpha, \beta) = \frac{\partial}{\partial \alpha}\, \rho(\alpha, \beta) = \frac{1}{3}\, \left[\xi(\alpha)\phi^\dagger(\alpha, \beta) + \phi(\alpha, \beta)\xi^\dagger(\alpha)\right] \tag{7.103}$$

is a tangent to the surface such that

$$\rho(\alpha, \beta) + \epsilon D(\alpha, \beta) = \rho(\alpha + \epsilon, \beta) + \mathcal{O}(\epsilon^2) \tag{7.104}$$

Now we can choose a two-dimensional section $\mathcal{Z}$ defined by the maximally mixed state $\rho_0 = I/3$ and the two matrices

$$\begin{aligned} \rho_1 = \rho(0, 0) &= \frac{1}{3}\, (e_1 + e_2 + e_3)(e_1 + e_2 + e_3)^\dagger \\ \rho_2 = \rho_1 + D(0, 0) &= \rho_1 + \frac{\mathrm{i}}{3}\, \left[e_2(e_1 + e_3)^\dagger - (e_1 + e_3)e_2^\dagger\right] \end{aligned} \tag{7.105}$$

We know that $\rho_1 = \rho(0,0)$ is a state on the surface at $\alpha = \beta = 0$, and that $D(0,0)$ is a tangent to the surface at the same point. This means that in general $\rho_2 \notin \mathcal{D}$, which is not important since we are interested in the plane $\mathcal{Z}$ defined by ρ_0, ρ_1 and ρ_2, and the intersection $\mathcal{Z} \cap \mathcal{D}$ will be non-empty. This region is mapped in Figure 7.14 using the Choi-Lam map $\mathbf{L}_C$. We see that in the point $\rho(0,0)$, the tangent to the red curve is identical to the tangent of the blue curve. Also, in Figure 7.15 we use $\rho_0 = I/3$, but here we have chosen two states $\rho_1 = \phi_1\phi_1^\dagger$ and $\rho_2 = \phi_2\phi_2^\dagger$ corresponding to two random product vectors $\phi_1 \otimes \phi_1^*$ and $\phi_2 \otimes \phi_2^*$ from the continuum of quartic zeros.

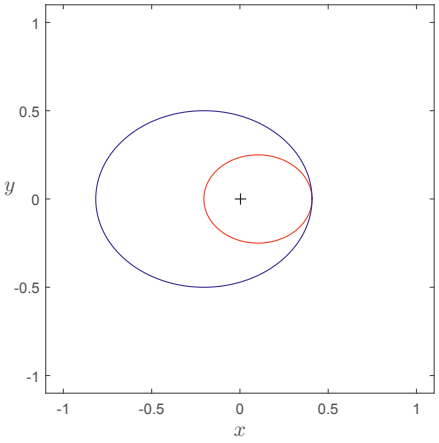


Figure 7.14: Here $\rho_0 = I/3$, $\rho_1 = \rho(0,0)$ and $\rho_2 = \rho_1 + D(0,0)$, as defined in (7.105). The Choi-Lam map $\mathbf{L}_C$ maps a curved surface into a curved surface, and the only state to be mapped onto the boundary is $\rho_1 = \rho(0,0)$.

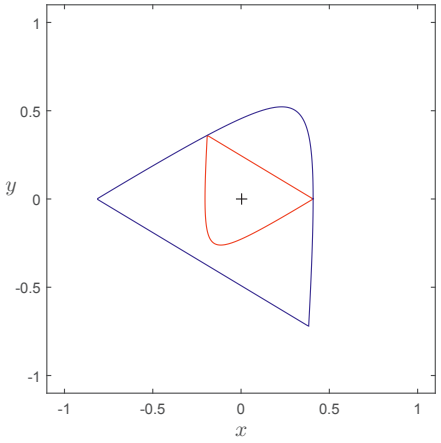


Figure 7.15: Here $\rho_0 = I/3$, and both $\rho_1 = \phi_1\phi_1^\dagger$ and $\rho_2 = \phi_2\phi_2^\dagger$ are two pure states corresponding to two randomly chosen product vectors $\phi_1 \otimes \phi_1^*$ and $\phi_2 \otimes \phi_2^*$ from the continuum of quartic zeros.

Chapter 8

Numerical methods

For many of the problems in which we have made numerical investigations, we have used various strategies to obtain solutions. Some of the essential background for the numerical methods are here summarized.

8.1 Real representations

There is a simple isomorphic relation between the set of $N \times N$ Hermitian matrices H_N, and the Euclidean real vector space $\mathbb{R}^{N^2}$. This simplifies our calculations considerably, and we have utilized this extensively throughout our numerical investigations. We can describe the N^2-dimensional set of Hermitian matrices H_N by the vector space $\mathbb{R}^{N^2}$, so that any Hermitian matrix $\rho \in H_N$ is written as a real $N^2 \times 1$ vector x, and any operator $\mathbf{U}$ on this space of Hermitian matrices is then represented by a real $N^2 \times N^2$ matrix $M \in \mathbb{R}^{N^4}$. So effectively the operation $\mathbf{U}\rho = U\rho U^\dagger = \rho'$ is replaced by $Mx = y$, where $y \in \mathbb{R}^{N^2}$ represents the state ρ'.

Likewise, any linear operator $\mathbf{L} : H_{N_A} \mapsto H_{N_B}$ that maps the state $\rho \in \mathcal{D}_{N_A}$ into the state $\tau \in \mathcal{D}_{N_B}$, can be represented by a real $N_B^2 \times N_A^2$ matrix L. The map then works by simple matrix multiplication with the vector $x \in \mathbb{R}^{N_A^2}$ corresponding to ρ. So we have $Lx = y$, where $y \in \mathbb{R}^{N_B^2}$ then represents $\tau \in \mathcal{D}_{N_B}$.

The matrix L corresponding to the map $\mathbf{L}$ can be calculated by transforming for example the fundamental matrix basis $E_i \in H_{N_A}$ with $1 \leq i \leq N_A^2$, with $\mathbf{L}$. With corresponding bases F_i for system B, we get

$$\mathbf{L}E_i = \sum_j L_{ji} F_j \tag{8.1}$$

With the basis vectors orthonormal we get for the real matrix L_{ji}

$$L_{ji} = \langle F_j, \mathbf{L}E_i \rangle \tag{8.2}$$

Then the transposed map $\mathbf{L}^T : H_{N_B} \mapsto H_{N_A}$ may be defined by the condition

$$\langle \mathbf{L}^T Y, X \rangle = \langle Y, \mathbf{L}X \rangle \qquad \forall\, X \in H_{N_A},\, Y \in H_{N_B} \tag{8.3}$$

The matrix representing the transposed map has matrix elements $L_{ij} = L^T_{ji}$, so transposition of the map $\mathbf{L}$ corresponds to transposition of the real matrix L_{ji}.

The above presents a practical and flexible way of conducting the action of the operator $\mathbf{L}$. This scheme is done in such a way that the inner product is preserved. For instance if $\mathbf{L}\rho = \tau$ and $Lx = y$ then $\langle \rho, \tau \rangle = \langle x, y \rangle$ where $\langle \rho, \tau \rangle = \mathrm{tr}(\rho\tau)$ is the Hilbert-Schmidt inner product and $\langle x, y \rangle$ is the standard Euclidean dot product.

8.2 General optimization

In a general optimization problem we try to formulate the problem in such a way that we are given a function $f : G \mapsto \mathbb{R}^+$ from some set G onto the positive real numbers, and seek an element x_0 in G such that $f(x_0) < f(x)$ for all other $x \in G$. Typically, and also in our cases the set G is represented as a subset of $\mathbb{R}^N$, often further specified by a set of constraints, equalities or inequalities on elements of G. The domain G of f is the search space and the elements of G are candidate solutions.

For most of our optimization problems we have used two alternative random search algorithms. The first is a completely local random method and the other is a more complex global optimization scheme, based on simulated annealing, which is an adaption of the Metropolis algorithm.

Random search

The local random method is rather simple, but surprisingly effective for a lot of problems, especially if the search space is convex. From a point x it chooses a random, but within the constraints allowed, direction A in the search space, and checks the point $x' = x + \epsilon A$, where ϵ is a step length, to see whether $f(x') < f(x)$. If this is the case, it chooses the new point x' as the starting point for the next iteration. If this is not the case, it will try a given (but large) number of times to choose another random, but within the constraints allowed, direction to achieve $f(x') < f(x)$. This procedure continues through a chain of iterations $x_0, x_1, \ldots x_n$, until the process stops because no further direction such that $f(x') < f(x)$ can be found within the numerical accuracy. As the value of $f(x_i)$ drops, the step length must be adjusted accordingly. The method is rather fast, since no gradients are calculated at any point.

Simulated annealing

For some cases the above method does not work very efficiently. This is especially the case when the function f has many local minima. Instead we have implemented a simulated annealing routine. The simulated annealing method is a Metropolis algorithm, and this means that it has a certain probability to take a step in a direction where the objective function f increases, so it is less likely to converge at a local minimum. In addition to the function f and state variable $x \in \mathbb{R}^N$, it requires a control parameter T. An annealing procedure that controls the reduction of T, and some schedule for how to make the random changes in the configuration, are also made.

We have used an approach suggested in [86]. We replace the state variable x by a simplex of $N+1$ points. The possible changes to the configuration are reflections, expansions and contractions of the simplex. We add a positive random variable, proportional to T, to the function value associated with each vertex, and subtract a similar random variable from the function value of every proposed replacement point. Thus the procedure will always accept a downhill move, but may sometimes accept uphill moves. At a temperature T the simplex will expand to cover the region reachable at that temperature. It then randomly samples points inside the region. As the temperature decreases, so will the simplex, which is likely to contain the lowest minimum encountered so far.

8.3 Finding product vectors in a subspace

Product vectors are completely essential to the study of entanglement, both in connection with PPT states and entanglement witnesses. The problem of finding product vectors in any generic subspace is in general a difficult one, so to be able to do this we resort to numerical procedures. We describe here two different approaches which both rely on solving an optimization problem.

Finding product vectors as a minimum double eigenvalue problem

One effective way of obtaining the product vectors in a given subspace $\mathcal{U}$ of $\mathbb{C}^N$, is to construct a positive matrix $M \in H_N$ that projects onto the orthogonal complement $\mathcal{V} = \mathcal{U}^\perp$, *i.e.* $\operatorname{Img} M = \mathcal{V}$. The matrix M could for instance, but not necessarily, be a genuine projection operator onto $\mathcal{V}$. We then find the product vectors $\psi = \phi \otimes \chi \in \mathbb{C}^N$ which have zero expectation value on M. They will then necessarily lie in $\mathcal{U}$. So, for a positive matrix $M \in H_N$ we want to have $\psi = \phi \otimes \chi \in \mathbb{C}^N$ such that

$$\psi^\dagger M \psi = 0 \tag{8.4}$$

and we want ψ, and therefore ϕ and χ, normalized to one. We introduce the Lagrange multipliers λ and μ, and define the function $f = f(\phi, \chi)$ as

$$f = \sum_{ij;kl} \phi_i^* \chi_j^* M_{ij;kl} \phi_k \chi_l - \lambda \left(\sum_k \phi_k^* \phi_k - 1 \right) - \mu \left(\sum_k \chi_k^* \chi_k - 1 \right) \tag{8.5}$$

We now want to find the local minima of the first term in f under the two constraints given by the Lagrange multipliers. Taking the derivative with respect to the components of ϕ and χ, and setting these equal to zero will ensure that we are dealing with a minimum. Since $\phi_i, \chi_j \in \mathbb{C}$ we would normally have to take the derivative of both ϕ_k and ϕ_k^*, but since f is a real valued function we have $\partial f / \partial \phi_k = (\partial f / \partial \phi_k^*)^*$, so when one of them is zero, so is the other. The two equations that must be satisfied in order to optimize f are then

$$\begin{aligned}\frac{\partial f}{\partial \phi_i^*} &= \sum_{jkl} \chi_j^* M_{ij;kl} \chi_l \phi_k - \lambda \phi_i = 0 \\ \frac{\partial f}{\partial \chi_j^*} &= \sum_{ikl} \phi_i^* M_{ij;kl} \phi_k \chi_l - \mu \chi_j = 0\end{aligned} \tag{8.6}$$

From the first equation in (8.6) we see that if we multiply from the left by ϕ_i^* and sum over i, we get

$$\sum_i \phi_i^* \frac{\partial f}{\partial \phi_i^*} = \sum_{ij;kl} \phi_i^* \chi_j^* M_{ij;kl} \phi_k \chi_l - \lambda \sum_i \phi_i^* \phi_i = \psi^\dagger M \psi - \lambda \phi^\dagger \phi \tag{8.7}$$

When we assume that ϕ and χ are normalized to one, and perform similar operations on the second equation in (8.6) with χ_j^*, we find that the values of λ and μ at an extremum are

$$\lambda = \psi^\dagger M \psi = \mu \tag{8.8}$$

We can introduce the matrices

$$B_{ik} = \sum_{jl} \chi_j^* M_{ij;kl} \chi_l \qquad C_{jl} = \sum_{ik} \phi_i^* M_{ij;kl} \phi_k \tag{8.9}$$

and write (8.6) simpler as

$$\begin{aligned}B\phi - \lambda \phi &= 0 \\ C\chi - \mu \chi &= 0\end{aligned} \tag{8.10}$$

The gradients of ϕ and χ are given by

$$\begin{aligned}g_\phi &\equiv B\phi - \lambda \phi \\ g_\chi &\equiv C\chi - \lambda \chi\end{aligned} \tag{8.11}$$

with λ given by (8.8) and using the starting vector for $\psi = \phi \otimes \chi$. Then we define new vectors ϕ' and χ' as

$$\begin{aligned}\phi' &= \phi + \epsilon g_\phi \\ \chi' &= \chi + \epsilon g_\chi\end{aligned} \tag{8.12}$$

To find the optimal step length ϵ we observe that to first order in ϵ we have

$$\psi' = \phi' \otimes \chi' \approx \psi + \epsilon w \equiv s(\epsilon) \tag{8.13}$$

with $w = \phi \otimes g_\chi + g_\phi \otimes \chi$. The step length ϵ with which we move along the gradients can be found by minimizing

$$h(\epsilon) = \frac{s^\dagger M s}{s^\dagger s} \tag{8.14}$$

Taking the derivative of h with respect to ϵ and demanding that it is zero, we get a second order equation $\epsilon^2 + b\epsilon + c = 0$, where the factors b and c depend on ψ, w and M. Since this equation may have two solutions we calculate both, and choose the one which is the minimum. The product vector is then updated according to (8.12), and the procedure is repeated (iterated) until a solution is found, *i.e.* ψ satisfies (8.4).

If the number of product vectors present in the subspace is finite, the procedure will generally find the same product vectors repeatedly, so the procedure must usually be done a large number of times with a sufficient spread in the starting points, in order to find the whole set of product vectors.

We can now take M to be the projection operator P onto the range of the density matrix ρ. The procedure will then find the product vectors in $\mathrm{Ker}\,\rho$. Equivalently $M = Q$, where Q is a projection operator onto the kernel, will return the product vectors in $\mathrm{Img}\,\rho$. We can use the same scheme to find product vectors in the range and kernel of ρ^P using the projection operators $\widetilde{P}$ and $\widetilde{Q}$, as defined in (5.21).

Finding product vectors by general optimization

It is also possible to search for product vectors in a given subspace by more general optimization methods. Empirically we experience that using the minimum double eigenvalue method outlined above, works quite well for most cases, but we have seen some cases where it is unable to find all the product vectors, and where more conventional optimization methods succeeds better.

We let the vectors $z_j \in \mathbb{C}^N$ with $j = 1, \ldots, m$ define an m-dimensional subspace $\mathcal{U}$ of $\mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$. If the vectors z_j is an orthogonal set we can construct a projection operator onto the space they span, as a matrix P which is idempotent and has m eigenvalues all equal to one. We can also likewise construct a projection operator Q onto the $(N-m)$-dimensional orthogonal complement $\mathcal{V} = \mathcal{U}^\perp$ in $\mathbb{C}^N$. Usually, but not necessarily, these are taken to be the range and kernel respectively, of some density matrix ρ. If z_j is not an orthogonal set then it can be made so by the Gram-Schmidt process. If we want to find product vectors in the subspace defined by P, the function f which must be minimized can be defined in two ways, and this gives rise to two optimization strategies.

- For the first case we choose $G = \mathbb{R}^{2(N_A+N_B)}$ and create a random starting point $x \in G$. From x we then construct two real $N_A \times 1$ vectors x_{A_1} and x_{A_2}, and two real $N_B \times 1$ vectors x_{B_1} and x_{B_2}. The vector

$$w = (x_{A_1} + \mathrm{i}x_{A_2}) \otimes (x_{B_1} + \mathrm{i}x_{B_2}) \tag{8.15}$$

  will then be a product vector in $\mathbb{C}^N$, but we will in general not have $wQw^\dagger = 0$, which it must be in order for $w \in \mathcal{U}$. So the function f used for the optimization is $f(x) = wQw^\dagger$ with w of the form (8.15) and Q as a fixed projection operator. So we seek to keep w on product form while at the same time minimize the projection onto $\mathcal{U}^\perp$.

- For the second strategy we choose $G = \mathbb{R}^{2N}$ and we create a random starting point $x \in G$ which then defines

$$w = w(x) = \begin{pmatrix} x_1 \\ \vdots \\ x_N \end{pmatrix} + \mathrm{i} \begin{pmatrix} x_{N+1} \\ \vdots \\ x_{2N} \end{pmatrix} \tag{8.16}$$

  which in general will not be a product vector. The first step in the next phase is to project the vector w onto the subspace $\mathcal{U}$ by using the projection operator P. The projected vector $w' = Pw$ will then be in $\mathcal{U}$, but it will still in general not be a product vector. The expression (3.52) gives a set of equations that the components of w_i must satisfy in order for w to be a product vector. So in this case we seek to satisfy this set of equations, whilst keeping $w \in \mathcal{U}$. So the function f becomes

$$f(x) = \sum_{ij;kl} |w_i w_j - w_k w_l| \tag{8.17}$$

  where the indices i, j, k, l are summed over the appropriate combinations. For the 3×3 system the set of equations that defines f is given by (6.2).

8.4 SL-equivalence of density matrices

Two generic states ρ_1 and ρ_2 will not, according to Definition 3.7, be SL-equivalent. To establish whether two states are SL-equivalent, is in general a very difficult problem.

PPT states of rank $(5, 5)$ in the 3×3 system

We have in Chapter 6 seen that any generic five-dimensional subspace of $\mathbb{C}^9$, which always contains six product vectors, can be transformed to the standard form (6.11). This means that all such subspaces with identical invariants p, q, r, s are SL-*equivalent subspaces*. For two density matrices ρ_1 and ρ_2 to be SL-equivalent according to Definition 3.7, their ranges $\mathrm{Img}\,\rho_1$ and $\mathrm{Img}\,\rho_2$ must be SL-equivalent. However, this condition is necessary but not sufficient, as density matrices with the same range are generally not SL-equivalent. So this scheme can only be used in a contrapositive way.

Note that though only $u_1, \ldots, u_5$ and $v_1, \ldots, v_5$ occur in (6.12), there will for generic subspaces always be a sixth product vector. Since the numbering is completely coincidental, all six product vectors must be taken into consideration when calculating the invariants. Different permutations of the six product vectors will in general give different s_i, so the invariants corresponding to all $6! = 720$ permutations must be calculated. In order to check if two states ρ_1 and ρ_2 can be SL-equivalent we must therefore find the product vectors in $\mathrm{Img}\,\rho_1$ and $\mathrm{Img}\,\rho_2$, then transform these to the standard form (6.11) and calculate the invariants s_i for all 720 permutations of the six product vectors. If the invariants s_i so calculated are identical for any permutation, then $\mathrm{Img}\,\rho_1$ and $\mathrm{Img}\,\rho_2$ can clearly be transformed by SL $\times$ SL-transformations to the form (6.11), with identical invariants.

General method

An option which might be used to establish whether two states are SL-equivalent in a positive way, is to actually try and find the transformations. Given two density matrices ρ_1 and ρ_2 which remain fixed, we can try to find a transformation $V = V_A \otimes V_B$ such that

$$\rho_1 = V\rho_2 V^\dagger \tag{8.18}$$

A linear transformation in $\mathrm{GL}(N_A, \mathbb{C})$ has $2N_A^2$ free parameters, a product transformation in $\mathrm{GL}(N_A, \mathbb{C}) \times \mathrm{GL}(N_B, \mathbb{C})$ will then have $2(N_A^2 + N_B^2)$ parameters. So for this optimization we use $G = \mathbb{R}^{2(N_A^2+N_B^2)}$ as search space. A total of $2N_A^2$ and $2N_B^2$ real parameters then define V_A and V_B respectively, and we calculate $V = V_A \otimes V_B$ for each step of the optimization scheme to ensure that V is a product transformation. The function f to be minimized is then

$$f(x) = \mathrm{tr}\left[(\rho_1 - V\rho_2 V^\dagger)^2\right] \tag{8.19}$$

which is the Hilbert-Schmidt norm of the distance between ρ_1 and ρ_2, apart from a square root which we avoid in order to ensure maximum numerical accuracy. We have utilized this method to check for product transformations between various states, with relatively good results. Since this optimization procedure gives the product transformations explicitly, presuming they exist, it is a good method for determining whether a state is genuinely SL-symmetric or not, because the transformations then must have to be trace preserving and have the special form (3.75), with $V_B^* = V_B^{-1}$.

8.5 Obtaining PPT states with specified rank

Producing PPT states with specified rank (m, n) for both ρ and ρ^P is an essential part of our work on PPT states. The search for such states can be done quite generally, that is without any constraints, or one may for instance fix the range of the states. Another possibility is to search specifically for PPT states that are SL-symmetric.

General method

There are different ways of doing this, but the most obvious ways are based on minimizing a specified number of eigenvalues for ρ and ρ^P, or more precisely obtaining values for these eigenvalues equal to zero. A normalized density matrix ρ on $\mathbb{C}^N$ can be represented by a vector $x \in \mathbb{R}^{N^2-1} = G$. The eigenvalues of ρ are then functions of this vector, *i.e.* $\lambda_k = \lambda_k(x)$. Since the partial transpose ρ^P is uniquely defined by ρ, its eigenvalues $\widetilde{\lambda}_k = \widetilde{\lambda}_k(x)$ will also depend on the same vector x. Assume that we want to construct a rank (m, n) state. Then $m_0 = N - m$ of the eigenvalues λ_k of ρ and $n_0 = N - n$ of the eigenvalues $\widetilde{\lambda}_k$ of ρ^P must be equal to zero. We may introduce an ordering where the index k sorts the eigenvalues λ_k and $\widetilde{\lambda}_k$ according to increasing values, $\lambda_1 < \ldots < \lambda_N$ and $\widetilde{\lambda}_1 < \ldots < \widetilde{\lambda}_N$. The fact that $\rho(x) \in H_N$ assures that the eigenvalues λ_k and $\widetilde{\lambda}_k$ are real, though some of them may be negative.

We then want the first $m_0 - 1$ eigenvalues of ρ and the first $n_0 - 1$ eigenvalues of ρ^P to be zero

$$\begin{aligned} \lambda_k &= 0 \qquad k = 1, \ldots, m_0 \\ \widetilde{\lambda}_k &= 0 \qquad k = 1, \ldots, n_0 \end{aligned} \tag{8.20}$$

We are also free to put conditions on the non-zero eigenvalues of ρ and ρ^P. We can for instance specify their values

$$\begin{aligned} \lambda_k &= \eta_k \qquad k = m_0 + 1, \ldots, N \\ \widetilde{\lambda}_k &= \widetilde{\eta}_k \qquad k = n_0 + 1, \ldots, N \end{aligned} \tag{8.21}$$

If we in particular chose $\eta_k = 1/m$ for all the non-zero eigenvalues of ρ, we obtain a density matrix proportional to a rank m projection operator. The minimization function f is now

$$f(x) = \sum_{k=1}^{m_0} |\lambda_k|^2 + \sum_{k=1}^{n_0} |\widetilde{\lambda}_k|^2 \tag{8.22}$$

Note that in theory we could end up with a density matrix with smaller ranks than intended, because we usually do not specifically demand that the eigenvalues λ_k and $\widetilde{\lambda}_k$ of (8.21) should be non-zero. In fact in many searches where there are constraints which cannot be met, or are very difficult to satisfy, there is a tendency that many of the eigenvalues λ_k and $\widetilde{\lambda}_k$ of (8.21) approach zero. Alternative methods for obtaining PPT states with specified ranks, that are based on calculation of gradients, have been constructed [55]. We have used the techniques described above to produce a large number of PPT states of various ranks.

PPT states with specified range

An m-dimensional subspace of $\mathbb{C}^N$ can be described by m linearly independent vectors $z_i \in \mathbb{C}^N$ with components expressed in the $N \times m$ matrix Z. The dimension of the set of complex rank m matrices is $2m^2$, so in order to find rank (m, n) PPT states we choose to use the search space $G = \mathbb{R}^{2m^2}$. We then construct a complex $m \times m$ matrix X out of the search vector $x \in \mathbb{R}^{2m^2}$, and define the $N \times m$ matrix $M = ZX$ which then contains general linear combinations of the five vectors z_j, with coefficients defined by X. The matrix

$$\rho = MM^\dagger = Z(XX^\dagger)Z^\dagger \tag{8.23}$$

will then be a positive and Hermitian matrix with a range spanned by z_j. It will in general not be normalized so we must ensure that $\mathrm{tr}(\rho) = 1$ for each iteration. Thus the constraints we need on ρ are always ensured. In general ρ^P will not be positive, so we seek to minimize the $n_0 = N - n$ smallest eigenvalues of ρ^P to zero

$$f(x) = \sum_{k=1}^{n_0} |\widetilde{\lambda}_k|^2 \tag{8.24}$$

SL-symmetric PPT states

PPT states with the property that they are SL-symmetric are in general difficult to obtain. A way of finding such states by random optimization is to combine the search spaces $G_1 = \mathbb{R}^{N_A^2 N_B^2 - 1}$ and $G_2 = \mathbb{R}^{2(N_A^2+N_B^2)}$. We then seek to minimize the function (8.19), with $\rho_1 = \rho$ and $\rho_2 = \rho^P$ governed by G_1, and V governed by G_2. Within the same procedure we must also minimize (8.22), in order to ensure that ρ is a PPT state of rank (m, n). A solution $x = x_0 \in \mathbb{R}^{N_A^2 N_B^2 + 2(N_A^2+N_B^2)-1}$ will then define a rank (m, n) PPT state ρ and an SL $\times$ SL-transformation that transforms ρ into ρ^P.

8.6 Finding extremal entanglement witnesses

The discussion in Section 7.4.2, along with Theorems 7.4 and 7.5, motivate a search algorithm for finding extremal entanglement witnesses. Starting from an initial witness $W = W_0 \in \mathrm{int}(\mathcal{S}^\circ)$, which might be the maximally mixed state I/N, we proceed down through a hierarchy of faces on $\mathcal{S}^\circ$ of decreasing dimensions. The search is in principle guaranteed to converge to an extremal witness in a finite number of iterations.

Algorithm 1: Finding an extremal entanglement witness

Precondition: Choose an initial witness $W = W_0 \in \mathrm{int}(\mathcal{S}^\circ)$ and construct $\mathbf{U}_W$

1. **while** $\dim \mathrm{Ker}(\mathbf{U}_W) > 1$
2. choose a $\Gamma \in \mathrm{Ker}(\mathbf{U}_W)$
3. **if** $\mathrm{tr}(\Gamma) \neq 0$
4. redefine $\Gamma \leftarrow \Gamma - \mathrm{tr}(\Gamma)W$
5. **endif**
6. find t_c as the maximal t such that $W + t\Gamma \in \mathcal{S}^\circ$
7. redefine $W \leftarrow W + t_c\Gamma$
8. locate all new zeros of W and construct an updated $\mathbf{U}_W$
9. **endwhile**

return W

The number of possible search directions Γ is reduced in each iteration as $\mathrm{Ker}(\mathbf{U}_W)$ is reduced. The main work in the algorithm is point number six, *i.e.* to find the boundary of the face defined by $\mathbf{U}_W$. Given the chosen direction Γ on the face, this is essentially finding the point t_c which minimizes $(\phi \otimes \chi)^\dagger W(t)(\phi \otimes \chi)$ for $W(t) = W_0 + t_c\Gamma$. This expectation value should turn to zero at the boundary of the face, in order to become negative once $W(t) \notin \mathcal{S}^\circ$.

8.7 Unital and trace preserving positive maps

In Section 7.7.1 we assumed that in order to transform a positive map $\mathbf{L}_W$ to a unital and trace preserving form $\widetilde{\mathbf{L}}_W$, we can transform the corresponding witness such that

$$W \mapsto \widetilde{W} = (U \otimes V)W(U \otimes V)^\dagger \tag{8.25}$$

Further, we derived a set of equations which the two operators U and V must satisfy

$$\mathbf{L}_W(U^\dagger U) = (V^\dagger V)^{-1} \qquad \mathbf{L}_W^T(V^\dagger V) = (U^\dagger U)^{-1} \tag{8.26}$$

In order to solve (8.26) we first find positive Hermitian matrices $X = U^\dagger U$ and $Y = V^\dagger V$ solving the equations

$$\mathbf{L}_W X = Y^{-1} \qquad \mathbf{L}_W^T Y = X^{-1} \tag{8.27}$$

Then we solve the equations

$$U^\dagger U = X \qquad V^\dagger V = Y \tag{8.28}$$

for U and V, and thus the transformation matrix $U \otimes V$. The general solutions will be

$$U = U_1 U_2 \qquad V = V_1 V_2 \tag{8.29}$$

where U_1 and V_1 are arbitrary unitary matrices, and $U_2 = \sqrt{X}$, $V_2 = \sqrt{Y}$ are the uniquely defined positive Hermitian square roots. Equation (8.27) makes sense because the matrices $X = U^\dagger U$ and $Y = V^\dagger V$ are strictly positive as long as U and V are non-singular, and the maps $\mathbf{L}_W$ and $\mathbf{L}_W^T$ as well as the inversions $X \mapsto X^{-1}$ and $Y \mapsto Y^{-1}$, transform strictly positive matrices into strictly positive matrices.

The method suggesting itself for solving Equation (8.27) is to iterate the equations. Given an approximate solution X_k for X we try to compute a better approximation X_{k+1} by a series of four transformations,

$$X_k \mapsto S_k = \mathbf{L}_W X_k \mapsto Y_k = S_k^{-1} \mapsto T_k = \mathbf{L}_W^T Y_k \mapsto X_{k+1} = T_k^{-1} \tag{8.30}$$

We start the iterations for example with $X_0 = I_A$, where I_A is the identity matrix in system A. The first two steps then calculate Y_0, and eventually X_1 as the first improved solution of X. A sufficient condition for the convergence of X_k and Y_k to unique limits X and Y is that the transformation $X_k \mapsto X_{k+1}$ is contractive. A small perturbation ΔX_k of X_k transforms linearly as

$$\Delta X_{k+1} = \mathbf{D}(\Delta X_k) \tag{8.31}$$

where the linear map $\mathbf{D}$ is the derivative of the non-linear transformation $X_k \mapsto X_{k+1}$. The transformation is contractive if all eigenvalues of $\mathbf{D}$ are smaller than one in absolute value. Now $\mathbf{D}$ is a composition of four linear maps,

$$\mathbf{D} = \mathbf{D}_4 \mathbf{L}_W^T \mathbf{D}_2 \mathbf{L}_W \tag{8.32}$$

where $\mathbf{D}_2$ and $\mathbf{D}_4$ are linearizations of the matrix inversions

$$\begin{aligned} \mathbf{D}_2(\Delta S_k) &= -S_k^{-1}(\Delta S_k)S_k^{-1} \\ \mathbf{D}_4(\Delta T_k) &= -T_k^{-1}(\Delta T_k)T_k^{-1} \end{aligned} \tag{8.33}$$

The examples of extremal positive maps $\mathbf{L}_W$ and $\mathbf{L}_W^T$ that we have studied numerically are strongly contractive, and we find in practice that this is enough to ensure that $\mathbf{D}$

is contractive, with eigenvalues typically no larger than about 0.5, even though $\mathbf{D}_2$ and $\mathbf{D}_4$ are not contractive.

We have used this iteration scheme on a large number of numerically produced extremal entanglement witnesses [71], and also on many non-extremal witnesses constructed as convex combinations of the extremal ones. Numerically our attempts, which are in the thousands, always converge, and it also appears that for a given witness W the solution X is unique, independent of the initial guess X_0.

Summary and outlook

Presented here is a summary of the four papers that make up the main body of this thesis. The summary consists of some background material, but the main emphasis is on the results, and mathematical details is avoided as much as possible. Also, some of the loose threads and still unresolved issues, are discussed.

Paper I

In this paper the focus is on the entangled PPT states in 3×3 dimensions, and mainly those of rank four and five.

We first review the UPB construction, based on orthogonal product vectors in the kernel, before we describe a different approach, in which the product vectors need not be orthogonal. This approach also throws some new light on a set of conditions that limit the selection of product vectors to be used for constructing entangled PPT states of rank $(4, 4)$.

The study of the entangled PPT states of rank $(5, 5)$ is then a natural progression, and the knowledge gained from the rank $(4, 4)$ states can be used as a starting point. We show how to construct entangled PPT states of rank $(5, 5)$ from the rank $(4, 4)$ states. We accomplish this by use of perturbation theory. In this way, we obtain rank $(5, 5)$ PPT states infinitesimally close to the boundary of the surface of such states. This construction also explains why states of rank $(5, 4)$ cannot be found in the generic case.

We discuss how to determine numerically the dimensions of surfaces of PPT states of fixed rank and range. For the surface of PPT states with rank $(5, 5)$, we find that the dimensions are consistent with a counting of the independent constraints involved in the above perturbation scheme. Though we find an eight-dimensional surface of rank $(5, 5)$ PPT states in every generic five-dimensional subspace, we have not found any general construction scheme for such states.

In order to construct generic PPT states of rank $(5, 5)$ which are not infinitesimally close to the rank $(4, 4)$ boundary states, we present a method for doing numerical integration of equations of motion for curves on the surface of rank $(5, 5)$ PPT states. In this way one may study for example the curvature of this surface, or how a curve on the surface approaches the boundary.

An interesting result is that for a generic PPT state defined by the UPB construction, the number of product vectors in the kernel needed to specify the state, is in general smaller than the dimension of the kernel. For instance, generic PPT states of

rank $(6,6)$ in the 4×4 system have 20 different product vectors in their ten-dimensional kernel, but the number of product vectors needed to specify the state is seven. This does not however apply to PPT states with no product vectors in the kernel, like the generic rank $(5,5)$ PPT states. This result raises new interesting questions to be answered by future research, for example, how to identify finite sets of product vectors that define PPT states with these product vectors in their kernel.

As mentioned, it is still an unresolved problem how to construct general rank five PPT states that are not close to the rank four states. We are even farther from a full understanding of higher rank PPT states in 3×3 dimensions.

The extremal PPT states of rank four are all generic, in the sense that there exists a standard form for these states, and their form is derived relatively straight from the UPB construction. For the extremal PPT states of rank five however, there also in addition to the generic cases, exist a whole plethora of non-generic types. To some extent our understanding of many of these non-generic forms is better than for the generic type. This is the main theme of Paper 4.

Paper II

In this paper we mainly study extremal entanglement witnesses in bipartite systems of dimensions $N = N_A N_B$. We define the cone of entanglement witnesses as the dual of the cone of the unnormalized separable density matrices, this means that $\mathrm{tr}(W\rho) = (\phi\otimes\chi)^\dagger W(\phi\otimes\chi) \geq 0$ whenever W is a witness and ρ is a pure product state $\rho = \psi\psi^\dagger$ with $\psi = \phi\otimes\chi$. A very essential feature is that the expectation value $\mathrm{tr}(W\rho)$ as a function of the vectors $\phi \in \mathbb{C}^{N_A}$ and $\chi \in \mathbb{C}^{N_B}$ is a positive semidefinite biquadratic form $f(\phi,\chi)$. Every zero of a positive biquadratic form, *i.e.* (ϕ_0,χ_0) for which $f(\phi_0,\chi_0) = 0$, imposes strong real and linear constraints on the form. In every direction at the zero the first derivative must vanish and the second derivative must be non-negative. If the second derivative vanishes, the third derivative must vanish and the fourth derivative must be non-negative. The real and symmetric Hessian matrix (which contains the second derivatives) at the zero, must be positive semidefinite. The eigenvectors of the Hessian with zero eigenvalue, if such exist, we call Hessian zeros. A zero of the biquadratic form $f(\phi,\chi)$ is called quadratic if it has no Hessian zeros, otherwise it is called quartic. We call a witness quadratic if it has only quadratic zeros, and quartic if it has at least one quartic zero. An important result we prove is that a witness is extremal if and only if no other witness has the same set, or a larger set, of zeros and Hessian zeros.

We use the complete set of constraints above, defined by the zeros and Hessian zeros of a witness, in order to test for extremality. If the witness is not extremal, the test gives all the directions in the space of Hermitian matrices in which we may search for witnesses that have more zeros or Hessian zeros, and hence are more nearly extremal. We have exploited this fact to construct an algorithm to find extremal witnesses. A finite number of iterated searches in random directions will then lead to an extremal witness. We find that the distinction between quadratic and quartic zeros is completely essential when we classify extremal witnesses. Nearly all extremal witnesses found in our

random searches by using Algorithm 1 from Section 8.6, are quadratic. We derive the minimum number of zeros a quadratic witness must have in order to be extremal, and nearly all of the quadratic extremal witnesses have this minimum number of zeros. To our knowledge, extremal witnesses of generic type have never been constructed before, even though they are by their very nature, far the most common. We have also used Algorithm 1 to find some extremal witnesses with more than the minimum number of zeros, and we regard these as non-generic.

The facial structure of the convex set of separable states and the convex set of witnesses are closely related, and this is clearly exposed in the algorithm we have constructed. The zeros of a witness, whether quadratic or quartic, always define an exposed face of the set of separable states. The other way around, an exposed face of the set of separable states is defined by a set of pure product states that are the zeros of some witness, in fact, they are the common zeros of all the witnesses in a face of the set of witnesses. The existence of witnesses with quartic zeros is the root of the existence of non-exposed faces of the set of witnesses. It is unknown whether the set of separable states has unexposed faces.

Related to the structure of faces on $\mathcal{S}$ are many important unsolved problems. We know that any state $\rho \in \mathcal{D}$ can be decomposed into a convex sum of maximally $r = \mathrm{rank}(\rho)$ states of rank one. This is a considerable progress from the maximum given by Carathéodorys theorem which is N^2, with $N = N_A N_B$ equal to the dimension of the Hilbert space. This fact is entirely due to the facial structure of $\mathcal{D}$. The maximum number of pure product states needed to write a separable state in the same manner is unknown, and this is due to the unknown structure of the faces of $\mathcal{S}$. Because the set of separable states $\mathcal{S}$ and the set of entanglement witnesses $\mathcal{S}^\circ$ are dual sets, the faces of these convex sets are related, and so a way of understanding the faces of $\mathcal{S}$ is to study the faces of $\mathcal{S}^\circ$. Another unsolved problem is the following: We have found a procedure for constructing an extremal witness from its zeros and Hessian zeros, but an arbitrarily chosen set of zeros and Hessian zeros does not in general define an extremal entanglement witness, *i.e.* they do not define an exposed face on $\mathcal{S}$. This is seen by studying the simplex faces that zeros of entanglement witnesses define, compared to the simplexes defined by random sets of zeros. The key question is how to choose a set of zeros and Hessian zeros, such that they are zeros of an entanglement witness.

We discuss some properties of optimal witnesses, and try to relate these to extremal witnesses. The minimum number of zeros of a quadratic extremal witness increases faster than the Hilbert space dimension. Since a witness is optimal if its zeros span the Hilbert space $\mathbb{C}^N$, this implies that in all but the lowest dimensions witnesses may be optimal and yet be very far from extremal. We also briefly discuss the so-called SPA separability conjecture in lieu of the results of our numerical investigations.

By studying the extremal entanglement witnesses we have made some progress towards understanding the convex sets of witnesses and separable mixed states. But this has also made the complexity of the problem even more clear than it was before. The main complication is the fact that zeros of extremal witnesses may be quartic, and since this opens up an almost unlimited range of variability among the quartic witnesses, it is a difficult problem. Nevertheless we believe that it is useful to pursue the study of

extremal witnesses in order to learn more about the geometry of the set of separable states, and it is rather clear that a combination of analytical and numerical methods will be needed also in future work.

Paper III

In this paper we are to some degree more occupied with positive maps rather than entanglement witnesses. There is a one-to-one correspondence between positive maps and entanglement witnesses, often called the Choi-Jamiolkowski isomorphism, which identifies a unique positive map $\mathbf{L}_W$ to each entanglement witness W.

A very useful property of maps is unitality, which means that a map always carries the identity into the identity, *i.e.* $\mathbf{L}I = I$. This is useful since it provides a fixed point for the map. A map may also be trace preserving, which means that for any state ρ we have $\mathrm{tr}(\rho) = \mathrm{tr}(\mathbf{L}\rho)$. Since a mixed state ρ intrinsically carries a probability distribution it is important that the mapped state $\mathbf{L}\rho$ also does. We first argue that every positive map $\mathbf{L}_W$ may be transformed into a form which is both unital and trace preserving through a product transformation of the corresponding entanglement witness W. If the witness lies in the interior of $\mathcal{S}^\circ$, which means that it has no zeros, the existence of such a product transformation has already been proved in the literature. It is however witnesses lying on the boundary $\partial\mathcal{S}^\circ$ which clearly represent the most interesting cases. We present an iteration scheme for computing the transformation numerically, and we find in practice for the 3×3 system, that the scheme works well even for extremal witnesses and other witnesses lying on $\partial\mathcal{S}^\circ$. We find numerically that the unital and trace preserving form of a positive map is unique up to unitary product transformations, and that they are all contractive.

In paper II, we studied extremal entanglement witnesses in dimension 3×3 by constructing numerical examples of generic extremal non-negative biquadratic forms. These are very complicated, and we do not know how to handle them other than by numerical methods. However, the corresponding extremal positive maps can be presented graphically, as we have done in this paper. We first present plots related to two different extremal entanglement witnesses in 3×3 dimensions. We produce the corresponding positive map $\mathbf{L} : H_3 \to H_3$, and then plot various two-dimensional sections in order to illustrate how the image $\mathbf{L}\mathcal{D}_3$ lies inside $\mathcal{D}_3$. The first example is a randomly chosen generic extremal entanglement witness with quadratic zeros, found in numerical searches done in paper II. We then repeat this scheme for a version of the Choi-Lam map with entanglement witness $W_C = \Omega(1, 0, 1; 0)$, where $\Omega(a, b, c; \theta)$ is the generalized Choi-Lam witness. This witness is extremal but highly non-generic, having only quartic zeros. It has three isolated quartic zeros, and one continuous two-dimensional set of zeros which are necessarily quartic. The most important feature of the plots is related to the zeros of the two witnesses. A zero (ϕ_0, χ_0) of W naturally defines a pure state $\phi_0\phi_0^\dagger$ in $\mathcal{D}_3$, which is mapped by the corresponding map $\mathbf{L}_W$ to the boundary of $\mathcal{D}_3$. In particular, a generic extremal witness in 3×3 dimensions, like the one presented here, has nine zeros (ϕ_i, χ_i), defining a simplex in $\mathcal{D}_3$ with nine vertices $\phi_i\phi_i^\dagger$ which are mapped to a simplex in $\mathcal{D}_3$ with nine vertices $\mathbf{L}_W(\phi_i\phi_i^\dagger)$ touching the

boundary of $\mathcal{D}_3$ from the inside.

One third example is based on the study of optimal witnesses by Lewenstein *et al.* [78]. They describe a method for constructing optimal witnesses in the 2×4 system. We have utilized this method to create such a witness, and then from this derived a structure of extremal quartic witnesses. From one of these extremal quartic witnesses we have constructed the corresponding positive map $\mathbf{L} : H_2 \mapsto H_4$, and then proceeded to perform various plots. It should be noted that while both maps in the 3×3 system are unital and trace preserving, this is not the case for this map in the 2×4 system.

We should emphasize that we have studied here only 3×3 and 2×4 dimensions, which are the simplest non-trivial cases. In higher dimensions, the symmetric dimensions $N_A \times N_B$ with $N_A = N_B > 3$ are clearly the most interesting. The complexity increases fast with the dimension, because the simplex of pure states in $\mathcal{D}_3$ corresponding to the quadratic zeros of extremal witnesses, becomes a non-simplex polytope, *i.e.* with a number of vertices larger than N_A^2.

We believe that the geometrical way of thinking illustrated here, may be a fruitful approach when one wants to construct examples of extremal maps and entanglement witnesses. It may be that the increase in complexity with increasing dimension, which is a well known phenomenon, is easier to handle geometrically than by other methods.

Paper IV

We continue our studies of PPT states of rank $(5, 5)$ in the 3×3 system. Though the classification of the extremal PPT states of rank four in these dimensions is believed to be complete and rather simple, the structure of the extremal PPT states of rank five appears much more complex, especially since there in this case exists a whole range of non-generic states.

The equivalence between PPT states under SL $\times$ SL-transformations is a very important concept. The special cases where ρ and its partial transpose ρ^P have this SL-equivalence are SL-symmetric states. We define a state ρ to be *genuinely* SL-symmetric if it is SL-equivalent to a state $\tau = \tau^P$. We then show that genuine SL-symmetry implies SL-symmetry, and that in the case of genuine SL-symmetry at least one SL-transformation from ρ to ρ^P must have a special diagonal block form and in addition be trace preserving. We argue that SL-symmetric states can be found numerically, but only by conducting specific searches.

We have randomly produced 50 SL-symmetric PPT states of rank $(5, 5)$, which apart from being SL-symmetric, are generic states. These searches are rather special in the sense that we look for product transformations that are trace preserving. This choice is motivated by the fact that we might in this way expect to find states that are genuinely SL-symmetric, and indeed about half of these 50 states turned out to be that.

Generic extremal PPT states of rank $(5, 5)$ have no product vectors in their kernel. How to construct PPT states that have a non-zero number n_{ker} of product vectors in the kernel is discussed in Paper 1 [57]. We here develop these matters further. These investigations is essentially a study of pairs of orthogonal *subspaces* of $\mathbb{C}^9$, rather than the states themselves. We construct non-generic orthogonal subspaces $\mathcal{U}$ and $\mathcal{V}$ with

$|\mathcal{U}| = 5$ and $|\mathcal{V}| = 4$, that allow product vectors which are orthogonal to $\mathcal{U}$, *i.e.* they lie in $\mathcal{V}$. We then construct PPT states of rank $(5,5)$ with $\mathrm{Img}\,\rho = \mathcal{U}$ and $\mathrm{Ker}\,\rho = \mathcal{V}$ with n_{ker} ranging from one to four. For a PPT state we always have that the number of product vectors in $\mathrm{Ker}\,\rho$ and $\mathrm{Ker}\,\rho^P$ are equal, and if the number of product vectors in $\mathrm{Img}\,\rho^P$ is $\widetilde{n}_{\mathrm{img}}$, we can characterize a PPT state by $\{n_{\mathrm{img}}, \widetilde{n}_{\mathrm{img}}; n_{\mathrm{ker}}\}$.

The $\{6, 6, n_{\mathrm{ker}}\}$ states we have constructed in this way are always extremal if $n_{\mathrm{ker}} = 1, 2$, while the cases $n_{\mathrm{ker}} = 3, 4$ always return non-extremal states. For the case $n_{\mathrm{ker}} = 2$ it is also possible to have $n_{\mathrm{img}} = \widetilde{n}_{\mathrm{img}} = \infty$, because this construction allows an infinite set of linear combinations of product vectors, that are again product vectors. These cases return both extremal and non-extremal states.

For the case $n_{\mathrm{ker}} = 4$ we find an interesting new analytical construction of all rank four extremal PPT states, up to SL-equivalence, where they appear as boundary states on one single five dimensional face on the set of normalized PPT states. The interior of the face consists of rank five states, a simplex of separable states surrounded by entangled PPT states. All these states are real matrices, symmetric under partial transposition.

Also, a very special subspace is collected from a small number of $\{2, 2; 1\}$ states found in random searches for SL-symmetric states. We describe analytically a set of states found by transformation to a standard form of one particular rank $(5,5)$ non-extremal PPT state in this subspace. After transformation to this standard form the rank $(5,5)$ state lies inside a circle bounded by extremal rank $(4,4)$ PPT states. The interior of the circle consists entirely of rank $(5,5)$ PPT states, each of which has exactly one product vector in its kernel, and this product vector is common to all the rank $(4,4)$ and rank $(5,5)$ states. Each of the rank $(4,4)$ states has five additional product vectors in its kernel, these are different for the different states. The rank $(4,4)$ states have no product vectors in their ranges, whereas all the rank $(5,5)$ states have one common range containing exactly two product vectors. All the rank $(4,4)$ and rank $(5,5)$ states are symmetric under partial transposition. We have also produced several extremal $\{2, 6; 1\}$ PPT states of rank $(5,5)$ in this subspace.

It is known that a subspace of $\mathbb{C}^9$ of dimension five cannot be an entangled subspace, *i.e.* it must contain at least one product vector [40]. As mentioned, the generic number n_{img} of product vectors in such a five-dimensional subspace is six. All our standard forms have either $n_{\mathrm{img}} = 6$ or $n_{\mathrm{img}} = \infty$. In order to gain a more complete understanding of the non-generic PPT states in the 3×3 system, it is important that the structure of orthogonal subspaces $\mathcal{U}$ and $\mathcal{V}$ of $\mathbb{C}^9$ is further analyzed, especially with regards to the number of product vectors that these subspaces contain.

For any five-dimensional subspace of $\mathbb{C}^9$ to contain less than six product vectors, the set of equations (3.52) must have degenerate solutions. An analysis regarding these matters is given in [56].

Bibliography

[1] E. Rutherford
"The Scattering of α and β particles by matter and the structure of the atom"
Philos. Mag. 6th series, **21**, 669 (1911)

[2] J. W. Strutt, 3rd Baron Rayleigh
"Remarks upon the law of complete radiation"
Philos. Mag. 5th series, **49**, 539 (1900)

[3] M. Planck
"Über eine verbesserung der Wienschen spektralgleichung"
Verh. Dtsch. Phys. Ges. **2**, 202 (1900)

[4] M. Planck
"Zur theorie des gesetzes der energieverteilung im normalspektrum"
Verh. Dtsch. Phys. Ges. **2**, 237 (1900)

[5] A. Einstein
"Über einen die erzeugung und verwandlung des lichtes betreffenden heuristischen gesichtspunkt"
Ann. Phys. (Berlin) **17** (6), 132 (1905)

[6] W. Heisenberg
"Über quantentheoretische umdeutung kinematischer und mechanischer beziehungen"
Z. Physik **33**, 879 (1925)

[7] L. de Broglie
"A tentative theory of light quanta"
Philos. Mag. 6th series, **47**, 446 (1924)

[8] E. Schrödinger
"Quantisierung als eigenwertproblem"
Ann. Phys. (Berlin) **79**, 361 (1926)

[9] E. Schrödinger
"Über das verhältnis der Heisenberg-Born-Jordanschen quantenmechanik zu der meinen"
Ann. Phys. (Berlin) **79**, 734 (1926)

[10] E. Schrödinger
"Die gegenwärtige situation in der quantenmechanik"
Naturwissenschaften **23**, 807 (1935)

[11] A. Einstein, B. Podolsky, and N. Rosen
"Can quantum mechanical description of physical reality be considered complete?"
Phys. Rev. **47**, 777 (1935)

[12] S. Gröblacher, T. Paterek, R. Kaltenbaek, Č. Brukner, *et al.*
"An experimental test of nonlocal realism"
Nature **446**, 871 (2007)

[13] D. Bohm
"A suggested interpretation of the quantum theory in terms of hidden variables"
Phys. Rev. **85**, 166 (1952)

[14] L. de Broglie
"La mécanique ondulatoire et la structure atomique de la matiére et du rayonnement"
J. Phys. Radium **8**, 225 (1927)

[15] J.S. Bell
"On the Einstein Podolsky Rosen paradox"
Physics **1**, 195 (1964)

[16] J.S. Bell
"Bertlmann's socks and the nature of reality"
J. de Phys. C2, **42**, 41 (1981)

[17] A. Aspect, P. Grangier, and G. Roger
"Experimental realization of Einstein-Podolsky-Rosen-Bohm gedanken experiment: A new violation of Bell's inequalities"
Phys. Rev. Lett. **49**, 91 (1982)

[18] M. Ansmann, H. Wang, R.C. Bialczak, M. Hofheinz, *et al.*
"Violation of Bell's inequality in Josephson phase qubits"
Nature **461**, 504 (2009)

[19] D. Salart, A. Baas, J.A.W. van Houwelingen, N. Gisin, *et al.*
"Spacelike separation in a Bell test assuming gravitationally induced collapses"
Phys. Rev. Lett. **100**, 220404 (2008)

[20] X.S. Ma, T. Herbst, T. Scheidl, D. Wang, *et al.*
"Quantum teleportation over 143 kilometres using active feed-forward"
Nature **489**, 269 (2012)

[21] R.F. Werner
"Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden variable model"
Phys. Rev. A **40**, 4277 (1989)

[22] C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, *et al.*
"Purification of noisy entanglement and faithful teleportation via noisy channels"
Phys. Rev. Lett. **76**, 722 (1996)

[23] E. Passaro, C. Vitelli, N. Spagnolo, F. Sciarrino, *et al.*
"Joining and splitting the quantum states of photons"
Phys. Rev. A **88**, 062321 (2013)

[24] R. Heilmann, M. Gräfe, S. Nolte, and A. Szameit
"Arbitrary photonic wave plate operations on chip: Realizing Hadamard, Pauli-X and rotation gates for polarization qubits"
Sci. Rep. **4**, 4118 (2014)

[25] C.H. Bennett and S.J. Wiesner
"Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states"
Phys. Rev. Lett. **69**, 2881 (1992)

[26] K. Mattle, H. Weinfurter, P.G. Kwiat, and A. Zeilinger
"Dense coding in experimental quantum communication"
Phys. Rev. Lett. **76**, 4656 (1996)

[27] B.M. Terhal
"Is entanglement monogamous?"
IBM. J. Res. Dev. **48**, 71 (2004)

[28] C.H. Bennett and G. Brassard
Quantum cryptography: Public key distribution and coin tossing
Proceedings of IEEE International Conference on Computers, Systems and Signal Processing, 8
IEEE Press, New York, 1984.

[29] L.M.K. Vandersypen, M. Steffen, G. Breyta, C.S. Yannoni, *et al.*
"Experimental realization of Shor's quantum factoring algorithm using NMR"
Nature **414**, 883 (2001)

[30] E. Schmidt
"Zur Theorie der linearen und nichtlinearen Integralgleichungen"
Math. Ann. **63**, 433 (1907)

[31] H.T. Lim, Y.S. Kim, Y.S. Ra, J. Bae, *et al.*
"Realizing physical approximation of the partial transpose"
Phys. Rev. Lett. **107**, 160401 (2011)

[32] A. Peres
"Separability criterion for density matrices"
Phys. Rev. Lett. **77**, 1413 (1996)

[33] M. Horodecki, P. Horodecki, and R. Horodecki
"Separability of mixed states: necessary and sufficient conditions"
Phys. Rev. Lett. **A223**, 1 (1996)

[34] P. Horodecki
"Separability criterion and inseparable mixed states with positive partial transpose"
Phys. Rev. Lett. **A232**, 333 (1997)

[35] C.H. Bennett, H.J. Bernstein, S. Popescu, and B. Schumacher
"Concentrating partial entanglement by local operations"
Phys. Rev. A **53**, 2046 (1996)

[36] M. Horodecki, P. Horodecki, and R. Horodecki
"Mixed-state entanglement and distillation: Is there bound entanglement in nature?"
Phys. Rev. Lett. **80**, 5239 (1998)

[37] D.P. DiVincenzo, P.W. Shor, J.A. Smolin, B.M. Terhal, *et al.*
"Evidence for bound entangled states with negative partial transpose"
Phys. Rev. A **61**, 062312 (2000)

[38] L. Pankowski, M. Piani, M. Horodecki, and P. Horodecki
A few steps more towards NPT bound entanglement
IEEE Transactions on Information Theory, 4085
IEEE Press, New York, (2010)

[39] R. Hartshorne
Algebraic Geometry
Springer, New York (2006)

[40] K.R. Parthasarathy
"On the maximal dimension of a completely entangled subspace for finite level quantum systems"
Proc. Math. Sci. **114**, 464 (2004)

[41] N.J. Cerf, C. Adami, and R.M. Gingrich
"Reduction criterion for separability"
Phys. Rev. A **60**, 898 (1999)

[42] M. Horodecki and P. Horodecki
"Reduction criterion of separability and limits for a class of distillation protocols"
Phys. Rev. A **59**, 4206 (1999)

[43] M.A. Nielsen and J. Kempe
"Separable states are more disordered globally than locally"
Phys. Rev. Lett. **86**, 5184 (2001)

[44] R. Horodecki and P. Horodecki
"Quantum redundancies and local realism"
Phys. Rev. Lett. **A194**, 147 (1994)

[45] K. Chen and L.A. Wu
"A matrix realignment method for recognizing entanglement"
Quantum. Inf. Comput. **3**, 193 (2003)

[46] H.K. Lo and S. Popescu
"Concentrating entanglement by local actions: Beyond mean values"
Phys. Rev. A **63**, 022301 (2001)

[47] W. Dür, G. Vidal, and J.I. Cirac
"Three qubits can be entangled in two inequivalent ways"
Phys. Rev. A **62**, 062314 (2000)

[48] C. Carathéodory
"Über den Variabilitältsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen"
Rend. Circ. Mat. Palermo **32**, 193 (1911)

[49] L. Gurvits and H. Barnum
"Largest separable balls around the maximally mixed bipartite quantum state"
Phys. Rev. A **66**, 062311 (2002)

[50] K.C. Ha and S.H. Kye
"Construction of entangled states with positive partial transposes based on indecomposable positive linear maps"
Phys. Rev. A **325**, 315 (2004)

[51] E. Alfsen and F. Shultz
"Unique decompositions, faces and automorphisms of separable states"
J. Math. Phys. **51**, 052201 (2010)

[52] H.S. Choi and S.H. Kye
"Facial structures and separable states"
J. Korean. Math. Soc. **49**, 623 (2012)

[53] P. Horodecki, M. Lewenstein, G. Vidal, and I. Cirac
"Operational criterion and constructive checks for the separability of low-rank density matrices"
Phys. Rev. A **62**, 032310 (2000)

[54] C.H. Bennett, D.P. DiVincenzo, T. Mor, P.W. Shor, *et al.*
"Unextendible product bases and bound entanglement"
Phys. Rev. Lett. **82**, 5385 (1999)

[55] J.M. Leinaas, J. Myrheim, and P.Ø. Sollid
"Low-rank extremal positive-partial-transpose states and unextendable product bases"
Phys. Rev. A **81**, 062330 (2010)

[56] L. Chen and D.Z. Djokovic
"Description of rank four PPT entangled states of two qutrits"
J. Math. Phys. **52**, 122203 (2011)

[57] L.O. Hansen, A. Hauge, J. Myrheim, and P.Ø. Sollid
"Low rank positive partial transpose states and their relation to product vectors"
Phys. Rev. A **85**, 022309 (2011)

[58] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, *et al.*
"Detection of entanglement with polarized photons: Experimental realization of an entanglement witness"
Phys. Rev. Lett. **91**, 227901 (2003)

[59] F.G.S.L. Brandao
"Quantifying entanglement with witness operators"
Phys. Rev. A **72**, 022310 (2005)

[60] D. Hilbert
"Mathematische probleme"
Arch. Math. Phys. **1**, 44 (1901)

[61] E. Artin
"Über die Zerlegung definiter funktionen in quadrate"
Abh. Math. Sem. Hamburg **5**, 110 (1927)

[62] T.S. Motzkin
The arithmetic-geometric inequality
Inequalities, 205
Academic Press, New York (1967)

[63] M.D. Choi
"Positive semidefinite biquadratic forms"
Linear Algebra Appl. **12**, 95 (1975)

[64] M.D. Choi and T.Y. Lam
"Extremal positive semidefinite forms"
Math. Ann. **231**, 1 (1977)

[65] A. Jamiolkowski
"Linear transformations which preserve trace and positive semidefiniteness of operators"
Rep. Math. Phys. **3**, 275 (1972)

[66] E. Størmer
"Positive linear maps of operator algebras"
Acta Math. **110**, 233 (1963)

[67] E. Størmer
"Decomposable linear maps on C^*-algebras"
Proc. Amer. Math. Soc. **86**, 402 (1982)

[68] M. Marciniak
"On extremal positive maps acting between type I factors"
Banach Center Publ. **89**, 221 (2010)

[69] J. Grabowski, M. Kuś, and G.Marmo
"Wigner's theorem and the geometry of extreme positive maps"
J. Phys. A: Math. Theor. **42**, 345301 (2009)

[70] S.L. Woronowicz
"Positive maps of low dimensional matrix algebras"
Rep. Math. Phys. **10**, 165 (1976)

[71] L.O. Hansen, A. Hauge, J. Myrheim, and P.Ø. Sollied
"Extremal entanglement witnesses"
Int. J. Quantum Inf. **13**, 1550060 (2015)

[72] K.C. Ha and S.H. Kye
"Entanglement witnesses arising from Choi type positive linear maps"
J. Phys. A: Math. Theor. **45**, 415305 (2012)

[73] E. Størmer
Positive Linear Maps of Operator Algebras
Monographs in Mathematics, Springer Verlag, Berlin (2007)

[74] J.M. Leinaas, J. Myrheim, and E. Ovrum
"Geometrical aspects of entanglement"
Phys. Rev. A **74**, 012313 (2006)

[75] A. Hauge
"A geometrical and computational study of entanglement witnesses"
Master thesis, Norwegian University of Science and Technology (2011)

[76] D. Chruściński and G. Sarbicki
"Exposed positive maps: A sufficient condition"
J. Phys. A: Math. Theor. **45**, 115304 (2012)

[77] D. Chruściński and G. Sarbicki
"A class of indecomposable positive maps"
J. Phys. A: Math. Theor. **46**, 015306 (2013)

[78] M. Lewenstein, B. Kraus, J.I. Cirac, and P. Horodecki
"Optimization of entanglement witnesses"
Phys. Rev. A **62**, 052310 (2000)

[79] M. Lewenstein, B. Kraus, J.I. Cirac, and P. Horodecki
"Characterization of separable states and entanglement witnesses"
Phys. Rev. A **63**, 044304 (2001)

[80] R. Augusiak, J. Tura, and M. Lewenstein
"A note on the optimality of decomposable entanglement witnesses and completely entangled subspaces"
J. Phys. A: Math. Theor. **44**, 212001 (2011)

[81] R. Augusiak, G. Sarbicki, and M. Lewenstein
"Optimal decomposable witnesses without the spanning property"
Phys. Rev. A **84**, 052323 (2011)

[82] P. Horodecki and A. Ekert
"Method for direct detection of quantum entanglement"
Phys. Rev. Lett. **89**, 127902 (2002)

[83] J.K. Korbicz, M.L. Almeida, J. Bae, M. Lewenstein, *et al.*
"Structural approximations to positive maps and entanglement breaking channels"
Phys. Rev. A **78**, 062105 (2008)

[84] D. Chruściński and G. Sarbicki
"Disproving the conjecture on the structural physical approximation to optimal decomposable entanglement witnesses"
J. Phys. A: Math. Theor. **47**, 195301 (2014)

[85] K.C. Ha and S.H. Kye
"The structural physical approximations and optimal entanglement witnesses"
J. Math. Phys. **53**, 102204 (2012)

[86] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery
Numerical Recipes: The Art of Scientific Computing, 3rd Edition
Cambridge University Press, New York (2007)

Papers

Paper I

L.O. Hansen, A. Hauge, J. Myrheim, and P.Ø. Sollid

"Low rank positive partial transpose states and their relation to product vectors"

Phys. Rev. A **85**, 022309 (2012)

PHYSICAL REVIEW A **85**, 022309 (2012)

Low-rank positive-partial-transpose states and their relation to product vectors

Leif Ove Hansen,[1] Andreas Hauge,[1] Jan Myrheim,[1] and Per Øyvind Sollid[2]
[1]*Department of Physics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway*
[2]*Department of Physics, University of Oslo, N-0316 Oslo, Norway*
(Received 14 April 2011; revised manuscript received 21 November 2011; published 6 February 2012)

Deciding whether a mixed quantum state is separable or entangled is a difficult problem in general. Separable states are positive under partial transposition [they are positive-partial-transpose (PPT) states], but this simple test does not exclude all entangled states. In order to understand the entangled PPT states, having so-called bound entanglement, we want to study the extremal PPT states. An extremal PPT state is either a pure product state, then it is separable, or it is entangled, in which case the state and its partial transpose must still be rather low-rank density matrices, although it is known that the rank must be at least 4. In a previous paper we presented a complete classification of the rank-4 entangled PPT states in dimension 3×3, generalizing the construction by Bennett *et al.* from unextendible sets of orthogonal product vectors. In the present paper we continue our investigations of the low-rank entangled PPT states, mostly in dimension 3×3, using a combination of analytical and numerical methods. We use perturbation theory in order to construct rank-5 entangled PPT states close to the known rank-4 states, and in order to compute dimensions and study the geometry of surfaces of low-rank PPT states. We exploit the close connection between low-rank PPT states and product vectors. In particular, we show how to reconstruct a low-rank PPT state from a sufficient number of product vectors in its kernel. It may seem surprising that the number of product vectors needed may be smaller than the dimension of the kernel.

DOI: 10.1103/PhysRevA.85.022309 PACS number(s): 03.67.Mn, 02.40.Ft, 03.65.Ud

I. INTRODUCTION

Quantum entanglement between subsystems of a composite physical system, independent of spatial separation, is a phenomenon which clearly distinguishes quantum physics from classical physics [1]. Entangled quantum states show correlations between local measurements on the subsystems which cannot be modeled within classical physics with local interactions, since they violate Bell inequalities [2,3], or even equalities such as the Greenberger-Horne-Zeilinger (GHZ) equalities in a three-particle system [4,5].

A pure classical state of a composite system has no correlations, since classical measurements are deterministic and disturb the system minimally. A statistical ensemble of pure classical states, what we may call a mixed classical state, can have correlations, but these correlations cannot violate Bell inequalities, by definition.

The only pure quantum states that are not entangled are the pure product states, which resemble pure classical states in that they have no correlations at all. By definition, a mixed quantum state is a statistical ensemble of pure quantum states, and it is said to be separable if it can be mixed entirely from pure product states. The separable mixed states are not entangled, since they cannot violate Bell inequalities, and the entangled mixed states are precisely those that are nonseparable.

The separability problem, how to characterize the set $\mathcal{S}$ of separable mixed states and decide whether a given mixed state is separable or entangled, is known to be a difficult mathematical problem [6]. It motivates our work presented here and in previous papers [7–10], although we have studied not so much the separable states directly as the larger class of mixed states called positive-partial-transpose (PPT) states.

The separable mixed states have the property that they remain positive after partial transposition; they are PPT states, for short. This is known as the Peres separability criterion [11]. It is a powerful separability test, which can be used, for example, to prove that any pure quantum state is either entangled or a pure product state. The set $\mathcal{P}$ of PPT states is in general larger than the set $\mathcal{S}$ of separable states, but the difference between the two sets is surprisingly small in low dimensions, and in the very lowest dimensions, 2×2, 2×3, and 3×2, there is no difference [12,13].

Entanglement is regarded as a physical resource which can be used for special tasks that cannot be performed otherwise, such as quantum teleportation and dense coding, quantum cryptography, or quantum computing. For such purposes one may want pure entangled states, and then it is useful to have methods for distilling out a number of entangled pure states from a larger number of entangled mixed states [14]. A special property of the entangled PPT states is that their entanglement is bound—it cannot be distilled into the entanglement of pure states [15,16].

Several groups are now studying experimentally the bound entanglement in entangled PPT states [17–20]. This is another motivation for our theoretical studies of the PPT states and the nature of their entanglement.

The set of PPT states is convex, and since all members of a convex set are convex combinations of the extremal points of the set, it is a natural idea to study the extremal PPT states. Our approach is to construct numerical examples and try to understand them by analytical means. The extremal PPT states are easily accessible for numerical study, since it is straightforward to test for both the PPT property and the extremality. The extremality condition for a PPT state implies that both the density matrix and its partial transpose must have rather low rank, and this motivates us to study more generally low-rank PPT states.

We know something about the extremal PPT states already from the fact that $\mathcal{S} \subset \mathcal{P} \subset \mathcal{D}$, where $\mathcal{S}$ is the set of separable states, $\mathcal{P}$ the set of PPT states, and $\mathcal{D}$ the full set of density matrices, and from the fact that all three sets are convex. By definition, the extremal points of $\mathcal{S}$ are the pure product

states, and since these are also extremal points of $\mathcal{D}$, they must be extremal points of $\mathcal{P}$. All other extremal PPT states are entangled, because a separable extremal PPT state must be extremal in $\mathcal{S}$ and hence a pure product state. Thus the extremal PPT states that are not pure product states are immediately interesting as examples of entangled PPT states.

A. The relation between PPT states and product vectors

There is a very close connection between PPT states and product vectors, which has previously been used, for example, to prove the separability of sufficiently low-rank PPT states [21]. In the present work we have exploited this connection further.

Bennett *et al.* [22,23] introduced a method for constructing low-rank mixed states that are obviously entangled PPT states, using what they called unextendible product bases (UPBs). From a UPB, defined as a maximal set of orthogonal product vectors which is not a complete basis of the Hilbert space, one constructs an orthogonal projection Q and the complementary projection $P = \mathbb{1} - Q$. Then $\rho = P/(\mathrm{Tr}\, P)$ is an entangled PPT state.

The UPB construction is most successful in the special case of rank-4 PPT states in 3×3 dimensions. In Ref. [10] we argued, based on evidence from numerical studies, that an extended version of the UPB construction, including nonunitary but nonsingular product transformations on the states, is general enough to produce all rank-4 entangled PPT states in 3×3 dimensions. This conjecture has since been proved independently by Skowronek [24] and by Chen and Djokovic [25].

Unfortunately, attempts to apply the UPB method directly in higher dimensions fail, even when the kernel contains product vectors, because there cannot exist a sufficient number of orthogonal product vectors. The orthogonality is essential in the construction by Bennett *et al.* of the PPT state as a projection operator. We would like to generalize the construction in such a way that it works without the orthogonality condition.

One possible generalization is to construct projection operators as more general convex combinations, or even as nonconvex linear combinations, of pure product states. This idea is explored in a separate paper [26].

In the present paper we discuss in general the constraints imposed on a PPT state ρ by the existence of product vectors in its kernel, and we show that these constraints are so strong that they actually determine the state uniquely. A surprising discovery is that in cases where the kernel contains a finite overcomplete set of product vectors, the state ρ can be reconstructed from only a subset of the product vectors, and the number of product vectors needed may even be smaller than the dimension of the kernel.

From this point of view, the important question is exactly what conditions the product vectors must satisfy in order for the constraint equations to have a solution for ρ. We can answer this question in the familiar special case of rank-4 PPT states in 3×3 dimensions, but not in other cases. We consider this an interesting problem for future research.

B. Outline of the paper

The contents of the present paper are organized as follows. First we review some linear algebra, in particular, degenerate perturbation theory, in Sec. II. The main purpose is to introduce notation and collect formulas for later reference. In Sec. III we point out the very stringent conditions that follow from the existence of product vectors in the kernel of the lowest rank PPT states. These conditions imply that not just any subspace containing product vectors may serve as the kernel of a PPT state, and the PPT state seems to be uniquely determined by the product vectors in its kernel. In fact, the state is determined as soon as a minimum number of product vectors is specified, and the required number may even be smaller than the dimension of the kernel.

In Sec. IV we discuss the rank-4 PPT states in 3×3 dimensions. We review the UPB construction, based on orthogonal product vectors in the kernel, before we describe an approach which is different in that the product vectors need not be orthogonal. The new approach also throws new light on a set of reality conditions that limit the selection of product vectors to be used for constructing rank-4 PPT states.

In Sec. V we discuss the rank-5 PPT states in 3×3 dimensions. We find an eight-dimensional surface of rank-5 PPT states in every generic five-dimensional subspace, but we have not found any general method to construct such states. However, we show how to construct rank-5 PPT states by perturbing rank 4-PPT states. Again, the product vectors in the kernel of the rank-4 state play an important role in our construction of the rank-5 states.

In Sec. VI we discuss rank-6 PPT states in 4×4 dimensions. The kernel of such a state has dimension 10 and contains 20 product vectors. The remarkable result we find is that the state can be constructed from only seven product vectors in the kernel. An arbitrary set of seven product vectors does not produce a rank-6 PPT state, but we do not know how to select sets of product vectors that can be used in such a construction.

In Sec. VII we discuss briefly how to determine numerically the dimensions of surfaces of PPT states of fixed rank. We find that the dimensions are given by a simple counting of independent constraints, except for the very lowest rank states, for which the constraints are not independent.

Finally, we discuss in Sec. VIII how to study a surface of PPT states by numerical integration of equations of motion for curves on the surface. In this way one may study, for example, the curvature of the surface, or how a curve on the surface approaches the boundary of the surface.

II. SOME BASIC LINEAR ALGEBRA

In this section we review some linear algebra, mainly in order to introduce notation and collect formulas for later use. In particular, degenerate perturbation theory is a central theme in the present paper. We define projection operators acting on the space of Hermitian matrices which we use in our numerical calculations.

A. Density matrices

Let H_N be the set of Hermitian $N \times N$ matrices. It has a natural structure as a real Hilbert space of dimension N^2 with the scalar product

$$(X, Y) = \mathrm{Tr}(XY). \tag{1}$$

In order to do numerical calculations we may choose an orthonormal basis of H_N and represent the Hermitian matrices as real vectors.

A mixed state, or density matrix, is a positive Hermitian matrix of unit trace. We define

$$\mathcal{D} = \mathcal{D}_N = \{\rho \in H_N \mid \rho \geqslant 0\,,\ \mathrm{Tr}\,\rho = 1\}. \tag{2}$$

For a given $\rho \in \mathcal{D}$ we define P and $Q = \mathbb{1} - P$ as the orthogonal projections onto $\mathrm{Img}\,\rho$, the image of ρ, and $\mathrm{Ker}\,\rho$, the kernel of ρ, respectively. The relations $P\rho = \rho P = P\rho P = \rho$ and $Q\rho = \rho Q = Q\rho Q = 0$ will be used in the following.

That ρ is positive, or positive semidefinite, written as $\rho \geqslant 0$, means that all the eigenvalues of ρ are non-negative. An equivalent condition is that $\psi^\dagger \rho \psi \geqslant 0$ for all $\psi \in \mathbb{C}^N$. It follows from the last inequality and the spectral representation

$$\rho = \sum_{i=1}^{N} \lambda_i\, \psi_i \psi_i^\dagger, \tag{3}$$

in terms of eigenvalues λ_i and orthonormal eigenvectors ψ_i of ρ, that $\psi^\dagger \rho \psi = 0$ if and only if $\rho\psi = 0$.

The definition of positive Hermitian matrices by inequalities of the form $\psi^\dagger \rho \psi \geqslant 0$ implies that $\mathcal{D}$ is a convex set. That is, if ρ is a convex combination of $\rho_1, \rho_2 \in \mathcal{D}$,

$$\rho = \lambda\rho_1 + (1-\lambda)\rho_2 \quad \text{with} \quad 0 < \lambda < 1, \tag{4}$$

then $\rho \in \mathcal{D}$. Furthermore, since $\mathrm{Ker}\,\rho = \{\psi \mid \psi^\dagger \rho \psi = 0\}$ when $\rho \geqslant 0$, it follows that

$$\mathrm{Ker}\,\rho = \mathrm{Ker}\,\rho_1 \cap \mathrm{Ker}\,\rho_2, \tag{5}$$

independent of λ, when ρ is a convex combination as above. Since $\mathrm{Ker}\,\rho$ is independent of λ, so is $\mathrm{Im}\,\rho = (\mathrm{Ker}\,\rho)^\perp$.

A convex set is defined by its extremal points: those points that are not convex combinations of other points. The extremal points of $\mathcal{D}$ are the pure states of the form $\rho = \psi\psi^\dagger$.

1. Finite perturbations

In the following, let $\rho \in \mathcal{D}$, and define projections P onto $\mathrm{Im}\,\rho$ and $Q = \mathbb{1} - P$ onto $\mathrm{Ker}\,\rho$. Consider a perturbation

$$\rho' = \rho + \epsilon A \tag{6}$$

where $A \neq 0$ is Hermitian, and $\mathrm{Tr}\,A = 0$ so that $\mathrm{Tr}\,\rho' = \mathrm{Tr}\,\rho$ [replace A by $A - (\mathrm{Tr}\,A)\,\rho$ if $\mathrm{Tr}\,A \neq 0$]. The real parameter ϵ may be finite or infinitesimal, so we first consider the case when ϵ is finite.

We observe that if $\mathrm{Im}\,A \subset \mathrm{Im}\,\rho$, or equivalently, if

$$PAP = A, \tag{7}$$

then there will be a finite range of values of ϵ, say $\epsilon_1 \leqslant \epsilon \leqslant \epsilon_2$ with $\epsilon_1 < 0 < \epsilon_2$, such that $\rho' \in \mathcal{D}$ and $\mathrm{Im}\,\rho' = \mathrm{Im}\,\rho$. This is so because the eigenvectors of ρ with zero eigenvalue will remain eigenvectors of ρ' with zero eigenvalue, and all the positive eigenvalues of ρ will change continuously with ϵ into eigenvalues of ρ'.

The other way around, if $\rho' \in \mathcal{D}$ for $\epsilon_1 \leqslant \epsilon \leqslant \epsilon_2$ with $\epsilon_1 < 0 < \epsilon_2$, then ρ' is a convex combination of $\rho + \epsilon_1 A$ and $\rho + \epsilon_2 A$ for every ϵ in the open interval $\epsilon_1 < \epsilon < \epsilon_2$. Hence $\mathrm{Im}\,\rho'$ is independent of ϵ in this interval, implying that $\mathrm{Im}\,A \subset \mathrm{Im}\,\rho$ and $PAP = A$.

This shows that ρ is extremal in $\mathcal{D}$ if and only if there exists no $A \neq 0$ with $\mathrm{Tr}\,A = 0$ and $PAP = A$. Another formulation of the condition is that there exists no $\rho' \in \mathcal{D}$ with $\rho' \neq \rho$ and $\mathrm{Im}\,\rho' = \mathrm{Im}\,\rho$. A third equivalent formulation of the extremality condition is that the equation $PAP = A$ without the condition $\mathrm{Tr}\,A = 0$ has $A = \rho$ as its only solution (up to proportionality).

2. Infinitesimal perturbations

Assume now that $\mathrm{Im}\,A \not\subset \mathrm{Im}\,\rho$. The question is how an infinitesimal perturbation affects the zero eigenvalues of ρ. When ρ is of low rank we need degenerate perturbation theory, which is well known from any textbook on quantum mechanics.

To first order in ϵ, the zero eigenvalues of ρ are perturbed into eigenvalues of ρ' that are ϵ times the eigenvalues of QAQ on the subspace $\mathrm{Ker}\,\rho$. Similarly, to first order in ϵ, the positive eigenvalues of ρ are perturbed into positive eigenvalues of ρ' in a way which is determined by how ρ and PAP act on $\mathrm{Im}\,\rho$.

It is clear from this that to first order in ϵ, the condition

$$QAQ = 0 \tag{8}$$

is necessary and sufficient to ensure that the rank of ρ' equals the rank of ρ, and that $\rho' \geqslant 0$ both for $\epsilon > 0$ and for $\epsilon < 0$.

More generally, to first order in ϵ, the rank of ρ' equals the rank of ρ plus the rank of QAQ. For example, if we want to perturb ρ in such a way that the rank increases by one, then we have to choose A such that

$$QAQ = \alpha\, \phi\phi^\dagger, \tag{9}$$

where $\phi \in \mathrm{Ker}\,\rho$ is a normalized eigenvector of QAQ with $\alpha \neq 0$ as an eigenvalue. Since QAQ is Hermitian, α must be real. If $\alpha > 0$, then $\rho' \geqslant 0$ for $\epsilon > 0$ but not for $\epsilon < 0$.

3. Projection operators on H_N

Using the projections P and Q defined above, we define projection operators on H_N, the real Hilbert space of Hermitian $N \times N$ matrices, as follows:

$$\begin{aligned} \mathbf{P}X &= PXP, \\ \mathbf{Q}X &= QXQ = X - PX - XP + PXP, \\ \mathbf{R}X &= (\mathbf{I} - \mathbf{P} - \mathbf{Q})X = PX + XP - 2PXP. \end{aligned} \tag{10}$$

Here $\mathbf{I}$ is the identity on H_N. It is straightforward to verify that these are complementary projections, with $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{Q}^2 = \mathbf{Q}$, $\mathbf{PQ} = \mathbf{QP} = \mathbf{0}$, and so on. They are symmetric with respect to the natural scalar product on H_N, for example,

$$(X, \mathbf{P}Y) = \mathrm{Tr}(XPYP) = \mathrm{Tr}(PXPY) = (\mathbf{P}X, Y). \tag{11}$$

Hence they project orthogonally, and relative to an orthonormal basis of H_N they are represented by symmetric matrices.

Relative to an orthonormal basis of $\mathbb{C}^N$ with the first basis vectors in $\mathrm{Im}\,\rho$ and the last basis vectors in $\mathrm{Ker}\,\rho$, a Hermitian matrix X takes the block form

$$X = \begin{pmatrix} U & V \\ V^\dagger & W \end{pmatrix}, \tag{12}$$

with $U^\dagger = U$ and $W^\dagger = W$. In this basis we have

$$P = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}, \tag{13}$$

and hence

$$\mathbf{P}X = \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{Q}X = \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix}, \quad \mathbf{R}X = \begin{pmatrix} 0 & V \\ V^\dagger & 0 \end{pmatrix}. \tag{14}$$

B. Composite systems

1. Product vectors

If $N = N_A N_B$ then the tensor product spaces $\mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ (a complex tensor product) and $H_N = H_{N_A} \otimes H_{N_B}$ (a real tensor product) may describe a composite quantum system with two subsystems A and B of Hilbert space dimensions N_A and N_B.

A vector $\psi \in \mathbb{C}^N$ then has components $\psi_I = \psi_{ij}$, where

$$I = 1,2,\ldots,N \leftrightarrow ij = 11,12,\ldots,1N_B,21,22,\ldots,N_A N_B. \tag{15}$$

A product vector $\psi = \phi \otimes \chi$ has components $\psi_{ij} = \phi_i \chi_j$. We see that ψ is a product vector if and only if its components satisfy the quadratic equations,

$$\psi_{ij}\psi_{kl} - \psi_{il}\psi_{kj} = 0. \tag{16}$$

These equations are not all independent; the number of independent complex equations is

$$K = (N_A - 1)(N_B - 1) = N - N_A - N_B + 1. \tag{17}$$

For example, if $\psi_{11} \neq 0$ we get a complete set of independent equations by taking $i = j = 1$ and $k = 2,3,\ldots,N_A$, $l = 2,3,\ldots,N_B$.

Since the equations are homogeneous, any solution $\psi \neq 0$ gives rise to a one-parameter family of solutions $c\psi$ with $c \in \mathbb{C}$. A vector ψ in a subspace of dimension n has n independent complex components. Since the most general nonzero solution must contain at least one free complex parameter, we conclude that a generic subspace of dimension n will contain nonzero product vectors if and only if

$$n \geqslant K + 1 = N - N_A - N_B + 2. \tag{18}$$

The limiting dimension

$$n = N - N_A - N_B + 2 \tag{19}$$

is particularly interesting. In this special case a nonzero solution will contain exactly one free parameter, which has to be a complex normalization constant.

Thus, up to proportionality, there will exist a finite set of product vectors in a generic subspace of dimension $n = N - N_A - N_B + 2$. The number of product vectors is [9]

$$p = \binom{N_A + N_B - 2}{N_A - 1} = \frac{(N_A + N_B - 2)!}{(N_A - 1)!\,(N_B - 1)!}. \tag{20}$$

A generic subspace of lower dimension will contain no nonzero product vector, whereas any subspace of higher dimension will contain a continuous infinity of different product vectors (different in the sense that they are not proportional).

2. Partial transposition

The following relation between matrix elements,

$$(X^P)_{ij;kl} = X_{il;kj}, \tag{21}$$

defines X^P, the partial transpose of the Hermitian matrix X with respect to the second subsystem. Partial transposition is an orthogonal linear transformation: it preserves the scalar product defined in Eq. (1).

A density matrix is called separable if it is a convex combination of the form

$$\rho = \sum_k p_k\, \sigma_k \otimes \tau_k, \tag{22}$$

with $\sigma_k \in \mathcal{D}_{N_A}$, $\tau_k \in \mathcal{D}_{N_B}$, $p_k > 0$, $\sum_k p_k = 1$. We denote by $\mathcal{S}$ the set of separable density matrices. Partial transposition obviously preserves the positivity of a separable density matrix. This is an easily testable condition for separability, known as the Peres criterion. The set of all PPT (positive partial transpose) matrices,

$$\mathcal{P} = \{\, \rho \in \mathcal{D} \mid \rho^P \geqslant 0 \,\} = \mathcal{D} \cap \mathcal{D}^P, \tag{23}$$

may be called the Peres set. A well-known result is that $\mathcal{P} = \mathcal{S}$ for $N = N_A N_B \leqslant 6$, whereas $\mathcal{P}$ is strictly larger than $\mathcal{S}$ in higher dimensions [12].

We will classify low-rank PPT states by the ranks (m,n) of ρ and ρ^P, respectively. Note that ranks (m,n) and (n,m) are equivalent for the purpose of classification because of the symmetric roles of ρ and ρ^P.

3. Product transformations

A product transformation of the form

$$\rho \mapsto \rho' = a V \rho V^\dagger \quad \text{with} \quad V = V_A \otimes V_B, \tag{24}$$

where a is a normalization factor and $V_A \in \mathrm{SL}(N_A,\mathbb{C})$, $V_B \in \mathrm{SL}(N_B,\mathbb{C})$, preserves positivity, rank, separability, and other interesting properties that the density matrix ρ may have. For example, it preserves positivity of the partial transpose, because

$$(\rho')^P = a \widetilde{V} \rho^P \widetilde{V}^\dagger \quad \text{with} \quad \widetilde{V} = V_A \otimes V_B^*. \tag{25}$$

The image and kernel of ρ and ρ^P transform in the following ways:

$$\mathrm{Im}\,\rho' = V\,\mathrm{Im}\,\rho, \quad \mathrm{Ker}\,\rho' = (V^\dagger)^{-1}\,\mathrm{Ker}\,\rho, \tag{26}$$

and

$$\mathrm{Im}\,(\rho')^P = \widetilde{V}\,\mathrm{Im}\,\rho^P, \quad \mathrm{Ker}\,(\rho')^P = (\widetilde{V}^\dagger)^{-1}\,\mathrm{Ker}\,\rho^P. \tag{27}$$

All these transformations are of product form and hence preserve the number of product vectors in a subspace.

We say that two density matrices ρ and ρ' related in this way are SL $\otimes$ SL equivalent, or simply Special linear (SL) equivalent. The concept of SL equivalence is important to us here because it simplifies very much our efforts to classify the low-rank PPT states.

C. Restricted perturbations

We have seen that Eq. (7) ensures that the perturbation $\rho' = \rho + \epsilon A$ preserves the image of ρ, so that $\mathrm{Im}\,\rho' = \mathrm{Im}\,\rho$

and $\rho' \geqslant 0$ for a finite range of values of the perturbation parameter ϵ. The weaker condition in Eq. (8) ensures only that rank $\rho' =$ rank ρ and $\rho' \geqslant 0$ for infinitesimal values of ϵ.

We will now discuss how to use perturbations with similar restrictions in order to study the convex set $\mathcal{P}$. In particular, we are interested in perturbations that either preserve the ranks (m,n) of ρ, or else change these ranks in controlled ways. We also state the conditions that identify the extremal points of $\mathcal{P}$.

In a similar way as we did for ρ, we define $\widetilde{P}$ and $\widetilde{Q} = \mathbb{1} - \widetilde{P}$ as the orthogonal projections onto Im ρ^P and Ker ρ^P. Then we define

$$\begin{aligned}
&\widetilde{\mathbf{P}}X = (\widetilde{P}X^P\widetilde{P})^P,\\
&\widetilde{\mathbf{Q}}X = (\widetilde{Q}X^P\widetilde{Q})^P = X - (\widetilde{P}X^P)^P - (X^P\widetilde{P})^P + (\widetilde{P}X^P\widetilde{P})^P,\\
&\widetilde{\mathbf{R}}X = (\mathbf{I} - \widetilde{\mathbf{P}} - \widetilde{\mathbf{Q}})X = (\widetilde{P}X^P)^P + (X^P\widetilde{P})^P - 2(\widetilde{P}X^P\widetilde{P})^P.
\end{aligned} \tag{28}$$

These are again projections on the real Hilbert space H_N, like $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$, again symmetric with respect to the natural scalar product on H_N. We may use these projections on H_N to impose various restrictions on the perturbation matrix A.

1. Testing for extremality in $\mathcal{P}$

The extremality condition for $\mathcal{P}$ is derived in a similar way as the extremality condition for $\mathcal{D}$ based on Eq. (7). Clearly ρ is extremal in $\mathcal{P}$ if and only if there exists no $\rho' \in \mathcal{P}$, $\rho' \neq \rho$, with both Im $\rho' =$ Im ρ and Im$(\rho')^P =$ Im ρ^P. Another way to formulate this condition is that $A = \rho$ is the only solution of the two equations $\mathbf{P}A = A$ and $\widetilde{\mathbf{P}}A = A$.

Since $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ are projections, the equations $\mathbf{P}A = A$ and $\widetilde{\mathbf{P}}A = A$ together are equivalent to the single eigenvalue equation

$$(\mathbf{P} + \widetilde{\mathbf{P}})A = 2A. \tag{29}$$

They are also equivalent to either one of the eigenvalue equations

$$\mathbf{P}\widetilde{\mathbf{P}}\mathbf{P}A = A, \quad \widetilde{\mathbf{P}}\mathbf{P}\widetilde{\mathbf{P}}A = A. \tag{30}$$

Note that the operators $\mathbf{P} + \widetilde{\mathbf{P}}$, $\mathbf{P}\widetilde{\mathbf{P}}\mathbf{P}$, and $\widetilde{\mathbf{P}}\mathbf{P}\widetilde{\mathbf{P}}$ are all real symmetric and therefore have complete sets of real eigenvalues and orthonormal eigenvectors. In fact, the eigenvalues are all non-negative, because the operators are positive.

When we diagonalize $\mathbf{P} + \widetilde{\mathbf{P}}$ we will always find $A = \rho$ as an eigenvector with eigenvalue 2. If it is the only such eigenvector, this proves that ρ is extremal in $\mathcal{P}$. If A is a solution not proportional to ρ, then we may impose the condition Tr $A = 0$ [replace A by $A - (\text{Tr}\, A)\rho$ if necessary], and we know that there exists a finite range of both positive and negative values of ϵ such that $\rho' = \rho + \epsilon A \in \mathcal{P}$; hence ρ is not extremal.

It should be noted that in our numerical calculations we may find eigenvalues of $\mathbf{P} + \widetilde{\mathbf{P}}$ that differ from 2 by less than 1% . However, since eigenvalues are calculated with a precision close to the internal precision of the computer, which is of order 10^{-16}, there is never any ambiguity as to whether an eigenvalue is equal to 2 or strictly smaller than 2.

2. Perturbations preserving the PPT property and ranks

The rank and positivity of ρ are preserved by the perturbation, to first order in ϵ, both for $\epsilon > 0$ and $\epsilon < 0$, if and only if $\mathbf{Q}A = 0$. Similarly, the rank and positivity of ρ^P are preserved if and only if $\widetilde{\mathbf{Q}}A = 0$. These two equations together are equivalent to the single-eigenvalue equation

$$(\mathbf{Q} + \widetilde{\mathbf{Q}})A = 0. \tag{31}$$

Again, $\mathbf{Q} + \widetilde{\mathbf{Q}}$ is real symmetric and has a complete set of real eigenvalues and eigenvectors.

In conclusion, the perturbations that preserve the PPT property, as well as the ranks (m,n) of ρ and ρ^P, to first order in ϵ, are the solutions of Eq. (31).

If we want to perturb in different ways, for example, such that Im $\rho' =$ Im ρ, but not necessarily Im$(\rho')^P =$ Im ρ^P, we only require $(\rho')^P$ and ρ^P to have the same rank. Then the conditions on A are that $\mathbf{P}A = A$ and $\widetilde{\mathbf{Q}}A = 0$, or equivalently,

$$(\mathbf{I} - \mathbf{P} + \widetilde{\mathbf{Q}})A = 0. \tag{32}$$

III. PRODUCT VECTORS IN THE KERNEL

Our purpose in this section is to state explicitly the very close connection between the lowest rank PPT states and product vectors. If the dimension of Ker ρ is equal to or larger than the limiting dimension n given by Eq. (19), or equivalently, if

$$\text{rank}\, \rho \leqslant N - n = N_A + N_B - 2, \tag{33}$$

then Ker ρ will always contain product vectors, and from these product vectors we obtain a set of linear equations for ρ. In numerical examples we find that these equations determine ρ uniquely.

Assume now that $\rho \in \mathcal{P}$. Recall that because $\rho \geqslant 0$ and $\rho^P \geqslant 0$, the equations $w^\dagger \rho w = 0$ and $\rho w = 0$ for $w \in \mathbb{C}^N$ are equivalent, and so are the equations $w^\dagger \rho^P w = 0$ and $\rho^P w = 0$. Taken together with the identity

$$(x \otimes y)^\dagger \rho\, (u \otimes v) = (x \otimes v^*)^\dagger \rho^P (u \otimes y^*), \tag{34}$$

this puts strong restrictions on ρ when we know a number of product vectors in Ker ρ.

Assume from now on that w is a product vector, $w = u \otimes v$. Defining $\widetilde{w} = u \otimes v^*$ we have the general relation

$$w^\dagger \rho w = \widetilde{w}^\dagger \rho^P \widetilde{w}, \tag{35}$$

which implies that the relations $w \in$ Ker ρ and $\widetilde{w} \in$ Ker ρ^P are equivalent. Assume that $w \in$ Ker ρ. For any $z \in \mathbb{C}^N$ we have the condition on ρ that $z^\dagger \rho w = 0$. In particular, when z is an arbitrary product vector, $z = x \otimes y$, we have the two conditions on ρ that

$$\begin{aligned}
&(x \otimes y)^\dagger \rho\, (u \otimes v) = 0,\\
&(x \otimes v)^\dagger \rho\, (u \otimes y) = (x \otimes y^*)^\dagger \rho^P (u \otimes v^*) = 0.
\end{aligned} \tag{36}$$

Assume that $w_i = u_i \otimes v_i \in$ Ker ρ for $i = 1,2,\ldots,p$. Then for arbitrary values of the indices i,j,k we have the following constraints on ρ:

$$(u_i \otimes v_j)^\dagger \rho\, (u_k \otimes v_k) = (u_i \otimes v_k)^\dagger \rho\, (u_k \otimes v_j) = 0. \tag{37}$$

Let us introduce matrices

$$A_{klij} = (u_k \otimes v_l)(u_i \otimes v_j)^\dagger \tag{38}$$

and Hermitian matrices

$$B_{klij} = A_{klij} + (A_{klij})^\dagger, \quad C_{klij} = \mathrm{i}\,(A_{klij} - (A_{klij})^\dagger), \tag{39}$$

then the constraints on ρ are of the form

$$\mathrm{Tr}(\rho B_{kkij}) = \mathrm{Tr}(\rho C_{kkij}) = \mathrm{Tr}(\rho B_{kjik}) = \mathrm{Tr}(\rho C_{kjik}) = 0. \tag{40}$$

Each equation $\mathrm{Tr}(\rho B) = 0$ with $B \neq 0$ or $\mathrm{Tr}(\rho C) = 0$ with $C \neq 0$ is one real valued constraint, but the counting of independent constraints is a little bit involved. We will use these constraints in the following.

IV. RANK-(4,4) EXTREMAL PPT STATES IN 3 × 3 DIMENSIONS

In this section we first review the important method for constructing entangled PPT states introduced by Bennett *et al.* in Ref. [22]. We have previously shown in Ref. [10] how this method may be used to parametrize rank (4,4) entangled PPT states in 3×3 dimensions. We argued there, from numerical evidence, that this parametrization is a complete classification of all rank-4 entangled PPT states. This conjecture has since been proved [24,25].

It is possible to reconstruct a rank-(4,4) entangled PPT state uniquely from five of the six product vectors in its kernel. If these product vectors are orthogonal, then the construction of Bennett *et al.* is directly applicable. If the product vectors are not orthogonal, then one may transform them to orthogonal form in the way described in Ref. [10], before using the Bennett construction. It is a nontrivial empirical fact that the transformation to orthogonal form always works for a set of five product vectors in the kernel of a rank-(4,4) entangled PPT state.

We describe here a more direct way to reconstruct a rank-4 entangled PPT state from the product vectors in its kernel, where we do not transform the product vectors to orthogonal form but use Eq. (40) instead. We give a parametrization of product vectors based on a standard form where they are not orthogonal. This standard form contains four parameters, which are complex in the case of a generic set of product vectors, but we give both numerical and analytical arguments which prove that the parameters have to be real when the product vectors lie in the kernel of a rank-4 entangled PPT state.

When we carry out the construction numerically, we find that with real parameters the generic case is that the constraint equations summarized in Eq. (40) have one solution for ρ, unique up to a proportionality constant. However, this unique ρ is often not a PPT state, because at least one of the conditions $\rho \geqslant 0$ and $\rho^P \geqslant 0$ is violated. These conditions impose further restrictions on the parameters, corresponding exactly to the condition that the product vectors may be transformed to orthogonal form. We describe in detail the subset of real parameter values producing PPT states.

We remark briefly on how a separable state of rank-(4,4) differs from an entangled PPT state of the same rank with respect to product vectors in its kernel.

A. The UPB construction of entangled PPT states

We now review briefly the construction of a rank-(4,4) entangled and extremal PPT state ρ in 3×3 dimensions from an unextendible orthonormal product basis (a UPB) of $\mathrm{Ker}\,\rho$ [22]. The UPB consists of five orthonormal product vectors $w_i = N_i u_i \otimes v_i$ with the property that there exists no product vector orthogonal to all of them. We include real normalization factors N_i here because we want to normalize such that $w_i^\dagger w_j = \delta_{ij}$ without necessarily normalizing the vectors u_i and v_i.

The orthogonality of the product vectors w_i follows from the cyclic orthogonality relations $u_1 \perp u_2 \perp u_3 \perp u_4 \perp u_5 \perp u_1$ and $v_1 \perp v_3 \perp v_5 \perp v_2 \perp v_4 \perp v_1$. There is the further condition that any three vectors u_i and any three v_i are linearly independent. The five-dimensional subspace spanned by these product vectors is the kernel of the density matrix,

$$\rho = \frac{1}{4}\left(\mathbb{1} - \sum_{i=1}^{5} w_i w_i^\dagger\right), \tag{41}$$

which is proportional to a projection operator. The partial transpose of ρ is

$$\rho^P = \frac{1}{4}\left(\mathbb{1} - \sum_{i=1}^{5} \widetilde{w}_i \widetilde{w}_i^\dagger\right), \tag{42}$$

with $\widetilde{w}_i = N_i u_i \otimes v_i^*$. Thus we have both $\rho \geqslant 0$ and $\rho^P \geqslant 0$ by construction. Note that if all the vectors v_i are real, $v_i^* = v_i$, then ρ is symmetric under partial transposition, $\rho^P = \rho$.

By a unitary product transformation as in Eq. (26) we may transform the above orthogonal UPB into the standard unnormalized form [10],

$$u = \begin{pmatrix} 1 & 0 & a & b & 0 \\ 0 & 1 & 0 & 1 & a \\ 0 & 0 & b & -a & 1 \end{pmatrix}, \quad v = \begin{pmatrix} 1 & d & 0 & 0 & c \\ 0 & 1 & 1 & c & 0 \\ 0 & -c & 0 & 1 & d \end{pmatrix}, \tag{43}$$

with a,b,c,d as positive real parameters. The following quantities determine these parameters:

$$\begin{aligned} s_1 &= -\frac{\det(u_1u_2u_4)\,\det(u_1u_3u_5)}{\det(u_1u_2u_5)\,\det(u_1u_3u_4)} = a^2, \\ s_2 &= -\frac{\det(u_1u_2u_3)\,\det(u_2u_4u_5)}{\det(u_1u_2u_4)\,\det(u_2u_3u_5)} = \frac{b^2}{a^2}, \end{aligned} \tag{44}$$

and

$$\begin{aligned} s_3 &= \frac{\det(v_1v_2v_3)\,\det(v_1v_4v_5)}{\det(v_1v_2v_5)\,\det(v_1v_3v_4)} = c^2, \\ s_4 &= \frac{\det(v_1v_3v_5)\,\det(v_2v_3v_4)}{\det(v_1v_2v_3)\,\det(v_3v_4v_5)} = \frac{d^2}{c^2}. \end{aligned} \tag{45}$$

These ratios of determinants are invariant under general $\mathrm{SL} \otimes \mathrm{SL}$ transformations, as well as independent of the normalization of the vectors.

In Ref. [10] we presented numerical evidence that every entangled rank-(4,4) PPT state is SL ⊗ SL equivalent to some state of the form of Eq. (41) with real product vectors as given in Eq. (43). This implies that the surface of all rank-(4,4) entangled PPT states has dimension 36. We count 32 degrees of freedom due to the SL(3,$\mathbb{C}$) ⊗ SL(3,$\mathbb{C}$) transformations, plus the four real SL ⊗ SL invariant parameters a,b,c,d in Eq. (43).

Using the method described in Sec. VII below, we have verified numerically that a rank-(4,4) entangled PPT state as described here lies indeed on a surface of dimension 36. This is another indication that our classification of such states is complete.

B. A different point of view

We present here the construction of an entangled PPT state ρ of rank (4,4), as seen from a different point of view. When ρ has rank 4, it means that Ker ρ has dimension 5. A generic five-dimensional subspace in $\mathbb{C}^9 = \mathbb{C}^3 \otimes \mathbb{C}^3$ has a basis of product vectors. In fact, it contains exactly six product vectors, any five of which are linearly independent. By Eq. (19), five is the limiting dimension for which the number of product vectors is nonzero and finite, and the number six is consistent with Eq. (20).

Any generic set of six product vectors $w_i = u_i \otimes v_i$ may be transformed by an SL ⊗ SL transformation to the standard un-normalized form:

$$u = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & p & p' \\ 0 & 0 & 1 & 1 & q & q' \end{pmatrix}, \qquad v = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & r & r' \\ 0 & 0 & 1 & 1 & s & s' \end{pmatrix}, \tag{46}$$

with p,q,r,s and p',q',r',s' as real or complex parameters. Genericity implies that any three u_i and any three v_i are linearly independent; this requires, for example, that $p,q \neq 0,1$ and $p \neq q$. With this parametrization, one may show that the six product vectors are linearly dependent and lie in a five-dimensional subspace if

$$\begin{aligned} p' &= f_1(p,q,r,s) \equiv \frac{(ps-qr)(1-s)}{(s-r)(q-s)}, \\ q' &= f_2(p,q,r,s) \equiv \frac{(ps-qr)(1-r)}{(s-r)(p-r)}, \\ r' &= f_3(p,q,r,s) \equiv \frac{(ps-qr)(q-1)}{(p-q)(q-s)}, \\ s' &= f_4(p,q,r,s) \equiv \frac{(ps-qr)(p-1)}{(p-q)(p-r)}. \end{aligned} \tag{47}$$

The product vectors w_5 and w_6 are interchangeable; thus the transformation given in Eq. (47) is its own inverse and is equivalent to the transformation

$$\begin{aligned} p &= f_1(p',q',r',s'), \\ q &= f_2(p',q',r',s'), \\ r &= f_3(p',q',r',s'), \\ s &= f_4(p',q',r',s'). \end{aligned} \tag{48}$$

We may choose to ignore w_6 when it is a linear combination of w_1 to w_5.

The values of the above invariants s_1,s_2,s_3,s_4 as functions of the new parameters p,q,r,s are

$$s_1 = -\frac{p}{q}, \quad s_2 = q-1, \quad s_3 = \frac{r-s}{s}, \quad s_4 = \frac{r}{1-r}. \tag{49}$$

The parameters p,q,r,s are actually new invariants; they cannot be changed by SL ⊗ SL transformations.

We want to construct a PPT state having the above product vectors in its kernel, and one possibility is to make a detour and use the UPB construction. If the values of p,q,r,s are such that the invariants s_1,s_2,s_3,s_4 are all real and strictly positive, then we use Eqs. (44) and (45), with u and v from Eq. (46), to find corresponding values of a,b,c,d, and we transform, by an SL ⊗ SL transformation, from the nonorthogonal standard form in Eq. (46) to the orthogonal standard form in Eq. (43). These orthogonal product vectors define the rank-(4,4) PPT state in Eq. (41), which we finally transform back to the desired PPT state.

The condition for the UPB detour to work immediately is that the invariants s_1,s_2,s_3,s_4 are all strictly positive, and we see from Eq. (49) that this happens if and only if p,q,r,s are all real, and $p < 0$, $q > 1$, $0 < r < 1$, $0 < s < r$. These inequalities define the regions marked 1 in the (p,q) and (r,s) planes as plotted in Fig. 1. If they do not hold, the next step we try is to permute the product vectors.

As discussed in Ref. [10] there are ten permutations of the product vectors w_i for $i = 1,2,\ldots,5$ which preserve the positivity of the invariants. These permutations form a group G which is the symmetry group of a regular pentagon, generated by the rotation, or cyclic permutation, $w_i \mapsto w_i'$ with

$$w_1' = w_5, \quad w_2' = w_1, \quad w_3' = w_2, \quad w_4' = w_3, \quad w_5' = w_4, \tag{50}$$

and the reflection

$$w_1' = w_4, \quad w_2' = w_3, \quad w_3' = w_2, \quad w_4' = w_1, \quad w_5' = w_5. \tag{51}$$

For short, we write the rotation as 51234 and the reflection as 43215.

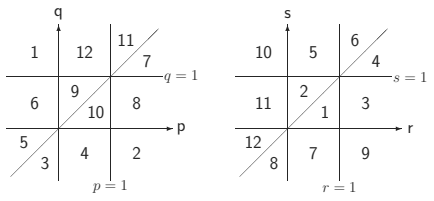


FIG. 1. Regions for the parameters p,q,r,s defined in Eq. (46) such that the product vectors $u_i \otimes v_i$ for $i = 1,2,\ldots,5$ are in the kernel of a rank-(4,4) extremal PPT state. The (p,q) plane is divided into 12 regions, with 12 corresponding regions in the (r,s) plane. The numbers 1–12 refer to the permutations of product vectors given in Eq. (52). For example, if (p,q) is in region 7, (r,s) must also be in region 7.

There are altogether $5! = 120$ permutations of the first five product vectors w_i, and they fall into 12 classes (left cosets of the group G as a subgroup of the permutation group S_5) which are not transformed into each other by G. We number the classes from 1 to 12, and pick one representative from each class as follows:

$$\begin{array}{llllll} 1\colon 1,2,3,4,5 & 2\colon 1,3,2,4,5 & 3\colon 2,1,3,4,5 & 4\colon 2,3,1,4,5 & 5\colon 3,1,2,4,5 & 6\colon 3,2,1,4,5 \\ 7\colon 1,2,4,3,5 & 8\colon 1,4,2,3,5 & 9\colon 2,1,4,3,5 & 10\colon 2,4,1,3,5 & 11\colon 1,3,4,2,5 & 12\colon 1,4,3,2,5 \end{array} \tag{52}$$

Each of these 12 classes defines a positivity region in each of the two parameter planes, where all four invariants s'_1,s'_2,s'_3,s'_4 computed from the permuted product vectors are positive. The 12 regions are disjoint and fill the planes completely, as shown in Fig. 1. On the border lines between the regions some genericity conditions of linear independence are violated.

To summarize, we have learned how to test whether a set of five product vectors $w_i = u_i \otimes v_i$, which is generic in the sense that any three u vectors are linearly independent and any three v vectors are also linearly independent, span the kernel of a rank-(4,4) PPT state. We transform to the standard form defined in Eq. (46), by a product transformation and normalization. Then the necessary and sufficient condition is that the parameters p,q,r,s are all real and that the parameter pairs (p,q) and (r,s) lie in corresponding regions in the parameter planes, as shown in Fig. 1.

It should be stressed that the parameter conditions illustrated in Fig. 1 are derived here by analytical means from the results reported in Ref. [10] but that those results were based partly on numerical evidence. Thus we have no analytical proof which is complete in every detail.

We discuss next a more direct method for constructing the PPT state from the product vectors in its kernel, but first we remark on the origin of the reality condition for the parameters p,q,r,s.

1. Why the parameters have to be real

As pointed out in Sec. III, when ρ is a rank-(4,4) state in dimension 3×3 and

$$w_i = u_i \otimes v_i \in \mathrm{Ker}\,\rho, \quad i = 1,2,\ldots,6, \tag{53}$$

then we must also have

$$\widetilde{w}_i = u_i \otimes v_i^* \in \mathrm{Ker}\,\rho^P, \quad i = 1,2,\ldots,6. \tag{54}$$

But the condition that the partial conjugation $w_i \mapsto \widetilde{w}_i$ preserves the linear dependence between the six product vectors is far from trivial. With product vectors of the form given in Eq. (46), linear dependence holds when the four equations in Eq. (47) [or equivalently Eq. (48)] hold. The corresponding equations for $\widetilde{w}_i$ are the following:

$$\begin{aligned} p' &= f_1(p,q,r^*,s^*), \\ q' &= f_2(p,q,r^*,s^*), \\ (r')^* &= f_3(p,q,r^*,s^*), \\ (s')^* &= f_4(p,q,r^*,s^*). \end{aligned} \tag{55}$$

By eliminating p',q',r',s' between the equations (47) and (55) we derive the equations

$$\begin{aligned} \frac{(ps-qr)(1-s)}{(s-r)(q-s)} &= \frac{(ps^*-qr^*)(1-s^*)}{(s^*-r^*)(q-s^*)}, \\ \frac{(ps-qr)(1-r)}{(s-r)(p-r)} &= \frac{(ps^*-qr^*)(1-r^*)}{(s^*-r^*)(p-r^*)}, \\ \frac{(ps-qr)(q-1)}{(p-q)(q-s)} &= \frac{(p^*s-q^*r)(q^*-1)}{(p^*-q^*)(q^*-s)}, \\ \frac{(ps-qr)(p-1)}{(p-q)(p-r)} &= \frac{(p^*s-q^*r)(p^*-1)}{(p^*-q^*)(p^*-r)}. \end{aligned} \tag{56}$$

The obvious solution is that the parameters p,q,r,s are all real. It takes some analysis to exclude complex parameter values.

If either r or s is complex, or both, then the first two equations in Eq. (56) give that

$$\begin{aligned} p &= q\,\frac{r(1-s)(s^*-r^*)(q-s^*)-r^*(1-s^*)(s-r)(q-s)}{s(1-s)(s^*-r^*)(q-s^*)-s^*(1-s^*)(s-r)(q-s)}, \\ q &= p\,\frac{s(1-r)(s^*-r^*)(p-r^*)-s^*(1-r^*)(s-r)(p-r)}{r(1-r)(s^*-r^*)(p-r^*)-r^*(1-r^*)(s-r)(p-r)}. \end{aligned} \tag{57}$$

Multiplying these equations together and dividing out pq we get an equation giving, for example, p as a function of q,r,s:

$$p = \frac{q\,\mathrm{Im}(-|r|^2 s + rs - r) + \mathrm{Im}(|r|^2 s - |s|^2 r + rs^*)}{q\,\mathrm{Im}(rs^* - r + s) + \mathrm{Im}(-|s|^2 r + rs - s)}. \tag{58}$$

This is one example of extra relations that must hold between the parameter values if they are complex. We see that in the generic case, with no special relations restricting the values of the parameters p,q,r,s, those values have to be real when the six product vectors given in Eq. (46) lie in the kernel of a rank-(4,4) PPT state.

C. Matrix representation relative to a nonorthonormal basis

We proceed next to construct the PPT state directly from nonorthogonal product vectors in its kernel. For our purpose it is useful to represent a Hermitian matrix A in the following way relative to a nonorthonormal basis.

Relative to an orthonormal basis $\{e_i\}$, a matrix A has matrix elements $\widetilde{A}_{ij}$ defined by the two equivalent formulas

$$\text{(I)} \quad \widetilde{A}_{ij} = e_i^\dagger A e_j, \quad \text{(II)} \quad A e_j = \sum_i \widetilde{A}_{ij}\, e_i. \tag{59}$$

If the basis is not orthonormal, then the two definitions are no longer equivalent, and we have to choose one of them. Here we choose definition (I), because it makes the matrix $\widetilde{A}$ Hermitian when A is Hermitian. The scalar products

$$g_{ij} = e_i^\dagger e_j \tag{60}$$

define the metric tensor g as a positive Hermitian matrix. In the usual way, we write the inverse matrix g^{-1} with upper indices so that

$$\sum_j g^{ij} g_{jk} = \delta_k^i. \tag{61}$$

We define the dual vectors

$$e^i = \sum_j g^{ji} e_j, \quad (e^i)^\dagger = \sum_j g^{ij} e_j^\dagger. \tag{62}$$

They satisfy the orthogonality relations

$$(e^i)^\dagger e_j = e_j^\dagger e^i = \delta_j^i \tag{63}$$

and the completeness relations

$$\mathbb{1} = \sum_{i,j} e_i g^{ij} e_j^\dagger = \sum_j e^j e_j^\dagger = \sum_i e_i (e^i)^\dagger. \tag{64}$$

Using these completeness relations and definition (I) above, we may write any matrix A as

$$A = \sum_{i,j} e^i \widetilde{A}_{ij} (e^j)^\dagger. \tag{65}$$

D. Conditions on ρ from product vectors in Ker ρ

It is possible to construct the rank-(4,4) PPT state ρ directly from five product vectors in Ker ρ using the constraints given in Eq. (40), without transforming the product vectors to the orthogonal form. We now describe this construction.

Given three product vectors $w_i = N_i u_i \otimes v_i$ in Ker ρ, with the restriction that all three u_i and all three v_i are linearly independent, we have the following product basis of $\mathbb{C}^9$, not necessarily orthonormal:

$$e_{ij} = u_i \otimes v_j, \quad ij = 11,12,13,21,22,23,31,32,33. \tag{66}$$

With respect to this basis we may define matrix elements of ρ as in Eq. (65):

$$\widetilde{\rho}_{ij;kl} = e_{ij}^\dagger \rho\, e_{kl}. \tag{67}$$

In order to count the independent constraints, it is convenient to use as example the standard form of the product vectors defined in Eq. (46). Now all the constraints from Eq. (40) imply that

$$\widetilde{\rho} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & a_1 & b_1 & 0 & 0 & 0 & 0 & b_2 & 0 \\ 0 & b_1^* & a_2 & 0 & 0 & b_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_3 & 0 & b_4 & b_5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & b_3^* & b_4^* & 0 & a_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & b_5^* & 0 & 0 & a_5 & b_6 & 0 \\ 0 & b_2^* & 0 & 0 & 0 & 0 & b_6^* & a_6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \tag{68}$$

with real diagonal elements $a_1, a_2, \ldots, a_6$ and complex off-diagonal elements $b_1, b_2, \ldots, b_6$. This Hermitian 9×9 matrix contains 18 real parameters, which means that there are altogether $81 - 18 = 63$ independent real constraints.

Including the fourth product vector from Eq. (46) gives additional constraints $\widetilde{\rho} w_4 = 0$, or explicitly written out,

$$\begin{aligned} a_1 + b_1 + b_2 &= 0, \\ b_1^* + a_2 + b_3 &= 0, \\ a_3 + b_4 + b_5 &= 0, \\ b_3^* + b_4^* + a_4 &= 0, \\ b_5^* + a_5 + b_6 &= 0, \\ b_2^* + b_6^* + a_6 &= 0. \end{aligned} \tag{69}$$

These are complex equations, to be split into real and imaginary parts. The real parts are six independent equations, whereas the complex parts are only five independent equations. However, we get another independent equation as the imaginary part of, for example, the complex equation

$$(u_1 \otimes v_4)^\dagger \rho\, (u_4 \otimes v_2) = a_1 + b_1^* + b_2 = 0. \tag{70}$$

Altogether, we get 12 independent real constraints, six from the real parts and six from the imaginary parts of the equations. The end result is that all the off-diagonal matrix elements b_i have to be real, and there are six relations between the 12 real coefficients a_i and b_i, for example,

$$a_1 + a_4 + a_5 = a_2 + a_3 + a_6 = -\sum_{i=1}^{6} b_i. \tag{71}$$

Thus including the fourth product vector in Ker ρ increases the number of independent real constraints from 63 to 75 and reduces the number of real parameters in ρ from 18 to 6.

The final step of including the fifth product vector from Eq. (46) is a calculation which we have done numerically, but not analytically. We find numerically that the generic case with five product vectors and complex parameters p,q,r,s is that there are 81 independent constraints, leaving only the trivial solution $\rho = 0$. In order to end up with one possible solution for ρ, we have to choose the parameters p,q,r,s to be real.

When we choose real values for p,q,r,s, we always find (generically) exactly one solution for ρ, that is, there are 80 independent constraints. The problem is that this uniquely determined matrix ρ, or its partial transpose, has in general both positive and negative eigenvalues.

The condition to ensure that both $\rho \geqslant 0$ and $\rho^P \geqslant 0$ (with the proper choice of sign for ρ), when the parameters p,q,r,s are real, is that the pair (p,q) and the pair (r,s) must lie in corresponding parameter regions, as shown in Fig. 1. We emphasize again that these results are supported by numerical evidence but not proved analytically.

E. Separable states of rank (4,4)

For comparison it may be of interest to consider a separable state of rank 4. It has the form

$$\rho = \sum_{i=1}^{4} \lambda_i \psi_i \psi_i^\dagger, \tag{72}$$

with $\lambda_i > 0$, $\sum_{i=1}^4 \lambda_i = 1$, ${\psi_i}^\dagger \psi_i = 1$, and $\psi_i = C_i\, \phi_i \otimes \chi_i$, with C_i as a normalization constant. In the generic case when any three vectors ϕ_i and any three χ_i are linearly independent, we may perform an SL $\otimes$ SL transformation and obtain the standard form

$$\phi = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \quad \chi = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \tag{73}$$

In this standard form the χ vectors are real, and hence $\rho^P = \rho$.

The kernel Ker ρ consists of the vectors that are orthogonal to all four product vectors ψ_i, and it contains exactly six product vectors $w_i = N_i\, u_i \otimes v_i$ as follows:

$$u = \begin{pmatrix} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & -1 & 0 & 0 & -1 \\ 1 & 0 & 1 & 0 & -1 & 0 \end{pmatrix},$$
$$v = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 & 1 & 0 \\ 0 & -1 & 0 & -1 & 0 & 1 \end{pmatrix}. \tag{74}$$

Note that the product vectors in the kernel of a separable rank-(4,4) PPT state are not generic, in that there are subsets of three linearly dependent vectors both among the u vectors and among the v vectors.

The surface of separable states of rank 4 has dimension $35 = 32 + 3$, where 32 is the number of parameters of the group SL$(3,\mathbb{C}) \otimes$ SL$(3,\mathbb{C})$, and 3 is the number of independent coefficients λ_i in Eq. (72).

V. RANK-(5,5) EXTREMAL PPT STATES IN 3 × 3 DIMENSIONS

Since we believe that we understand completely the rank-(4,4) entangled states in dimension 3×3, a natural next step is to try to understand the (5,5) states in the same dimension. The results presented in this section are to a large extent obtained numerically, but the numerical results lead to an analytical understanding of how to perturb rank-(4,4) extremal PPT states into rank-(5,5) PPT states.

Unfortunately, we do not know any explicit procedure for constructing the most general rank-(5,5) PPT states. However, we have searched numerically and found a large number of such states, and the states we find are typically extremal PPT states [9]. In the present section we show how to use perturbation theory, first in order to study surfaces of rank-(5,5) PPT states, and next in order to construct rank-(5,5) PPT states close to the known rank-(4,4) PPT states. By perturbing a rank-(4,4) PPT state we are able to produce a set of rank-(5,5) extremal PPT states, but this method can never lead to rank (4,5) or (5,4), as we show explicitly.

We have seen above that rank-(4,4) entangled PPT states exist only in special four-dimensional subspaces. The rank-(5,5) extremal and entangled PPT states, on the other hand, are found in generic five-dimensional subspaces. The empirical fact is that there is an eight-dimensional surface of such states in the 24-dimensional space of normalized density matrices on a generic five-dimensional subspace. We find that the surface of all rank-(5,5) extremal and entangled PPT states in the 80-dimensional space of normalized density matrices in 3×3 dimensions has dimension 48. The dimensions 8 and 48 are consistent, since the set of five-dimensional subspaces has dimension 40.

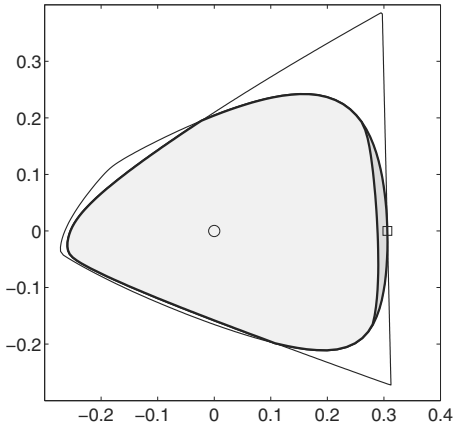


FIG. 2. Two-dimensional section through $\mathcal{D}$, the set of density matrices in 3×3 dimensions. The boundaries of $\mathcal{S}$, the set of separable states, and of $\mathcal{P} = \mathcal{D} \cap \mathcal{D}^P$, the set of PPT states, are both drawn as thick lines. The boundaries of $\mathcal{D}$ and of $\mathcal{D}^P$, the set of partially transposed density matrices, cross at two points. They are drawn as thin lines where they lie outside $\mathcal{P}$. The origin, marked by a small circle, is the maximally mixed state. The point marked by a small square is an extremal rank-(5,5) PPT state. The boundary of $\mathcal{D}$ is drawn through this point as a thin straight line. The boundary of $\mathcal{D}^P$ is drawn thick at this point, because it is also the boundary of $\mathcal{P}$. The lightly shaded region around the origin is $\mathcal{S}$. The small and more darkly shaded region close to the (5,5) state is the difference between $\mathcal{P}$ and $\mathcal{S}$. Away from this small region, the boundaries of $\mathcal{P}$ and $\mathcal{S}$ are indistinguishable in the plot. The coordinates in the plot are defined by the Hilbert-Schmidt metric and are unnamed and dimensionless.

When we perturb from rank (4,4) to rank (5,5), we find that from one rank-(4,4) state there are 12 directions into the surface of (5,5) states. Since the surface of (4,4) states has dimension 36 and the surface of (5,5) states has dimension 48, these numbers are consistent with the hypothesis that it is possible to reach all (5,5) states continuously from the (4,4) states.

In order to illustrate the geometry we have plotted in Fig. 2 a two-dimensional section through the set of density matrices. The section is defined by the maximally mixed state $\mathbb{1}/9$, by a randomly selected rank-(5,5) entangled and extremal PPT state ρ, and by a direction A through ρ such that the perturbed state $\rho' = \rho + \epsilon A$ is a rank-(5,5) PPT state for infinitesimal positive and negative ϵ and has Im $\rho' =$ Im ρ even for finite ϵ. The straight line section plotted through ρ represents a finite range of values of ϵ such that $\rho' \geqslant 0$, rank $\rho' = 5$, and the partial transpose $(\rho')^P$ for finite $\epsilon \neq 0$ has rank 9 and several negative eigenvalues. The line is the intersection of the plotted plane with the set of density matrices on the five-dimensional subspace Im ρ. The figure illustrates the fact that the difference between the sets $\mathcal{P}$ and $\mathcal{S}$ is small, but it is largest close to extremal entangled PPT states.

A. The surface of rank-(5,5) PPT states

As discussed in the previous section, a generic five-dimensional subspace in 3×3 dimensions contains exactly six product vectors which can be transformed by SL $\otimes$ SL transformations, as in Eqs. (26) and (27), into the standard form given in Eqs. (46) and (47), with SL $\otimes$ SL invariant complex parameters p,q,r,s. Thus each such subspace belongs to an equivalence class under SL $\otimes$ SL transformations, and the equivalence classes are parametrized by eight real parameters. There is a discrete ambiguity in the parametrization, since it depends on the ordering of the six product vectors.

In one given generic five-dimensional subspace we may construct a five-dimensional set of rank-(5,5) separable states as convex combinations of the six pure product states in the subspace. However, we find numerically that the dimension of the surface of rank-(5,5) PPT states with the given subspace as image is not five but eight. We determine this dimension numerically in the following way.

As a point of departure we need one rank-(5,5) PPT state ρ_0 in the given five-dimensional subspace, but such a state is easy to find numerically. Let P be the orthogonal projection on this subspace and take X to be a Hermitian matrix; then $\rho_0 = PX^2P/\operatorname{Tr}(PX^2P)$ is a density matrix on the subspace. The remaining problem is to vary X so as to minimize the sum of squares of the four lowest eigenvalues of ρ_0^{P}. Then these four lowest eigenvalues will all become zero so that ρ_0 is a rank-(5,5) PPT state, usually extremal and hence entangled.

At this' stage the projections P and $Q = \mathbb{1} - P$ are already defined, and we proceed to compute the projections $\widetilde{P}$ on $\operatorname{Ker}\rho_0^{P}$ and $\widetilde{Q}$ on $\operatorname{Im}\rho_0^{P}$. We then look for perturbations

$$\rho = \rho_0 + \epsilon A, \tag{75}$$

with $\operatorname{Tr} A = 0$, where A satisfies both equations $\mathbf{P}A = PAP = A$ and $\widetilde{\mathbf{Q}}A = (\widetilde{Q}A^{P}\widetilde{Q})^{P} = 0$, or equivalently Eq. (32),

$$(\mathbf{I} - \mathbf{P} + \widetilde{\mathbf{Q}})A = 0. \tag{76}$$

The number of linearly independent solutions for A with $\operatorname{Tr} A = 0$ is the dimension of the surface of rank-(5,5) PPT states at the point ρ_0. We always find the dimension to be 8.

We guess that the dimension 8 can be understood as follows. If we try to find a PPT state ρ with fixed five-dimensional subspaces $\operatorname{Im}\rho$ and $\operatorname{Im}\rho^{P}$, this means that we fix the projections P and $\widetilde{P}$ and determine ρ as a solution of the equation

$$(\mathbf{P} + \widetilde{\mathbf{P}})\rho = 2\rho, \tag{77}$$

with $\operatorname{Tr}\rho = 1$. Then there is typically no solution at all for ρ; solutions exist only for special pairs of subspaces. If now the two subspaces are chosen in such a way that a solution exists, then the solution is (typically) unique, and the uniqueness means that ρ is an extremal point of $\mathcal{P}$.

We may fix instead $\operatorname{Im}\rho$ but not $\operatorname{Im}\rho^{P}$, only the rank of ρ^{P}. This is the case described by Eq. (76), where the set of solutions has dimension 8. Our guess is that the role of these eight parameters is to specify the SL $\otimes$ SL equivalence class to which the five-dimensional subspace $\operatorname{Im}\rho^{P}$ belongs.

In fact, when we fix $\operatorname{Im}\rho$ there is no degree of freedom left corresponding to SL $\otimes$ SL transformations. This is so because the set of product vectors in $\operatorname{Im}\rho$ is discrete and cannot be transformed continuously by an SL $\otimes$ SL transformation within the fixed subspace $\operatorname{Im}\rho$. Hence the only way to vary the subspace $\operatorname{Im}\rho^{P}$ without varying $\operatorname{Im}\rho$ is to vary the equivalence class of $\operatorname{Im}\rho^{P}$.

If we allow both $\operatorname{Im}\rho$ and $\operatorname{Im}\rho^{P}$ to vary but still require the ranks of ρ and ρ^{P} to be five, the equation to be solved for the perturbation A is

$$(\mathbf{Q} + \widetilde{\mathbf{Q}})A = 0. \tag{78}$$

In this case the number of linearly independent solutions for A is found numerically to be 48, and this is the dimension of the surface of all rank-(5,5) PPT states.

We understand the dimension 48 as follows. There are $8 + 8 = 16$ parameters for the SL $\otimes$ SL equivalence classes of the subspaces $\operatorname{Im}\rho$ and $\operatorname{Im}\rho^{P}$, and there are 32 parameters for the $\mathrm{SL}(3,\mathbb{C}) \otimes \mathrm{SL}(3,\mathbb{C})$ transformations.

To summarize, an extremal and hence entangled rank-(5,5) PPT state ρ is uniquely determined by Eq. (77), as soon as we specify the five-dimensional subspaces $\operatorname{Im}\rho$ and $\operatorname{Im}\rho^{P}$. We may choose arbitrarily one of these subspaces, but not the other one simultaneously. Each five-dimensional subspace is determined by eight real SL $\otimes$ SL invariant parameters and an SL $\otimes$ SL transformation. According to our understanding, which is so far only a plausible hypothesis based on numerical studies, the eight invariant parameters can be chosen independently for $\operatorname{Im}\rho$ and $\operatorname{Im}\rho^{P}$, but the SL $\otimes$ SL transformations cannot be chosen independently.

Thus, there exists an unknown relation between the subspaces $\operatorname{Im}\rho$ and $\operatorname{Im}\rho^{P}$ belonging to a rank (5,5) PPT state ρ. If we had known that relation, we would have been able to construct such states directly.

B. Perturbing from rank (4,4) to rank (5,5)

Since we do not know how to construct the most general rank-(5,5) PPT states, we turn next to the more restricted problem of constructing (5,5) states that are infinitesimally close to (4,4) states.

Consider once more an infinitesimal perturbation $\rho' = \rho + \epsilon A$, this time with ρ as the rank-(4,4) state defined in Eq. (41), involving the standard real product vectors defined in Eq. (43). The most general case is equivalent to this special case by some SL $\otimes$ SL transformation.

An extra bonus of this special choice of ρ is that $\rho^{P} = \rho$. In the notation used above we have projections $P = \widetilde{P}$ on $\operatorname{Im}\rho = \operatorname{Im}\rho^{P}$ and $Q = \widetilde{Q}$ on $\operatorname{Ker}\rho = \operatorname{Ker}\rho^{P}$.

1. Conditions on the perturbation matrix A

By Eq. (9), the condition for ρ' to have rank 5 is that

$$QAQ = \alpha\, w w^{\dagger}, \tag{79}$$

where α is a real number, $\alpha \neq 0$, and

$$w = \sum_{i=1}^{5} c_i w_i, \tag{80}$$

with complex coefficients c_i such that

$$w^\dagger w = \sum_{i=1}^{5} |c_i|^2 = 1. \tag{81}$$

Similarly, the condition for $(\rho')^P$ to have rank 5 is that

$$\widetilde{Q} A^P \widetilde{Q} = Q A^P Q = \beta z z^\dagger, \tag{82}$$

where β is real, $\beta \neq 0$, and

$$z = \sum_{i=1}^{5} d_i w_i \quad \text{with} \quad z^\dagger z = \sum_{i=1}^{5} |d_i|^2 = 1. \tag{83}$$

Note that the possibilities that ρ' has rank either (4,5) or (5,4) are included if we allow either α or β to be zero.

By Eq. (20), there is one extra product vector in $\mathrm{Ker}\,\rho = \mathrm{Ker}\,\rho^P$, which may be written as

$$w_6 = \sum_{i=1}^{5} a_i w_i, \tag{84}$$

this time with real coefficients a_i. Since $w_i = N_i u_i \otimes v_i$ with N_i real and v_i real for $i = 1,2,\ldots,6$, we have for any Hermitian matrix A that

$$w_i^\dagger A w_i = w_i^\dagger A^P w_i. \tag{85}$$

By the definition of the projection Q we have that $Q w_i = w_i$ for $i = 1,2,\ldots,6$. It follows then from Eq. (79) that

$$w_i^\dagger A w_i = w_i^\dagger Q A Q w_i = \alpha\, |w_i^\dagger w|^2 \tag{86}$$

and from Eq. (82) that

$$w_i^\dagger A^P w_i = w_i^\dagger Q A^P Q w_i = \beta\, |w_i^\dagger z|^2. \tag{87}$$

Together with Eq. (85) this gives the equations

$$\alpha\, |c_i|^2 = \beta\, |d_i|^2 \tag{88}$$

for $i = 1,2,\ldots = 5$, and the sixth equation

$$\alpha \left| \sum_{i=1}^{5} a_i c_i \right|^2 = \beta \left| \sum_{i=1}^{5} a_i d_i \right|^2. \tag{89}$$

It follows further that

$$\alpha = \sum_{i=1}^{5} \alpha\, |c_i|^2 = \sum_{i=1}^{5} \beta\, |d_i|^2 = \beta \tag{90}$$

and that

$$|c_i| = |d_i| \quad \text{for} \quad i = 1,2,\ldots,5. \tag{91}$$

Thus the coefficient d_i can differ from c_i only by a phase factor. The total of five phase factors are reduced to four independent phase factors by the extra equation

$$\left| \sum_{i=1}^{5} a_i c_i \right| = \left| \sum_{i=1}^{5} a_i d_i \right|. \tag{92}$$

For infinitesimal values of ϵ, both ρ' and $(\rho')^P$ will have four eigenvalues infinitesimally close to 1/4 and one eigenvalue close to zero, which is $\epsilon\alpha$ for ρ' and $\epsilon\beta$ for $(\rho')^P$. This eigenvalue is the same for ρ' and $(\rho')^P$, since $\alpha = \beta$. With $\alpha > 0$ this means that both $\rho' \geqslant 0$ and $(\rho')^P \geqslant 0$ for $\epsilon > 0$, but not for $\epsilon < 0$. Thus, we get automatically a PPT state of rank (5,5). We never get rank (5,4) or (4,5), and it never happens that ρ' is not a PPT state because one of the ρ' or $(\rho')^P$ has a negative eigenvalue.

For a more general rank-(4,4) state ρ, which is obtained by some SL $\otimes$ SL transformation from a state of the special type discussed here, the smallest positive eigenvalues of ρ' and $(\rho')^P$ are no longer equal. But they are still tied together in such a way that they go to zero simultaneously when we move along the surface of (5,5) states and approach its boundary. The boundary must therefore consist of (4,4) states.

2. *Computing the perturbation A*

Define $W = w w^\dagger$ and $Z = z z^\dagger$ in the same notation as above. These are both projections, $W^2 = W$ and $Z^2 = Z$, with $QW = WQ = W$ and $QZ = ZQ = Z$. It follows from Eqs. (79) and (82) that

$$\begin{aligned} WAW &= WQAQW = \alpha W^3 = \alpha W = QAQ, \\ ZA^PZ &= ZQA^PQZ = \beta Z^3 = \beta Z = QA^PQ. \end{aligned} \tag{93}$$

As in Eqs. (10) and (28), we define

$$\begin{aligned} \mathbf{P}X &= PXP, \quad \mathbf{Q}X = QXQ, \\ \widetilde{\mathbf{P}}X &= (PX^PP)^P, \quad \widetilde{\mathbf{Q}}X = (QX^PQ)^P, \end{aligned} \tag{94}$$

and furthermore,

$$\mathbf{W}X = WXW, \quad \widetilde{\mathbf{Z}}X = (ZX^PZ)^P. \tag{95}$$

We may also define $\mathbf{S} = \mathbf{Q} - \mathbf{W}$ and $\widetilde{\mathbf{S}} = \widetilde{\mathbf{Q}} - \widetilde{\mathbf{Z}}$. These are again orthogonal projections on H_N.

Equation (93) now takes the forms $\mathbf{S}A = 0$ and $\widetilde{\mathbf{S}}A = 0$. These are the least restrictive conditions we may impose on A. Equivalently,

$$(\mathbf{S} + \widetilde{\mathbf{S}})A = 0. \tag{96}$$

To compute A from this equation we introduce an orthonormal basis in the real Hilbert space H_N. Relative to this basis, the operator $\mathbf{S} + \widetilde{\mathbf{S}}$ is represented by a real symmetric positive semidefinite matrix, which has a complete set of real eigenvectors with real eigenvalues. We choose A as an eigenvector of $\mathbf{S} + \widetilde{\mathbf{S}}$ with eigenvalue zero.

Apart from the trivial solution $A = \rho$, which is eliminated when we impose the usual condition that $\mathrm{Tr}\,A = 0$, we find 37 linearly independent solutions of Eq. (96). Out of these 37 solutions, 36 are perturbations that give $\rho' = \rho + \epsilon A$ as a rank-(4,4) state both for $\epsilon > 0$ and $\epsilon < 0$. They do not depend on either vector w or z, since they satisfy the conditions $\mathbf{Q}A = 0$ and $\widetilde{\mathbf{Q}}A = 0$. But because $\mathbf{W} = \mathbf{WQ}$ and $\widetilde{\mathbf{Z}} = \widetilde{\mathbf{Z}}\widetilde{\mathbf{Q}}$, they also satisfy the conditions

$$\mathbf{W}A = \mathbf{WQ}A = 0, \quad \widetilde{\mathbf{Z}}A = \widetilde{\mathbf{Z}}\widetilde{\mathbf{Q}}A = 0, \tag{97}$$

and hence Eq. (96). The number 36 is the dimension of the surface of rank-(4,4) extremal PPT states, as noted in Sec. IV A. The 37th independent solution is the one giving a rank-(5,5) extremal PPT state.

A more restricted class of perturbations consists of those where we fix the five-dimensional subspace $\mathrm{Im}\,\rho'$ to be the

direct sum of the four-dimensional subspace Im ρ and the one-dimensional subspace of the vector w. The projection on Im ρ' is then

$$P_5 = P + W, \tag{98}$$

and the partial condition on A is that $\mathbf{P}_5 A = A$ when we define

$$\mathbf{P}_5 X = P_5 X P_5. \tag{99}$$

The full condition on A, replacing Eq. (96), is that

$$(\mathbf{P}_5 - \widetilde{\mathbf{S}})A = A. \tag{100}$$

Again apart from the trivial solution $A = \rho$, we find five linearly independent solutions of Eq. (100), of which four give ρ' as a rank-(4,4) state both for $\epsilon > 0$ and $\epsilon < 0$. The fifth independent solution is the one giving ρ' as a rank-(5,5) extremal PPT state.

The four directions that give only new (4,4) states are easily identified, since they do not depend on the vector z. To find them we simply repeat the calculation with a "wrong" z, violating the conditions (91) and (92). In this way we find no (5,5) state, but we find the same set of perturbations into (4,4) states. The number 4 is the dimension of the surface of (4,4) states with image within the fixed five-dimensional subspace projected out by the projector P_5.

There is a natural explanation of why this surface has dimension 4. In fact, when we fix P_5 and look for (4,4) states with an image within this fixed five-dimensional subspace, we eliminate all degrees of freedom corresponding to SL $\otimes$ SL transformations. But we still allow variations of the four real SL $\otimes$ SL invariant parameters that are needed to define a rank-(4,4) state.

We conclude that for fixed vectors w and z there is one direction away from the surface of rank-(4,4) extremal PPT states and into the surface of rank-(5,5) extremal PPT states.

For a fixed vector w there is a four-parameter family of acceptable vectors z. Recall that these four parameters determine the five relative phases between the coefficients c_i in Eq. (80) and the corresponding coefficients d_i in Eq. (83).

The vector w is an arbitrary vector in the five-dimensional kernel of the unperturbed state ρ; hence it contains four complex parameters, or eight real parameters, after we take out an uninteresting complex normalization factor. Altogether, there are $8 + 4 = 12$ independent directions away from the 36-dimensional surface of rank-(4,4) PPT states and into the surface of rank-(5,5) PPT states.

When we perturb an arbitrary rank-(5,5) PPT state in such a way that we preserve the ranks of the state and its partial transpose, we find numerically that the surface of rank-(5,5) PPT states has dimension 48. The fact that $48 = 36 + 12$ is consistent with the hypothesis that we can reach every rank-(5,5) PPT state if we start from a rank-(4,4) PPT state and move continuously along the surface of rank-(5,5) PPT states.

VI. RANK-(6,6) EXTREMAL PPT STATES IN 4 × 4 DIMENSIONS

We discuss in some detail one more example of the relation between PPT states and product vectors. According to Eq. (19), the rank-(6,6) PPT states in 4×4 dimensions represent just the limiting case with a finite number of product vectors in the kernel. In this respect they are similar to the rank-(4,4) states in 3×3 dimensions.

The kernel of a rank-6 state in 16 dimensions has dimension 10, and the generic case, according to Eq. (20), is that it contains exactly 20 product vectors, any ten of which are linearly independent. We will see here that, by Eq. (40), the product vectors in the kernel put such strong restrictions on the state that the rank-(6,6) PPT state may be reconstructed uniquely from only seven product vectors in its kernel. As usual, we arrive at this conclusion by doing the construction numerically.

To see how it works, take a set of product vectors in 4×4 dimensions. We may take random product vectors, or else a set of product vectors with the special property that they belong to Ker ρ, where ρ is a rank-(6,6) PPT state. We find numerically that the number of constraints generated by fewer than seven product vectors is the same in both cases. We find the following numbers.

From four product vectors assumed to lie in Ker ρ for an unknown ρ, or actually lying in Ker ρ for a known ρ, we get 172 independent constraints on ρ of the form given in Eq. (40). These constraints leave 84 free real parameters in ρ, before we normalize and set Tr $\rho = 1$.

From five product vectors in Ker ρ we get 205 independent constraints, leaving 51 parameters in ρ.

From six product vectors in Ker ρ we get 234 independent constraints, and 22 parameters in ρ.

Finally, seven product vectors in Ker ρ give either 255 or 256 independent constraints, and either 1 or 0 real parameters in ρ. If there is one parameter left, it is a proportionality constant to be fixed by the normalization condition Tr $\rho = 1$.

The standard form of seven generic product vectors in 4×4 dimensions, generalizing Eq. (46), is the following:

$$u = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & p_1 & p_4 \\ 0 & 0 & 1 & 0 & 1 & p_2 & p_5 \\ 0 & 0 & 0 & 1 & 1 & p_3 & p_6 \end{pmatrix},$$
$$v = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & p_7 & p_{10} \\ 0 & 0 & 1 & 0 & 1 & p_8 & p_{11} \\ 0 & 0 & 0 & 1 & 1 & p_9 & p_{12} \end{pmatrix}. \tag{101}$$

There are 12 complex parameters $p_1, p_2, \ldots, p_{12}$, that is, 24 real parameters. These are invariant in the sense that we cannot change them by SL$(4,\mathbb{C}) \otimes$ SL$(4,\mathbb{C})$ transformations.

Not just any arbitrary set of seven product vectors defines a rank-(6,6) PPT state. We arrive at this conclusion not only because we find numerically that seven generic product vectors allow only $\rho = 0$ as a solution of all the constraint equations, but also because the following dimension counting shows that we need less than 24 invariant parameters in order to parametrize the rank-(6,6) PPT states.

Take one known rank-(6,6) PPT state ρ and perturb it into another rank-(6,6) PPT state $\rho' = \rho + \epsilon A$ with ϵ infinitesimal. Here A must be a solution of Eq. (31), with operators $\mathbf{Q}$ and $\widetilde{\mathbf{Q}}$ defined relative to ρ as explained. The number of linearly independent solutions for A, found numerically, is

76, including the trivial solution $A = \rho$. This shows that the surface of rank-(6,6) PPT states has 75 real dimensions.

Of these 75 dimensions, 60 dimensions result from product transformations $\rho \mapsto V\rho V^\dagger$ with $V = V_A \otimes V_B$ and $V_A, V_B \in \mathrm{SL}(4,\mathbb{C})$. The remaining 15 dimensions must correspond to 15 SL $\otimes$ SL invariant parameters of the seven product vectors. Thus u and v in Eq. (101) contain nine parameters too many.

It is also worth noting that by the counting explained in the next section, the set of six-dimensional subspaces of $\mathbb{C}^{16}$ has real dimension $16^2 - 6^2 - 10^2 = 120$, much larger than the dimension 75 of the surface of (6,6) states. Thus, not every six-dimensional subspace of $\mathbb{C}^{16}$ is the host of a rank-(6,6) PPT state. This conclusion is just as one would expect from the analogy to the case of the rank-(4,4) PPT states in $\mathbb{C}^9$.

VII. DIMENSION COUNTING

We will describe in this section how to compute numerically the dimensions of surfaces of PPT states of given ranks. We list some numerical results and discuss how they may be understood in most cases by a simple counting of constraints, assuming the constraints to be independent. It is only for the lowest rank PPT states, with the largest number of constraints, that the constraints start to overlap.

A useful exercise to start with is to compute the real (as opposed to complex) dimension of the set of all r-dimensional subspaces of an N-dimensional complex Hilbert space.

First note that the unitary group $\mathrm{U}(k)$ has k^2 real dimensions. Take an orthonormal basis of the Hilbert space. The first r basis vectors span an r-dimensional subspace and define also its orthogonal complement spanned by the last $s = N - r$ basis vectors. A $\mathrm{U}(N)$ transformation transforms this basis into another orthonormal basis, but the $\mathrm{U}(r)$ transformations within the first r basis vectors, and the $\mathrm{U}(s)$ transformations within the last s basis vectors, do not change either subspace. It follows that the dimension of the set of r-dimensional subspaces, equal to the dimension of the set of s-dimensional subspaces, is

$$d = N^2 - r^2 - s^2 = 2rs. \tag{102}$$

Assuming that we have found a PPT state ρ of rank (m,n), it lies on a surface of rank (m,n) PPT states. We compute the dimension of the surface at point ρ by counting the number of independent solutions A of Eq. (31),

$$(\mathbf{Q} + \widetilde{\mathbf{Q}})A = 0, \tag{103}$$

equivalent to the two equations

$$\mathbf{Q}A = QAQ = 0, \quad \widetilde{\mathbf{Q}}A = (\widetilde{Q}A^P\widetilde{Q})^P = 0. \tag{104}$$

We have to throw away the trivial solution $A = \rho$, or equivalently, impose the condition $\mathrm{Tr}\,A = 0$. We get a lower bound for the dimension if we assume that the constraints on A from the two equations in Eq. (104) are independent.

The equation $QAQ = 0$ represents $(N-m)^2$ real constraints, since Q is the orthogonal projection on the $N-m$ dimensional subspace $\mathrm{Ker}\,\rho$. Equation (14) illustrates this counting of constraints. Similarly, the equation $\widetilde{Q}A^P\widetilde{Q} = 0$ represents $(N-n)^2$ real constraints, since $\widetilde{Q}$ is the orthogonal projection on the $N-n$ dimensional subspace $\mathrm{Ker}\,\rho^P$. Because the constraints are not necessarily independent, we get the following lower bound for the dimension:

$$d \geqslant N^2 - (N-m)^2 - (N-n)^2 - 1. \tag{105}$$

Take $N = 3 \times 3 = 9$ as an example. We find numerically that Eq. (105) holds with equality for all ranks from the full rank $(m,n) = (9,9)$ down to $(m,n) = (5,5)$. In particular, for rank (5,5) the dimension of the surface is

$$d = 9^2 - 4^2 - 4^2 - 1 = 48. \tag{106}$$

By Eq. (102) the set of five-dimensional subspaces has dimension 40; hence we might expect to find an eight-dimensional surface of rank-(5,5) PPT states in every five-dimensional subspace. And that is actually what we find.

For rank (4,4) the constraints are not all independent, and we have the strict inequality

$$d = 36 > 9^2 - 5^2 - 5^2 - 1 = 30. \tag{107}$$

The set of four-dimensional subspaces has again dimension 40; hence there cannot exist rank-(4,4) PPT states in every four-dimensional subspace. There are $40 - 36 = 4$ constraints restricting the four-dimensional subspaces supporting rank-(4,4) PPT states, and we observe numerically that in one given four-dimensional subspace there can exist at most one unique such state. The four constraints are the conditions that the four parameters a,b,c,d in Eq. (43), or p,q,r,s in Eq. (46), have to be real.

If we want to compute the dimension of the surface of rank-(m,n) PPT states with fixed image space, we have to count the independent solutions of Eq. (32),

$$(\mathbf{I} - \mathbf{P} + \widetilde{\mathbf{Q}})A = 0, \tag{108}$$

equivalent to the two equations

$$\mathbf{P}A = PAP = A, \quad \widetilde{\mathbf{Q}}A = (\widetilde{Q}A^P\widetilde{Q})^P = 0. \tag{109}$$

The equation $\mathbf{P}A = A$ leaves m^2 real parameters in A and represents $N^2 - m^2$ real constraints, as is visualized in Eq. (14). The lower bound on the dimension is therefore

$$d \geqslant m^2 - (N-n)^2 - 1. \tag{110}$$

In the above example with $N = 9$ and $(m,n) = (5,5)$ we find numerically $d = 8$, as already mentioned, so that the inequality in Eq. (110) holds as an equality. With $(m,n) = (4,4)$, on the other hand, we get

$$d = 0 > 4^2 - 5^2 - 1 = -10. \tag{111}$$

VIII. NUMERICAL INTEGRATION

In this section we describe a numerical method for tracing curves on a surface of PPT states of fixed ranks (m,n). This is a tool for studying the geometry of the surface, for example, by tracing geodesics to see how they curve, or studying how the surface approaches a boundary consisting of states of lower ranks.

We use the example of rank-(5,5) PPT states in dimension 3×3 to illustrate how these methods may be applied. The numerical results verify some of our conclusions obtained by

other methods. They also indicate that the geometry of surfaces of extremal PPT states may be rather complicated.

A. Equations of motion

Let $\rho = \rho(t)$ be a curve described by a continuous "time" parameter t and denote the time derivative by $\dot{\rho} = \mathrm{d}\rho/\mathrm{d}t$. The perturbation expansion $\rho(t+\epsilon) = \rho(t) + \epsilon A$ is equivalent to the differential equation

$$\dot{\rho} = A. \tag{112}$$

This becomes a well-defined equation of motion as soon as we define the right-hand side as a function of ρ and t, $A = A(\rho,t)$.

The most general case considered here is a curve lying on a surface of PPT states of fixed ranks (m,n). Then Eq. (31) is the necessary and sufficient condition that A has to satisfy:

$$(\mathbf{Q} + \widetilde{\mathbf{Q}})A = 0. \tag{113}$$

The projection operators $\mathbf{Q}$ and $\widetilde{\mathbf{Q}}$ on H_N are defined relative to ρ in exactly the same way as before. They will obviously become t dependent when ρ becomes t dependent.

We consider also the more restricted case when not only the ranks of ρ and ρ^P are fixed, but $\mathrm{Im}\,\rho$ is fixed as well. The stronger necessary and sufficient condition on A is then Eq. (32):

$$(\mathbf{I} - \mathbf{P} + \widetilde{\mathbf{Q}})A = 0. \tag{114}$$

In this equation $\mathbf{I}$ and $\mathbf{P}$ are t independent, and only $\widetilde{\mathbf{Q}}$ is t dependent.

The conditions (113) and (114) are both of the form

$$\mathbf{T}A = 0, \tag{115}$$

with either $\mathbf{T} = \mathbf{Q} + \widetilde{\mathbf{Q}}$ or $\mathbf{T} = \mathbf{I} - \mathbf{P} + \widetilde{\mathbf{Q}}$. In either case, $\mathbf{T}$ is a positive real symmetric operator with respect to the natural scalar product on H_N. The effect of Eq. (115) is only to restrict the curve to lie on the given surface, and this still leaves many degrees of freedom available for moving around on the surface. Let us mention three examples of how to generate a curve.

One possibility is a random walk on the surface. Then at each point on the curve we choose the direction A more or less randomly within the kernel of the operator $\mathbf{T}$ at that point.

A second possibility is to trace out a geodesic curve on the surface. Remember that our surface of PPT states is embedded in the Euclidean space H_N. By definition, a geodesic on an embedded surface (a simple example is a great circle on the two-dimensional surface of a three-dimensional sphere) is a curve which does not change its length and direction on the surface. Hence it changes direction in the embedding space only as much as it has to in order to stay on the surface. In order to generate a geodesic we choose a starting point ρ_0 and a starting direction A_0 at ρ_0, and as we move along the curve we keep the length of the tangent vector A constant, while changing its direction in the embedding space minimally.

A third possibility is to take a differentiable real function $f = f(\rho)$ and choose A at the point ρ as the (positive or negative) gradient of f at ρ. In this way we trace out flow lines of that function on the surface and find local maxima or minima of f. If we know that the maximum or minimum is on the boundary of our surface, we may use this as a method for locating the boundary.

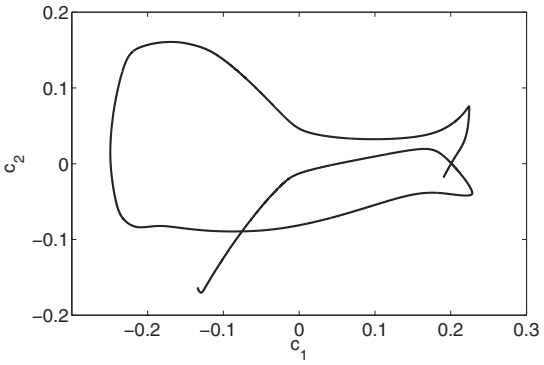


FIG. 3. A projection of a geodesic curve on the eight-dimensional surface of rank-(5,5) PPT states with a fixed image space. We have made a principal component analysis and plotted the two largest principal components c_1 and c_2. The curve starts middle right and ends lower left.

B. Numerical examples

We have implemented numerically the two methods for generating geodesics and for generating flow lines, using a classical fourth-order Runge-Kutta integration method. With ρ and A of order one and time steps of order 10^{-4}, this gives a precision of order 10^{-16}, which is the machine precision.

Figure 3 shows a two-dimensional projection of a geodesic curve $\rho = \rho(t)$ on the eight-dimensional curved surface of rank-(5,5) PPT states with $\mathrm{Im}\,\rho$ constant. The Hilbert space dimension is 3×3. The two dimensions plotted are the two most important ones as given by a principal component analysis of the full curve. The apparent kinks in the curve are only effects of the projection and do not reflect sharp bends of the surface. The most important conclusion to be drawn from this figure is that the geometry of the surface is not a simple one.

Figure 4 shows the five nonzero eigenvalues of ρ and ρ^P along the path of Fig. 3. The condition that one eigenvalue of either ρ or ρ^P goes to zero defines the boundary of the surface. We see that the curve approaches the boundary twice but turns around each time and continues in the interior. The eigenvalue spectra of ρ and ρ^P are remarkably similar, yet they are not identical. The fact that the smallest eigenvalues approach zero collectively may be understood from the conclusions of Sec. V B. When both ρ and ρ^P simultaneously get one dominant eigenvalue, we interpret it as an indication that ρ approaches a pure product state.

It is quite natural that a geodesic chosen at random will not hit the boundary, since the boundary consists of rank-(4,4) PPT states and has dimension 4, while the surface itself has dimension 8. On the other hand, the geometry of these surfaces certainly stretches our imagination. We imagine that the two-dimensional surface of a rain drop may be a fair analogy—it has a zero-dimensional boundary, a tip which is a single point, and there is zero probability of hitting the tip if one

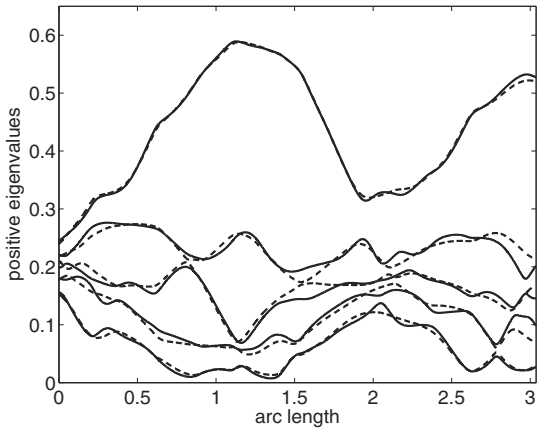


FIG. 4. Variation along the curve in Fig. 3 of the five nonzero eigenvalues of the density matrix (fully drawn lines) and its partial transpose (broken lines). The abscissa is the arc length along the curve.

moves along a geodesic starting at a random point in a random direction.

If we want to hit the boundary we cannot follow a geodesic. We have to integrate the equation $\dot{\rho} = A$ and choose the direction A in such a way that the smallest positive eigenvalue of either ρ or ρ^P goes to zero. Thus we define a function $f = f(\rho)$ as the smallest one of the ten nonzero eigenvalues of ρ and ρ^P. This function is differentiable almost everywhere.

Defining A as the projection of the negative gradient of f onto the kernel of $\mathbf{T}$, we trace paths from random (5,5) states to the boundary of the surface of (5,5) states. In accordance with the results of Sec. V B, we observe that the smallest nonzero eigenvalue of ρ and ρ^P approach zero collectively, so that we only find (4,4) states as boundary points and never (4,5) or (5,4) states. With adaptive step lengths in the Runge-Kutta method we obtain machine precision of eigenvalues and are therefore confident that we find true (4,4) states.

IX. SUMMARY AND OUTLOOK

The work presented here is part of an ongoing program to study quantum entanglement in mixed states. We have studied low-rank entangled PPT states, especially extremal PPT states, using a combination of analytical and numerical techniques. In particular, we have used perturbation theory and exploited the close connection between low-rank PPT states and product vectors. The concept of SL $\otimes$ SL equivalence, or SL equivalence for short, is central in the classification of PPT states. One use for perturbation theory is in the numerical calculation of the dimensions of various surfaces of PPT states. Another use is in more detailed studies of the geometry of such a surface by the tracing of curves on the surface.

We have revisited the rank-(4,4) extremal PPT states in 3×3 dimensions. A state of this kind has six product vectors in its kernel, and as we have previously discussed it is always related by SL $\otimes$ SL equivalence to the special class of rank-(4,4) extremal PPT states constructed by Bennett *et al.* from unextendible sets of orthogonal product vectors. We have demonstrated here a more direct way of constructing the state from the product vectors in its kernel without transforming them to orthogonal form. The transformation to orthogonal form works only for the rank-(4,4) states in 3×3 dimensions, whereas our direct method works just as well for similar cases in higher dimensions.

We find numerically that the surface of rank-(4,4) extremal PPT states in dimension 3×3 has dimension 36. Since the group $\mathrm{SL}(3,\mathbb{C}) \otimes \mathrm{SL}(3,\mathbb{C})$ has 32 real parameters, this confirms the known result that four real parameters are needed to specify an equivalence class of such states under SL $\otimes$ SL equivalence.

To a rank-(4,4) state ρ in 3×3 dimensions correspond four-dimensional subspaces $\mathrm{Im}\,\rho$ and $\mathrm{Im}\,\rho^P$ and five-dimensional subspaces $\mathrm{Ker}\,\rho = (\mathrm{Im}\,\rho)^\perp$ and $\mathrm{Ker}\,\rho^P = (\mathrm{Im}\,\rho^P)^\perp$. The product vectors in $\mathrm{Ker}\,\rho$ are related to those in $\mathrm{Ker}\,\rho^P$ by a complex conjugation of the second factor in the direct product; we may call this operation a partial conjugation. This direct relation between $\mathrm{Ker}\,\rho$ and $\mathrm{Ker}\,\rho^P$, not involving ρ explicitly, means that all the four subspaces $\mathrm{Im}\,\rho$, $\mathrm{Im}\,\rho^P$, $\mathrm{Ker}\,\rho$, and $\mathrm{Ker}\,\rho^P$ are uniquely given as soon as one of them is given.

The set of four-dimensional subspaces and the set of five-dimensional subspaces both have dimension 40. Thus, a four-dimensional subspace has to satisfy $40 - 36 = 4$ constraints in order to support a rank-(4,4) extremal PPT state. There is never more than one unique rank (4,4) extremal PPT state on any given four-dimensional subspace. The four constraints may be understood as certain reality conditions on the six product vectors in $\mathrm{Ker}\,\rho$, or equivalently in $\mathrm{Ker}\,\rho^P$. They arise because the partial conjugation of product vectors, relating $\mathrm{Ker}\,\rho$ and $\mathrm{Ker}\,\rho^P$, must preserve the linear dependence between six vectors in a five-dimensional subspace.

We have not been able to understand and classify the rank-(5,5) extremal PPT states in 3×3 dimensions as completely as the rank-(4,4) states. However, we have counted the number of real parameters needed to describe the rank-(5,5) states. We find numerically a smooth 48-dimensional surface of such states on the full nine-dimensional Hilbert space. On every generic five-dimensional subspace we find an eight-dimensional surface of such states. The dimension 8 is easily understood, keeping in mind that the set of five-dimensional subspaces has dimension 40.

The 48-parameter set of rank-(5,5) extremal PPT states splits into SL $\otimes$ SL equivalence classes described by $48 - 32 = 16$ parameters. Similarly, the 40-parameter set of five-dimensional subspaces splits into equivalence classes described by $40 - 32 = 8$ parameters. It is a natural guess that the 16 parameters needed to describe the SL $\otimes$ SL equivalence class of a rank-(5,5) PPT state ρ are precisely the eight parameters describing the equivalence class of $\mathrm{Im}\,\rho$ plus the eight parameters describing the equivalence class of $\mathrm{Im}\,\rho^P$. In other words, we guess that the equivalence class of $\mathrm{Im}\,\rho$ and that of $\mathrm{Im}\,\rho^P$ may vary independently.

Failing in our attempt to solve the general problem of constructing rank-(5,5) extremal PPT states in 3×3 dimensions, we have solved the more restricted problem of constructing such states that are infinitesimally close to the known rank-(4,4) extremal PPT states.

The set of six product vectors in the kernel of a rank-(4,4) state ρ plays an important role in imposing constraints on a perturbation A to ensure that $\rho' = \rho + \epsilon A$ is a rank-(5,5) PPT state for ϵ infinitesimal and positive. The constraints admit solutions for A such that ρ' has rank (4,4), independent of the sign of ϵ, then A is a tangent to the 36-dimensional surface of (4,4) states. In addition, there is a 12-parameter nonlinear set of solutions for A such that ρ' has rank (5,5) and such that A is a direction orthogonal to the surface of (4,4) states. The general solution for A is a linear combination of a direction along the surface of rank-(4,4) states and a direction orthogonal to this surface pointing into the surface of (5,5) states.

Thus our construction by perturbation accounts completely for the dimension $48 = 36 + 12$ of the surface of (5,5) states. It is rather remarkable that the construction is guaranteed to produce PPT states of ranks (5,5), because when the perturbed state ρ' gets an infinitesimal strictly positive eigenvalue, there is a corresponding infinitesimal eigenvalue of $(\rho')^P$ which is also strictly positive, never zero or negative.

In dimension 3×3, the rank-(4,4) PPT states are set apart by the special property that the kernel of a state is always spanned by product vectors and contains a finite number of product vectors. This property is shared more generally by PPT states of rank (n,n) in dimension $N_A \times N_B$ with $n = N_A + N_B - 2$.

As another example from this interesting general class of PPT states we have considered the rank-(6,6) extremal PPT states in dimension 4×4. In this case the kernel of a state has dimension 10 and contains 20 product vectors. Our method for reconstructing a PPT state from the product vectors in its kernel works very well here again. In our numerical examples we find that the state can be reconstructed uniquely from an arbitrary subset of only seven product vectors in its kernel.

One remarkable feature of this result is that seven product vectors uniquely determine a ten-dimensional subspace. It is clear that a six-dimensional subspace in dimension 4×4 has to satisfy some set of very strict conditions in order to support a rank-(6,6) extremal PPT state. An obvious nontrivial set of conditions is that the partial conjugation of the 20 product vectors in its ten-dimensional orthogonal complement must lie in a ten-dimensional subspace. How to handle these constraints is one of the many open problems left for future research.

In conclusion, we have made some small progress in our efforts to understand the low-rank entangled PPT states, but much remains to be done. How to construct general rank-(5,5) PPT states that are not close to rank-(4,4) states is still an unsolved problem, and we are even further from a full understanding of higher rank extremal PPT states in 3×3 dimensions, or in higher dimensions.

ACKNOWLEDGMENTS

We acknowledge gratefully research grants from The Norwegian University of Science and Technology (L.O.H.) and from The Norwegian Research Council (P.Ø.S.). L.O.H., J.M., and P.Ø.S. also want to thank NORDITA for hospitality during the workshop and conference on quantum information in Stockholm. We thank the referee for useful comments regarding the first version of our manuscript.

[1] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[2] J. S. Bell, Physics **1**(3), 195 (1964).
[3] L. Masanes, Y.-C. Liang, and A. C. Doherty, Phys. Rev. Lett. **100**, 090403 (2008).
[4] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory, and Conceptions of the Universe*, edited by M. Kafatos (Kluwer, Dordrecht, The Netherlands, 1989), p. 69; reproduced as e-print arXiv:0712.0921v1 (2007).
[5] N. D. Mermin, Phys. Today **43**, 9 (1990).
[6] L. Gurvits, in *Proceedings of the Thirty-Fifth ACM Symposium on Theory of Computing* (ACM, New York, 2003), pp. 10–19.
[7] J. M. Leinaas, J. Myrheim, and E. Ovrum, Phys. Rev. A **74**, 012313 (2006).
[8] J. M. Leinaas, J. Myrheim, and E. Ovrum, Phys. Rev. A **76**, 034304 (2007).
[9] J. M. Leinaas, J. Myrheim, and P. Ø. Sollid, Phys. Rev. A **81**, 062329 (2010).
[10] J. M. Leinaas, J. Myrheim, and P. Ø. Sollid, Phys. Rev. A **81**, 062330 (2010).
[11] A. Peres, Phys. Rev. Lett. **77**, 1413 (1996).
[12] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Lett. A **223**, 1 (1996).
[13] P. Horodecki, Phys. Lett. A **232**, 333 (1997).
[14] C. H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J. A. Smolin, and W. K. Wootters, Phys. Rev. Lett. **76**, 722 (1996); **78**, 2031(E) (1996).
[15] M. Horodecki, P. Horodecki, and R. Horodecki, Phys. Rev. Lett. **80**, 5239 (1998).
[16] R. Horodecki, Europhys. News **41/6**, 21 (2010).
[17] E. Amselem and M. Bourennane, Nat. Phys. **5**, 748 (2009).
[18] H. Kampermann, D. Bruss, X. Peng, and D. Suter, Phys. Rev. A **81**, 040304(R) (2010).
[19] J. Lavoie, R. Kaltenbaek, M. Piani, and K. J. Resch, Phys. Rev. Lett. **105**, 130501 (2010).
[20] J. T. Barreiro, P. Schindler, O. Gühne, T. Monz, M. Chwalla, C. F. Roos, M. Hennrich, and R. Blatt, Nat. Phys. **6**, 943 (2010).
[21] P. Horodecki, M. Lewenstein, G. Vidal, and I. Cirac, Phys. Rev. A **62**, 032310 (2000).
[22] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, Phys. Rev. Lett. **82**, 5385 (1999).
[23] D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, Commun. Math. Phys. **238**, 379 (2003).
[24] Ł. Skowronek, J. Math. Phys. **52**, 122202 (2011).
[25] L. Chen and D. Z. Djokovic, J. Math. Phys. **52**, 122203 (2011).
[26] P. Ø. Sollid, J. M. Leinaas, and J. Myrheim, Phys. Rev. A **84**, 042325 (2011).

Paper II

L.O. Hansen, A. Hauge, J. Myrheim, and P.Ø. Sollid

"Extremal entanglement witnesses"

Preprint of paper published in Int. J. Quantum Inf. **13**, 1550060 (2015)

Preprint of paper published in
International Journal of Quantum Information
Vol. 13, No. 8 (2015) 1550060
DOI: 10.1142/S0219749915500604

http://www.worldscientific.com/worldscinet/ijqi

Extremal entanglement witnesses

Leif Ove Hansen, Andreas Hauge† and Jan Myrheim
Department of Physics, Norwegian University of Science and Technology, N–7491 Trondheim, Norway

Per Øyvind Sollid
Department of Physics, University of Oslo, N–0316 Oslo, Norway

April 29th, 2016

We present a study of extremal entanglement witnesses on a bipartite composite quantum system. We define the cone of witnesses as the dual of the set of separable density matrices, thus $\mathrm{Tr}\,\Omega\rho \geq 0$ when Ω is a witness and ρ is a pure product state, $\rho = \psi\psi^\dagger$ with $\psi = \phi \otimes \chi$. The set of witnesses of unit trace is a compact convex set, uniquely defined by its extremal points. The expectation value $f(\phi, \chi) = \mathrm{Tr}\,\Omega\rho$ as a function of vectors ϕ and χ is a positive semidefinite biquadratic form. Every zero of $f(\phi, \chi)$ imposes strong real-linear constraints on f and Ω. The real and symmetric Hessian matrix at the zero must be positive semidefinite. Its eigenvectors with zero eigenvalue, if such exist, we call Hessian zeros. A zero of $f(\phi, \chi)$ is quadratic if it has no Hessian zeros, otherwise it is quartic. We call a witness quadratic if it has only quadratic zeros, and quartic if it has at least one quartic zero. A main result we prove is that a witness is extremal if and only if no other witness has the same, or a larger, set of zeros and Hessian zeros. A quadratic extremal witness has a minimum number of isolated zeros depending on dimensions. If a witness is not extremal, then the constraints defined by its zeros and Hessian zeros determine all directions in which we may search for witnesses having more zeros or Hessian zeros. A finite number of iterated searches in random directions, by numerical methods, lead to an extremal witness which is nearly always quadratic and has the minimum number of zeros. We discuss briefly some topics related to extremal witnesses, in particular the relation between the facial structures of the dual sets of witnesses and separable states. We discuss the relation between extremality and optimality of witnesses, and a conjecture of separability of the so called structural physical approximation (SPA) of an optimal witness. Finally, we discuss how to treat the entanglement witnesses on a complex Hilbert space as a subset of the witnesses on a real Hilbert space.

Keywords: Entanglement witnesses; positive maps; convex sets.

1. Introduction

Entanglement is the quintessence of non-classicality in composite quantum systems. As a physical resource it finds many applications especially in quantum information theory, see

†Deceased 31 May 2015

Refs. 1, 2 and references therein. Entangled states are exactly those states that can not be modelled within the classical paradigm of locality and realism.[3-6] Contrary to states of classical composite systems, they allow better knowledge of the system as a whole than of each component of the system.[7,8] The notion of entanglement is therefore of interest within areas of both application and interpretation of quantum mechanics.

The problem of distinguishing entangled states from non-entangled (separable) states is a fundamental issue. Pure separable states of a bipartite system are exactly the pure product states, they are easily recognized by singular value decomposition, also known as Schmidt decomposition.[2] The situation is much more complicated for mixed states. A separable mixed state is a state mixed from pure product states.[5]

The separability problem of distinguishing mixed separable states from mixed entangled states has received attention along two lines. On the one hand, operational or computational tests for separability are constructed.[9] On the other hand, the geometry of the problem is studied, with the aim of clarifying the structures defining separability. [10-18] Two results, one from each approach, deserve special attention. An operationally simple necessary condition for separability was given by Peres.[9] He pointed out that the partial transpose of a separable state is again a separable state, therefore we know that a state is entangled if its partial transpose is not positive semidefinite. States with positive partial transpose, so called PPT states, are interesting in their own right.[19] A significant geometrical result is that of Życzkowski *et al.* that the volume of separable states is non-zero.[12] They prove the existence of a ball of separable states surrounding the maximally mixed state. The maximal radius of this ball was provided by Gurvits and Barnum.[13]

The well known necessary and sufficient separability condition in terms of positive linear maps or entanglement witnesses connects these two lines of thought.[11,20] There exists a set of linear maps on states, or equivalently a set of observables called entanglement witnesses, that may detect, or witness, entanglement of a state. This result is the point of departure for almost all research into the problem, be it computational, operational or geometrical. The major obstacle in the application of positive linear maps and entanglement witnesses is that they are difficult to identify, in principle as difficult as are the entangled states themselves.

Our contribution falls in the line of geometrical studies of entanglement. The ultimate goal is to classify the extremal points of the convex set of entanglement witnesses, the building blocks of the set, in a manner useful for purposes in quantum mechanics. Positive linear maps, and the extremal positive maps, were studied already in 1963 by Størmer in the context of partially ordered vector spaces and C^* algebras.[14,15] Since then the extremal positive maps and extremal witnesses have only received scattered attention.[21-28]

Here we combine and apply basic notions from convex geometry and optimization theory in order to study the extremal entanglement witnesses. We obtain computationally useful necessary and sufficient conditions for extremality. We use these ideas to analyse previously known examples and construct new numerical examples of extremal witnesses. The extremal witnesses found in random searches are generic, by definition. It turns out that the previously known examples of extremal witnesses are very far from being generic, and the generic extremal witness is of a type not yet published in the literature.

We discuss what we learn about the geometry of the set of witnesses, and also the set of separable states, by studying extremal witnesses. Other topics we comment on are the

relation between extremal and optimal witnesses,[26,29] and the so called structural physical approximation (SPA) of witnesses relating them to physical maps of states.[30-33]

Numerical methods play a central role in our work. An important reason is that the entanglement witnesses we want to study are so complicated that there is little hope of treating them analytically. Numerical work is useful for illustrating the theory, examining questions of interest and guiding our thoughts in questions we cannot answer rigorously.

The methods we use for studying extremal witnesses are essentially the same as we have used previously for studying extremal states with positive partial transpose, *i.e.* PPT states.[18,34-37] The work presented here is also based in part on the master's thesis of one of us.[38]

Outline of the article

This article consists of three main parts. Section 2 is the first part, where we review material necessary for the appreciation of later sections. The basic concepts of convex geometry are indispensable. We review the geometry of entanglement witnesses, and formulate the study of extremal witnesses as an optimization problem where we represent the witness as a positive semidefinite biquadratic form and study its zeros.

The second part consists of Sec. 3 and Sec. 4, where we develop a necessary and sufficient condition for the extremality of a witness. The basic idea is that every zero of a positive biquadratic form representing a witness imposes strong linear constraints on the form. The extremality condition in terms of the zeros of a witness and the associated constraints is our most important result, not only because it gives theoretical insight, but because it is directly useful for numerical computations.

In the third, last, and most voluminous part, the Secs. 5 to 11, we apply the extremality condition from Sec. 4. As a first application we study decomposable witnesses in Sec. 5. These are well understood and not very useful as witnesses, since they can only detect the entanglement of a state in the trivial case when its partial transpose is not positive. But they serve to illustrate the concepts, and we find them useful as stepping stones towards non-decomposable witnesses.

In Sec. 6 we study examples of extremal non-decomposable witnesses. First we study two examples known from the literature, and confirm their extremality by our numerical method. Then we construct numerically random examples of extremal witnesses in order to learn about their properties. We find that the generic extremal witnesses constructed numerically belong almost without exception to a completely new class, not previously noticed in the literature. They have a fixed number of isolated quadratic zeros, whereas the previously published extremal witnesses have continuous sets of zeros. The number of zeros depends on the dimensions of the Hilbert spaces of the two subsystems.

In Sec. 7 we show examples of a special D-shaped type of faces of the set of witnesses in the lowest non-trivial dimensions, 2×4 and 3×3. These faces are "next to extremal", bordered by extremal witnesses plus a straight section of decomposable witnesses.

In Sec. 8 we study faces of the set of separable states, using the duality between separable states and witnesses. We classify two different families of faces, generalizing results of Alfsen and Shultz,[16] and present some statistics on randomly generated maximal simplex faces. We point out that the facial structure is relevant for the question of how many pure product states are needed in the convex decomposition of an arbitrary separable

state, and suggest that it may be possible to improve substantially the trivial bound given by the dimension of the set.[39,40]

In Sec. 9 we compare the notions of optimal and extremal witnesses.[26,29] In low dimensions we find that almost every non-decomposable optimal witness is extremal, whereas an abundance of non-decomposable and non-extremal optimal witnesses exist in higher dimensions.

In Sec. 10 we comment on a separability conjecture regarding structural physical approximations (SPAs) of optimal witnesses.[31,32] This conjecture has since been refuted.[33,41] We have tried to test numerically a modified separability conjecture. It holds within the numerical precision of our separability test, but this test with the available precision is clearly not a definitive proof.

In Sec. 11 we point out that it is possible, and maybe even natural from a certain point of view, to treat the entanglement witnesses on a complex tensor product space as a subset of the witnesses on a real tensor product space.

In Sec. 12 we summarize our work and suggest some possible directions for future efforts. Some details regarding numerical methods are found in the Appendix.

2. Background material

In this section we review material necessary as background for the following sections. Convex geometry is the mathematical basis of the theory of mixed quantum states, and is equally basic in the entanglement theory for mixed states. We introduce the concepts of dual convex cones, entanglement witnesses, positive maps, and biquadratic forms. See *e.g.* introductory sections of Refs. 10, 42 for further details on convexity. Concepts from optimization theory are useful in the numerical treatment of a special minimization problem, see Refs. 42, 43.

2.1. *Convexity*

The basic concepts of convex geometry are useful or even essential for describing mixed quantum states as probabilistic mixtures of pure quantum states. A convex subset of a real affine space is defined by the property that any convex combination

$$x = (1-p)x_1 + px_2 \qquad \text{with} \qquad 0 \le p \le 1 \tag{1}$$

of members x_1, x_2 is a member. If $x_1 \neq x_2$ and $0 < p < 1$ we say that x is a proper convex combination of x_1, x_2, and x is an interior point of the line segment between x_1 and x_2. The dimension of a convex set is the dimension of the smallest affine space containing it. We will be dealing here only with finite dimensional sets. Closed and bounded subsets of finite dimensional Euclidean spaces are compact, according to the Heine–Borel theorem. A compact convex set has extremal points that are not convex combinations of other points. It is completely described by its extremal points, since any point in the set can be decomposed as a convex combination involving no more than $n+1$ extremal points, where n is the dimension of the set.[40] It is called a polytope if it has a finite number of extremal points, and a simplex if the number of extremal points is one more than the dimension.

Definition 1. *A bidirection at x in the convex set $\mathcal{K}$ is a direction vector $v \neq 0$ such that $x + tv \in \mathcal{K}$ for t in some interval $[t_1, t_2]$ with $t_1 < 0 < t_2$.*

A face $\mathcal{F}$ of $\mathcal{K}$ is a convex subset of $\mathcal{K}$ with the property that if $x \in \mathcal{F}$ then every bidirection in $\mathcal{K}$ at x is a bidirection in $\mathcal{F}$ at x. An equivalent condition defining $\mathcal{F}$ as a face is that if $x \in \mathcal{F}$ is a proper convex combination of $x_1, x_2 \in \mathcal{K}$, then $x_1, x_2 \in \mathcal{F}$.

Intuitively, one might say that a set $\mathcal{F}$ is a face of $\mathcal{K}$ if it is "self reliant" up to convex combinations.

The empty set and $\mathcal{K}$ itself are faces of $\mathcal{K}$, by definition. All other faces are called proper faces. The extremal points of $\mathcal{K}$ are the zero dimensional faces of $\mathcal{K}$. A face of dimension $n-1$ where n is the dimension of $\mathcal{K}$ is called a facet.

A point x in a convex set $\mathcal{K}$ is an interior point of $\mathcal{K}$ if every direction at x (in the minimal affine space containing $\mathcal{K}$) is a bidirection, otherwise x is a boundary point of $\mathcal{K}$. Every point $x \in \mathcal{K}$ is either an extremal point of $\mathcal{K}$ or an interior point of a unique face $\mathcal{F}_x$ of dimension one or higher. This face is the intersection of $\mathcal{K}$ with the affine space

$$\mathcal{A}_x = \{\, x + tv \mid t \in \mathbb{R}\,,\ v \in \mathcal{B}_x \,\} \tag{2}$$

where $\mathcal{B}_x$ is the set of all bidirections in $\mathcal{K}$ at x. Thus, every face $\mathcal{F}$ of $\mathcal{K}$ is an intersection $\mathcal{F} = \mathcal{A} \cap \mathcal{K}$ of $\mathcal{K}$ with some affine space $\mathcal{A}$. We take $\mathcal{A}$ to be a subspace of the minimal affine space containing $\mathcal{K}$. The minimum dimension of $\mathcal{A}$ is the dimension of $\mathcal{F}$, but it may also be possible to choose $\mathcal{A}$ as an affine space of higher dimension than $\mathcal{F}$. A proper face $\mathcal{F} = \mathcal{A} \cap \mathcal{K}$ is said to be exposed if $\mathcal{A}$ has dimension $n-1$ where n is the dimension of $\mathcal{K}$. Note that an exposed face may have dimension less than $n-1$, thus it is not necessarily a facet, and it may be just a point. Every exposed point is extremal, since it is a zero dimensional face, but a convex set may have extremal points that are not exposed.

A proper face of $\mathcal{K}$ is part of the boundary of $\mathcal{K}$. The faces of a face $\mathcal{F}$ of $\mathcal{K}$ are the faces of $\mathcal{K}$ contained in $\mathcal{F}$. In particular, the following result is useful for understanding the face structure of a convex set when its extremal points are known.

Theorem 1. *Let $\mathcal{F}$ be a face of the convex set $\mathcal{K}$. Then a point in $\mathcal{F}$ is an extremal point of $\mathcal{F}$ if and only if it is an extremal point of $\mathcal{K}$.*

We state here another useful result which follows directly from the definitions.

Theorem 2. *An intersection $\mathcal{K} = \mathcal{K}_1 \cap \mathcal{K}_2$ of two convex sets $\mathcal{K}_1$ and $\mathcal{K}_2$ is again a convex set, and every face $\mathcal{F}$ of $\mathcal{K}$ is an intersection $\mathcal{F} = \mathcal{F}_1 \cap \mathcal{F}_2$ of faces $\mathcal{F}_1$ of $\mathcal{K}_1$ and $\mathcal{F}_2$ of $\mathcal{K}_2$.*

The facial structure of $\mathcal{K} = \mathcal{K}_1 \cap \mathcal{K}_2$ follows from Def. 1 because the bidirections at any point $x \in \mathcal{K}$ are the common bidirections at x in $\mathcal{K}_1$ and $\mathcal{K}_2$.

A cone C in a real vector space is a set such that if $x \in C$, $x \neq 0$, then $tx \in C$ but $-tx \notin C$ for every $t > 0$. The concept of dual convex cones, to be defined below, has found a central place in the theory of quantum entanglement.

2.2. *Quantum states*

We are concerned here with a bipartite quantum system with Hilbert space $\mathcal{H} = \mathcal{H}_a \otimes \mathcal{H}_b$ of finite dimension $N = N_a N_b$. The real vector space H of observables on $\mathcal{H}$ has dimension N^2, a natural Euclidean inner product $\langle A, B\rangle = \operatorname{Tr} AB$ and the corresponding Hilbert–Schmidt norm $\|A\| = \sqrt{\operatorname{Tr} A^2}$. We take $\mathcal{H}_a = \mathbb{C}^{N_a}$, $\mathcal{H}_b = \mathbb{C}^{N_b}$, so H is the set of Hermitian matrices in the matrix algebra $\mathrm{M}_N(\mathbb{C})$. We write the components of a vector $\psi \in \mathbb{C}^N$ as

$\psi_I = \psi_{ij}$ with $I = 1, 2, \ldots, N$ or $ij = 11, 12, \ldots, N_a N_b$. For any $A \in H$ we define A^P as the partial transpose of A with respect to the second subsystem, that is,

$$(A^P)_{ij;kl} = A_{il;kj}. \tag{3}$$

The set of positive semidefinite matrices in H is a closed convex cone which we denote by $\mathcal{D}$. The state space of the quantum system is the intersection $\mathcal{D}_1$ between $\mathcal{D}$ and the hyperplane defined by $\mathrm{Tr}A = 1$. The matrices in $\mathcal{D}_1$ are called density matrices, or mixed quantum states. A vector $\psi \in \mathcal{H}$ with $\psi^\dagger\psi = 1$ defines a pure state $\psi\psi^\dagger \in \mathcal{D}_1$. $\mathcal{D}_1$ is a compact convex set of dimension $N^2 - 1$. The pure states are the extremal points of $\mathcal{D}_1$, in fact this is an alternative definition of $\mathcal{D}_1$.

A pure state $\psi\psi^\dagger$ is separable if ψ is a product vector, $\psi = \phi \otimes \chi$, thus a separable pure state is a tensor product of pure states,

$$\psi\psi^\dagger = (\phi \otimes \chi)(\phi \otimes \chi)^\dagger = (\phi\phi^\dagger) \otimes (\chi\chi^\dagger). \tag{4}$$

The set of separable states $\mathcal{S}_1$ is the smallest convex subset of $\mathcal{D}_1$ containing all the separable pure states. Since the separable pure states are extremal points of $\mathcal{D}_1$ containing $\mathcal{S}_1$, they are also extremal points of $\mathcal{S}_1$, and they are all the extremal points of $\mathcal{S}_1$. $\mathcal{S}_1$ is compact and defines a convex cone $\mathcal{S} \subset \mathcal{D}$. The dimension of $\mathcal{S}_1$ is $N^2 - 1$, the same as the dimension of $\mathcal{D}_1$, hence every separable state may be written as a convex combination of N^2 or fewer pure product states.[39,40]

The fact that $\mathcal{S}$ and $\mathcal{D}$ have the same dimension N^2 is not quite as trivial as one might be tempted to think. It is a consequence of the fundamental fact that we use complex Hilbert spaces in quantum mechanics, as the following argument shows. In a quantum mechanics based on real Hilbert spaces every separable state would have to be symmetric under partial transposition, and the dimension of $\mathcal{S}$ would be much smaller than the dimension of $\mathcal{D}$. With complex Hilbert spaces, a generic real symmetric matrix which is a separable state is not symmetric under partial transposition, and the remarkable conclusion is that its representation as a convex combination of pure product states must necessarily involve complex product vectors. The basic reason that $\mathcal{S}$ and $\mathcal{D}$ have the same dimension is the relation $H = H_a \otimes H_b$ between the real vector spaces H, H_a and H_b of observables on $\mathcal{H}$, $\mathcal{H}_a$ and $\mathcal{H}_b$. We construct $\mathcal{S}$ explicitly as a subset of $H_a \otimes H_b$, in such a way that the dimension of $\mathcal{S}$ is the full dimension of $H_a \otimes H_b$.

We define the set of positive partial transpose states (PPT states) as $\mathcal{P}_1 = \mathcal{D}_1 \cap \mathcal{D}_1^P$. Partial transposition is an invertible linear operation preserving the convex structure of $\mathcal{D}_1$, hence $\mathcal{P}_1$ is also a compact convex set defining a convex cone $\mathcal{P}$. We have that $\mathcal{S}_1 \subseteq \mathcal{P}_1 \subset \mathcal{D}_1$. The fact that every separable state has a positive partial transpose is obvious, and provides a simple and powerful test for separability, known as the Peres criterion. [9] In dimensions 2×2 and 2×3 the converse statement is also true, that every PPT state is separable.[11] The entanglement of PPT states is not distillable into entanglement of pure states, and it is believed that the entangled PPT states alone possess this special "bound" type of entanglement.[19]

2.3. *Dual cones, entanglement witnesses*

The existence of entangled PPT states motivates the introduction of entanglement witnesses that may reveal their entanglement. We define the dual cone of $\mathcal{S}$ as

$$\mathcal{S}^\circ = \{\Omega \in H \mid \mathrm{Tr}\,\Omega\rho \geq 0 \ \forall\ \rho \in \mathcal{S}\}. \tag{5}$$

The members of $\mathcal{S}^\circ$ will here be called entanglement witnesses. The usual convention is that a witness is required to have at least one negative eigenvalue, but since our focus is on the geometry of $\mathcal{S}^\circ$ this restriction is not important to us here.

The definition of $\mathcal{S}^\circ$ as a dual cone implies that it is closed and convex. Since $\mathcal{S}$ is a closed convex cone the dual of $\mathcal{S}^\circ$ is $\mathcal{S}$ itself [11,42] (the fact that $\mathcal{S}^{\circ\circ} = \mathcal{S}$ is given as an exercise in Ref. 42). Thus, a witness having a negative expectation value in a state proves the state to be entangled, and given any entangled state there exists a witness having a negative expectation value in that state. This two-way implication makes entanglement witnesses powerful tools for detecting entanglement, both theoretically and experimentally.[20,26,44-49]

From the experimental point of view the testimony of an entanglement witness is of a statistical nature. A positive or negative result of one single measurement of a witness gives little information about whether the state in question is separable or entangled. A positive or negative average over many measurements will give a more reliable answer, but some statistical uncertainty must always remain.

Since the extremal points of $\mathcal{S}_1$ are the pure product states, a matrix $\Omega \in H$ is a witness if and only if its expectation value is non-negative in every pure product state. Furthermore, since the dimension of $\mathcal{S}$ is the full dimension of H, every witness $\Omega \neq 0$ has strictly positive expectation values in some pure product states. Since every product vector is a member of some basis of orthogonal product vectors, we conclude that every witness $\Omega \neq 0$ has $\mathrm{Tr}\,\Omega > 0$.[12] Accordingly the set $\mathcal{S}_1^\circ$ of normalized (unit trace) witnesses completely describes all of $\mathcal{S}^\circ$.

Theorem 3. *$\mathcal{S}_1^\circ$ is a bounded set.*

Proof. The maximally mixed state $\rho_0 = I/N$ is in $\mathcal{S}_1$. Define for $\theta > 0$ and $\Gamma \in H$ with $\mathrm{Tr}\,\Gamma = 0$, $\mathrm{Tr}\,\Gamma^2 = 1$,

$$f(\theta, \Gamma) = \min_{\rho \in \mathcal{S}_1} \mathrm{Tr}\,\rho\,(\rho_0 + \theta\Gamma) = \frac{1}{N} + \theta \min_{\rho \in \mathcal{S}_1} \mathrm{Tr}\,\rho\Gamma. \tag{6}$$

Since $\mathcal{S}_1$ is compact, f is well defined. Since $\mathrm{Tr}\,\Gamma = 0$, we have $\Gamma \notin \mathcal{S}^\circ$ and the minimum of $\mathrm{Tr}\,\rho\Gamma$ is strictly negative. Therefore, given any Γ we can always find the $\theta(\Gamma)$ which makes $f(\theta, \Gamma) = 0$. Since Γ lies in a compact set there exists a Γ^* with $\theta^* = \theta(\Gamma^*) = \max_\Gamma \theta(\Gamma)$. This θ^* defines an $N^2 - 1$ dimensional Euclidean ball $\mathcal{B}_1$ centered at ρ_0 containing $\mathcal{S}_1^\circ$. □

Note that the dual set $\mathcal{B}_1^\circ$ is an $N^2 - 1$ dimensional ball contained in $\mathcal{S}_1$, also centered at ρ_0.[12] From now on, when we talk about entanglement witnesses we will usually assume that they are normalized and lie in $\mathcal{S}_1^\circ$. Since $\mathcal{S}_1^\circ$ is a compact convex set it is completely described by its extremal points. The ultimate objective of the present work is to characterize these extremal witnesses.

The concept of dual cones applies of course also to the cone $\mathcal{D}$ of positive semidefinite matrices and the cone $\mathcal{P}$ of PPT matrices. The cone $\mathcal{D}$ is self-dual, $\mathcal{D}^\circ = \mathcal{D}$. Since $\mathcal{P} = \mathcal{D} \cap \mathcal{D}^P$, the dual cone $\mathcal{P}^\circ$ is the convex hull of $\mathcal{D} \cup \mathcal{D}^P$. Hence, the extremal points of $\mathcal{P}_1^\circ$ are the pure states $\psi\psi^\dagger$ and the partially transposed pure states $(\psi\psi^\dagger)^P$. These are also

extremal in $\mathcal{S}_1^\circ$.[14,15,24,25] A witness $\Omega \in \mathcal{P}^\circ$ is called decomposable, because it has the form

$$\Omega = \rho + \sigma^P \quad \text{with} \quad \rho, \sigma \in \mathcal{D}. \tag{7}$$

This terminology comes from the mathematical theory of positive maps.[14] Decomposable witnesses are in a sense trivial, and not very useful as witnesses, since they can not be used for detecting entangled PPT states.

Altogether, we have the following sequence of compact convex sets,

$$\mathcal{S}_1 \subset \mathcal{P}_1 \subset \mathcal{D}_1 = \mathcal{D}_1^\circ \subset \mathcal{P}_1^\circ \subseteq \mathcal{S}_1^\circ. \tag{8}$$

All these sets, with the exception of $\mathcal{D}_1$, are invariant under partial transposition. The sets $\mathcal{S}_1$, $\mathcal{D}_1$, and $\mathcal{P}_1^\circ$ are very simply described in terms of their extremal points. The extremal points of $\mathcal{P}_1$ are not fully understood, though some progress has been made in this direction.[18,34-37] The extremal points of $\mathcal{S}_1^\circ$ are what we investigate here, they include the extremal points of $\mathcal{S}_1$, $\mathcal{D}_1$, and $\mathcal{P}_1^\circ$.

A face of $\mathcal{D}_1$ is a complete set of density matrices on some subspace of $\mathcal{H}$.[10] Faces of $\mathcal{S}_1$ have recently been studied by Alfsen and Shultz.[16,50] We also comment on faces of $\mathcal{S}_1$ in Sec. 8. As we develop our results regarding extremal witnesses we will simultaneously obtain a classification of faces of $\mathcal{S}_1^\circ$.

2.4. *Positive maps*

The study of the set of separable states through the dual set of entanglement witnesses was started by Michał, Paweł, and Ryszard Horodecki.[11] They pointed out the relation between witnesses and positive maps, and used known results from the mathematical theory of positive maps to throw light on the separability problem.[14] In particular, the fundamental result that there exist entangled PPT states is equivalent to the existence of non-decomposable positive maps.

We describe here briefly how entanglement witnesses are related to positive maps, but do not aim at developing the perspectives of positive maps in great detail. In Sec. 10 we return briefly to the use of positive maps for detecting entanglement.

We use a matrix $A \in H$ to define a real linear map $\mathbf{L}_A : H_a \to H_b$ such that $Y = \mathbf{L}_A X$ when

$$Y_{jl} = \sum_{i,k} A_{ij;kl} X_{ki}. \tag{9}$$

The correspondence $A \leftrightarrow \mathbf{L}_A$ is a vector space isomorphism between H and the space of real linear maps $H_a \to H_b$. A slightly different version of this isomorphism is more common in the literature, this is the Jamiołkowski isomorphism $A \leftrightarrow \mathbf{J}_A$ by which $\mathbf{J}_A X = \mathbf{L}_A(X^T)$.[51]

The transposed real linear map $\mathbf{L}_A^T : H_b \to H_a$ is defined such that $X = \mathbf{L}_A^T Y$ when

$$X_{ik} = \sum_{j,l} A_{ij;kl} Y_{lj}. \tag{10}$$

It is the transpose of $\mathbf{L}_A$ with respect to the natural scalar products $\langle U, V \rangle = \mathrm{Tr}\, UV$ in H_a

and H_b. In fact, for any $X \in H_a$, $Y \in H_b$ we have

$$\langle \mathbf{L}_A X, Y\rangle = \mathrm{Tr}\,((\mathbf{L}_A X)Y) = \sum_{i,j,k,l} A_{ij;kl} X_{ki} Y_{lj} = \mathrm{Tr}\,(X(\mathbf{L}_A^T Y)), \tag{11}$$

and therefore

$$\langle \mathbf{L}_A X, Y\rangle = \langle X, \mathbf{L}_A^T Y\rangle. \tag{12}$$

The maps $\mathbf{L}_A$ and $\mathbf{L}_A^T$ act on one dimensional projection operators $\phi\phi^\dagger \in H_a$ and $\chi\chi^\dagger \in H_b$ according to

$$\begin{aligned} \mathbf{L}_A(\phi\phi^\dagger) &= (\phi \otimes I_b)^\dagger A(\phi \otimes I_b), \\ \mathbf{L}_A^T(\chi\chi^\dagger) &= (I_a \otimes \chi)^\dagger A(I_a \otimes \chi). \end{aligned} \tag{13}$$

Note that $\phi \otimes I_b$ is an $N \times N_b$ matrix such that $(\phi \otimes I_b)\chi = \phi \otimes \chi$, whereas $I_a \otimes \chi$ is an $N \times N_a$ matrix such that $(I_a \otimes \chi)\phi = \phi \otimes \chi$. It follows that

$$\chi^\dagger \left(\mathbf{L}_A(\phi\phi^\dagger)\right)\chi = \phi^\dagger \left(\mathbf{L}_A^T(\chi\chi^\dagger)\right)\phi = (\phi \otimes \chi)^\dagger A(\phi \otimes \chi). \tag{14}$$

If Ω is an entanglement witness then Eq. (14) implies that $\mathbf{L}_\Omega$ is a positive linear map $H_a \to H_b$, mapping positive semidefinite matrices in H_a to positive semidefinite matrices in H_b. Similarly, $\mathbf{L}_\Omega^T$ is a positive linear map $H_b \to H_a$. The correspondence $\Omega \leftrightarrow \mathbf{L}_\Omega$ is a vector space isomorphism between the set of entanglement witnesses and the set of positive maps.

A positive map $\mathbf{M}$ is said to be completely positive if every map $\mathbf{I} \otimes \mathbf{M}$, where $\mathbf{I}$ is the identity map in an arbitrary dimension, is positive. It is easily shown that $\mathbf{J}_A$ is completely positive if and only if A is a positive matrix. Thus $\mathbf{L}_A$ is completely positive if and only if A^P is a positive matrix.

2.5. *Biquadratic forms and optimization*

The expectation value of an observable $A \in H$ in a pure product state $\psi\psi^\dagger$ with $\psi = \phi \otimes \chi$ is a biquadratic form

$$f_A(\phi, \chi) = (\phi \otimes \chi)^\dagger A(\phi \otimes \chi). \tag{15}$$

The condition for $\Omega \in H$ to be an entanglement witness is that the corresponding biquadratic form f_Ω is positive semidefinite, *i.e.* that Ω has non-negative expectation value in the set of pure product vectors $\psi = \phi \otimes \chi$. Our approach here is to study witnesses through the associated biquadratic forms.

In the present work we study especially boundary witnesses, corresponding to biquadratic forms that are marginally positive. This leads naturally to an alternative definition of entanglement witnesses in terms of an optimization problem.

Definition 2. *A matrix $A \in H$ is an entanglement witness if and only if the minimum value p^* of the problem*

$$\begin{aligned} \textit{minimize: } & f_A(\phi, \chi) = (\phi \otimes \chi)^\dagger A(\phi \otimes \chi), \quad \phi \in \mathcal{H}_a,\ \chi \in \mathcal{H}_b, \\ & \textit{subject to: } \|\phi\| = \|\chi\| = 1, \end{aligned} \tag{16}$$

is non-negative.

The normalization demand on ϕ and χ ensures that $p^* > -\infty$ when $A \notin \mathcal{S}^\circ$, and that $p^* > 0$ for every A in the interior of $\mathcal{S}^\circ$. It is natural to use here the Hilbert–Schmidt norm $\|A\| = \sqrt{\operatorname{Tr} A^2}$, which was introduced above, but in principle any norm would serve the same purpose.

The boundary $\partial\mathcal{S}^\circ$ of $\mathcal{S}^\circ$ consists of those $\Omega \in \mathcal{S}^\circ$ for which there exists a separable state ρ orthogonal to Ω, *i.e.* with $\operatorname{Tr}\Omega\rho = 0$. In terms of problem (16), $A \in \partial\mathcal{S}^\circ$ if and only if $p^* = 0$. Since we know that $\mathcal{S}^{\circ\circ} = \mathcal{S}$ we can similarly understand the boundary $\partial\mathcal{S}$ of the set of separable states $\mathcal{S}$, as the subset of $\mathcal{S}$ which are orthogonal to some witness. We will return to this duality between the boundaries of $\mathcal{S}$ and $\mathcal{S}^\circ$ in much more detail in Sec. 8.

2.6. *Equivalence under SL ⊗ SL transformations*

It is very useful to observe that all the main concepts discussed in the present article are invariant under what we call $\mathrm{SL}\otimes\mathrm{SL}$ transformations, in which a matrix A is transformed into $VAV^\dagger$ with an invertible product matrix $V = V_a \otimes V_b$.

For example, since such a transformation is linear in A, it preserves convex combinations, extremal points, and in general all the convexity properties of different sets. It preserves the positivity of matrices, the tensor product structure of vectors and matrices, the number of zeros of witnesses, and in general all properties related to entanglement, except that it may increase or decrease entanglement as measured quantitatively if either V_a or V_b is not unitary.

Thus, for our purposes it is useful to consider two density matrices or two entanglement witnesses to be equivalent if they are related by an $\mathrm{SL}\otimes\mathrm{SL}$ transformation. This sorting into equivalence classes helps to reduce the problem of understanding and classifying entangled states and entanglement witnesses.

3. Secondary constraints at zeros of witnesses

By definition, a witness Ω satisfies the following infinite set of inequalities, all linear in Ω,

$$f_\Omega(\phi,\chi) \geq 0 \qquad \text{with} \qquad \phi \in \mathcal{H}_a,\ \chi \in \mathcal{H}_b, \quad \|\phi\| = \|\chi\| = 1. \tag{17}$$

These constraints on Ω, represented here as a biquadratic form f_Ω, are the primary constraints defining the set $\mathcal{S}_1^\circ$, apart from the trivial linear constraint $\operatorname{Tr}\Omega = 1$.

If Ω is situated on the boundary of $\mathcal{S}_1^\circ$ it means that at least one of these primary inequalities is an equality. We will call the pair (ϕ_0,χ_0) a zero of Ω if $f_\Omega(\phi_0,\chi_0) = 0$. We count $(a\phi_0, b\chi_0)$ with $a, b \in \mathbb{C}$ as the same zero. The primary constraints (17) with (ϕ,χ) close to the zero (ϕ_0,χ_0) lead to rather stringent constraints on Ω, which we introduce as secondary constraints to be imposed at the zero. These secondary constraints are both equalities and inequalities, and they are linear in Ω, like the primary constraints from which they are derived. They are summarized in explicit form in Appendix A. In the next section we apply all the secondary constraints to the problem of constructing and classifying extremal witnesses.

The analysis of secondary constraints that we have presented here in this section is essentially the same as the one carried out by Lewenstein *et al.* in their study of optimal witnesses.[26]

3.1. *Positivity constraints on polynomials*

A model example may illustrate how we treat constraints. Let $f(t)$ be a real polynomial in one real variable t, of degree four and with a strictly positive quartic term, satisfying the primary constraints $f(t) \geq 0$ for all t. These are constraints on the coefficients of the polynomial. The equation $f(t) = 0$ can have zero, one or two real roots for t. Assume that $f(t_0) = 0$. Because this is a minimum, we must have $f'(t_0) = 0$ and $f''(t_0) \geq 0$. In the limiting case $f''(t_0) = 0$ we must have also $f^{(3)}(t_0) = 0$.

Thus, if $f(t_0) = 0$ we get secondary constraints $f'(t_0) = 0$, and either $f''(t_0) > 0$ or $f''(t_0) = 0$, $f^{(3)}(t_0) = 0$. If there is a second zero t_1, similar secondary constraints must hold there. It should be clear that we may replace the infinite set of primary constraints $f(t) \geq 0$ for every t by the finite set of secondary constraints at the zeros t_0 and t_1.

Because the zeros of a witness Ω are roots of a polynomial equation in several variables, they have the following property.

Theorem 4. *The set of zeros of a witness consists of at most a finite number of components, where each component is either an isolated point or a continuous connected surface.*

Proof. Choose (ϕ_2, χ_2) such that $f_\Omega(\phi_2, \chi_2) > 0$, and define $f(t) = f_\Omega(\phi_1 + t\phi_2, \chi_1 + t\chi_2)$. For given (ϕ_1, χ_1) and (ϕ_2, χ_2) this is a non-negative polynomial in the real variable t of degree four, hence it has zero, one or two real roots for t. By varying (ϕ_1, χ_1) we reach all the zeros of Ω. The zeros of $f(t)$ will move continuously when we vary (ϕ_1, χ_1), except that they may appear or disappear. This construction should result in a set of zeros as described in the theorem. □

Now let (ϕ_0, χ_0) be a zero of Ω, with $\|\phi_0\| = \|\chi_0\| = 1$. Since the constraints (17) on the polynomial f_Ω at $(\phi, \chi) = (\phi_0 + \xi, \chi_0 + \zeta)$ are actually independent of the normalization conditions $\|\phi\| = \|\chi\| = 1$, we choose to abandon these non-linear normalization conditions (non-linear in the Hilbert–Schmidt norm) and replace them by the linear constraints $\phi_0^\dagger \xi = 0$, $\chi_0^\dagger \zeta = 0$. Strictly speaking, even these orthogonality conditions are not essential, the important point is that we vary ϕ and χ in directions away from ϕ_0 and χ_0.

We find that some of the secondary constraints on f_Ω are intrinsically real equations rather than complex equations, therefore we introduce real variables $x \in \mathbb{R}^{2N_a-2}$, $y \in \mathbb{R}^{2N_b-2}$ and write $\xi = J_0 x$, $\zeta = K_0 y$ with $\phi_0^\dagger J_0 = 0$, $\chi_0^\dagger K_0 = 0$. Our biquadratic form is then a real inhomogeneous polynomial quadratic in x and quadratic in y,

$$f(x, y) = f_\Omega(\phi, \chi) = ((\phi_0 + \xi) \otimes (\chi_0 + \zeta))^\dagger \, \Omega \, ((\phi_0 + \xi) \otimes (\chi_0 + \zeta)). \tag{18}$$

The linear term of the polynomial is

$$\begin{aligned} f_1(x, y) &= 2\,\mathrm{Re}\, ((\xi \otimes \chi_0)^\dagger \Omega (\phi_0 \otimes \chi_0) + (\phi_0 \otimes \zeta)^\dagger \Omega (\phi_0 \otimes \chi_0)) \\ &= x^T \mathrm{D}_x f(0, 0) + y^T \mathrm{D}_y f(0, 0), \end{aligned} \tag{19}$$

in terms of the gradients

$$\begin{aligned} \mathrm{D}_x f(0, 0) &= 2\,\mathrm{Re}\, (J_0 \otimes \chi_0)^\dagger \Omega (\phi_0 \otimes \chi_0), \\ \mathrm{D}_y f(0, 0) &= 2\,\mathrm{Re}\, (\phi_0 \otimes K_0)^\dagger \Omega (\phi_0 \otimes \chi_0). \end{aligned} \tag{20}$$

The quadratic term is

$$\begin{aligned} f_2(x,y) &= (\xi\otimes\chi_0)^\dagger\,\Omega\,(\xi\otimes\chi_0) + (\phi_0\otimes\zeta)^\dagger\,\Omega\,(\phi_0\otimes\zeta) \\ &\quad + 2\,\mathrm{Re}\,((\phi_0\otimes\zeta)^\dagger\,\Omega\,(\xi\otimes\chi_0) + (\phi_0\otimes\chi_0)^\dagger\,\Omega\,(\xi\otimes\zeta)) \\ &= z^T G_\Omega\, z \end{aligned} \tag{21}$$

where $z^T = (x^T, y^T)$, and $2\,G_\Omega = \mathrm{D}^2 f(0,0)$ is the second derivative, or so called Hessian matrix, which is a always real and symmetric matrix,

$$G_\Omega = \mathrm{Re}\begin{bmatrix} g_{xx} & g_{yx}^T \\ g_{yx} & g_{yy}\end{bmatrix}, \tag{22}$$

where the elements are

$$\begin{aligned} g_{xx} &= (J_0\otimes\chi_0)^\dagger\Omega(J_0\otimes\chi_0), \\ g_{yx} &= (\phi_0\otimes K_0)^\dagger\Omega(J_0\otimes\chi_0) + (\phi_0\otimes K_0^*)^\dagger\Omega^P(J_0\otimes\chi_0^*), \\ g_{yy} &= (\phi_0\otimes K_0)^\dagger\Omega(\phi_0\otimes K_0). \end{aligned} \tag{23}$$

The cubic term is like the linear term but with $\phi_0 \leftrightarrow \xi$ and $\chi_0 \leftrightarrow \zeta$,

$$f_3(x,y) = 2\,\mathrm{Re}\,((\phi_0\otimes\zeta)^\dagger\Omega(\xi\otimes\zeta) + (\xi\otimes\chi_0)^\dagger\Omega(\xi\otimes\zeta)). \tag{24}$$

The quartic term is simply

$$f_4(x,y) = f_\Omega(\xi,\zeta) = (\xi\otimes\zeta)^\dagger\Omega(\xi\otimes\zeta). \tag{25}$$

The fact that the constant term of the polynomial vanishes, $f(0,0) = 0$, is one real linear constraint on Ω,

$$\mathbf{T}_0 : H \to \mathbb{R}\,, \quad \mathbf{T}_0\Omega = (\phi_0\otimes\chi_0)^\dagger\Omega(\phi_0\otimes\chi_0) = 0. \tag{26}$$

Because $(x,y) = (0,0)$ is a minimum of the polynomial the linear term must also vanish identically, and this results in another linear system of constraints,

$$\mathbf{T}_1 : H \to \mathbb{R}^{2(N_a+N_b-2)}, \quad \mathbf{T}_1\Omega = \begin{bmatrix}\mathrm{D}_x f(0,0)\\ \mathrm{D}_y f(0,0)\end{bmatrix} = 0. \tag{27}$$

Note that the equality constraints $\mathbf{T}_0$ and $\mathbf{T}_1$ are the same for every witness with the zero (ϕ_0,χ_0), they are uniquely defined by the zero alone. The total number of constraints in $\mathbf{T}_0$ and $\mathbf{T}_1$ is

$$M_2 = 2(N_a + N_b) - 3. \tag{28}$$

All of these M_2 constraints are linearly independent. They are given an explicitly real representation in Appendix A.

It is worth noting that the vanishing of the constant and linear terms of the polynomial $f(x,y)$ is equivalent to the conditions that

$$\begin{aligned} (\phi\otimes\chi_0)^\dagger\Omega(\phi_0\otimes\chi_0) = 0 \quad &\forall\,\phi\in\mathcal{H}_a, \\ (\phi_0\otimes\chi)^\dagger\Omega(\phi_0\otimes\chi_0) = 0 \quad &\forall\,\chi\in\mathcal{H}_b. \end{aligned} \tag{29}$$

Hence the $2(N_a + N_a) - 3$ real constraints $\mathbf{T}_0\Omega = 0$ and $\mathbf{T}_1\Omega = 0$ may be expressed more simply as the following set of $N_a + N_b$ complex constraints, equivalent to $2(N_a + N_b)$ real constraints that are then not completely independent,

$$\begin{aligned} \mathbf{L}_\Omega^T(\chi_0\chi_0^\dagger)\,\phi_0 &= (I_a \otimes \chi_0)^\dagger\Omega(\phi_0 \otimes \chi_0) = 0, \\ \mathbf{L}_\Omega(\phi_0\phi_0^\dagger)\,\chi_0 &= (\phi_0 \otimes I_b)^\dagger\Omega(\phi_0 \otimes \chi_0) = 0. \end{aligned} \tag{30}$$

The quadratic term of the polynomial has to be non-negative, again because $(x, y) = (0, 0)$ is a minimum. The inequalities

$$z^T G_\Omega\, z \geq 0 \qquad \forall\, z \in \mathbb{R}^{2(N_a+N_b-2)} \tag{31}$$

are secondary inequality constraints, linear in Ω, equivalent to the non-linear constraints that all the eigenvalues of the Hessian matrix G_Ω must be non-negative.

There are now two alternatives. If the inequalities (31) hold with strict inequality for all $z \neq 0$, in other words, if all the eigenvalues of the Hessian are strictly positive, then we call (ϕ_0, χ_0) a quadratic zero. In this case $\mathbf{T}_0$ and $\mathbf{T}_1$ are the only equality constraints placed on Ω by the existence of the zero (ϕ_0, χ_0). We can immediately state the following important result.

Theorem 5. *A quadratic zero is isolated: there is a finite distance to the next zero. Hence, a witness can have at most a finite number of quadratic zeros.*

3.2. *Hessian zeros*

The second alternative is that the Hessian has K zero eigenvalues with $K \geq 1$. We will call $z \neq 0$ a Hessian zero at the zero (ϕ_0, χ_0) if $G_\Omega z = 0$. It is then a real linear combination

$$z = \sum_{i=1}^{K} a_i z_i \tag{32}$$

of basis vectors $z_i \in \operatorname{Ker} G_\Omega$. The K linearly independent eigenvectors z_i define the following system of linear constraints on Ω,

$$\mathbf{T}_2 : H \to \mathbb{R}^{2K(N_a+N_b-2)}, \quad (\mathbf{T}_2\Omega)_i = G_\Omega z_i = 0\,, \quad i = 1, \ldots, K. \tag{33}$$

These constraints ensure that $G_\Omega z = 0$ and hence $f_2(x, y) = z^T G_\Omega z = 0$. The vanishing of the quadratic term of the polynomial implies that the cubic term must vanish as well. We therefore call (ϕ_0, χ_0) a quartic zero, and the direction z at (ϕ_0, χ_0) a quartic direction. The vanishing cubic term is

$$f_3(x, y) = 2 \sum_{l,m,n} a_l a_m a_n \mathrm{Re}\left((\phi_0 \otimes \zeta_l)^\dagger\Omega(\xi_m \otimes \zeta_n) + (\xi_l \otimes \chi_0)^\dagger\Omega(\xi_m \otimes \zeta_n)\right) = 0, \tag{34}$$

where $\xi_i = J_0 x_i$ and $\zeta_i = K_0 y_i$. Since the product $a_l a_m a_n$ is completely symmetric in the indices l, m, n we should symmetrize the expression it multiplies. All the different properly symmetrized expressions must vanish. In this way we obtain a linear system

$$(\mathbf{T}_3\Omega)_{lmn} = 0 \qquad \text{with} \qquad 1 \leq l \leq m \leq n \leq K. \tag{35}$$

The number of equations is the binomial coefficient $\binom{K+2}{3}$, but it is likely that these constraints on Ω will in general not be independent.

The total number of constraints in $\mathbf{T}_2$ and $\mathbf{T}_3$ is therefore

$$M_4(K) = 2K(N_a + N_b - 2) + \binom{K+2}{3}. \tag{36}$$

Note that $M_4(1) = M_2 = 2(N_a + N_b) - 3$. Since the constraints in $\mathbf{T}_2$ and $\mathbf{T}_3$ address other terms in the polynomial than those in $\mathbf{T}_0$ and $\mathbf{T}_1$, the two sets of constraints should be independent from each other. Furthermore, in the case $K = 1$, all $M_4(1)$ equations in the total system $\mathbf{T}_2, \mathbf{T}_3$ should be linearly independent. In the case of $K > 1$ we expect overlapping constraints.

3.3. *Summary of constraints*

Let Ω be a witness, and let Z be the complete set of zeros of Ω, with $Z' \subseteq Z$ as the subset of quartic zeros. Each zero in Z defines zeroth and first order equality constraints $\mathbf{T}_0$ and $\mathbf{T}_1$. The combination of a set of constraints is the direct sum. Thus we define $\mathbf{U}_0$ as the direct sum of the constraints $\mathbf{T}_0$ over all the zeros in Z, and similarly $\mathbf{U}_1$ as the direct sum of the constraints $\mathbf{T}_1$. We define $\mathbf{U}_{01} = \mathbf{U}_0 \oplus \mathbf{U}_1$. Each zero in Z' introduces additional second and third order constraints $\mathbf{T}_2$ and $\mathbf{T}_3$. We denote the direct sums of these by $\mathbf{U}_2$ and $\mathbf{U}_3$, respectively, and we define $\mathbf{U}_{23} = \mathbf{U}_2 \oplus \mathbf{U}_3$. Denote by $\mathbf{U}_\Omega$ the full system of constraints on Ω, $\mathbf{U}_\Omega = \mathbf{U}_{01} \oplus \mathbf{U}_{23}$.

It is an important observation that all the constraints are completely determined by the zeros and Hessian zeros of Ω. Thus they depend only indirectly on Ω. In particular, $\mathbf{U}_{01}$ depends only on Z, whereas $\mathbf{U}_{23}$ depends on Z' and on the kernel of the Hessian at each quartic zero.

A boundary witness with only quadratic zeros will be called here a quadratic boundary witness, or simply a quadratic witness, since we are talking most of the time about boundary witnesses. A boundary witness with at least one quartic zero will be called a quartic witness. Note that if Ω is quadratic then Z' is empty so $\mathbf{U}_\Omega = \mathbf{U}_{01}$. This classification of boundary witnesses as quadratic and quartic is fundamental, since the two classes have rather different properties. The quadratic witnesses turn up in much greater numbers in random searches, and are also simpler to understand theoretically.

4. Extremal witnesses

In this section we apply the constraints developed in the previous section to study extremal witnesses. We approach the problem through the basic definition that a non-extremal witness is one which can be written as a convex combination of other witnesses. In the language of convex sets it is an interior point of a face of $\mathcal{S}_1^\circ$ of dimension one or higher.

In Sec. 4.1 we characterize the extremal witnesses in terms of their zeros. In Secs. 4.2 and 4.3 we characterize them in terms of the constraints related to their zeros. and present a search algorithm for finding them numerically. The geometrical interpretation of this algorithm is that we start from a given witness, reconstruct the unique face of $\mathcal{S}_1^\circ$ of which it is an interior point, go to the boundary of this face, and repeat the process. In Sec. 4.4 we discuss the expected number of zeros of extremal witnesses.

Theorems 8 and 10 give necessary and sufficient conditions for a witness to be extremal, and they are main results of our work.

4.1. *Zeros of witnesses, convexity and extremality*

Let Ω be a convex combination of two different witnesses Λ and Σ,

$$\Omega = (1-p)\,\Lambda + p\,\Sigma, \tag{37}$$

with $0 < p < 1$. Then the corresponding biquadratic form is a convex combination

$$f_\Omega(\phi,\chi) = (1-p)\,f_\Lambda(\phi,\chi) + p\,f_\Sigma(\phi,\chi), \tag{38}$$

and the Hessian matrix at any zero (ϕ_0,χ_0) of Ω is also a convex combination

$$G_\Omega = (1-p)\,G_\Lambda + p\,G_\Sigma. \tag{39}$$

The facts that the biquadratic form of an entanglement witness is positive semidefinite, and that the Hessian matrix at any zero of a witness is also positive semidefinite, imply the following result.

Theorem 6. *If Ω is a convex combination of two witnesses Λ and Σ, as above, then (ϕ_0,χ_0) is a zero of Ω if and only if it is a zero of both Λ and Σ.*

Similarly, z is a Hessian zero of Ω at the zero (ϕ_0,χ_0) if and only if it is a Hessian zero at (ϕ_0,χ_0) of both Λ and Σ.

Thus, all witnesses in the interior of the line segment between Λ and Σ have exactly the same zeros and Hessian zeros.

Assume further that Λ and Σ are extremal points of the intersection of the straight line with $\mathcal{S}_1^\circ$, so that $\Omega \notin \mathcal{S}_1^\circ$ when $p < 0$ and when $p > 1$. To be specific, consider the case $p > 1$. We want to show that Σ has at least one zero or Hessian zero in addition to the zeros and Hessian zeros of the witnesses in the interior of the line segment.

By assumption, the set of negative points

$$X_-(p) = \{(\phi,\chi) \mid f_\Omega(\phi,\chi) < 0\} \tag{40}$$

is non-empty for $p > 1$. Clearly $X_-(p) \subset X_-(q)$ for $1 < p < q$, since $X_-(p)$ is empty for $0 \le p \le 1$, and $f_\Omega(\phi,\chi)$ is a linear function of p for fixed (ϕ,χ) as given by Eq. (38). The closure $\overline{X}_-(p)$ is a compact set when we normalize so that $\|\phi\| = \|\chi\| = 1$. It follows that the limit of $\overline{X}_-(p)$ as $p \to 1+$ is a non-empty set X_0 of zeros of the witness Σ,

$$X_0 = \bigcap_{p>1} \overline{X}_-(p). \tag{41}$$

Every zero $(\phi_0,\chi_0) \in X_0$ is of the kind we are looking for. If it is not a zero of Ω for $p < 1$, then it is a new zero of Σ as compared to the zeros of the witnesses in the interior of the line segment. If it is a zero of Ω for $p < 1$, then Σ has at least one Hessian zero at (ϕ_0,χ_0) which is not a Hessian zero at (ϕ_0,χ_0) of the witnesses in the interior of the line segment.

To prove the last statement, assume that $(\phi_0, \chi_0) \in X_0$ is a zero of Ω for $0 < p < 1$. Then it is a zero of Ω for any p, again because $f_\Omega(\phi, \chi)$ is a linear function of p. Similarly, if z is a Hessian zero of Ω at (ϕ_0, χ_0) for $0 < p < 1$, it is a Hessian zero of Ω at (ϕ_0, χ_0) for any p.

From the assumption that $(\phi_0, \chi_0) \in X_0$ follows that $f_\Omega(\phi, \chi)$ for any $p > 1$ takes negative values for some (ϕ, χ) arbitrarily close to (ϕ_0, χ_0). The only way this may happen is that for $p > 1$ the second derivative $z^T G_\Omega z$ is negative in some direction z, meaning that G_Ω for $p > 1$ has one or more negative eigenvalues $\lambda_i(p)$, with eigenvectors $z_i(p)$ that are orthogonal to each other and orthogonal to the p independent part of $\operatorname{Ker} G_\Omega$.

In the limit $p \to 1+$ the negative eigenvalues $\lambda_i(p)$ of G_Ω go to zero, and the corresponding eigenvectors $z_i(p)$ go to eigenvectors $z_i(1)$ of G_Σ with zero eigenvalues. These eigenvectors are then Hessian zeros of Σ at (ϕ_0, χ_0) that are not Hessian zeros of Ω for p values in the range $0 < p < 1$.

We summarize the present discussion as follows.

Theorem 7. *If a line segment in $\mathcal{S}_1^\circ$ with end points Λ and Σ can not be prolonged within $\mathcal{S}_1^\circ$ in either direction, then Σ has the same zeros and Hessian zeros as the interior points of the line segment, and at least one additional zero or Hessian zero.*

The same holds for Λ, with additional zeros and Hessian zeros that are different from those of Σ.

These theorems lead to the following extremality criterion for witnesses.

Theorem 8. *A witness Ω is extremal if and only if no witness $\Lambda \neq \Omega$ has a set of zeros and Hessian zeros including the zeros and Hessian zeros of Ω.*

An equivalent condition is that there can exist no witness $\Lambda \neq \Omega$ satisfying the constraints $\mathbf{U}_\Omega \Lambda = 0$.

Proof. The "if" part follows directly from Thm. 6.

To prove the "only if" part, assume that the set of zeros and Hessian zeros of some witness $\Lambda \neq \Omega$ include the zeros and Hessian zeros of Ω. Then by Thm. 6 the interior points of the line segment with Λ and Ω as end points have exactly the same zeros and Hessian zeros as Ω. By Thm. 7 this line segment can be prolonged within $\mathcal{S}_1^\circ$ so that it gets Ω as an interior point. Hence Ω is not extremal. □

4.2. *How to search for extremal witnesses*

Once the zeros and Hessian zeros of a witness Ω are known, it is a simple computational task to find a finite perturbation of Ω within the unique face of $\mathcal{S}_1^\circ$ where Ω is an interior point. The most general direction for such a perturbation is a traceless $\Gamma \in \operatorname{Ker} \mathbf{U}_\Omega$. Note that we only need to find some $\Gamma' \in \operatorname{Ker} \mathbf{U}_\Omega$ not proportional to Ω, then $\Gamma = \Gamma' - (\operatorname{Tr} \Gamma')\, \Omega$ is non-zero and traceless and lies in $\operatorname{Ker} \mathbf{U}_\Omega$.

Theorem 9. *Let Ω be a witness, and let $\Gamma \in H$, $\Gamma \neq 0$, $\operatorname{Tr} \Gamma = 0$. Then $\Lambda = \Omega + t\Gamma$ is a witness for all t in some interval $[t_1, t_2]$, with $t_1 < 0 < t_2$, if and only if $\Gamma \in \operatorname{Ker} \mathbf{U}_\Omega$.*

The maximal value of t_2 is the value of t where Λ acquires a new zero or Hessian zero. The minimal value of t_1 is determined in the same way.

Proof. To prove the "only if" part, assume that $\Lambda = \Omega + t\Gamma$ is a witness for $t_1 \leq t \leq t_2$. Then by Thm. 6, Λ has the same zeros and Hessian zeros as Ω for $t_1 < t < t_2$. Since

the constraints $\mathbf{U}_\Omega$ depend only on the zeros and Hessian zeros of Ω, we conclude that $\mathbf{U}_\Omega\Lambda = 0$ for $t_1 < t < t_2$, and hence $\mathbf{U}_\Omega\Gamma = 0$.

To prove the "if" part, assume that $\mathbf{U}_\Omega\Gamma = 0$. Since $\mathbf{U}_\Omega\Omega = 0$, it follows that $\mathbf{U}_\Omega\Lambda = 0$ for any value of t. Consider the set of negative points,

$$X_-(t) = \{(\phi, \chi) \mid f_\Lambda(\phi, \chi) < 0\}. \tag{42}$$

We want to argue that $X_-(t)$ must be empty for t in some interval $[t_1, t_2]$ with $t_1 < 0 < t_2$. In fact, since $\mathcal{S}_1^\circ$ is a compact set, there must exist some $t_2 \geq 0$ such that $X_-(t)$ is empty for $0 \leq t \leq t_2$ and non-empty for $t > t_2$. Then the limit of $\overline{X}_-(t)$ as $t \to t_2+$ is a non-empty set X_0 of zeros of $\Lambda_2 = \Omega + t_2\Gamma$. A similar reasoning as the one leading to Thm. 7 now leads us to the conclusion that Λ_2 must have at least one zero or Hessian zero which is not a zero or Hessian zero of Ω. This proves that $\Lambda_2 \neq \Omega$ and $t_2 > 0$.

By a similar argument we deduce the existence of a lower limit $t_1 < 0$. □

Figure 1 shows a model for how a new isolated quadratic zero of Λ appears, or how an existing quadratic zero turns into a quartic zero, as the parameter t increases in the function

$$f_t(u) = f_{\Omega+t\Gamma}(\phi + u\phi', \chi + u\chi'). \tag{43}$$

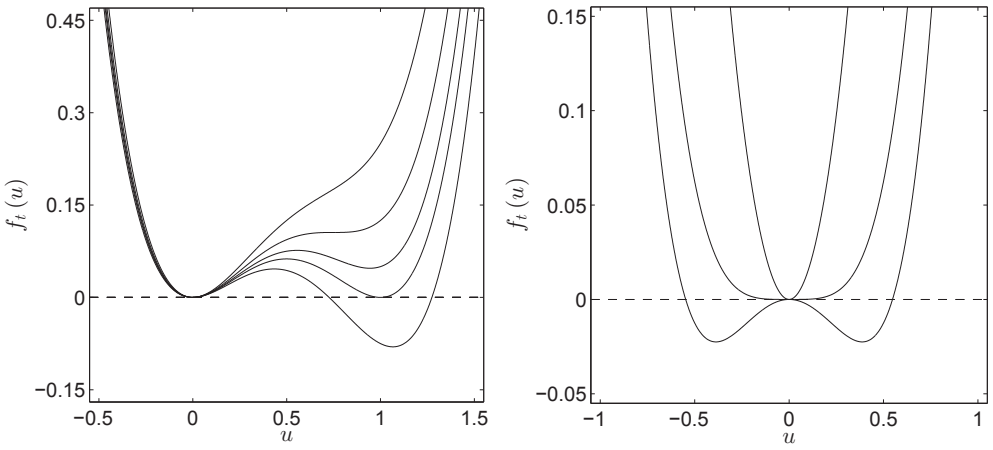


Fig. 1. Models for how the positivity limit is reached. Left: a new local minimum appears, turning into a zero and then a negative minimum. The model function plotted for five different values of t is $f_t(u) = u^2((u-1)^2 + 1 - t)$, depending on a parameter t with the critical value $t_c = 1$. Right: a quadratic zero turns into a quartic zero, and then new negative minima branch off. The model function is $f_t(u) = u^2(u^2 + 1 - t)$, again with the critical parameter value $t_c = 1$.

The next theorem is an immediate corollary. It is a slightly stronger version of Thm. 8. It is interesting for the theoretical understanding, and together with Thm. 9 it is directly useful for numerical calculations.

Theorem 10. *An entanglement witness Ω is extremal if and only if* $\operatorname{Ker}\mathbf{U}_\Omega$ *is one dimensional (spanned by Ω).*

Thus, once the zeros and Hessian zeros of the witness Ω are known we test for extremality by computing the dimension of $\mathrm{Ker}\,\mathbf{U}_\Omega$, for example by a singular value decomposition of $\mathbf{U}_\Omega$. This allows for a simple numerical implementation of the extremality criterion.

These theorems motivate Alg. 1, a search algorithm for finding extremal witnesses. Obviously, any extremal witness might be reached by this algorithm already in the first iteration, starting for example from the maximally mixed state. The search is guaranteed to converge to an extremal witness in a finite number of iterations, since the number of possible search directions is reduced in each iteration. We here summarize this in Alg. 1. Starting from an initial witness $\Omega = \Omega_0 \in \mathrm{int}(\mathcal{S}^\circ)$, we proceed down through a hierarchy of faces on $\mathcal{S}^\circ$ of decreasing dimensions. The number of possible search directions Γ on each face, is reduced in each iteration as $\mathrm{Ker}\,\mathbf{U}_\Omega$ is reduced.

Algorithm 1: Finding an extremal entanglement witness

Precondition: Choose an initial witness $\Omega = \Omega_0 \in \mathrm{int}(\mathcal{S}^\circ)$ and construct $\mathbf{U}_\Omega$

1. **while** $\dim \mathrm{Ker}\,(\mathbf{U}_\Omega) > 1$
2. choose a $\Gamma \in \mathrm{Ker}\,(\mathbf{U}_\Omega)$
3. **if** $\mathrm{Tr}(\Gamma) \neq 0$
4. redefine $\Gamma \leftarrow \Gamma - \mathrm{Tr}(\Gamma)\Omega$
5. **endif**
6. find t_c as the maximal t such that $\Omega + t\Gamma$ is still a witness
7. redefine $\Omega \leftarrow \Omega + t_c\Gamma$
8. locate all zeros of Ω and construct an updated $\mathbf{U}_\Omega$
9. **endwhile**

return Ω

4.3. *Faces of the set $\mathcal{S}_1^\circ$ of normalized witnesses*

Theorem 9 has the following geometrical meaning. Define

$$\mathcal{F}_\Omega = (\mathrm{Ker}\,\mathbf{U}_\Omega) \cap \mathcal{S}_1^\circ. \tag{44}$$

Equivalently, define $\mathcal{F}_\Omega$ as the set of all witnesses of the form $\Omega + t\Gamma$ with $\Gamma \in \mathrm{Ker}\,\mathbf{U}_\Omega$ and $\mathrm{Tr}\,\Gamma = 0$. If Ω is extremal in $\mathcal{S}_1^\circ$ then $\mathcal{F}_\Omega$ consists of the single point Ω. Otherwise, $\mathcal{F}_\Omega$ is the unique face of $\mathcal{S}_1^\circ$ having Ω as an interior point.

Thus, Alg. 1 produces a decreasing sequence of faces of $\mathcal{S}_1^\circ$, $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \ldots \supset \mathcal{F}_n$, where every face $\mathcal{F}_j$ is a face of every $\mathcal{F}_i$ with $i < j$, and the extremal point found is an extremal point of all these faces.

The theorems in Sec. 4.1 imply that every face $\mathcal{F}$ of $\mathcal{S}_1^\circ$ is uniquely characterized by a set of zeros and Hessian zeros that is the complete set of zeros and Hessian zeros of every witness in the interior of $\mathcal{F}$. Every boundary point of $\mathcal{F}$ is a witness having the zeros and Hessian zeros characteristic of the interior points, plus at least one more zero or Hessian zero. In general, the boundary of $\mathcal{S}_1^\circ$ is a hierarchy of faces, faces of faces, faces of faces of faces, and so on. The number of zeros and Hessian zeros increases each time we step from one face onto a face of the face, and the descent through the hierarchy from one face to the next always ends in an extremal witness.

4.4. *Zeros of extremal witnesses*

A natural question regards the number of zeros a witness must have in order to be extremal. We provide two lower bounds for quadratic witnesses, one obtained by comparison with a pure state as a witness and the other by counting constraints. Similar bounds for quartic witnesses are not easy to obtain.

The first of these bounds reveals a "double spanning" property of zeros of quadratic extremal witnesses. We define the partial conjugate of (ϕ, χ) as (ϕ, χ^*).

Theorem 11. *The zeros of a quadratic extremal witness Ω span the Hilbert space. The partially conjugated zeros also span the Hilbert space.*

Proof. Assume that the zeros span less than the whole Hilbert space, so that there exists a vector ψ orthogonal to all the zeros. Then the projection $P = \psi\psi^\dagger$ is a witness with a set of zeros including all the zeros of Ω. But P has a continuum of zeros, all quartic, see Thm. 14. Hence $P \neq \Omega$ and Ω is not extremal, by Thm. 8.

The partial conjugates of the zeros of Ω span the Hilbert space because they are the zeros of the quadratic extremal witness Ω^P. □

This proof does not hold when Ω is a quartic witness, because Ω may then have Hessian zeros that are not Hessian zeros of P. A pure state $\psi\psi^\dagger$ as a quartic extremal witness is a counterexample, where the zeros do not span the Hilbert space, although the partially conjugated zeros may span the Hilbert space. The partial transpose $(\psi\psi^\dagger)^P$ is a counterexample of the opposite kind, where the zeros may span the Hilbert space but the partially conjugated zeros do not. Furthermore, even if neither the zeros nor the partially conjugated zeros span the Hilbert space, it is still possible for a quartic witness to be extremal, because a quartic zero leads to more constraints than a quadratic zero. In Sec. 6.3.2 we describe the numerical construction of an extremal quartic witness in dimension 3×3 with only eight zeros, and with one Hessian zero at one of the zeros.

Counting constraints from quadratic zeros gives a different lower bound for the number of zeros of a quadratic extremal witness. There are $M_2 = 2(N_a + N_b) - 3$ constraints per quadratic zero. With n quadratic zeros there are a total of nM_2 constraints, which in the generic case are linearly independent until nM_2 approaches $N^2 - 1$, the dimension of $\mathcal{S}_1^\circ$. A lower bound on the number of zeros of a quadratic extremal witness is then given by

$$n_c = \text{ceil}\,\frac{(N_a N_b)^2 - 1}{2(N_a + N_b) - 3}, \tag{45}$$

where ceil rounds upwards to the nearest integer.

In the special case $N_a = 2$ this formula gives the lower bound $n_c = N - 1$, which is weaker than the lower bound N given by Thm. 11. When $N_a \geq 3$ and $N_b \geq 3$ the formula gives a lower bound $n_c \geq N$, consistent with Thm. 11.

Table 1 lists the numerically computed number of linearly independent constraints m_{ind} arising from n_c randomly chosen pairs (ϕ, χ). n_c is the estimated minimum number of zeros required for a quadratic witness to be extremal, assuming maximal independence of constraints. M_2 is the number of linearly independent constraints from each zero. In order to compute m_{ind} we interpret each (ϕ, χ) as a quadratic zero, build $\mathbf{U}_{01}$ and do a singular value decomposition. The table shows that in many cases the number of independent

constraints from a random set of product vectors is equal to the dimension of the real vector space H of Hermitian matrices. In such a case there exists no biquadratic form which is positive semidefinite and has these zeros. We conclude that the zeros of a positive semidefinite biquadratic form have to satisfy some relations that reduce the number of independent constraints.

Table 1. Numbers related to quadratic zeros of witnesses in dimension $N = N_a N_b$.

N_a,N_b	N^2	M_2	n_c	m_{ind}
2,2	16	5	3	14
2,3	36	7	5	34
2,4	64	9	7	62
2,5	100	11	9	98
3,3	81	9	9	81
3,4	144	11	13	143
3,5	225	13	18	225
4,4	256	13	20	256
4,5	400	15	27	400
5,5	625	17	37	625

Note that for $N_a = 2$ there is an intrinsic degeneracy causing the number of constraints from n_c zeros to be $n_c M_2 - 1$, so the actual minimum number of zeros is $n_c + 1 = N$. So the number of constraints is reduced by one. For example, in the 2×4 system the actual number of constraints from $n_c = 7$ random product vectors is 62 instead of $63 = n_c M_2$. This degeneracy implies that one extra zero is needed, thus saving Thm. 11. The proof of Thm. 11 indicates the origin of the degeneracy. With $N - 1$ zeros there exists a vector ψ orthogonal to all the zeros and a vector η orthogonal to all the partially conjugated zeros, and because $\psi\psi^\dagger$ and $(\eta\eta^\dagger)^P$ both lie in $\mathrm{Ker}\,\mathbf{U}_{01}$ we must have $\dim \mathrm{Ker}\,\mathbf{U}_{01} \geq 2$. For other dimensions there is no similar degeneracy.

The 3×4 system is special in that the number of constraints from a generic set of $n_c = 13$ product vectors add up to exactly the number needed to define a unique $A \in H$ with $\mathbf{U}_{01} A = 0$. This does not however imply that A defined in this way from a random set of product vectors will be an extremal witness, in fact it will usually not be a witness, because it will have both positive and negative expectation values in product states.

A central question in general, well worth further attention, is how to choose a set of product vectors such that they may serve as the zeros of an extremal witness. According to Thm. 18, the zeros of a witness must define pure states lying on a face of the set $\mathcal{S}$ of separable states. However, this statement leaves open the practical question of how to test numerically whether a set of pure product states lies on a face of $\mathcal{S}$.

In $2 \times N_b$ and 3×3 systems, the minimum number of zeros of a quadratic extremal witness is exactly equal to the dimension $N = N_a N_b$ of the composite Hilbert space $\mathcal{H}$. This implies the following result, which we state as a theorem.

Theorem 12. *In dimensions $2 \times N_b$ and 3×3 a generic witness with doubly spanning zeros will be extremal.*

We do not know whether this is a general theorem in these dimensions, valid for all witnesses and not only for the generic witnesses. It may be that some construction like the one described in Sec. 6.3.3 will lead to non-generic counterexamples.

In higher dimensions the minimum number of zeros to define a quadratic extremal witness is strictly larger than the dimension of $\mathcal{H}$. This difference between $2 \times N_b$ and 3×3 systems on the one hand, and higher dimensional systems on the other hand, has important consequences discussed in Secs. 8 and 9.

5. Decomposable witnesses

A decomposable witness is called so because it corresponds to a decomposable positive map, and because it has the form $\Omega = \rho + \sigma^P$ with $\rho, \sigma \in \mathcal{D}$, possibly $\rho = 0$ or $\sigma = 0$. Although the decomposable witnesses are useless for detecting entangled PPT states, they are useful in other ways, for example as stepping stones in some of our numerical methods for constructing extremal non-decomposable witnesses.

In this section we will summarize some basic properties of decomposable witnesses. This is a natural place to start when we want to understand entanglement witnesses in general. In particular, we are interested in the relation between a witness and its zeros.

Since the set $\mathcal{P}_1^\circ$ of decomposable witnesses is the convex hull of $\mathcal{D}_1$ and $\mathcal{D}_1^P$, an extremal point of $\mathcal{P}_1^\circ$ must be an extremal point of either $\mathcal{D}_1$ or $\mathcal{D}_1^P$. That is, it must be either a pure state $\psi\psi^\dagger$ or a partially conjugated pure state $(\psi\psi^\dagger)^P$, or both if ψ is a product vector. In the present section we verify the results in Refs. 14, 24, 25, that witnesses of the forms $\psi\psi^\dagger$ and $(\psi\psi^\dagger)^P$ are extremal in $\mathcal{S}_1^\circ$. Since they are extremal in $\mathcal{S}_1^\circ$, they are also extremal in the subset $\mathcal{P}_1^\circ \subseteq \mathcal{S}_1^\circ$. Thus they are precisely the extremal points of $\mathcal{P}_1^\circ$.

5.1. *Zeros of a decomposable witness*

The biquadratic form corresponding to the decomposable witness $\Omega = \rho + \sigma^P$ is

$$f_\Omega(\phi, \chi) = f_\rho(\phi, \chi) + f_\sigma(\phi, \chi^*). \tag{46}$$

It is positive semidefinite because f_ρ and f_σ are positive semidefinite.

Assume now that (ϕ_0, χ_0) is a zero of Ω. The above decomposition of f_Ω shows that (ϕ_0, χ_0) must be a zero of ρ, and the partial conjugate (ϕ_0, χ_0^*) must be a zero of σ. But because $\rho, \sigma \in \mathcal{D}$ it follows that

$$\rho(\phi_0 \otimes \chi_0) = 0\,, \qquad \sigma(\phi_0 \otimes \chi_0^*) = 0. \tag{47}$$

This proves the following theorem.

Theorem 13. *The zeros of a decomposable witness $\Omega = \rho + \sigma^P$ span the Hilbert space only if $\rho = 0$. The partially conjugated zeros span the Hilbert space only if $\sigma = 0$. Hence, by Thm. 11, a quadratic extremal witness is non-decomposable.*

Note that if (ϕ_0, χ_0) is a zero of Ω and $\phi \otimes \chi$ is any product vector, then

$$(\phi \otimes \chi_0)\sigma^P(\phi_0 \otimes \chi) = (\phi \otimes \chi^*)\sigma(\phi_0 \otimes \chi_0^*) = 0, \tag{48}$$

and hence

$$(\phi \otimes \chi_0)\Omega(\phi_0 \otimes \chi) = (\phi \otimes \chi_0)\rho(\phi_0 \otimes \chi). \tag{49}$$

Similarly,

$$(\phi \otimes \chi_0^*)\rho^P(\phi_0 \otimes \chi^*) = (\phi \otimes \chi)\rho(\phi_0 \otimes \chi_0) = 0, \tag{50}$$

and hence

$$(\phi \otimes \chi_0^*)\Omega^P(\phi_0 \otimes \chi^*) = (\phi \otimes \chi_0^*)\sigma(\phi_0 \otimes \chi^*). \tag{51}$$

These equations may be useful for computing ρ and σ if Ω is known but its decomposition $\Omega = \rho + \sigma^P$ is unknown. We will return to this problem in Sec. 5.3.

5.2. *Pure states and partially transposed pure states*

Let P be a pure state, $P = \psi\psi^\dagger$, and let Q be the partial transpose of a pure state, $Q = (\eta\eta^\dagger)^P$, with $\psi, \eta \in \mathcal{H}$. The corresponding positive maps $\mathbf{L}_P$ and $\mathbf{L}_Q$ are rank one preservers: they map matrices of rank one to matrices of rank one or zero, because

$$\begin{aligned} \mathbf{L}_P(\phi\phi^\dagger) &= (\phi \otimes I_b)^\dagger P(\phi \otimes I_b) = \zeta\zeta^\dagger \\ \mathbf{L}_Q(\phi\phi^\dagger) &= (\phi \otimes I_b)^\dagger Q(\phi \otimes I_b) = \theta\theta^\dagger \end{aligned} \tag{52}$$

with $\zeta = (\phi \otimes I_b)^\dagger\psi$ and $\theta = (\phi^* \otimes I_b)^\dagger\eta^*$. The corresponding biquadratic forms are positive semidefinite because they are absolute squares,

$$\begin{aligned} f_P(\phi,\chi) = (\phi \otimes \chi)^\dagger P(\phi \otimes \chi) = |\psi^\dagger(\phi \otimes \chi)|^2, \\ f_Q(\phi,\chi) = (\phi \otimes \chi)^\dagger Q(\phi \otimes \chi) = (\phi \otimes \chi^*)^\dagger Q^P(\phi \otimes \chi^*) = |\eta^\dagger(\phi \otimes \chi^*)|^2. \end{aligned} \tag{53}$$

The zeros of P are the product vectors orthogonal to ψ. A singular value decomposition (Schmidt decomposition) gives orthonormal bases $\{u_i\}$ in $\mathcal{H}_a$ and $\{v_j\}$ in $\mathcal{H}_b$ such that

$$\psi = \sum_{i=1}^{m} c_i\, u_i \otimes v_i \qquad \text{with} \quad c_i > 0. \tag{54}$$

Here m is the Schmidt number of ψ, $1 \le m \le \min(N_a, N_b)$. The condition for a product vector

$$\phi \otimes \chi = \sum_{i=1}^{N_a}\sum_{j=1}^{N_b} a_i b_j\, u_i \otimes v_j \tag{55}$$

to be orthogonal to ψ is that

$$\sum_{i=1}^{m} c_i a_i b_i = 0. \tag{56}$$

For any dimensions $N_a \ge 2$, $N_b \ge 2$ and any Schmidt number m the set of zeros is continuous and connected. All the zeros are quartic, since quadratic zeros are isolated. The zeros do not span the whole Hilbert space, only the subspace orthogonal to ψ. However, the partially conjugated zeros span the Hilbert space, except when $m = 1$ so that ψ is a product vector.[52] This almost completes the proof of the following theorem, what remains is to prove the extremality.

Theorem 14. *A pure state as a witness has only quartic zeros, all in one continuous connected set, and it is extremal in $\mathcal{S}_1^\circ$. The same is true for the partial transpose of a pure state.*

Proof. We want to prove that P is extremal, the proof for Q is similar. We show that the only witness Ω satisfying the constraints $\mathbf{U}_P\Omega = 0$ is $\Omega = P$, from this we conclude that P is extremal, using Thm. 8.

Even though the zeros of P are all quartic, they are so many that we need not use the secondary constraints coming from the quartic nature of the zeros. According to Eq. (30), the constraints $\mathbf{U}_P\Omega = 0$ include the constraints $\mathbf{L}_\Omega(\phi\phi^\dagger)\chi = 0$ for every zero (ϕ, χ) of P. The zeros of P are the solutions of the equation $\psi^\dagger(\phi \otimes \chi) = 0$. For every $\phi \in \mathcal{H}_a$ this is one linear equation for $\chi \in \mathcal{H}_a$, the solutions of which form a subspace of $\mathcal{H}_a$ of dimension either $N_b - 1$ or N_b. But this means that $\mathbf{L}_\Omega(\phi\phi^\dagger)$ has rank either one or zero, so that the positive map $\mathbf{L}_\Omega$ is a rank one preserver. Hence, either $\Omega = \omega\omega^\dagger$ or $\Omega = (\widetilde{\omega}\widetilde{\omega}^\dagger)^P$ for some $\omega, \widetilde{\omega} \in \mathcal{H}$.[24]

In the first case, $\Omega = \omega\omega^\dagger$, we must have $\omega^\dagger(\phi \otimes \chi) = 0$ for every (ϕ, χ) such that $\psi^\dagger(\phi \otimes \chi) = 0$. But then ω must be proportional to ψ, and $\Omega = P$.

In the second case, $\Omega = (\widetilde{\omega}\widetilde{\omega}^\dagger)^P$, we must have $\widetilde{\omega}^\dagger(\phi \otimes \chi^*) = 0$ for every (ϕ, χ) such that $\psi^\dagger(\phi \otimes \chi) = 0$. This is possible, but only if ψ and $\widetilde{\omega}$ are product vectors, with $\widetilde{\omega}$ the partial conjugate of ψ. Then we again have $\Omega = P$. □

The fact that every zero (ϕ_0, χ_0) of P is quartic can also be seen directly from Eq. (21). With $\Omega = P = \psi\psi^\dagger$ and $\psi^\dagger(\phi_0 \otimes \chi_0) = 0$ this equation takes the form

$$z^T G_\Omega\, z = f_2(x, y) = |\psi^\dagger(\xi \otimes \chi_0 + \phi_0 \otimes \zeta)|^2. \tag{57}$$

Thus, $z^T G_\Omega\, z = 0$ unless the vector $(\xi \otimes \chi_0 + \phi_0 \otimes \zeta) \in \mathcal{H}$ has a component along ψ. This means that the Hessian matrix G_Ω at the zero (ϕ_0, χ_0) has rank at most two.

5.3. *Decomposing a decomposable witness*

If we want to prove that a given witness Ω is decomposable, the definitive solution is of course to decompose it as $\Omega = \rho + \sigma^P$ with $\rho, \sigma \in \mathcal{D}$. We want to discuss here methods for doing this, based on a knowledge of zeros of Ω. Unfortunately, the present discussion does not lead to a complete solution of the problem.

Assume that we know a finite set of k zeros of Ω, $Z = \{\phi_i \otimes \chi_i\}$, with partial conjugates $\widetilde{Z} = \{\phi_i \otimes \chi_i^*\}$. Z may be the complete set of zeros of Ω, or only a subset. The orthogonal complement $Z^\perp$ is a d_1 dimensional subspace of $\mathcal{H}$, and $\widetilde{Z}^\perp$ is a d_2 dimensional subspace, with $d_1 \geq N - k$ and $d_2 \geq N - k$. Let P and $\widetilde{P}$ be the orthogonal projections onto $Z^\perp$ and $\widetilde{Z}^\perp$, respectively.

Define orthogonal projections $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ on the real Hilbert space H such that $\mathbf{P}X = PXP$ and $\widetilde{\mathbf{P}}X = (\widetilde{P}X^P\widetilde{P})^P$ for $X \in H$. Define the overlap of these two projections as the orthogonal projection $\mathbf{O}$ onto the subspace $(\mathbf{P}H) \cap (\widetilde{\mathbf{P}}H)$. Since $\mathbf{O}X = X$ if and only if $\mathbf{P}X = X$ and $\widetilde{\mathbf{P}}X = X$, we may compute $\mathbf{O}$ numerically by picking out the eigenvectors of $\mathbf{P} + \widetilde{\mathbf{P}}$ with eigenvalue two. It follows that

$$\mathbf{O} = \mathbf{P}\mathbf{O} = \mathbf{O}\mathbf{P} = \widetilde{\mathbf{P}}\mathbf{O} = \mathbf{O}\widetilde{\mathbf{P}}, \tag{58}$$

and that $\mathbf{P}' = \mathbf{P} - \mathbf{O}$ and $\widetilde{\mathbf{P}}' = \widetilde{\mathbf{P}} - \mathbf{O}$ are orthogonal projections.

If $\Omega = \rho + \sigma^P$ with $\rho, \sigma \in \mathcal{D}$, then we must have $Z \subset \operatorname{Ker}\rho$ and $\widetilde{Z} \subset \operatorname{Ker}\sigma$, hence $\mathbf{P}\rho = \rho$ and $\widetilde{\mathbf{P}}\sigma^P = (\widetilde{P}\sigma\widetilde{P})^P = \sigma^P$. Defining $\rho_1 = \mathbf{P}'\rho$, $\rho_2 = \mathbf{O}\rho$, $\sigma_1^P = \widetilde{\mathbf{P}}'\sigma^P$, $\sigma_2^P = \mathbf{O}\sigma^P$ we have that

$$\Omega = \rho_1 + \rho_2 + \sigma_2^P + \sigma_1^P \tag{59}$$

with $\rho_1 \in \mathbf{P}'H$, $\sigma_1^P \in \widetilde{\mathbf{P}}'H$, and $\rho_2 + \sigma_2^P \in \mathbf{O}H$. It follows that

$$\rho_1 + \sigma_1^P = \Omega - \mathbf{O}\Omega, \qquad \rho_2 + \sigma_2^P = \mathbf{O}\Omega. \tag{60}$$

The decomposition of $\Omega - \mathbf{O}\Omega$ into ρ_1 and σ_1^P is unique and easily computed.

If the overlap $\mathbf{O}$ is zero, then $\rho_2 = 0$, $\sigma_2 = 0$, and this is the end of the story, except that we should check that both ρ_1 and σ_1 are positive semidefinite. Otherwise it remains to decompose $\mathbf{O}\Omega$ into ρ_2 and σ_2^P in such a way that $\rho = \rho_1 + \rho_2$ and $\sigma = \sigma_1 + \sigma_2$ are both positive semidefinite. This is the difficult part of the problem, and the solution, if it exists, need not be unique.

We have not pursued the problem further. To do so one should use Eqs. (49) and (51), which put strong and presumably useful restrictions on ρ and σ. For any product vector $\phi \otimes \chi$ and any $\phi_i \otimes \chi_i \in Z$ it is required that

$$\begin{aligned}(\phi \otimes \chi_i)\rho(\phi_i \otimes \chi) &= (\phi \otimes \chi_i)\Omega(\phi_i \otimes \chi),\\ (\phi \otimes \chi_i^*)\sigma(\phi_i \otimes \chi^*) &= (\phi \otimes \chi_i^*)\Omega^P(\phi_i \otimes \chi^*).\end{aligned} \tag{61}$$

5.4. *Decomposable witnesses with prescribed zeros*

One use of decomposable witnesses is that they provide examples of witnesses with prescribed zeros. We will describe now how this works, and we use the same notation as in the previous subsection.

Let $\Omega = \rho + \sigma^P$ be a decomposable witness, as before. The necessary and sufficient conditions for Ω to have Z as a predefined set of zeros is that $\rho = UU^\dagger$ and $\sigma = VV^\dagger$ with $X, Y \in H$, $U = PX$, $V = \widetilde{P}Y$. We choose ρ and σ to have the maximal ranks

$$\operatorname{rank}\rho = d_1, \qquad \operatorname{rank}\sigma = d_2, \tag{62}$$

where $d_1 \geq N - k$ and $d_2 \geq N - k$ are the dimensions of $Z^\perp$ and $\widetilde{Z}^\perp$ respectively. In the generic case, when $k < N$ there will be no linear dependencies between the zeros, or between their partial conjugates, so that we will have $d_1 = d_2 = N - k$.

The set of unnormalized decomposable witnesses of this form has dimension

$$\begin{aligned}d_D &= \operatorname{rank}\mathbf{P}' + \operatorname{rank}\widetilde{\mathbf{P}}' + \operatorname{rank}\mathbf{O}\\ &= \operatorname{rank}\mathbf{P} + \operatorname{rank}\widetilde{\mathbf{P}} - \operatorname{rank}\mathbf{O}\\ &= d_1^2 + d_2^2 - \operatorname{rank}\mathbf{O}.\end{aligned} \tag{63}$$

The corresponding set of normalized witnesses is a convex subset $\mathcal{F}_D \subset \mathcal{F}_\Omega$ of dimension $d_D - 1$, consisting of witnesses of the form $\Omega + t\Gamma$ with

$$\Gamma = A + B^P, \quad \mathbf{P}A = A, \quad \widetilde{\mathbf{P}}(B^P) = B^P, \quad \operatorname{Tr}A = \operatorname{Tr}B = 0. \tag{64}$$

These conditions ensure that $\rho + tA$ and $\sigma + tB$ will remain positive for small enough positive or negative finite values of the real parameter t.

The witness Ω is an interior point of a unique face of $\mathcal{S}_1^\circ$, which we call $\mathcal{F}_\Omega$, as defined in Eq. (44). The dimension of this face is $d_\Omega - 1$ when we define

$$d_\Omega = \dim \mathrm{Ker}\, \mathbf{U}_\Omega. \tag{65}$$

For a quadratic witness Ω a lower limit is

$$d_\Omega \geq N^2 - kM_2 \qquad \text{with} \qquad M_2 = 2(N_a + N_b) - 3. \tag{66}$$

This inequality will usually be an equality, especially when k is small.

Since $\mathcal{F}_D \subset \mathcal{F}_\Omega$ we must have $d_D \leq d_\Omega$, implying a lower bound for the overlap $\mathbf{O}$,

$$\mathrm{rank}\, \mathbf{O} \geq d_1^2 + d_2^2 - d_\Omega \geq 2(N-k)^2 - d_\Omega. \tag{67}$$

In particular, when the inequality in Eq. (66) is an equality we have the non-trivial lower bound

$$\mathrm{rank}\, \mathbf{O} \geq N^2 - k(4N - 2k - M_2). \tag{68}$$

The generic case when k is small is that $d_D = d_\Omega$, which means that $\mathrm{rank}\, \mathbf{O}$ has the minimal value allowed by this inequality. See Table 2 where we have tabulated some values for d_D and d_Ω found numerically.

When we prescribe k zeros of the decomposable witness Ω, there is a possibility that the actual number of zeros k', is larger than k. A simple counting exercise indicates how many zeros to expect. A product vector $\phi \otimes \chi$ is a zero if and only if $\rho(\phi \otimes \chi) = 0$ and $\sigma(\phi \otimes \chi^*) = 0$. These are $2(N-k)$ complex equations for $\phi \otimes \chi$. There are $N_a + N_b - 2$ complex degrees of freedom in a product vector, up to normalization of each factor. The critical value $k = k_c$ is the number of zeros for which the number of equations equals the number of variables, *i.e.*,

$$k_c = N + 1 - \frac{N_a + N_b}{2}. \tag{69}$$

We expect to find

$$\begin{aligned} k' &= k \quad \text{if} \quad k < k_c, \\ k' &\geq k \quad \text{if} \quad k = k_c, \\ k' &= \infty \quad \text{if} \quad k > k_c. \end{aligned} \tag{70}$$

Whether Ω will be quadratic or quartic can be estimated as follows. The Hessian G_Ω at a zero of Ω is a real square matrix of dimension $2(N_a + N_b - 2)$. If $\mathrm{rank}\, \rho = d_1$ and $\mathrm{rank}\, \sigma = d_2$, then we may write $\Omega = \rho + \sigma^P$ as a convex combination of $d_1 + d_2$ or fewer extremal decomposable witnesses. Each extremal decomposable witness contributes at most two non-zero eigenvalues to G_Ω. Thus, if we define d_H as the minimal dimension of the kernel of G_Ω at any zero, we have a lower bound

$$d_H \geq 2(N_a + N_b - d_1 - d_2 - 2). \tag{71}$$

In Table 2 we list some numbers related to decomposable witnesses, for different dimensions $N_a \times N_b$. The numbers presented were obtained numerically as follows. We choose k random product vectors. When $k < N$ there will exist decomposable witnesses with these as zeros, and d_D is the computed dimension of the set of such witnesses. k' is the actual

number of zeros they have. The listed value of d_Ω, the dimension of the kernel of $\mathbf{U}_\Omega$ when Ω is a quadratic witness with these zeros, is the lower bound given in Eq. (66). The listed value of the dimension d_H of the kernel of the Hessian is the lower bound from Eq. (71).

Table 2. Numbers related to decomposable witnesses. k and k' is the prescribed and actual number of zeros respectively, and k_c is the critical number of prescribed zeros. The parameters d_D, d_Ω and d_H are the dimensions of sets, as explained in the text.

$N_a \times N_b$	k	k'	d_D	d_Ω	d_H
2×4	1	1	55	55	0
$k_c = 6$	2	2	46	46	0
	3	3	37	37	0
	4	4	28	28	0
	5	5	18	19	2
	6	8	8	10	4
	7	∞	2	2	6
3×3	1	1	72	72	0
$k_c = 7$	2	2	63	63	0
	3	3	54	54	0
	4	4	44	45	0
	5	5	32	36	0
	6	6	18	27	2
	7	10,14	8	18	4
	8	∞	2	9	6
3×4	1	1	133	133	0
$k_c = 9.5$	2	2	122	122	0
	3	3	111	111	0
	4	4	100	100	0
	5	5	88	89	0
	6	6	72	78	0
	7	7	50	67	2
	8	8	32	56	4
	9	9	18	45	6
	10	∞	8	34	8
	11	∞	2	23	10
	12	–	0	12	–

We see from the table that with only a few zeros, the set of decomposable witnesses in the face has the same dimension as the face itself. But $\mathcal{F}_\Omega$ can not consist entirely of decomposable witnesses, because in that case our numerical searches for extremal witnesses, where we search for witnesses with increasing numbers of zeros, would only give decomposable witnesses, exactly the opposite of what happens. Hence we conclude that $\mathcal{F}_D$ is a closed subset of $\mathcal{F}_\Omega$, of the same dimension as $\mathcal{F}_\Omega$ but strictly smaller than $\mathcal{F}_\Omega$.

On a face defined by a higher number of zeros the set of decomposable witnesses has a lower dimension. In some cases the decomposable witnesses will lie on the boundary of

the face in question. This will either be because they have more than k zeros, or because some of the k zeros are quartic. In other cases the decomposable witnesses make up a low dimensional part of the interior of the quadratic face. This is the case *e.g* for faces $\mathcal{F}_4$ in 3×3 systems, where the set of decomposable witnesses is 44 dimensional, the face is 45 dimensional, and the decomposable witnesses have four quadratic zeros and hence are situated in the interior of the face.

6. Examples of extremal witnesses

In this section we apply Thm. 10 and Alg. 1 to study examples of extremal witnesses. We study first examples known from the literature, in particular the Choi–Lam witness[21,22] and the Robertson witness.[23] In Secs. 6.3.1, and 6.3.2 we construct numerical examples of generic witnesses, study these and report some observations. Finally, in Sec. 6.3.3 we present an example of a non-generic extremal witness, numerically constructed, having more than the minimum number of zeros.

6.1. *The Choi–Lam witness*

Define, like in Ref. 27,

$$\Omega_K(a,b,c;\theta) = \begin{pmatrix} a & . & . & . & -\mathrm{e}^{\mathrm{i}\theta} & . & . & . & -\mathrm{e}^{-\mathrm{i}\theta} \\ . & c & . & . & . & . & . & . & . \\ . & . & b & . & . & . & . & . & . \\ . & . & . & b & . & . & . & . & . \\ -\mathrm{e}^{-\mathrm{i}\theta} & . & . & . & a & . & . & . & -\mathrm{e}^{\mathrm{i}\theta} \\ . & . & . & . & . & c & . & . & . \\ . & . & . & . & . & . & c & . & . \\ . & . & . & . & . & . & . & b & . \\ -\mathrm{e}^{\mathrm{i}\theta} & . & . & . & -\mathrm{e}^{-\mathrm{i}\theta} & . & . & . & a \end{pmatrix}. \tag{72}$$

We write dots instead of zeros in the matrix to make it more readable. The special case $\Omega_C = \Omega_K(1,0,1;0)$ is the Choi–Lam witness, one of the first examples of a non-decomposable witness.[21,22] The set of zeros of the entanglement witness Ω_C consists of three isolated quartic zeros

$$e_{13} = e_1 \otimes e_3, \quad e_{21} = e_2 \otimes e_1, \quad e_{32} = e_3 \otimes e_2, \tag{73}$$

where e_1, e_2, e_3 are the natural basis vectors in $\mathbb{C}^3$, and a continuum of zeros $\phi \otimes \chi$ where $\alpha, \beta \in \mathbb{R}$ and

$$\phi = e_1 + \mathrm{e}^{i\alpha}\, e_2 + \mathrm{e}^{i\beta}\, e_3, \qquad \chi = \phi^*. \tag{74}$$

The product vectors defined in Eq. (74) span a seven dimensional subspace consisting of all vectors $\psi \in \mathbb{C}^9$ with components $\psi_1 = \psi_5 = \psi_9$. The three product vectors defined in Eq. (73) have $\psi_1 = \psi_5 = \psi_9 = 0$ and lie in the same subspace.

The Hessian has a doubly degenerate kernel at each of these zeros. Hence a single zero of $\Omega = \Omega_C$ contributes 29 equations in $\mathbf{U}_\Omega$: nine from $\mathbf{T}_0$ and $\mathbf{T}_1$, $2 \cdot 8 = 16$ from $\mathbf{T}_2$, and four from $\mathbf{T}_3$. However, by a singular value decomposition of $\mathbf{T} = \mathbf{T}_0 \oplus \mathbf{T}_1 \oplus \mathbf{T}_2 \oplus \mathbf{T}_3$ at

one zero at a time we find numerically that the number of independent equations is 24 for each of the isolated zeros and 28 for any randomly chosen non-isolated zero. We again see that redundant equations appear when the kernel of the Hessian is more than one dimensional. We also observe that the redundancies depend on the nature of the zero.

Choosing increasingly many non-isolated zeros, only 67 linearly independent equations are obtained, out of the 80 needed for proving extremality. These 67 equations are obtained with the quadratic and quartic constraints from three zeros, or with only the quadratic constraints from nine zeros. With all three isolated zeros and any single zero from the continuum, $\operatorname{Ker} \mathbf{U}_\Omega$ is one dimensional and uniquely defines the Choi–Lam witness. This verifies numerically that it is extremal. We need the quartic constraints from the three isolated zeros in order to prove extremality, because the quadratic constraints from all the zeros provide only 76 independent equations, leaving a four dimensional face of witnesses having the same set of zeros as the Choi–Lam witness Ω_C. This proves that Ω_C is not exposed, but is an extremal point of a four dimensional exposed face of $\mathcal{S}_1^\circ$.

Equation (72) with $\theta = 0$ and the restrictions $0 \le a \le 1$, $a + b + c = 2$, $bc = (1-a)^2$, defines more generally a one parameter family of extremal witnesses considered by Ha and Kye.[53,54] They prove that $\Omega_K(a,b,c;0)$ is both extremal and exposed for $0 < a < 1$. The original Choi–Lam witness Ω_C is extremal but not exposed, it is the limiting case $a = c = 1$, $b = 0$. We will return in Sec. 8 to a more detailed discussion of the facial structure of the set of separable states and the set of witnesses.

We have verified by our numerical methods, for several values of a with $0 < a < 1$, that $\Omega_K(a,b,c;0)$ is indeed extremal. As explained in Ref. 54 there are four classes of zeros. The zeros in one of these classes have Hessians with two dimensional kernels, while the Hessians of the zeros in the other three classes have one dimensional kernels. It turns out that a set of four zeros, one from each class, uniquely defines $\Omega_K(a,b,c;0)$ as the only solution to the constraints imposed by the zeros when utilizing both quadratic and quartic constraints. This shows numerically that the witness is extremal.

The more general case with $\theta \neq 0$ has been treated as an example of optimal, and in fact extremal, witnesses with structural physical approximations that are entangled PPT states.[33] We will return to these concepts in Sec. 10.

6.2. *The Robertson witness*

Another example we have studied is the extremal positive map in dimension 4×4 introduced by Robertson,[23]

$$X \to \begin{pmatrix} X_{33}+X_{44} & 0 & X_{13}+X_{42} & X_{14}-X_{32} \\ 0 & X_{33}+X_{44} & X_{23}-X_{41} & X_{24}+X_{31} \\ X_{31}+X_{24} & X_{32}-X_{14} & X_{11}+X_{22} & 0 \\ X_{41}-X_{23} & X_{42}+X_{13} & 0 & X_{11}+X_{22} \end{pmatrix}. \tag{75}$$

The corresponding entanglement witness Ω_R, has the biquadratic form

$$\begin{aligned} f_R(\phi,\chi) = {} & (|\phi_1|^2+|\phi_2|^2)(|\chi_3|^2+|\chi_4|^2) + (|\phi_3|^2+|\phi_4|^2)(|\chi_1|^2+|\chi_2|^2) \\ & + 2\,\mathrm{Re}\,[(\phi_1^*\phi_3+\phi_4^*\phi_2)(\chi_3^*\chi_1+\chi_2^*\chi_4) + (\phi_2^*\phi_3-\phi_4^*\phi_1)(\chi_3^*\chi_2-\chi_1^*\chi_4)]. \end{aligned} \tag{76}$$

Every product vector $\phi \otimes \chi$ with $\phi_1 = \phi_2 = \chi_1 = \chi_2 = 0$ or $\phi_3 = \phi_4 = \chi_3 = \chi_4 = 0$ is a zero.

More generally, $\phi \otimes \chi$ is a zero if

$$(|\phi_1|^2 + |\phi_2|^2)(|\chi_3|^2 + |\chi_4|^2) = (|\phi_3|^2 + |\phi_4|^2)(|\chi_1|^2 + |\chi_2|^2) \tag{77}$$

and

$$\begin{aligned} \phi_1^*\phi_3 + \phi_4^*\phi_2 = &\quad -\chi_1^*\chi_3 - \chi_4^*\chi_2, \\ \phi_2^*\phi_3 - \phi_4^*\phi_1 = &\quad -\chi_2^*\chi_3 + \chi_4^*\chi_1. \end{aligned} \tag{78}$$

For any given ϕ we obtain a continuum of zeros in the following way. Define

$$a = \phi_1^*\phi_3 + \phi_4^*\phi_2, \qquad b = \phi_2^*\phi_3 - \phi_4^*\phi_1. \tag{79}$$

Then choose χ_1, χ_2 at random and define

$$\chi_3 = \frac{-a\chi_1 - b\chi_2}{|\chi_1|^2 + |\chi_2|^2}, \qquad \chi_4 = \frac{b^*\chi_1 - a^*\chi_2}{|\chi_1|^2 + |\chi_2|^2}. \tag{80}$$

This solves Eq. (78). In order to solve also Eq. (77), rescale $\chi_i \to c\chi_i$ for $i = 1, 2$ and $\chi_i \to \chi_i/c$ for $i = 3, 4$ with a suitably chosen constant $c > 0$.

The entanglement witness Ω_R itself is a 16×16 Hermitian matrix, and by Eq. (9) it corresponds to the witness

$$\Omega_R = \begin{pmatrix}
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . & 1 & . & . & -1 & . \\
. & . & 1 & . & . & . & . & . & 1 & . & . & . & . & . & . & . \\
. & . & . & 1 & . & . & . & . & . & . & . & . & 1 & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . & -1 & . & . & 1 & . \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & 1 & . & . & 1 & . & . & . & . & . & . \\
. & . & . & . & . & . & . & 1 & . & . & . & . & . & 1 & . & . \\
. & . & 1 & . & . & . & . & . & 1 & . & . & . & . & . & . & . \\
. & . & . & . & . & . & 1 & . & . & 1 & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\
. & 1 & . & . & -1 & . & . & . & . & . & . & . & . & . & . & . \\
. & . & . & 1 & . & . & . & . & . & . & . & . & 1 & . & . & . \\
. & . & . & . & . & . & . & 1 & . & . & . & . & . & 1 & . & . \\
. & -1 & . & . & 1 & . & . & . & . & . & . & . & . & . & . & . \\
. & . & . & . & . & . & . & . & . & . & . & . & . & . & . & .
\end{pmatrix}. \tag{81}$$

We have verified numerically that Ω_R is extremal. Four zeros determine the witness uniquely through the quadratic and quartic constraints, *e.g.* the zeros

$$e_{ij} = e_i \otimes e_j, \;\; ij = 11, 12, 33, 34. \tag{82}$$

It appears that all zeros have Hessians with eight dimensional kernels. The witness has a continuum of zeros, and 20 randomly chosen of these turn out to also uniquely determine the witness through only quadratic constraints. The fact that the quadratic constraints are sufficient to determine the witness uniquely proves that it is exposed. See Sec. 8 for a discussion of exposed faces.

As a related example we have looked at a witness Ω_Z belonging to a new class of witnesses in dimension $N \times (2K)$ introduced by Zwolak and Chruściński.[28,55,56] We take $N = 2$, $K = 1$, and $|z_{12}| = 1$ in their notation. Structurally similar to Ω_R, Ω_Z also has the zeros defined in Eq. (82), and again the quadratic and quartic constraints from these four zeros determine Ω_Z uniquely. This verifies numerically that Ω_Z is extremal, and exemplifies the fact that the same set of zeros with different Hessian zeros can uniquely specify two different witnesses. This witness, like Ω_R, has a continuous set of zeros and can be determined uniquely through only the quadratic constraints from a randomly chosen set of 20 of these zeros. Hence it is exposed.

6.3. *Numerical examples, generic and non-generic*

We have successfully implemented Alg. 1 and used it to locate numerical examples of extremal witnesses in 2×4, 3×3 and 3×4 dimensions. In the process of searching for an extremal witness we produce witnesses situated on a hierarchy of successively lower dimensional faces of $\mathcal{S}_1^\circ$.

We define extremal witnesses to be generic if such witnesses can be found with non-zero probability by means of Alg. 1 for random search directions Γ in every iteration. A vast majority of the extremal witnesses found in numerical random searches are quadratic, but a small number of quartic witnesses are found. There may be numerical problems in locating a zero which is quartic or close to quartic. In such cases our implementation of the algorithm stops prematurely. We discuss the quartic witnesses in Sec. 6.3.2.

Line 5 of Alg. 1 regards locating the boundary of a face. Appendix B describes how this can be formulated as a problem of locating a simple root of a special function, and also mentions other possible approaches.

6.3.1. *Quadratic extremal witnesses*

We make the following comments concerning the quadratic extremal witnesses found.

- We have experienced premature stops due to numerical problems with zeros that are close to quartic. See further comments in Sec. 6.3.2.
- When no existing zero becomes close to quartic, a single new quadratic zero appears in every iteration of Alg. 1 when the boundary of the current face is reached. For a small number of zeros, there is no redundancy between constraints from the existing zeros and constraints from the new zero. A redundancy appears typically with the seventh zero in 2×4, with the ninth zero in 3×3, and never in 3×4. This gives a hierarchy of faces of $\mathcal{S}_1^\circ$ of dimension $N^2 - 1 - kM_2$, where k is the number of zeros of a witness in the interior of the face, and $M_2 = 2(N_a + N_b) - 3$.
- The extremal witnesses have the expected number of zeros as listed in Table 1. The zeros and the partially conjugated zeros span $\mathcal{H}$, as required by Thm. 11.
- A quadratic extremal witness has at least one negative eigenvalue, and the same is true for its partial transpose. We do not find extremal witnesses with more than three negative eigenvalues in 3×3 or more than four negative eigenvalues in 3×4. See Fig. 2 for the details.
- Every witness and its partial transpose have full rank.

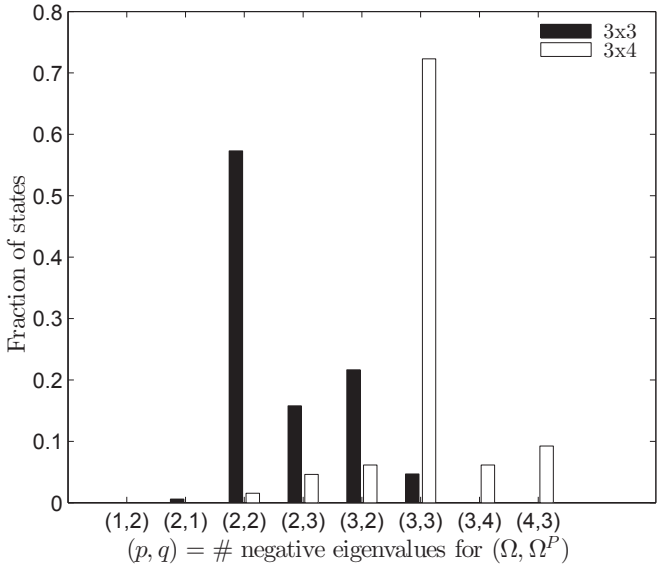


Fig. 2. Classification of generic quadratic extremal witnesses found numerically by Alg. 1 in dimensions 3×3 and 3×4. A witness Ω of type (p, q) has p negative eigenvalues, and its partial transpose Ω^P has q negative eigenvalues.

6.3.2. *Quartic extremal witnesses*

In each iteration of Alg. 1 the critical parameter value t_c is reached when either a new quadratic zero appears, or a new zero eigenvalue of the Hessian matrix appears at one of the existing zeros. If the second alternative occurs at least once during a search, then the extremal witness found will be quartic, otherwise it will be quadratic. Our experience is that a random search most often produces a quadratic extremal witness and only rarely a quartic extremal witness.

In order to make this observation more quantitative we have made random searches for quartic witnesses in dimension 3×3 as follows. We take 58 hierarchies of faces of quadratic witnesses generated by Alg. 1, and generate 100 random perturbations Γ away from the quadratic witness found on each face. Since there are eight faces in each of the 58 hierarchies, this is a total of 46400 tests. For each test we compute t_c' as the smallest t resulting in a zero eigenvalue of the Hessian at a zero. t_c' is thus an upper bound for t_c. At $t = t_c'$ we test whether $\Omega + t\Gamma$ is still a witness, in which case it has got a quartic zero, or if it is no longer a witness, which will mean that there has appeared a new quadratic zero for some $t < t_c'$. Three runs failed, probably due to some bug in the algorithm. With a tolerance of $\pm 10^{-14}$ on function values we found that only 91 out of 46397 successful tests resulted in quartic witnesses. Thus, in this quantitative test the probability for finding a quartic witness was $\sim 0.2\,\%$. We conclude that the fraction of quartic extremal witnesses among the generic extremal witnesses is small but non-zero.

We constructed an explicit example of a quartic extremal witness in 3×3 with eight zeros in the following way. Starting at a quadratic witness on a face $\mathcal{F}_4$ generated by Alg. 1 we found a quartic witness on the boundary of $\mathcal{F}_4$ by choosing one quadratic

zero to become quartic and perturbing the quadratic witness accordingly (this may not succeed, in which case we choose a different zero). Continuing Alg. 1 from this quartic witness resulted in an extremal witness with eight zeros, of which one is a quartic zero with a one dimensional kernel of the Hessian. Running Alg. 1 succeeded since we knew a priori which zero was quartic, so problems with the numerical precision could be overcome. We observe that as $M_2 = 9$ constraints from each of the six first quadratic zeros are added to $\mathbf{U}_{01}$, and further as from (28) and (36) the $M_2 + M_4(1) = 18$ constraints from the quartic zero are added to form $\mathbf{U}_\Omega$, all of these are linearly independent. Hence $\mathbf{U}_\Omega$ has rank 72 as expected, and a single new quadratic zero is located in the final step, adding eight new linearly independent constraints so as to make $\dim \mathrm{Ker}\, \mathbf{U}_\Omega = 1$.

6.3.3. *non-generic quadratic extremal witnesses*

In order to demonstrate that a quadratic extremal witness may have more than the minimum number of zeros, we have constructed quadratic extremal witnesses in 3×3 dimensions with ten rather than the expected number of nine zeros.

One method for finding such witnesses is illustrated in Fig. 4, where the kink in the curve to the right represents a witness with ten zeros. All witnesses on the face have eight zeros in common, whereas the curved lines on both sides of the kink consist of quadratic extremal witnesses with nine zeros. We will return to this example in the next section.

An entirely different method is described here. The basic idea is to use Thm. 11 and try to construct nine product vectors $\psi_i = \phi_i \otimes \chi_i$ that are linearly dependent, and for good measure such that also the partially conjugated product vectors $\widetilde{\psi}_i = \phi_i \otimes \chi_i^*$ are linearly dependent. Then a quadratic witness with these nine zeros, and no more, can not be extremal, but it might lead to a quadratic extremal witness with ten zeros.

A first attempt is to choose the nine product vectors directly, by minimizing the sum of the smallest singular value of the 9×9 matrix $\psi = [\psi_1, \psi_2, \ldots, \psi_9]$ and the smallest singular value of the corresponding 9×9 matrix $\widetilde{\psi}$. Since singular values are non-negative by definition, minimization will make both these singular values zero. The minimization problem is solved *e.g.* by the Nelder–Mead algorithm, or by a random search.

Let Z be a set of nine product vectors generated in this way, then any witness Ω with these as zeros has to satisfy the constraints $\mathbf{U}_{01}\Omega = 0$. The generic result we find is that $\dim \mathrm{Ker}\, \mathbf{U}_{01} = 2$. Hence Ω has to be a decomposable witness,

$$\Omega = p\, \eta\eta^\dagger + (1-p)\, (\widetilde{\eta}\widetilde{\eta}^\dagger)^P \quad \text{with} \quad 0 \le p \le 1, \tag{83}$$

and $\eta, \widetilde{\eta} \in \mathcal{H}$, $\eta^\dagger \psi_i = \widetilde{\eta}^\dagger \widetilde{\psi}_i = 0$ for $i = 1, \ldots, 9$. This decomposable Ω is not what we are looking for, in fact it has a continuum of quartic zeros. The extremal decomposable witnesses $\eta\eta^\dagger$ and $(\widetilde{\eta}\widetilde{\eta}^\dagger)^P$ each contribute two non-zero eigenvalues to the 8×8 Hessian matrix G_Ω at every zero, hence $\dim \mathrm{Ker}\, G_\Omega = 4$ at every zero, unless $p = 0$ or $p = 1$ in which case $\dim \mathrm{Ker}\, G_\Omega = 6$.

A second attempt is to choose eight product vectors ψ_i that are linearly dependent and have linearly dependent partial conjugates $\widetilde{\psi}_i$. This almost works, but not quite. We may construct a decomposable witness Ω with these zeros from two pure states orthogonal to all ψ_i and two partially transposed pure states orthogonal to all $\widetilde{\psi}_i$. As a convex combination of four extremal decomposable witnesses, each contributing two non-zero eigenvalues to

each 8×8 Hessian matrix, Ω will be quadratic. Ω is determined by eight real parameters, including a normalization constant, since $\Omega = \rho + \sigma^P$ where ρ and σ are positive matrices of rank two, with $\rho\psi_i = \sigma\widetilde{\psi}_i = 0$. The maximum number of independent constraints from eight quadratic zeros is 72, giving $\dim \mathrm{Ker}\, \mathbf{U}_{01} = 9$, one more than the dimension of the set of decomposable witnesses having these zeros. However, we find that Ω generically has more than the eight prescribed zeros, increasing the number of independent constraints to 73. Hence, again the constraints leave only the decomposable witnesses, none of which are quadratic and extremal.

The third and successful attempt is to choose seven product vectors with similar linear dependencies. Denote by $\mathbf{U}_{01}^{(7)}$ the corresponding linear system of constraints. The kernel of $\mathbf{U}_{01}^{(7)}$ is found, in two different cases, to have dimension 21 or 22. A decomposable witness with the given product vectors as zeros has the form $\Omega_7 = \rho + \sigma^P$ where ρ and σ are positive matrices of rank three. Hence, the set of such decomposable witnesses is 18 dimensional, so that there are three or four dimensions in $\mathrm{Ker}\, \mathbf{U}_{01}^{(7)}$ orthogonal to the face of decomposable witnesses. Defining Γ_7 to lie in these three or four dimensions one can walk towards the boundary of the face $\mathcal{F}_7 = (\mathrm{Ker}\, \mathbf{U}_{01}^{(7)}) \cap \mathcal{S}_1^\circ$, finding $\Omega_8 = \Omega_7 + t_c\Gamma_7$ with eight zeros. Ω_8 is now guaranteed to be non-decomposable. Let $\mathbf{U}_{01}^{(8)}$ be the system defined by these eight zeros, defining the face $\mathcal{F}_8 = (\mathrm{Ker}\, \mathbf{U}_{01}^{(8)}) \cap \mathcal{S}_1^\circ$. We find that $\mathrm{Ker}\, \mathbf{U}_{01}^{(8)}$ has dimension nine less than $\mathrm{Ker}\, \mathbf{U}_{01}^{(7)}$. Defining $\Gamma_8 \in \mathrm{Ker}\, \mathbf{U}_{01}^{(8)}$, we locate an $\Omega_9 \in \mathcal{F}_9$, on the boundary of $\mathcal{F}_8$, with nine zeros. The kernel of $\mathbf{U}_{01}^{(9)}$ has dimension nine less than $\mathrm{Ker}\, \mathbf{U}_{01}^{(8)}$, *i.e.* three or four, hence there is still freedom to move along $\mathcal{F}_9$. Doing so produces a quadratic extremal witness in 3×3 with ten zeros rather than nine.

7. D-shaped faces of the set of witnesses in low dimensions

In this section we reveal a special geometry of next-to-extremal faces of $\mathcal{S}_1^\circ$ in 2×4 and 3×3 systems, related to the presence of decomposable witnesses.

Let $\mathcal{F}_k$ denote a face of $\mathcal{S}_1^\circ$ with interior points that are quadratic witnesses with k zeros. This is typically what we find in the k-th iteration of Alg. 1. A face $\mathcal{F}_7$ in dimension 2×4, or $\mathcal{F}_8$ in dimension 3×3, is the last face found before an extremal quadratic witness is reached. These particular faces have a special geometry, because the number of zeros is one less than the dimension of the Hilbert space, and as a result part of the boundary is a line segment of decomposable witnesses.

A decomposable witness on such a face has the form

$$\Omega = (1-p)\,\psi\psi^\dagger + p\,(\eta\eta^\dagger)^P \quad \text{with} \quad 0 \le p \le 1, \tag{84}$$

where ψ is orthogonal to the $N-1$ product vectors $\phi_i \otimes \chi_i$ that are the zeros of all the witnesses in the interior of the face, and η is orthogonal to the partially conjugated product vectors $\phi_i \otimes \chi_i^*$. We take ψ and η to be normalized vectors, $\psi^\dagger\psi = \eta^\dagger\eta = 1$.

When we apply Alg. 1 in 2×4 dimensions and find a quadratic extremal witness, the generic case is that the face $\mathcal{F}_7$ is two dimensional. An example is shown in Fig. 3. The line segment of decomposable witnesses must be part of the boundary of the face $\mathcal{F}_7$, because the interior of the face consists of quadratic witnesses with a fixed set of seven quadratic zeros, whereas the decomposable witnesses have additional zeros and Hessian zeros, in fact infinitely many quartic zeros. The rest of the boundary of the face is curved,

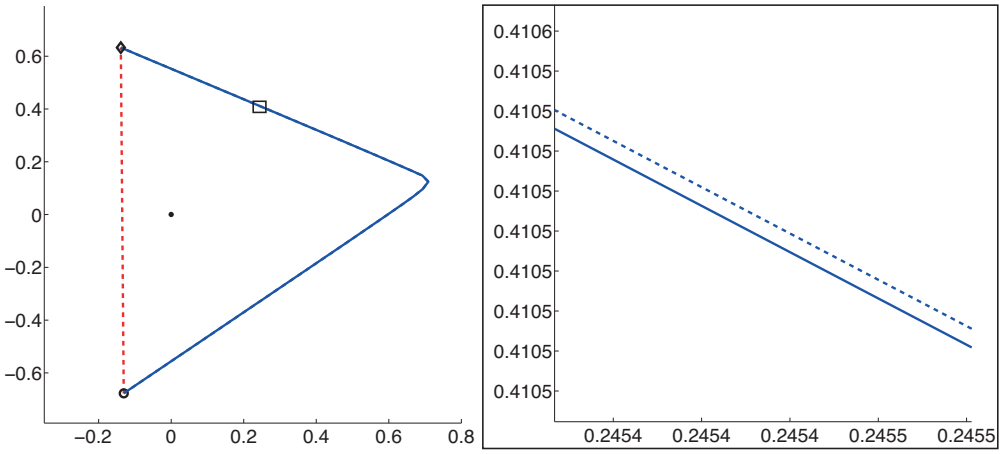


Fig. 3. A two dimensional face $\mathcal{F}_7 \subset \mathcal{S}_1^\circ$ obtained by applying Alg. 1 in dimension 2×4. Distances are defined by the Hilbert–Schmidt metric. The straight line segment (dashed, in red) consists of decomposable witnesses, its upper end point is a pure state, marked by a diamond, and its lower end point is the partial transpose of a pure state, marked by a circle. Starting from a quadratic witness with seven zeros, marked by the black dot, the curved boundary of the face (in blue) was located by perturbing in all directions. Every point on this curve represents an extremal witness with eight quadratic zeros, except at the kink in the middle of the curved boundary where the witness has nine quadratic zeros. These extremal witnesses are very nearly quartic. We show this by drawing another curved line (dashed, in blue) where the first Hessian zero appears, that is, where one zero becomes quartic. This dashed curve is only visible in the enlarged part of the figure.

and consists of quadratic extremal witnesses with eight zeros. There is one exception, however, seen in the figure as a kink in the curved part of the boundary, and this is a quadratic extremal witness with nine zeros.

The interesting explanation is as follows. As we go along the curve consisting of quadratic witnesses with eight zeros, seven zeros are fixed, they are the zeros defining the face. But the eighth zero has to change along the curve, because any two points on this part of the boundary can be joined by a line segment passing through the interior of the face, hence these two boundary points can have only the seven zeros in common. Starting from the two extremal decomposable witnesses $\psi\psi^\dagger$ and $(\eta\eta^\dagger)^P$ we get two curved sections of the boundary where the eighth zero changes continuously. These two sections meet in one point which is then a witness with two quadratic zeros in addition to the seven zeros defining the face.

Accordingly, this next-to-extremal face $\mathcal{F}_7$ is two dimensional and has the shape of a "D", where the straight edge is the line segment of decomposable witnesses and the round part consists of quadratic extremal witnesses. Figure 3 is a numerically produced example of such a face. Similarly we expect that next-to-extremal faces in any $2 \times N_b$ systems will be either D-shaped, or if we by accident hit the straight edge of the D, line segments.

In the case of $\mathcal{F}_8$ in 3×3 we can also construct the line segment of decomposable witnesses on the boundary of $\mathcal{F}_8$. The remaining boundary will again consist of quadratic extremal witnesses, this time with nine or ten zeros. This part of the boundary is a curved seven dimensional surface, since the face itself is eight dimensional. Any two dimensional section of $\mathcal{F}_8$ passing through the line segment of decomposable witnesses is shaped as a

D. See Fig. 4 for a numerically computed example.

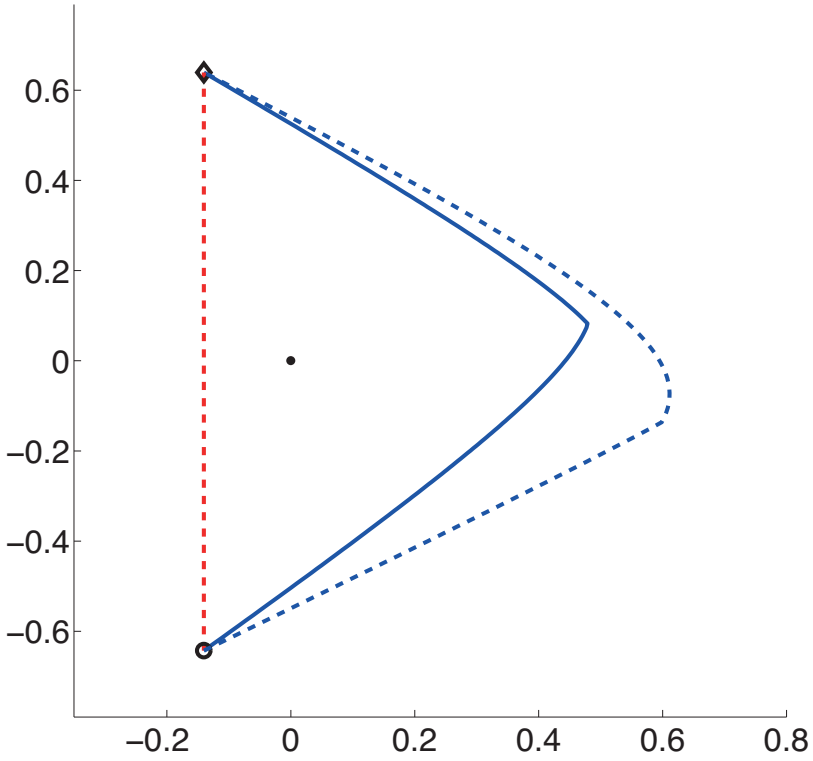


Fig. 4. A special two dimensional section of an eight dimensional face $\mathcal{F}_8 \subset \mathcal{S}_1^\circ$ obtained from output of Alg. 1 in dimension 3×3. This section passes through a straight line segment of decomposable witnesses (dashed, in red) which is part of the boundary of the face. The upper end point of the line segment is a pure state, marked by a diamond, and the lower end point is the partial transpose of a pure state, marked by a circle. Starting from a quadratic witness with eight zeros, marked by the black dot, the curved boundary of the face (in blue) was located by perturbing in all directions in the plane. Every point on this curve represents an extremal witness with nine quadratic zeros, except at the kink where the witness has ten zeros. The curved dashed line (in blue) is where one zero becomes quartic.

Note that we can not guarantee that any choice of eight random product vectors gives rise to such a D. Many choices may give rise to only the line segment of decomposable witnesses, since there are inequality constraints that are not automatically satisfied, even if we are able to satisfy the equality constraints that we have discussed here.

In other dimensions line segments of decomposable witnesses constructed from $N-1$ zeros also define D-shaped faces, but the round part of the "D" will in those cases consist not of extremal witnesses but of lower dimensional faces.

8. Faces of the set of separable states

Our understanding of witnesses as exposed in Sec. 4 translates into an understanding of faces of $\mathcal{S}_1$, the set of separable states. A face of a compact convex set is defined by the extremal points it contains, and in the case of $\mathcal{S}_1$ the extremal points are the pure product

states. Alfsen and Shultz[16] describe faces of $\mathcal{S}_1$ in two categories, either simplexes defined by at most $\max(N_a, N_b)$ pure product states, or direct convex sums of faces isomorphic to matrix algebras. According to our understanding these two categories correspond to quadratic and certain quartic witnesses respectively. It has been known for some time that the set of entanglement witnesses has unexposed faces. These questions have been studied by Chruściński and collaborators.[28,57] In our understanding the unexposed faces of $\mathcal{S}_1^\circ$ are those containing quartic witnesses as interior points. The facial structure of various sets related to quantum entanglement has been studied especially by Kye and collaborators.[58]

In this section we describe faces of $\mathcal{S}_1$ defined by different types of witnesses. We state some basic results which may be well known and which are actually quite generally valid for any pair of dual convex cones. The distinction between exposed and unexposed faces is central, and it would be interesting to know whether all the faces of the set of separable states are exposed. We believe that this is true, although we have no proof. Some support for our conjecture may be drawn from the known facts that both the set of density matrices and the set of PPT states have only exposed faces.

To finish this section we point out how the facial structure of the set of separable states is related to a question which is of practical importance when we want to test whether a given state is separable. The question is how many pure product states we need, if the state is separable, in order to write it as a convex combination of pure product states.

8.1. *Duality of faces*

Given any subset $\mathcal{X} \subset \mathcal{S}_1$ we define its dual in $\mathcal{S}_1^\circ$ as

$$\mathcal{X}^\circ = \{\, \Omega \in \mathcal{S}_1^\circ \mid \operatorname{Tr} \Omega\rho = 0 \ \ \forall \rho \in \mathcal{X} \,\}. \tag{85}$$

Similarly, given any subset $\mathcal{Y} \subset \mathcal{S}_1^\circ$ we define its dual in $\mathcal{S}_1$ as

$$\mathcal{Y}^\circ = \{\, \rho \in \mathcal{S}_1 \mid \operatorname{Tr} \Omega\rho = 0 \ \ \forall \Omega \in \mathcal{Y} \,\}. \tag{86}$$

We will assume here that $\mathcal{X}^\circ$ and $\mathcal{Y}^\circ$ are non-empty, a minimum requirement is that $\mathcal{X} \subset \partial\mathcal{S}_1$ and $\mathcal{Y} \subset \partial\mathcal{S}_1^\circ$. Then also $\mathcal{X}^\circ \subset \partial\mathcal{S}_1^\circ$ and $\mathcal{Y}^\circ \subset \partial\mathcal{S}_1$.

There always exists one single $\rho_0 \in \partial\mathcal{S}_1$ such that

$$\mathcal{X}^\circ = \{\rho_0\}^\circ = \{\, \Omega \in \mathcal{S}_1^\circ \mid \operatorname{Tr} \Omega\rho_0 = 0 \,\}. \tag{87}$$

In fact, every $\rho \in \mathcal{X}$ gives one linear constraint $\operatorname{Tr} \Omega\rho = 0$ as part of the definition of $\mathcal{X}^\circ$. In finite dimension at most a finite number of linear constraints can be independent, hence

$$\mathcal{X}^\circ = \{\, \Omega \in \mathcal{S}_1^\circ \mid \operatorname{Tr} \Omega\rho_i = 0 \ \text{ for } \ i = 1, 2, \ldots, k \,\} \tag{88}$$

for some states $\rho_1, \rho_2, \ldots, \rho_k \in \mathcal{X}$. Define for example

$$\rho_0 = \frac{1}{k} \sum_{i=1}^{k} \rho_i. \tag{89}$$

If $\mathcal{X}$ is not convex, it may happen that $\rho_0 \notin \mathcal{X}$. Because $\operatorname{Tr} \Omega\rho_i \geq 0$ for $i = 1, 2, \ldots, k$ the equation $\operatorname{Tr} \Omega\rho_0 = 0$ implies that $\operatorname{Tr} \Omega\rho_i = 0$ for $i = 1, 2, \ldots, k$ and hence $\Omega \in \mathcal{X}^\circ$.

Taking three times the dual we get that $\mathcal{X}^\circ = \mathcal{F}^\circ$ where $\mathcal{F} = \mathcal{X}^{\circ\circ}$ is the double dual of $\mathcal{X}$. Clearly $\mathcal{F}$ contains $\mathcal{X}$, and by our next theorem $\mathcal{F}$ is an exposed face of $\mathcal{S}_1$.

Reasoning in the same way we conclude that it is always possible to find one single witness Ω_0 such that

$$\mathcal{Y}^\circ = \{\Omega_0\}^\circ = \{\, \rho \in \mathcal{S}_1 \mid \mathrm{Tr}\, \Omega_0 \rho = 0 \,\}, \tag{90}$$

and an exposed face $\mathcal{G}$ of $\mathcal{S}_1^\circ$ containing $\mathcal{Y}$, in fact $\mathcal{G} = \mathcal{Y}^{\circ\circ}$, such that $\mathcal{Y}^\circ = \mathcal{G}^\circ$.

Theorem 15. *$\mathcal{X}^\circ$ is an exposed face of $\mathcal{S}_1^\circ$, and $\mathcal{Y}^\circ$ is an exposed face of $\mathcal{S}_1$.*

Proof. Assume that $\Omega \in \mathcal{X}^\circ$ is a proper convex combination of $\Omega_1, \Omega_2 \in \mathcal{S}_1^\circ$,

$$\Omega = (1-p)\,\Omega_1 + p\,\Omega_2 \qquad \text{with} \qquad 0 < p < 1. \tag{91}$$

We have to prove that $\Omega_1, \Omega_2 \in \mathcal{X}^\circ$. The assumption that $\Omega_1, \Omega_2 \in \mathcal{S}_1^\circ$ means that $\mathrm{Tr}\, \Omega_1 \rho \geq 0$ and $\mathrm{Tr}\, \Omega_2 \rho \geq 0$ for every $\rho \in \mathcal{S}_1$. For every $\rho \in \mathcal{X}$ we have in addition that

$$0 = \mathrm{Tr}\, \Omega \rho = (1-p)\, \mathrm{Tr}\, \Omega_1 \rho + p\, \mathrm{Tr}\, \Omega_2 \rho, \tag{92}$$

and from this we conclude that $\mathrm{Tr}\, \Omega_1 \rho = \mathrm{Tr}\, \Omega_2 \rho = 0$. This proves that $\Omega_1, \Omega_2 \in \mathcal{X}^\circ$, so that $\mathcal{X}^\circ$ is a face.

It is exposed because it is dual to one single $\rho_0 \in \mathcal{S}_1$. In fact, the equation $\mathrm{Tr}\, \Lambda \rho_0 = 0$ for Λ defines a hyperplane of dimension $N^2 - 2$ in the $N^2 - 1$ dimensional affine space of Hermitian $N \times N$ matrices of unit trace.

The proof that $\mathcal{Y}^\circ$ is an exposed face is entirely similar. □

This theorem has the following converse.

Theorem 16. *An exposed face of $\mathcal{S}_1^\circ$ is the dual of a separable state, and an exposed face of $\mathcal{S}_1$ is the dual of a witness.*

Proof. We prove the second half of the theorem, the first half is proved in a similar way.

An exposed face $\mathcal{F}$ of $\mathcal{S}_1$ is the intersection of $\mathcal{S}_1$ with a hyperplane given by an equation for $\rho \in H$ of the form $\mathrm{Tr}\, \Lambda \rho = 0$ where $\Lambda \in H$ is fixed. The maximally mixed state $\rho_0 = I/N$ is an interior point of $\mathcal{S}_1$, hence $\mathrm{Tr}\, \Lambda = N \mathrm{Tr}\, \Lambda \rho_0 \neq 0$ and we may impose the normalization condition $\mathrm{Tr}\, \Lambda = 1$. We have to prove that $\Lambda \in \mathcal{S}_1^\circ$, which means that $\mathrm{Tr}\, \Lambda \rho \geq 0$ for all $\rho \in \mathcal{S}_1$.

Assume that there exists some $\rho_1 \in \mathcal{S}_1$ with $\mathrm{Tr}\, \Lambda \rho_1 < 0$. Choose any $\rho_2 \in \mathcal{S}_1$ with $\mathrm{Tr}\, \Lambda \rho_2 > 0$, for example $\rho_2 = \rho_0$. Then $\rho_1, \rho_2 \notin \mathcal{F}$, since $\mathcal{F}$ is defined as the set of all $\rho \in \mathcal{S}_1$ having $\mathrm{Tr}\, \Lambda \rho = 0$. But there exists a proper convex combination $\rho = (1-p)\rho_1 + p\rho_2$ with $0 < p < 1$ such that $\mathrm{Tr}\, \Lambda \rho = 0$, and hence $\rho \in \mathcal{F}$, contradicting the assumption that $\mathcal{F}$ is a face of $\mathcal{S}_1$. □

We summarize Thms. 15 and 16 as follows.

Theorem 17. *There is a one to one correspondence between exposed faces of $\mathcal{S}_1$ and exposed faces of $\mathcal{S}_1^\circ$. The faces in each pair are dual (orthogonal) to each other.*

Since the extremal points of $\mathcal{S}_1$ are the pure product states, by Thm. 1 the extremal points of a face $\mathcal{F} \in \mathcal{S}_1$ are the pure product states contained in $\mathcal{F}$. It follows that the

extremal points of the face $\mathcal{Y}^\circ$ are the common zeros of all the witnesses in $\mathcal{Y}$. We have actually proved the following result.

Theorem 18. *A set of product vectors in $\mathcal{H}$ is the complete set of zeros of some witness if and only if they are the extremal points of an exposed face of $\mathcal{S}_1$ when regarded as states in $\mathcal{S}_1$.*

The remaining question, to which we do not know the answer, is whether $\mathcal{S}_1$ has unexposed faces. We state the following theorem, which actually holds not only for $\mathcal{S}_1$ but for any compact convex set.

Theorem 19. *Every proper face of $\mathcal{S}_1$ is contained in an exposed face of $\mathcal{S}_1$.*

Proof. Given a face $\mathcal{F}$ of $\mathcal{S}_1$, we have to prove that there exists a witness Ω such that $\mathcal{F}$ is contained in the dual face Ω°.

Choose $\rho \in \mathcal{F}$, in the interior of $\mathcal{F}$ if $\mathcal{F}$ contains more than one point. Choose also a separable state $\sigma \notin \mathcal{F}$, and define $\tau = (1+t)\rho - t\sigma$. Since $\mathcal{F}$ is a face, and $\tau \in \mathcal{S}_1$ for $-1 \le t \le 0$, we know that $\tau \notin \mathcal{S}_1$ for every $t > 0$, and the set

$$\mathcal{Y}(t) = \{\, \Lambda \in \mathcal{S}_1^\circ \mid \operatorname{Tr} \Lambda\tau \le 0 \,\} \tag{93}$$

is non-empty for every $t > 0$. Every $\mathcal{Y}(t)$ is a compact set, and $\mathcal{Y}(t_1) \subset \mathcal{Y}(t_2)$ for $0 < t_1 < t_2$. Hence, the intersection of all sets $\mathcal{Y}(t)$ for $t > 0$ is non-empty and contains at least one witness Ω such that $\operatorname{Tr}\Omega\tau \le 0$ for every $t > 0$. Clearly we must then have $\operatorname{Tr}\Omega\rho = 0$. Since $\operatorname{Tr}\Omega\rho = 0$ for one point ρ in the interior of the face $\mathcal{F}$, it follows that $\operatorname{Tr}\Omega\rho = 0$ for every $\rho \in \mathcal{F}$. □

Unfortunately, this does not amount to a proof that $\mathcal{F}$ is exposed, because it might happen that $\operatorname{Tr}\Omega\sigma = 0$ even though $\sigma \notin \mathcal{F}$.

8.2. *Faces of $\mathcal{D}_1$ and $\mathcal{P}_1$*

When discussing faces of $\mathcal{S}_1$, the set of separable states, it may be illuminating to consider the simpler examples of faces of $\mathcal{D}_1$, the set of density matrices, and faces of $\mathcal{P}_1$, the set of PPT states. Recall that $\mathcal{D}_1$ is selfdual, $\mathcal{D}_1^\circ = \mathcal{D}_1$. The fact that all faces of $\mathcal{D}_1$ and $\mathcal{P}_1$ are exposed may indicate that the same is true for all faces of $\mathcal{S}_1$.

A face $\mathcal{F}$ of $\mathcal{D}_1$ is a complete set of density matrices on a subspace $\mathcal{U} \subset \mathcal{H}$, thus there is a one to one correspondence between faces of $\mathcal{D}_1$ and subspaces of $\mathcal{H}$. A density matrix ρ belongs to the face $\mathcal{F}$ when $\operatorname{Img}\rho \subset \mathcal{U}$, and it is an interior point of $\mathcal{F}$ when $\operatorname{Img}\rho = \mathcal{U}$.

It is straightforward to show that when $\rho, \sigma \in \mathcal{D}_1$ we have $\operatorname{Tr}\rho\sigma = 0$ if and only if $\operatorname{Img}\rho \perp \operatorname{Img}\sigma$. Hence, the dual, or opposite, face $\mathcal{F}^\circ$ is the set of density matrices on the subspace $\mathcal{U}^\perp$, the orthogonal complement of $\mathcal{U}$. The double dual of $\mathcal{F}$ is $\mathcal{F}$ itself, $\mathcal{F}^{\circ\circ} = \mathcal{F}$, since $(\mathcal{U}^\perp)^\perp = \mathcal{U}$. Every proper face $\mathcal{F}$ of $\mathcal{D}_1$ is exposed, since it is the dual of an arbitrarily chosen interior point $\sigma \in \mathcal{F}^\circ$.

The definition $\mathcal{P}_1 = \mathcal{D}_1 \cap \mathcal{D}_1^P$ implies, by Thm. 2, that every face $\mathcal{G}$ of $\mathcal{P}_1$ is an intersection $\mathcal{G} = \mathcal{E} \cap \mathcal{F}^P$, where $\mathcal{E}$ and $\mathcal{F}$ are faces of $\mathcal{D}_1$. This is the geometrical meaning of the procedure for finding extremal PPT states introduced in Ref. 34.

It follows that every face of $\mathcal{P}_1$ is exposed. In fact, the face $\mathcal{G}$ is dual to a decomposable witness $\rho + \sigma^P$ where $\rho, \sigma \in \mathcal{D}$, such that $\mathcal{E}$ is dual to ρ and $\mathcal{F}$ is dual to σ.

8.3. *Unexposed faces of* $\mathcal{S}_1^\circ$

The Choi–Lam witness Ω_C, as given in Eq. (72) with $\theta = 0$, $a = c = 1$, $b = 0$, is an example of an extremal witness, a zero dimensional face of $\mathcal{S}_1^\circ$, which is not exposed. The three isolated product vectors defined in Eq. (73) and the continuum of product vectors defined in Eq. (74), interpreted as states in $\mathcal{S}_1$, are the extremal points of the dual face $\mathcal{F}_C = \{\Omega_C\}^\circ \subset \mathcal{S}_1$. One may check both numerically and analytically that the face $\mathcal{F}_C$ has dimension 21. The separable states in the interior of $\mathcal{F}_C$ have rank seven, since they are constructed from product vectors in a seven dimensional subspace.

We find numerically that the constraints $\mathbf{U}_0$ and $\mathbf{U}_1$ associated with the zeros of Ω_C define a four dimensional face of $\mathcal{S}_1^\circ$. This is then the dual face $\mathcal{F}_C^\circ$, the double dual of Ω_C. The last four constraints needed to prove that Ω_C is extremal come from the constraints $\mathbf{U}_2$ and $\mathbf{U}_3$ expressing the quartic nature of the zeros.

This is one example showing the mechanism for how faces of $\mathcal{S}_1^\circ$ may avoid being exposed. In general, a witness Ω having one or more isolated quartic zeros will be an interior point of an unexposed face. This unexposed face is then a face of a larger exposed face consisting of witnesses having the same zeros as Ω, but such that all the isolated zeros are quadratic.

8.4. *Simplex faces of* $\mathcal{S}_1$

Theorem 18 expresses the relation between zeros of entanglement witnesses and exposed faces of $\mathcal{S}_1$, the set of separable states. We do not know whether $\mathcal{S}_1$ has unexposed faces. We consider first faces that are simplexes, having only a finite number of extremal points.

8.4.1. *Faces of* $\mathcal{S}_1$ *dual to quadratic extremal witnesses*

The exposed faces of $\mathcal{S}_1$ defined by extremal witnesses with only quadratic zeros are simplexes, and as such are particularly simple to study. Given the product vectors $\psi_i = \phi_i \otimes \chi_i$ for $i = 1, 2, \ldots, k$, with $k = n + 1$, as the zeros of a quadratic extremal witness Ω. The corresponding product states $\rho_i = \psi_i \psi_i^\dagger$ are the vertices of an n-simplex, which is the exposed face Ω° dual to Ω.

Let ρ be an interior point of this face,

$$\rho = \sum_{i=1}^{k} p_i \, \rho_i, \qquad \sum_{i=0}^{k} p_i = 1, \qquad p_i > 0. \tag{94}$$

According to Thm. 11, both ρ, constructed from the zeros $\phi_i \otimes \chi_i$, and its partial transpose ρ^P, constructed in the same way from the partially conjugated zeros $\phi_i \otimes \chi_i^*$, have full rank $N = N_a N_b$. Thus, ρ lies not only in the interior of $\mathcal{D}_1$, the set of density matrices, but also in the interior of $\mathcal{P}_1$, the set of PPT states.

This geometric fact has the following interesting consequence. Let λ denote the smallest one among the eigenvalues of ρ and ρ^P, we know that $0 < \lambda < 1/N$. Then we may construct entangled PPT states from the separable state ρ and the maximally mixed state ρ_0 as

$$\sigma = p\rho + (1 - p)\rho_0, \qquad 1 < p \leq \frac{1}{1 - N\lambda}. \tag{95}$$

The number of zeros of generic quadratic extremal witnesses presented in Table 1 has consequences also for the geometry of $\mathcal{S}_1$. In dimensions 2×4 and 3×3 such a witness will have N zeros and define a face of $\mathcal{S}_1$ in the interior of $\mathcal{P}_1 \subset \mathcal{D}_1$. The boundary of this face will consist of other faces of $\mathcal{S}_1$ defined by less than N zeros, accordingly these faces are located inside faces of $\mathcal{D}_1$ on the boundary of $\mathcal{D}_1$. In higher dimensions such a witness has more than N zeros, and defines a face of $\mathcal{S}_1$ also in the interior of $\mathcal{D}_1$, but such that its boundary contains faces still in the interior of $\mathcal{D}_1$. In this way the geometry of $\mathcal{S}_1$ in relation to $\mathcal{D}_1$ becomes more and more complicated as the dimensions increase.

8.4.2. *Some numerical results*

In an attempt to learn more about the geometry of simplex faces we have studied numerically in dimension 3×3 the faces of $\mathcal{S}_1$ dual to generic quadratic extremal witnesses. They are 8-simplexes, each defined by nine linearly independent pure product states that are the zeros of the witness. Our sample consisted of 171 extremal witnesses found in random numerical searches.

Volumes

Define the edge length factors $f_{ij} = d_{ij}^2 = \|\rho_i - \rho_j\|^2$ in the Hilbert–Schmidt norm, and the symmetric $(n+2) \times (n+2)$ Cayley–Menger matrix

$$D = \begin{pmatrix} 0 & 1 & 1 & \cdots & 1 \\ 1 & 0 & f_{12} & \cdots & f_{1k} \\ 1 & f_{21} & 0 & \cdots & f_{2k} \\ \vdots & \vdots & \vdots & \ddots & \cdots \\ 1 & f_{k1} & f_{k2} & \cdots & 0 \end{pmatrix}. \tag{96}$$

The volume of this n-simplex is then given by

$$V = \frac{\sqrt{|\det(D)|}}{2^n\, n!}. \tag{97}$$

The volume of the regular n-simplex with edge length s is

$$V_{\mathrm{reg}} = \frac{s^n}{n!}\sqrt{\frac{n+1}{2^n}}. \tag{98}$$

The maximal distance between two pure states is $\sqrt{2}$, when the states are orthogonal. For $n = 8$ and $s = \sqrt{2}$ we have $V_{\mathrm{reg}} = 3/8! \approx 7.440 \times 10^{-5}$.

Due to the highly irregular shapes we find volumes of the faces varying over five orders of magnitude. Since we regard density matrices related by $\mathrm{SL} \otimes \mathrm{SL}$ transformations as equivalent, a natural question is how regular these simplexes can be made by such transformations, in other words, what is the maximum volume V^* we may obtain. Figure 5 shows our data for the ratio V^*/V_{reg}. There is some variation left in this ratio, indicating that the 8-simplexes are genuinely irregular.

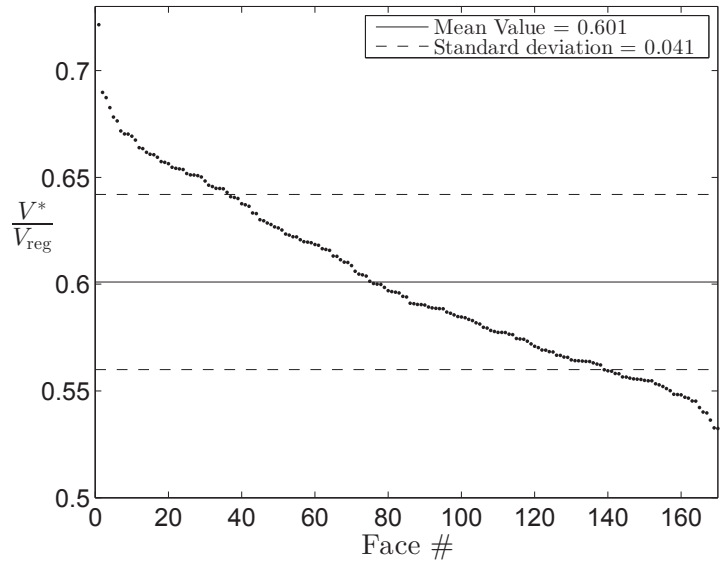


Fig. 5. Ratio of volumes V^* and $V_{\rm reg}$ of 8-simplex faces of $\mathcal{S}_1$ in 3×3. V^* is the simplex volume maximized under $\mathrm{SL} \otimes \mathrm{SL}$ transformations, and $V_{\rm reg}$ is the volume of the regular simplex of orthogonal pure product states.

Positions relative to the maximally mixed state

The second geometric property we have studied is the distance from the maximally mixed state to the center of each simplex face, defined as the average of the vertices,

$$\rho_c = \frac{1}{k} \sum_{i=1}^{k} \rho_i. \tag{99}$$

We compare this distance to the radius of the maximal ball of separable states centered around the maximally mixed state,[13]

$$R_m = \frac{1}{\sqrt{N(N-1)}}. \tag{100}$$

For $N = 3 \times 3$ we have $R_m = 1/\sqrt{72} \approx 0.1179$. The distance $d_c = \|\rho_c - \rho_0\|$ can be minimized by $\mathrm{SL} \otimes \mathrm{SL}$ transformations on ρ_c, resulting in a unique minimal distance d_c^* for each equivalence class of faces. Figure 6 shows our data for the ratio d_c^*/R_m. We see that d_c^* does not saturate the lower bound R_m, and this is another indication of the intrinsic irregularity of the 8-simplexes.

The third property studied is the orientation of each simplex relative to the maximally mixed state ρ_0. On a given face there exists a unique state $\rho_{\rm min}$ closest to ρ_0, the minimum distance is $d_{\rm min} = \|\rho_{\rm min} - \rho_0\|$. In dimension $N = N_a N_b$ the distance from any pure state to the maximally mixed state ρ_0 is $\sqrt{(N-1)/N}$, and in our case this gives an upper limit

$$d_{\rm min} < \sqrt{\frac{8}{9}} \approx 0.9428. \tag{101}$$

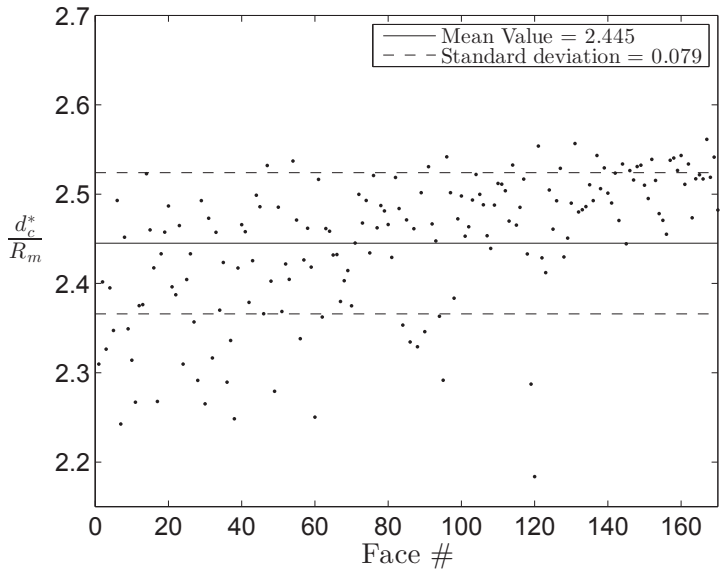


Fig. 6. Ratio of distances d_c^* and R_m for 8-simplex faces of $\mathcal{S}_1$ in dimension 3×3. d_c^* is the distance from the center of the face minimized under SL $\otimes$ SL transformations, and R_m is the radius of the maximal ball of separable states. The points are plotted in the same order as in Fig. 5.

With our sample of 171 faces it happens in only four cases, or $\sim 2\,\%$, that $\rho_{\min}$ lies in the interior of the face and has the full rank nine. In the remaining $\sim 98\,\%$ of the cases, $\rho_{\min}$ lies on the boundary of the face and has lower rank.

Figure 7 shows the rank and distance for the state $\rho_{\min}$ in each face in our sample. We see a tendency that the minimum distance $d_{\min}$ is smaller when the rank of $\rho_{\min}$ is higher. This confirms our expectation that the most regular simplex faces are positioned most symmetrically relative to the maximally mixed state, and also come closest to this state.

We find numerically that if we generate an 8-simplex by generating a random set of nine pure product states, then $\rho_{\min}$ will lie in the interior of the simplex and have full rank in about 90 % of the cases. Also we observe no ranks smaller than six. Thus, Fig. 7 proves that a simplex face generated by a random search for an extremal witness looks very different from a randomly generated simplex. A random set of product vectors that define an 8-simplex will in general not be the zeros of an entanglement witness, and the simplex will not be a face of $\mathcal{S}_1$.

8.5. *Other types of faces of $\mathcal{S}_1$*

To every entanglement witness there corresponds an exposed face of $\mathcal{S}_1$, and the great variety of entanglement witnesses implies a similar variety of faces of $\mathcal{S}_1$. The class of simplex faces, having only a finite number of extremal points, is rather special, other faces may have only continuous sets of extremal points, or both discrete and continuous subsets of extremal points. Some special examples in dimension 3×3 may serve as illustrations.

The Choi–Lam witness is a good example. It has three discrete zeros and one continuous set of zeros, and these zeros in their role as pure product states are the extremal points

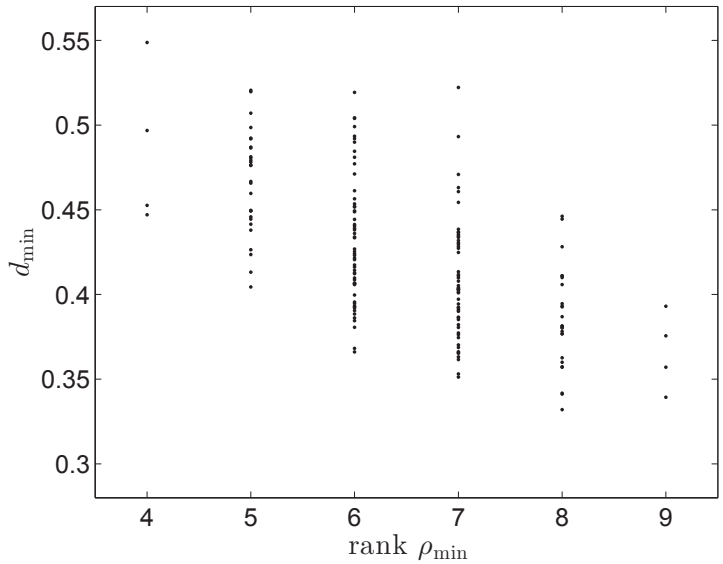


Fig. 7. The rank of $\rho_{\min}$ and its distance $d_{\min}$ from the maximally mixed state for 8-simplex faces in dimension 3×3. The faces are derived from our set of quadratic extremal entanglement witnesses found in random searches using Alg. 1.

of a face of $\mathcal{S}_1$.

An extremal decomposable witness, *i.e.* a pure state or the partial transpose of a pure state, defines a face of $\mathcal{S}_1$ with one single continuous set of extremal points. To be more specific, consider a pure state $\Omega = \psi\psi^\dagger$, then the zeros of Ω are the product vectors orthogonal to ψ. We consider two typical cases.

One example is the product vector $\psi = e_1 \otimes e_1$. Then the zeros are the product vectors

$$\phi \otimes \chi = (a_1 e_1 + a_2 e_2 + a_3 e_3) \otimes (b_1 e_1 + b_2 e_2 + b_3 e_3) \tag{102}$$

with either $a_1 = 0$ or $b_1 = 0$. These product vectors produce pure product states belonging to a full matrix algebra in dimension 2×3 if $a_1 = 0$, and in dimension 3×2 if $b_1 = 0$. They are the extremal points of a face of the type called in Ref. 16 a convex combination of matrix algebras.

Another example is the Bell type entangled pure state

$$\psi = \frac{1}{\sqrt{3}} (e_1 \otimes e_1 + e_2 \otimes e_2 + e_3 \otimes e_3). \tag{103}$$

The zeros in this case are the product vectors as in Eq. (102) with

$$a_1 b_1 + a_2 b_2 + a_3 b_3 = 0. \tag{104}$$

For every vector $\phi = a_1 e_1 + a_2 e_2 + a_3 e_3$ there is a two dimensional subspace of vectors $\chi = b_1 e_1 + b_2 e_2 + b_3 e_3$ satisfying this orthogonality condition. And vice versa, for every vector χ there is a two dimensional subspace of vectors ϕ. These pure product states are

the extremal points of a face which is neither a simplex nor a convex combination of matrix algebras.

8.6. *Minimal convex decomposition*

We know that any point in a compact convex set may be written as a convex combination of extremal points, but if we want to do the decomposition in practice an important question is how many extremal points we need. We would like to know a minimal number of extremal points that is sufficient for all points. The facial structure of the set is then relevant.

The theorem of Carathéodory gives a sufficient number of extremal points as $n+1$ where n is the dimension of the set.[40] However, the set $\mathcal{D}_1$ of normalized density matrices is an example where one can do much better. The dimension of the set is $n = N^2 - 1$, where N is the dimension of the Hilbert space, but the spectral representation of a density matrix in terms of its eigenvalues and eigenvectors is a decomposition using only N extremal points. In order to see how this number N is related to the facial structure of $\mathcal{D}_1$, consider the following proof of Carathéodory's theorem.

Any point x in a compact convex set $\mathcal{C}$ is either extremal or an interior point of a unique face $\mathcal{F}_1$, which might be the whole set $\mathcal{C}$. If x is not extremal it may be written as a convex combination

$$x = (1 - p_1)x_1 + p_1 y_1, \tag{105}$$

where x_1 is an arbitrary extremal point of $\mathcal{F}_1$ and y_1 is another boundary point of $\mathcal{F}_1$. If y_1 is extremal we define $x_2 = y_1$, otherwise y_1 is an interior point of a proper face $\mathcal{F}_2$ of $\mathcal{F}_1$, and we write

$$y_1 = (1 - p_2)x_2 + p_2 y_2, \tag{106}$$

where x_2 is an extremal point of $\mathcal{F}_2$ and y_2 is a boundary point of $\mathcal{F}_2$. Continuing this process we obtain a decomposition of x as a convex combination of extremal points $x_1, x_2, \ldots, x_k \in \mathcal{C}$, and a sequence $\mathcal{F}_1 \supset \mathcal{F}_2 \supset \ldots \supset \mathcal{F}_k$ of faces of decreasing dimensions

$$n \geq n_1 > n_2 > \ldots > n_k = 0. \tag{107}$$

The length of the sequence is k, and the obvious inequality $k \leq n + 1$ is Carathéodory's theorem.

For the set $\mathcal{D}_1$ of normalized density matrices the longest possible sequence of face dimensions has length N, it is

$$N^2 - 1 > (N-1)^2 - 1 > \ldots > (N-j)^2 - 1 > \ldots > 8 > 3 > 0. \tag{108}$$

In general, we decompose an arbitrary density matrix ρ of rank r as a convex combination of r pure states that may be eigenvectors of ρ, but need not be.

We may apply this procedure to the decomposition of an arbitrary separable state as a convex combination of pure product states. As we see, the number of pure product states we have to use depends very much on the facial structure of the set of separable states. It might happen, for example, that the face $\mathcal{F}_2$ is the dual of a quadratic witness, and then it contains only a finite number of pure product states, much smaller than the dimension N^2-1 of the set $\mathcal{S}_1$ of normalized separable states. Recall that the pure product

states in such a face are the zeros of the witness, and the number of zeros is largest when the witness is extremal. In the generic case the number of zeros of a quadratic extremal witness is n_c, as given by Eq. (45) and tabulated in Table 1.

It seems likely therefore that also in the decomposition of separable states one could do much better than the N^2 pure product states guaranteed by Carathéodory's theorem. We consider this an interesting open problem for further study.

9. Optimal and extremal witnesses

The notion of optimal entanglement witnesses was developed by Lewenstein *et al.*[26,29] In this section we comment briefly on the relation between optimal and extremal witnesses. In systems of dimension $2 \times N_b$ and 3×3 the two concepts are, at least generically, closely related. This has significant geometric consequences.

Recall the definition that a witness Ω is extremal if and only if there does not exist two other witnesses Ω_0 and Ω_1 such that

$$\Omega = (1-p)\,\Omega_0 + p\,\Omega_1 \quad \text{with} \quad 0 < p < 1. \tag{109}$$

Lewenstein *et al.* introduce two different concepts of optimality and give two necessary and sufficient conditions for optimality, similar to the condition for extremality.[26] By their Thm. 1, a witness Ω is optimal if and only if it is not a convex combination of the form (109), where Ω_1 is a positive semidefinite matrix. By their Thm. 1(b), Ω is non-decomposable optimal if and only if it is not a convex combination of this form where Ω_1 is a decomposable witness. Obviously, extremal witnesses are optimal by both optimality criteria.

Ha and Kye give examples to show that an optimal witness which is non-decomposable need not be non-decomposable optimal according to the definition of Lewenstein *et al.*, hence they propose to replace the ambiguous term "non-decomposable optimal witness" by "optimal PPTES witness", where PPTES refers to entangled states with positive partial transpose.[59] Here we will use the original terminology.

Note that the criterion for optimality is not invariant under partial transposition, hence a witness Ω may be optimal while its partial transpose Ω^P is not optimal, and vice versa. The criterion for non-decomposable optimality, on the other hand, is invariant under partial transposition, so that Ω is non-decomposable optimal if and only if Ω^P is non-decomposable optimal.

Apparently the condition of being extremal is stricter than those of being optimal or non-decomposable optimal. Accordingly there should, in general, exist plenty of non-extremal witnesses that are either optimal or non-decomposable optimal. We now investigate this question further in the spirit of Thm. 11. The following sufficient optimality conditions were proved by Lewenstein *et al.*[26]

Theorem 20. *A witness Ω is optimal if its zeros span the Hilbert space.*

Ω is non-decomposable optimal if its zeros span the Hilbert space, and at the same time its partially conjugated zeros span the Hilbert space.

Proof. Equation (109) implies that the zeros and Hessian zeros of Ω are also zeros and Hessian zeros of Ω_0 and Ω_1. But if Ω_1 is positive semidefinite and its zeros span the Hilbert

space, then $\Omega_1 = 0$. Similarly, if Ω_1 is decomposable and its zeros and partially conjugated zeros both span the Hilbert space, then $\Omega_1 = 0$. □

For a quadratic witness these conditions are not only sufficient but also necessary.

Theorem 21. *Let Ω be a quadratic witness. Then the zeros of Ω span the Hilbert space, and Ω is non-decomposable optimal if and only if its zeros span the Hilbert space and its partially conjugated zeros also span the Hilbert space.*

Proof. We have to prove the "only if" part, and we reason like in the proofs of Thms. 8 and 11. If the zeros do not span the Hilbert space, there exists a vector ψ orthogonal to all zeros, and we define $\Omega_1 = \psi\psi^\dagger$. Then the zeros of Ω_1 include the zeros of Ω, and we know that $\Omega_1 \neq \Omega$ because Ω is quadratic and Ω_1 is quartic. The line segment from Ω_1 to Ω consists of witnesses having exactly the same zeros and Hessian zeros as Ω (Ω has no Hessian zeros since it is quadratic). By Thm. 7 this line segment can be prolonged within $\mathcal{S}_1^\circ$ so that it gets Ω as an interior point. Hence Ω is neither optimal nor non-decomposable optimal. If the partially conjugated zeros do not span the Hilbert space, then there exists a vector η orthogonal to all the partially conjugated zeros. Now we define $\Omega_1 = (\eta\eta^\dagger)^P$, and use it in the same way as a counterexample to prove that Ω is not non-decomposable optimal. □

Following Thms. 13 and 21, we see that if Ω is a quadratic witness which is non-decomposable optimal, then it is non-decomposable.

It is known that the equivalence between optimality and the spanning property holds also in the case of decomposable witnesses in $2 \times N_b$ dimensions. Augusiak, Tura and Lewenstein proved the following result [60]

Theorem 22. *Let Ω be a decomposable witness in $2 \times N_b$ dimensions, then the following conditions are equivalent:*

(1) *Ω is optimal;*
(2) *The zeros of Ω span the Hilbert space;*
(3) *$\Omega = \sigma^P$, where $\sigma > 0$ and $\operatorname{Img}\sigma$ contains no product vectors.*

A subspace containing no product vectors is often called a completely entangled subspace. Obviously, Ω is not extremal unless σ is a pure state. It is also known that for dimensions 3×3 and higher, there exist optimal decomposable witnesses without the spanning property. [61]

Lewenstein *et al.* also give an optimality test for quartic witnesses without the spanning property. Since we do not have a complete understanding of constraints from quartic zeros, we leave the relation between non-decomposable optimality and extremality of quartic witnesses as an open problem.

In dimensions 3×3 and $2 \times N_b$, a witness with doubly spanning zeros will generically be extremal, by Thm. 12, and hence both optimal and non-decomposable optimal. From Table 1 we see that in higher dimensions the number of zeros of a quadratic extremal witness must be larger than the Hilbert space dimension $N = N_a N_b$. Hence non-extremal witnesses with doubly spanning zeros must be very common, so that non-decomposable optimality is a significantly weaker property than extremality. In particular, when we search for generic quadratic extremal witnesses and reach the stage where the witnesses

have N or more zeros, then every face of $\mathcal{S}_1^\circ$ we encounter will consist entirely of non-decomposable optimal witnesses, most of which are not extremal.

The observations made above give a picture of increasingly complicated geometry as dimensions increase. In dimensions 2×2 and 2×3 every witness is decomposable.[11,14] In $2\times N_b$ and 3×3 a generic quadratic witness is either non-decomposable optimal and extremal, or it is neither. In higher dimensions there exist an abundance of non-decomposable optimal non-extremal witnesses.

10. The SPA separability conjecture

Horodecki and Ekert introduced the concept of a "structural physical approximation" (SPA) of a non-physical map that can be used as a mathematical operation for detecting entanglement.[30,31] The SPA is a physical map and may be implemented in an experimental setup. An example is partial transposition, where the transposition map is applied to one subsystem. The transposition map is unphysical, since it is not completely positive, hence the need for an SPA.

The Jamiołkowski isomorphism relates positive maps and entanglement witnesses, and the SPA of a witness Ω is usually defined as the nearest positive matrix that is a convex combination of Ω and the maximally mixed state I/N. The witness itself corresponds to a map which is positive but not completely positive, whereas its SPA corresponds to a completely positive map. Define

$$\Sigma(p) = (1-p)\Omega + \frac{p}{N}\,I, \tag{110}$$

and let

- p_0 be the smallest value of p such that $\Sigma(p)$ is a separable density matrix;
- p_1 be the smallest value of p such that $\Sigma(p)$ is a density matrix;
- p_2 be the smallest value of p such that $(\Sigma(p))^P$ is a density matrix.

If $\lambda_1 \le 0$ and $\lambda_2 \le 0$ are the lowest eigenvalues of Ω and Ω^P respectively, then

$$p_1 = -\frac{N\lambda_1}{1-N\lambda_1}, \qquad p_2 = -\frac{N\lambda_2}{1-N\lambda_2}. \tag{111}$$

The SPA of Ω is defined in Ref. 32 as $\widetilde{\Omega} = \Sigma(p_1)$.

Clearly $p_0 \ge p_1$ and $p_0 \ge p_2$, but we may have either $p_1 < p_2$, $p_1 = p_2$, or $p_1 > p_2$. The SPA of Ω is a PPT state if and only if $p_1 \ge p_2$. It follows directly from the definitions that the SPA of the witness Ω^P is $(\Sigma(p_2))^P$, and this is a PPT state if and only if $p_2 \ge p_1$.

The so called SPA separability conjecture was put forward by Korbicz *et al.*, and states that the SPA of an optimal entanglement witness is always a separable state.[32] Thus for optimal witnesses we should have $p_0 = p_1 \ge p_2$. One reason for the interest in this conjecture is that separability simplifies the physical implementation of the corresponding map.

The conjecture was supported by several explicit constructions.[52,56,62,63] In particular, it is true for extremal decomposable witnesses, that is, for witnesses of the form $\Omega = (\psi\psi^\dagger)^P$ where ψ is an entangled vector.[32,52] Note that in this case $\Omega^P = \psi\psi^\dagger$ is not an entanglement witness, according to the standard definition that a witness must have at

least one negative eigenvalue. According to our definition a witness is not required to have negative eigenvalues, thus we define $\psi\psi^\dagger$ to be an extremal decomposable witness.

Counterexamples to the SPA conjecture are known. It was shown recently by Chruściński and Sarbicki that the SPA of an optimal but non-extremal decomposable witness, of the form $\Omega = \sigma^P$ with $\sigma \geq 0$ of rank three, may be entangled, although it is automatically a PPT state.[64] This counterexample leaves open the possibility that the conjecture may hold for all such witnesses in dimensions $2 \times N_b$ with $N_b \geq 4$.

Counterexamples of the form of the generalized Choi witness, Eq. (72), have been given by Ha and Kye and by Størmer.[33,41] With the parameter values considered by Størmer we find that the witnesses are not extremal, and their SPAs are entangled because they are not PPT states. More interesting are the parameter values considered by Ha and Kye, because they give extremal witnesses with SPAs that are entangled PPT states. We describe these witnesses in some more detail below.

The SPA conjecture applies directly to extremal entanglement witnesses, since they are optimal. If Ω is an extremal witness, then so is the partial transpose Ω^P. However, if the SPA of Ω is a PPT state, then the SPA of Ω^P is usually not, and vice versa. This means that the original SPA conjecture can not be true both for Ω and Ω^P, simply because separable states are PPT states. Ha and Kye gave essentially this argument in Ref. 33. What saves the SPA conjecture in the case of an extremal decomposable witness $\Omega = (\psi\psi^\dagger)^P$ and its partial transpose $\Omega^P = \psi\psi^\dagger$ is that the latter is excluded by the standard definition.

An obvious modification of the conjecture would be to redefine the SPA as the nearest PPT matrix instead of the nearest positive matrix. In the case of a pair of extremal witnesses Ω and Ω^P the modified conjecture holds either for both or for none. Since the original SPA conjecture holds for extremal decomposable witnesses $\Omega = (\psi\psi^\dagger)^P$, the modified SPA conjecture holds for $\Omega^P = \psi\psi^\dagger$, which is an extremal decomposable witness according to our definition.

The counterexamples cited above disprove also this modified separability conjecture in its full generality, including some extremal non-decomposable witnesses. However, we want to point out that the known counterexamples are rather non-generic. Therefore we have made a small numerical investigation of how the modified conjecture works for generic quadratic extremal witnesses. We are unable to draw a definitive conclusion because of the limited numerical precision of our separability test.

The counterexample of Ha and Kye

The generalized Choi witness, Eq. (72), contains four parameters a, b, c, θ. It is required here that $2/3 \leq a < 1$. Each value of a gives twelve different witnesses, depending on two arbitrary signs and one arbitrary angle $\theta_1 = 0, \pm 2\pi/3$. We define

$$\begin{aligned} a_1 &= \sqrt{\frac{1-(1-a)^2}{2}}, & a_2 &= \sqrt{\frac{1-9(1-a)^2}{2}}, \\ b &= \frac{a_1 \pm a_2}{2}, & c &= a_1 - b, \\ \theta_0 &= \arccos\left(\frac{a+a_1}{2}\right), & \theta &= \pm\theta_0 + \theta_1. \end{aligned} \tag{112}$$

The unnormalized SPA of this witness Ω is $\rho = \Omega + a_1 I$, of rank eight, and the SPA of Ω^P is ρ^P, of rank six. Thus ρ is a PPT state. The kernel of ρ is spanned by the vector $\psi^{(0)}$, and the kernel of ρ^P is spanned by $\psi^{(1)}, \psi^{(2)}, \psi^{(3)}$ where

$$(\psi^{(0)}, \psi^{(1)}, \psi^{(2)}, \psi^{(3)}) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \mathrm{e}^{\mathrm{i}\theta} & 0 & 0 \\ 0 & 0 & c + a_1 & 0 \\ 0 & c + a_1 & 0 & 0 \\ \mathrm{e}^{-\mathrm{i}\theta_1} & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathrm{e}^{\mathrm{i}\theta} \\ 0 & 0 & \mathrm{e}^{\mathrm{i}\theta} & 0 \\ 0 & 0 & 0 & c + a_1 \\ \mathrm{e}^{\mathrm{i}\theta_1} & 0 & 0 & 0 \end{pmatrix}. \tag{113}$$

A separable state σ satisfies the range criterion that $\mathrm{Img}\,\sigma$ is spanned by product vectors such that the partially conjugated product vectors span $\mathrm{Img}\,\sigma^P$. The present PPT state ρ is entangled, in fact it is a so called edge state, violating the range criterion maximally. It is straightforward to verify that there exists no product vector $\varphi \otimes \chi$ orthogonal to $\psi^{(0)}$ such that $\varphi \otimes \chi^*$ is orthogonal to all of $\psi^{(1)}, \psi^{(2)}, \psi^{(3)}$.

Generic quadratic extremal witnesses

When we find an extremal witness Ω in a random numerical search by Alg. 1, presumably the probabilities of finding Ω and Ω^P are equal. Hence we expect to have about fifty per cent probability of finding an extremal witness with an SPA which is a PPT state. This is also what we see in practice.

For each extremal witness Ω found in dimensions 3×3 and 3×4 we have computed the SPA of Ω and of Ω^P. As expected, in each case one SPA is a PPT state and the other one is not. The number of PPT states in our numerical searches originating from the quadratic witnesses Ω and Ω^P are 81 and 90 respectively for the 3×3 system, and 41 and 24 for the 3×4 system.

Since the original separability conjecture of Korbicz *et al.* fails both in the generic case and by non-generic counterexamples, a natural question is whether the SPA is always separable when it is a PPT state. As mentioned there are counterexamples in Refs. 33, 41. On the other hand, these counterexamples are rather special, and there remains a possibility that the SPA of a generic extremal witness with only quadratic zeros could be separable if it is a PPT state.

For this reason, in our numerical examples we have tried to check the separability of the SPA whenever it is a PPT state. We find that it is always very close to being separable, in no case are we able to conclude that it is certainly not separable. Unfortunately, our numerical separability test is not sufficiently precise that we may state with any conviction the opposite conclusion that the state is separable.

In view of the known examples and counterexamples the status of the modified SPA separability conjecture, which defines the SPA as a PPT state, is not completely clear. It would be interesting to know some simple criterion for exactly when it is true. This may also be regarded as one aspect of the more general question of which border states of the set of separable states are also border states of the set of PPT states.

On the other hand, if we regard separability as an essential property of the SPA, it seems that the natural solution would be to define the SPA by the parameter value $p = p_0$, where $\Sigma(p)$ first becomes separable, rather than by the value $p = p_1$, where $\Sigma(p)$ first becomes positive, or by $p = \max(p_1, p_2)$, where $\Sigma(p)$ first becomes a PPT state.

11. A real problem in a complex cloak

We want to outline here a different approach to the problem of describing the convex cone $\mathcal{S}^\circ$ of entanglement witnesses, where we regard it as a real rather than a complex problem. This means that we treat $\mathcal{S}_1^\circ$ as a subset of a higher dimensional compact convex set. We are motivated by the observation that the problem is intrinsically real, since H is a real vector space, f_Ω is a real valued function, and some of the constraints discussed in Sec. 3 are explicitly real. In particular, the second derivative matrix, the Hessian, defining constraints at a quartic zero of a witness is a real and not a complex matrix.

Another motivation is the possibility of expressing the extremal points of $\mathcal{S}_1^\circ$ as convex combinations of extremal points of the larger set. We show how a pure state as a witness is a convex combination of real witnesses that are not of the complex form.

In order to arrive at an explicitly real formulation of the problem we proceed as follows. Define $J : \mathbb{R}^{2N_a} \to \mathbb{C}^{N_a}$ and $K : \mathbb{R}^{2N_b} \to \mathbb{C}^{N_b}$ as

$$J = (I_a, \mathrm{i}I_a), \qquad K = (I_b, \mathrm{i}I_b), \tag{114}$$

and define

$$Z = (J \otimes K)^\dagger \Omega (J \otimes K). \tag{115}$$

Then if $\phi = Jx$ and $\chi = Ky$ we have that

$$g_Z(x, y) = (x \otimes y)^T Z (x \otimes y) = (\phi \otimes \chi)^\dagger \Omega (\phi \otimes \chi) = f_\Omega(\phi, \chi). \tag{116}$$

Since Z is a Hermitian matrix, its real part is symmetric and its imaginary part anti-symmetric. The expectation value $g_Z(x, y)$ is real because only the symmetric part of Z contributes. Furthermore, only the part of Z that is symmetric under partial transposition contributes. Hence $g_Z(x, y) = g_W(x, y)$ when we define

$$W = \frac{1}{2} \mathrm{Re}\, (Z + Z^P). \tag{117}$$

Since $J \otimes K = (I_a \otimes K, \mathrm{i}I_a \otimes K)$ is an $N \times (4N)$ matrix, Z is a $(4N) \times (4N)$ matrix,

$$Z = \begin{pmatrix} X & \mathrm{i}X \\ -\mathrm{i}X & X \end{pmatrix}, \tag{118}$$

where X is a $(2N) \times (2N)$ matrix,

$$X = (I_a \otimes K)^\dagger \Omega (I_a \otimes K). \tag{119}$$

The partial transposition of Z transposes submatrices of size $(2N_b) \times (2N_b)$.

We may now replace the complex biquadratic form f_Ω by the real form g_W, and repeat the analysis in Sec. 3 almost unchanged. We therefore see that complex witnesses in dimension $N = N_a \times N_b$ correspond naturally to real witnesses in $(2N_a) \times (2N_b) = 4N$.

Write the matrix elements of Ω as

$$(\Omega_{ik})_{jl} = \Omega_{ij;kl}. \tag{120}$$

In this notation, Ω_{ik} is an $N_b \times N_b$ submatrix of the $N \times N$ matrix Ω. Then we have that

$$X_{ik} = K^\dagger \Omega_{ik} K = \begin{pmatrix} \Omega_{ik} & \mathrm{i}\Omega_{ik} \\ -\mathrm{i}\Omega_{ik} & \Omega_{ik} \end{pmatrix}. \tag{121}$$

The complex matrix Ω_{ik} may always be decomposed as

$$\Omega_{ik} = A_{ik} + B_{ik} + \mathrm{i}\,(C_{ik} + D_{ik}), \tag{122}$$

where $A_{ik}, B_{ik}, C_{ik}, D_{ik}$ are real matrices, and A_{ik}, C_{ik} are symmetric, B_{ik}, D_{ik} are antisymmetric. The Hermiticity of Ω means that

$$\Omega_{ik} = (\Omega_{ki})^\dagger = A_{ki} - B_{ki} - \mathrm{i}\,(C_{ki} - D_{ki}), \tag{123}$$

or equivalently,

$$A_{ki} = A_{ik}, \quad B_{ki} = -B_{ik}, \quad C_{ki} = -C_{ik}, \quad D_{ki} = D_{ik}. \tag{124}$$

In particular,

$$B_{ii} = C_{ii} = 0. \tag{125}$$

We finally arrive at the following explicit relation between the complex witness Ω and the real witness W,

$$W = \begin{pmatrix} U & -V \\ V & U \end{pmatrix}, \tag{126}$$

with

$$\begin{aligned} U_{ik} &= \frac{1}{2}\,\mathrm{Re}\,(X_{ik} + X_{ik}^T) = \begin{pmatrix} A_{ik} & -D_{ik} \\ D_{ik} & A_{ik} \end{pmatrix}, \\ V_{ik} &= \frac{1}{2}\,\mathrm{Im}\,(X_{ik} + X_{ik}^T) = \begin{pmatrix} C_{ik} & B_{ik} \\ -B_{ik} & C_{ik} \end{pmatrix}. \end{aligned} \tag{127}$$

Note that the $(2N_b) \times (2N_b)$ matrices U_{ik} and V_{ik} are symmetric by construction, $U_{ik} = (U_{ik})^T$ and $V_{ik} = (V_{ik})^T$. This means that W is symmetric under partial transposition, $W^P = W$.

It also follows from Eqs. (124) and (127) that $U_{ki} = U_{ik}$ and $V_{ki} = -V_{ik}$. This means that the $(2N) \times (2N)$ matrix U is symmetric, $U^T = U$ because

$$(U^T)_{ik} = (U_{ki})^T = U_{ki} = U_{ik}, \tag{128}$$

whereas V is antisymmetric, $V^T = -V$ because

$$(V^T)_{ik} = (V_{ki})^T = V_{ki} = -V_{ik}. \tag{129}$$

This means that W is symmetric, $W^T = W$.

11.1. *A pure state as a witness*

If Ω is a pure state, $\Omega = \psi\psi^\dagger$ with $\psi^\dagger\psi = 1$, then

$$f_\Omega(\phi,\chi) = (\phi\otimes\chi)^\dagger\Omega(\phi\otimes\chi) = |z|^2 = (\mathrm{Re}\, z)^2 + (\mathrm{Im}\, z)^2, \tag{130}$$

with

$$z = (\phi\otimes\chi)^\dagger\psi = (x\otimes y)^T(J\otimes K)^\dagger\psi. \tag{131}$$

We have seen that Ω is extremal in $\mathcal{S}_1^\circ$, but we see now that it is not extremal in the larger set of real witnesses that are not of the complex form given in Eq. (126). Write $\psi = a + \mathrm{i}b$ with a, b real, and

$$p = (J\otimes K)^\dagger\psi = \begin{pmatrix} q \\ -\mathrm{i}q \end{pmatrix}, \tag{132}$$

with

$$q = (I_a\otimes K)^\dagger\psi. \tag{133}$$

Here q is a $(2N)\times 1$ matrix, and if we write q_i with $i = 1, 2, \ldots, N_a$ for the submatrices of size $(2N_b)\times 1$, we have that

$$q_i = K^\dagger\psi_i = \begin{pmatrix} \psi_i \\ -\mathrm{i}\psi_i \end{pmatrix} = \begin{pmatrix} a_i + \mathrm{i}b_i \\ b_i - \mathrm{i}a_i \end{pmatrix} = r_i + \mathrm{i}s_i, \tag{134}$$

with

$$r_i = \begin{pmatrix} a_i \\ b_i \end{pmatrix}, \qquad s_i = \begin{pmatrix} b_i \\ -a_i \end{pmatrix}. \tag{135}$$

We now have that

$$(\mathrm{Re}\, z)^2 = (x\otimes y)^T Z_r (x\otimes y), \qquad (\mathrm{Im}\, z)^2 = (x\otimes y)^T Z_i (x\otimes y), \tag{136}$$

with

$$\begin{aligned} Z_r &= (\mathrm{Re}\, p)(\mathrm{Re}\, p)^T = \begin{pmatrix} rr^T & rs^T \\ sr^T & ss^T \end{pmatrix}, \\ Z_i &= (\mathrm{Im}\, p)(\mathrm{Im}\, p)^T = \begin{pmatrix} ss^T & -sr^T \\ -rs^T & rr^T \end{pmatrix}. \end{aligned} \tag{137}$$

These two matrices are symmetric, $Z_r^T = Z_r$ and $Z_i^T = Z_i$, but we should also make them symmetric under partial transposition. Therefore we replace them by the matrices

$$\begin{aligned} W_r &= \frac{1}{2}\,(Z_r + Z_r^P) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}, \\ W_i &= \frac{1}{2}\,(Z_i + Z_i^P) = \begin{pmatrix} D & -C \\ -B & A \end{pmatrix}, \end{aligned} \tag{138}$$

where

$$
\begin{aligned}
A_{ik} &= \frac{1}{2}\left(r_i r_k^T + r_k r_i^T\right) = \frac{1}{2}\begin{pmatrix} a_i a_k^T + a_k a_i^T & a_i b_k^T + a_k b_i^T \\ b_i a_k^T + b_k a_i^T & b_i b_k^T + b_k b_i^T \end{pmatrix}, \\
B_{ik} &= \frac{1}{2}\left(r_i s_k^T + s_k r_i^T\right) = \frac{1}{2}\begin{pmatrix} a_i b_k^T + b_k a_i^T & -a_i a_k^T + b_k b_i^T \\ b_i b_k^T - a_k a_i^T & -b_i a_k^T - a_k b_i^T \end{pmatrix}, \\
C_{ik} &= \frac{1}{2}\left(s_i r_k^T + r_k s_i^T\right) = \frac{1}{2}\begin{pmatrix} b_i a_k^T + a_k b_i^T & b_i b_k^T - a_k a_i^T \\ -a_i a_k^T + b_k b_i^T & -a_i b_k^T - b_k a_i^T \end{pmatrix}, \\
D_{ik} &= \frac{1}{2}\left(s_i s_k^T + s_k s_i^T\right) = \frac{1}{2}\begin{pmatrix} b_i b_k^T + b_k b_i^T & -b_i a_k^T - b_k a_i^T \\ -a_i b_k^T - a_k b_i^T & a_i a_k^T + a_k a_i^T \end{pmatrix}.
\end{aligned}
\tag{139}
$$

12. Summary and outlook

In this article we have approached the problem of distinguishing entangled quantum states from separable states through the related problem of classifying and understanding entanglement witnesses. We have studied in particular the extremal entanglement witnesses from which all other witnesses may be constructed. We give necessary and sufficient conditions for a witness to be extremal, in terms of the zeros and Hessian zeros of the witness, and in terms of linear constraints on the witness imposed by the existence of zeros. The extremality conditions lead to systematic methods for constructing numerical examples of extremal witnesses.

We find that the distinction between quadratic and quartic zeros is all important when we classify extremal witnesses. Nearly all extremal witnesses found in random searches are quadratic, having only quadratic zeros. There is a minimum number of zeros a quadratic witness must have in order to be extremal, and most of the quadratic extremal witnesses have this minimum number of zeros. To our knowledge, extremal witnesses of this type have never been observed before, even though they are by far the most common.

The number of zeros of a quadratic extremal witness increases faster than the Hilbert space dimension. Since a witness is optimal if its zeros span the Hilbert space, this implies that in all but the lowest dimensions witnesses may be optimal and yet be very far from extremal.

The facial structures of the convex set of separable states and the convex set of witnesses are closely related. The zeros of a witness, whether quadratic or quartic, always define an exposed face of the set of separable states. The other way around, an exposed face of the set of separable states is defined by a set of pure product states that are the zeros of some witness, in fact, they are the common zeros of all the witnesses in a face of the set of witnesses. The existence of witnesses with quartic zeros is the root of the existence of non-exposed faces of the set of witnesses. It is unknown to us whether the set of separable states has unexposed faces.

One possible way to test whether or not a given state is separable is to try to decompose it as a convex combination of pure product states. In such a test it is of great practical importance to know the number of pure product states that is needed in the worst case. The general theorem of Carathéodory provides a limit of $n+1$ where n is the dimension of the set, but this is a rather large number, and an interesting unsolved problem is to improve this limit, if possible. We point out that this is a question closely related to the facial structure of the set of separable states, which in turn is related to the set of

entanglement witnesses.

Another unsolved problem we may mention here again is the following. We have found a procedure for constructing an extremal witness from its zeros and Hessian zeros, but an arbitrarily chosen set of zeros and Hessian zeros does not in general define an extremal entanglement witness. How do we construct an extremal witness by choosing a set of zeros and Hessian zeros we want it to have? Clearly this is a much more complicated problem than the trivial problem of constructing a polynomial in one variable from its zeros.

In conclusion, by studying the extremal entanglement witnesses we have made some small progress towards understanding the convex sets of witnesses and separable mixed states. But this has also made the complexity of the problem even more clear than it was before. The main complication is the fact that zeros of extremal witnesses may be quartic, since this opens up an almost unlimited range of variability among the quartic witnesses. A full understanding of this variability seems a very distant goal. Nevertheless we believe that it is useful to pursue the study of extremal witnesses in order to learn more about the geometry of the set of separable states, and it is rather clear that a combination of analytical and numerical methods will be needed also in future work.

Acknowledgments

We inform in deep sorrow that our coauthor Andreas Hauge passed away on 31st May 2015. Also, we acknowledge gratefully research grants from The Norwegian University of Science and Technology (Leif Ove Hansen) and from The Norwegian Research Council (Per Øyvind Sollid).

Appendix A. Explicit expressions for positivity constraints

We show here explicit expressions for the constraints developed in Sec. 3.1.

Given a zero (ϕ_0, χ_0), let j_m for $m = 1, \ldots, 2(N_a - 1)$ be the column vectors of the matrix J_0. We construct them for example from an orthonormal basis $\phi_0, \phi_1, \ldots, \phi_{N_a-1}$ of $\mathcal{H}_a$, taking $j_{2l-1} = \phi_l$, $j_{2l} = \mathrm{i}\phi_l$ for $l = 1, \ldots, N_a - 1$. Similarly, let k_n for $n = 1, \ldots, 2(N_b - 1)$ be the column vectors of K_0. Then the equations $\mathbf{T}_0\Omega = 0$, $\mathbf{T}_1\Omega = 0$ may be written as

$$\mathrm{Tr}\,(E^{(0)}\Omega) = 0, \quad \mathrm{Tr}\,(E^{(1)}_m\Omega) = 0, \quad \mathrm{Tr}\,(E^{(2)}_n\Omega) = 0, \tag{A.1}$$

with

$$\begin{aligned} E^{(0)} &= \phi_0\phi_0^\dagger \otimes \chi_0\chi_0^\dagger, \\ E^{(1)}_m &= (j_m\phi_0^\dagger + \phi_0 j_m^\dagger) \otimes \chi_0\chi_0^\dagger, \\ E^{(2)}_n &= \phi_0\phi_0^\dagger \otimes (k_n\chi_0^\dagger + \chi_0 k_n^\dagger). \end{aligned} \tag{A.2}$$

With $z_i \in \mathrm{Ker}\, G_\Omega$, $z_i^T = (x_i^T, y_i^T)$ and $\xi_i = J_0 x_i$, $\zeta_i = K_0 y_i$, for $i = 1, \ldots, K$, the constraints $(\mathbf{T}_2\Omega)_i = 0$ given in Eq. (33) may be written explicitly as follows,

$$\mathrm{Tr}\,(F^{(1)}_{im}\Omega) = 0, \quad \mathrm{Tr}\,(F^{(2)}_{in}\Omega) = 0, \tag{A.3}$$

with

$$\begin{aligned} F^{(1)}_{im} &= (\xi_i j_m^\dagger + j_m\xi_i^\dagger) \otimes \chi_0\chi_0^\dagger + (\phi_0 j_m^\dagger + j_m\phi_0^\dagger) \otimes (\zeta_i\chi_0^\dagger + \chi_0\zeta_i^\dagger), \\ F^{(2)}_{in} &= \phi_0\phi_0^\dagger \otimes (\zeta_i k_n^\dagger + k_n\zeta_i^\dagger) + (\xi_i\phi_0^\dagger + \phi_0\xi_i^\dagger) \otimes (\chi_0 k_n^\dagger + k_n\chi_0^\dagger). \end{aligned} \tag{A.4}$$

System $\mathbf{T}_3$ can be written as follows. For any linear combinations

$$\xi = \sum_{i=1}^{K} a_i \xi_i, \qquad \zeta = \sum_{i=1}^{K} a_i \zeta_i, \tag{A.5}$$

with real coefficients a_i and ξ_i, ζ_i as above, we have

$$\operatorname{Tr} G\Omega = 0, \quad G = (\xi\phi_0^\dagger + \phi_0\xi^\dagger) \otimes \zeta\zeta^\dagger + \xi\xi^\dagger \otimes (\zeta\chi_0^\dagger + \chi_0\zeta^\dagger). \tag{A.6}$$

This is one single constraint for each set of coefficients a_i. The minimum number of different linear combinations that we have to use is given by the binomial coefficient

$$\binom{K+2}{3} = \frac{K(K+1)(K+2)}{6}. \tag{A.7}$$

Appendix B. Numerical implementation of Algorithm 1

In this appendix we first describe possibilities for solving the optimization problem (16) and then, assuming that this problem can be solved, we describe how to localize the boundary of a face of quadratic witnesses in $\mathcal{S}_1^\circ$.

Solving the optimization problem

We first describe some possibilities for solving the optimization problem (16). This question has received little attention in the physics community, and in the optimization literature focus shifted towards a rigorous rather than towards an "applied" approach. Two comments are in place. In Refs. 35, 65 simple algorithms were presented which in our experience work fine in most situations. However, when applied to Alg. 1 we repeatedly observe poor convergence, presumably due to degeneracies of eigenvalues. In Ref. 66 it is stated that a straight forward parameterization of problem (16) can not be solved in an efficient manner. Instead they present a formalism which again, in our view, is unnecessarily complicated for low dimensions. The conceptually simplest approach is to perform repeated local minimization from random starting points, thus heuristically (though sufficiently safe) obtaining all global optimal points. Below we describe different possible approaches for local minimization, either applying available optimization algorithms or implementing simple ones ourselves.

In the case that only the optimal value p^* and not the optimal solution (ϕ^*, χ^*) is sought for, a positive maps approach is fruitful. Given an operator $A \in H$ one can define the linear map $\mathbf{L}_A$ such that

$$(\phi \otimes \chi)^\dagger A(\phi \otimes \chi) = \chi^\dagger \mathbf{L}_A(\phi\phi^\dagger)\chi. \tag{B.1}$$

Problem (16) is then equivalent to minimizing the smallest eigenvalue of $\mathbf{L}_A(\phi\phi^\dagger)$ as a function of normalized ϕ, and hence to the unconstrained minimization of

$$\widetilde{f}_\Omega(\phi) = \min\left[\text{spectrum } \frac{\mathbf{L}_A(\phi\phi^\dagger)}{\phi^\dagger\phi}\right]. \tag{B.2}$$

With a real parametrization $\phi = Jx$ as in Eq.(114) this function is readily minimized by the Nelder–Mead downhill simplex algorithm, listed *e.g.* in Ref. 67 and implemented *e.g.* in MATLAB's "fminsearch" function[68] or Mathematica's "NMinimize".[69]

In the case that also the optimal point (ϕ^*, χ^*) is wanted, several possibilities exist. An approximate solution is easily found by feeding the real problem given by (117) to a SQP-algorithm,[43,70] *e.g.* the one provided in MATLAB's "fmincon" function.[68] Another possibility is to apply the infinity norm in (16) and perform box-constrained minimization in each face of the ∞-ball. Since the Hessian of the function is readily available trust-region algorithms are suitable (also provided in MATLAB's "fmincon"). If A is known to be an entanglement witness, even the box-constraints on each face of the ∞-ball can be removed, since in that case the objective function is bounded below. The problem has then been reduced to unconstrained minimization in a series of affine spaces, one for each face of the ∞-ball.

With this latter formulation, we have had success with a quasi-Newton approach[43] performing exact line search, see Ref. 38 for details. Further, we have had promising results with a complex conjugate gradient algorithm also performing exact line search,[70] though further development is necessary.

Finding the boundary of a face

Here we describe how to localize the boundary of a face of quadratic witnesses in $\mathcal{S}_1^\circ$. We assume we have some routine for solving problem (16).

Given a quadratic witness Ω situated in the interior of a face defined by k product vectors, and a perturbation Γ, we define a family of functions as follows. Consider problem (16) with $A(t) = \Omega + t\Gamma$. Given all local minima, sort these in ascending order according to function value. Then the function $\kappa_0(t)$ is the value of local minimum number $k+1$ if this exists and 0 otherwise. Function $\kappa_i(t)$, $i = 1, \ldots, k$ is the smallest positive eigenvalue of $\mathrm{D}^2(f_\Omega + t f_\Gamma)$ evaluated at zero number i. Now each $\kappa_i(t)$, $i = 0, \ldots, k$ is a function with a single simple root in $t \in [0, \infty)$ which can be easily located using any standard rootfinding technique.

Another possible approach to finding the boundary of the face is as follows. Rather than locating all local minima of problem (16), aim only at locating the global minimum value. This global minimum value as a function of Ω will be called the face function. The face function is zero on a face, positive in the interior of $\mathcal{S}^\circ$ and negative outside of $\mathcal{S}^\circ$. Finding some initial $\theta > \theta_c$ then allows for one-sided extrapolation towards θ_c.

If we employ the 2-norm for the constraints in (16) the face function is concave as a function of $\theta > \theta_0$. To see this, note that the optimal value of problem (16) is equal to the optimal value of

$$\widetilde{f}(\phi, \chi) = \frac{(\phi \otimes \chi)^\dagger \Omega (\phi \otimes \chi)}{(\phi^\dagger \phi)(\chi^\dagger \chi)}. \tag{B.3}$$

Let $p^*(\Omega)$ denote the optimal value. Since $p^*(\Omega)$ is defined as a point-wise minimum $p^*(\Omega)$ is a concave function of Ω.[42] Projection of Ω on some face is an affine operation, an operation preserving concavity. Accordingly the face function is a concave function once the new local minimum exists. This fact makes either approach for locating the boundary of a face simpler, since the qualitative form of the function whose root to find is known.

References

1. R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki,
Quantum entanglement,
Rev. Mod. Phys. **81** (2009) 865.
2. M. A. Nielsen and I. L. Chuang,
Quantum Computation and Quantum Information,
Cambridge University Press, 2000.
3. A. Einstein, B. Podolsky, and N. Rosen,
Can quantum-mechanical description of physical reality be considered complete?,
Phys. Rev. **47** (1935) 777.
4. J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt,
Proposed experiment to test local hidden-variable theories,
Phys. Rev. Lett. **23** (1969) 880.
5. R. F. Werner,
Quantum states with Einstein–Podolsky–Rosen correlations admitting a hidden-variable model,
Phys. Rev. A **40** (1989) 4277.
6. L. Masanes, Y.-C. Liang, and A. C. Doherty,
All bipartite entangled states display some hidden non-locality,
Phys. Rev. Lett. **100** (2008) 090403.
7. E. Schrödinger,
Die gegenwärtige situation in der quantenmechanik,
Naturwissenschaften, **23** (1935) 807,823,844.
8. R. Horodecki and P. Horodecki,
Quantum redundancies and local realism,
Phys. Lett. A **194** (1994) 147.
9. A. Peres,
Separability criterion for density matrices,
Phys. Rev. Lett. **77** (1996) 1413.
10. I. Bengtsson and K. Życzkowski,
Geometry of Quantum States,
Cambridge University Press, 2006.
11. M. Horodecki, P. Horodecki, and R. Horodecki,
Separability of mixed states: necessary and sufficient conditions,
Phys. Lett. A **223** (1996) 1.
12. K. Życzkowski, P. Horodecki, A. Sanpera, and M. Lewenstein,
Volume of the set of separable states,
Phys. Rev. A **58** (1998) 883.
13. L. Gurvits and H. Barnum,
Largest separable balls around the maximally mixed bipartite quantum state,
Phys. Rev. A **66** (2002) 062311.
14. E. Størmer,
Positive linear maps of operator algebras,
Acta Mathematica **110** (1963) 233.
15. E. Størmer,
Positive Linear Maps of Operator Algebras,
Springer, 2013.
16. E. Alfsen and F. Shultz,
Unique decompositions, faces, and automorphisms of separable states,
J. Math. Phys. **51** (2010) 052201.
17. J. M. Leinaas, J. Myrheim, and E. Ovrum,
Geometrical aspects of entanglement,
Phys. Rev. A **74** (2006) 012313.
18. L. O. Hansen, A. Hauge, J. Myrheim, and P. Ø. Sollid,
Low rank positive partial transpose states and their relation to product vectors,
Phys. Rev. A **85** (2012) 022309.
19. M. Horodecki, P. Horodecki, and R. Horodecki,
Mixed-State Entanglement and Distillation: Is there a "Bound" Entanglement in Nature?,
Phys. Rev. Lett. **80** (1998) 5239.
20. B. M. Terhal,
Bell inequalities and the separability criterion,
Phys. Lett. A **271** (2000) 319.

21. M.-D. Choi,
Positive Semidefinite Biquadratic Forms,
Linear Alg. Appl. **12** (1975) 95.
22. M.-D. Choi and T.-Y. Lam,
Extremal Positive Semidefinite Forms,
Math. Ann. **231** (1977) 1.
23. A. Guyan Robertson,
Positive Projections on C^*-algebras and an Extremal Positive Map,
J. London Math. Soc. **2** (1985) 133.
24. J. Grabowski, M. Kuś, and G. Marmo,
Wigner's theorem and the geometry of extreme positive maps,
J. Phys. A: Math. Theor. **42** (2009) 345301.
25. M. Marciniak,
On extremal positive maps acting between type 1 factors,
Banach Center Publ. **89** (2010) 201.
26. M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki,
Optimization of entanglement witnesses,
Phys. Rev. A **62** (2000) 052310.
27. K.-C. Ha and S.-H. Kye,
Entanglement witnesses arising from Choi type positive linear maps,
J. Phys. A: Math. Theor. **45** (2012) 415305.
28. G. Sarbicki and D. Chruściński,
A class of exposed indecomposable positive maps,
J. Phys. A: Math. Theor. **46** (2013) 015306.
29. M. Lewenstein, B. Kraus, P. Horodecki, and J. I. Cirac,
Characterization of separable states and entanglement witnesses,
Phys. Rev. A **63** (2001) 044304.
30. P. Horodecki,
From limits of quantum operations to multicopy entanglement witnesses and state-spectrum estimation,
Phys. Rev. A **68** (2003) 052101.
31. P. Horodecki and A. Ekert,
Method for direct detection of quantum entanglement,
Phys. Rev. Lett. **89** (2002) 127902.
32. J. K. Korbicz, M. L. Almeida, J. Bae, M. Lewenstein, and A. Acín,
Structural approximations to positive maps and entanglement-breaking channels,
Phys. Rev. A **78** (2008) 062105.
33. K.-C. Ha and S.-H. Kye,
The structural physical approximations and optimal entanglement witnesses,
J. Math. Phys. **53** (2012) 102204.
34. J. M. Leinaas, J. Myrheim, and E. Ovrum,
Extreme points of the set of density matrices with positive partial transpose,
Phys. Rev. A **76** (2007) 034304.
35. J. M. Leinaas, J. Myrheim, and P. Ø. Sollid,
Numerical studies of entangled positive-partial-transpose states in composite quantum systems,
Phys. Rev. A **81** (2010) 062329.
36. J. M. Leinaas, J. Myrheim, and P. Ø. Sollid,
Low rank extremal positive-partial-transpose states and unextendible product bases,
Phys. Rev. A **81** (2010) 062330.
37. P. Ø. Sollid, J. M. Leinaas, and J. Myrheim,
Unextendible product bases and extremal density matrices with positive partial transpose,
Phys. Rev. A **84** (2011) 042325.
38. A. Hauge,
A geometrical and computational study of entanglement witnesses,
Master's thesis, Norwegian University of Science and Technology, Dept. of Physics, 2011.
39. P. Horodecki,
Separability criterion and inseparable mixed states with positive partial transposition,
Phys. Lett. A **232** (1997) 333.
40. C. Carathéodory,
Über den Variabilitätsbereich der Fourierschen Konstanten von positiven harmonischen Funktionen,
Rendiconti del Circolo Matematico di Palermo **32** (1911) 193.
41. E. Størmer,
Separable states and the structural physical approximation of a positive map,

J. Funct. Anal. **264** (2013) 2197.
42. S. Boyd and L. Vandenberghe,
Convex Optimization,
Cambridge University Press, 2004.
43. J. Nocedal and S. J. Wright,
Numerical Optimization,
Springer, 2006.
44. F. G. S. L. Brandão,
Quantifying entanglement with witness operators,
Phys. Rev. A **72** (2005) 022310.
45. H.-P. Breuer,
Optimal entanglement criterion for mixed quantum states,
Phys. Rev. Lett. **97** (2006) 080501.
46. A. C. Doherty, P. A. Parrilo, and F. M. Spedalieri,
Complete family of separability criteria,
Phys. Rev. A **69** (2004) 022308.
47. L. M. Ioannou and B. C. Travaglione,
Quantum separability and entanglement detection via entanglement-witness search and global optimization,
Phys. Rev. A **73** (2006) 052314.
48. A. Sanpera, D. Bruß, and M. Lewenstein,
Schmidt-number witnesses and bound entanglement,
Phys. Rev. A **63** (2001) 050301.
49. G. Tóth and O. Gühne,
Detecting genuine multipartite entanglement with two local measurements,
Phys. Rev. Lett. **94** (2005) 060501.
50. E. Alfsen and F. Shultz,
Finding decompositions of a class of separable states,
Lin. Alg. Appl. **437** (2012) 2613.
51. A. Jamiołkowski,
Linear transformations which preserve trace and positive semidefiniteness of operators,
Reports on mathematical physics, **3** (1972) 275.
52. R. Augusiak, J. Bae, L. Czekaj, and M. Lewenstein,
On structural physical approximations and entanglement breaking maps,
J. Phys. A: Math. Theor. **44** (2011) 185308.
53. S.-J. Cho, S.-H. Kye, and S.-G. Lee,
Generalized Choi maps in three-dimensional matrix algebra,
Lin. Alg. Appl. **171** (1992) 213.
54. K.-C. Ha and S.-H. Kye,
Entanglement Witnesses Arising from Exposed Positive Linear Maps,
Open Systems & Information Dynamics, **18** (2011) 323.
55. J. P. Zwolak and D. Chruściński,
New tools for investigating positive maps in matrix algebras,
Rep. Math. Phys **71** (2013) 163.
56. D. Chruściński and J. Pytel,
Constructing optimal entanglement witnesses. II. Witnessing entanglement in $4N \otimes 4N$ systems,
Phys. Rev. A **82** (2010) 052310.
57. D. Chruściński and G. Sarbicki,
Exposed positive maps: a sufficient condition,
J. Phys. A: Math. Theor. **45** (2012) 115304.
58. S.-H. Kye,
Facial structures for various notions of positivity and applications to the theory of entanglement,
Rev. Math. Phys. **25** (2013) 1330002.
59. K.-C. Ha and S.-H. Kye,
Optimality for indecomposable entanglement witnesses,
Phys. Rev. A, **86** (2012) 034301.
60. R. Augusiak, J. Tura, and M. Lewenstein,
A note on the optimality of decomposable entanglement witnesses and completely entangled subspaces,
J. Phys. A: Math. Theor. **44** (2011) 212001.
61. R. Augusiak, G. Sarbicki, and M. Lewenstein,
Optimal decomposable witnesses without the spanning property,
Phys. Rev. A **84** (2011) 052323.
62. D. Chruściński, J. Pytel, and G. Sarbicki,

Constructing optimal entanglement witnesses,
Phys. Rev. A **80** (2009) 062314.
63. D. Chruściński and J. Pytel,
Optimal entanglement witnesses from generalized reduction and Robertson maps,
J. Phys. A: Math. Theor. **44** (2011) 165304.
64. D. Chruściński and G. Sarbicki,
Disproving the conjecture on the structural physical approximation to optimal decomposable entanglement witnesses,
J. Phys. A: Math. Theor. **47** (2014) 195301.
65. G. Dahl, J. M. Leinaas, J. Myrheim, and E. Ovrum,
A tensor product matrix approximation problem in quantum physics,
Lin. Alg. Appl., **420** (2007) 711.
66. J. Eisert, P. Hyllus, O. Gühne, and M. Curty,
Complete hierarchies of efficient approximations to problems in entanglement theory,
Phys. Rev. A **70** (2004) 062316.
67. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery,
Numerical Recipes,
Cambridge University Press, 2007.
68. MATLAB R2011b Optimization Toolbox Documentation.
`http://www.mathworks.se/help/toolbox/optim/index.html`.
69. Mathematica 8 Optimization Documentation.
`http://reference.wolfram.com/mathematica/guide/Optimization.html`.
70. Ø. S. Garberg,
Numerical optimization of expectation values in product states,
Project thesis, Norwegian University of Science and Technology, Dept. of Physics, 2011.

Paper III

L.O. Hansen and J. Myrheim

"Visualising extremal positive maps in unital and trace preserving form"

Phys. Rev. A **92**, 042306 (2015)

PHYSICAL REVIEW A **92**, 042306 (2015)

Visualizing extremal positive maps in unital and trace-preserving form

Leif Ove Hansen and Jan Myrheim
Department of Physics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway
(Received 21 May 2015; published 5 October 2015)

We define an entanglement witness in a composite quantum system as an observable having a nonnegative expectation value in every separable state. Then a state is entangled if and only if it has a negative expectation value of some entanglement witness. Equivalent representations of entanglement witnesses are nonnegative biquadratic forms or positive linear maps of Hermitian matrices. As reported elsewhere, we have studied extremal entanglement witnesses in dimension 3×3 by constructing numerical examples of generic extremal nonnegative forms. These are so complicated that we do not know how to handle them other than by numerical methods. However, the corresponding extremal positive maps can be presented graphically, as we attempt to do in the present paper. We understand that a positive map is extremal when the image of $\mathcal{D}$, the set of density matrices, fills out $\mathcal{D}$ maximally, in a certain sense. For the graphical presentation of a map we transform it to a standard form where it is unital and trace preserving. We present an iterative algorithm for the transformation, which converges rapidly in all our numerical examples and presumably works for any positive map. This standard form of an entanglement witness is unique up to unitary product transformations.

DOI: 10.1103/PhysRevA.92.042306 PACS number(s): 03.67.Mn, 03.65.Ud, 02.40.Ft

I. INTRODUCTION

The phenomenon of quantum entanglement [1,2] is central to the development and understanding of quantum information theory and computation [3]. For a general bipartite mixed quantum state it is not known how to determine in an efficient way whether the state is entangled or separable. An experimental setup usually involves mixed quantum states, and the need to control, manipulate, and quantify entanglement in such states is therefore of fundamental importance. The characterization of entangled states of a quantum system composed of two subsystems thus remains one of the key unsolved problems in the theory of quantum information. We will consider here only the finite-dimensional case.

Entanglement of a quantum state may be revealed experimentally through negative expectation values of certain observables, called entanglement witnesses, that have positive expectation values in all separable states [4,5]. The Choi-Jamiołkowski isomorphism relates the entanglement witnesses to positive maps on matrix algebras [6,7]. In this way the study of entanglement witnesses is closely related to the study of positive maps. This study was pioneered by Størmer [8], who has recently given an extensive review [9].

In order to investigate the separability problem using positive maps we need to understand the structure of the set of positive maps. Unfortunately, this is a problem which seems just as difficult as the original separability problem. Since the set of positive maps is a compact convex set, a classification of the extremal positive maps would entail a classification of all positive maps. This is our motivation for studying the extremal positive maps.

Outline of this paper

The contents of the paper are organized in the following manner. In Sec. II we review some basic theory regarding positive maps and entanglement witnesses. We define the link between entanglement witnesses and positive maps, known as the Choi-Jamiołkowski isomorphism. An important class of positive maps, the completely positive maps, is defined. These maps are quantum operations since they map quantum states into quantum states. The important distinction between decomposable and nondecomposable maps is made, and the relation to the Peres separability criterion [10] is shown. Furthermore, the unital and trace-preserving properties of positive maps are defined, and the relation between these two properties is established. These two concepts are further discussed in Sec. III. Finally, the expectation values of entanglement witnesses in pure product states are introduced. These are positive semidefinite biquadratic forms, the zeros of which play an important role. The importance of the zeros for the understanding of the extremal entanglement witnesses and the corresponding extremal positive maps is outlined.

In Sec. III we argue that any entanglement witness may be transformed into a unital and trace-preserving form through a product transformation. We use the fact that a map is trace preserving if and only if the transposed map is unital to define conditions on this product transformation. We then suggest an iteration scheme to solve the resulting equations, and we define and investigate the conditions under which this iteration scheme will converge.

Finally, in Sec. IV we use the unital and trace-preserving form of a generic extremal entanglement witness in the 3×3 system that we have produced numerically [11] to create various plots showing how the corresponding positive map acts. We present similar plots illustrating the action of the Choi-Lam extremal positive map. We also discuss briefly an example of extremal witnesses in dimension 2×4.

II. PRELIMINARY THEORY

In this section we introduce our notation and review some mathematical background. We write H_k for the set of Hermitian $k \times k$ matrices, which is a real Hilbert space with the natural scalar product

$$\langle X, Y \rangle = \mathrm{Tr}\,(XY)\,. \tag{1}$$

The mathematical description of a finite-dimensional composite quantum system involves the complex tensor product $\mathbb{C}^N = \mathbb{C}^m \otimes \mathbb{C}^n$ and the real tensor product $H_N = H_m \otimes H_n$ with $N = mn$. We write the components of $\psi \in \mathbb{C}^N$ as $\psi_I = \psi_{ij}$, where

$$I = 1,2,\ldots,N \leftrightarrow ij = 11,12,\ldots,1n,21,\ldots,mn\ . \quad (2)$$

The matrix elements of an $N \times N$ matrix A are $A_{IK} = A_{ij;kl}$, with $I \leftrightarrow ij$ and $K \leftrightarrow kl$. A product vector $\psi = \phi \otimes \chi$ has components $\psi_{ij} = \phi_i \chi_j$, and a product matrix $A = B \otimes C$ has components $A_{ij;kl} = B_{ik} C_{jl}$. We define the partial transpose A^P as the transpose with respect to the second factor in the tensor product,

$$(A^P)_{ij;kl} = A_{il;kj}. \quad (3)$$

We use *density matrices*, positive matrices of unit trace, to represent physical states of the most general kind, so-called mixed states. The set of all $N \times N$ density matrices we call $\mathcal{D}_N$, or $\mathcal{D}$ if we need not specify the dimension. It is a compact convex set of dimension $N^2 - 1$, completely determined by its extremal points, which are the pure states $\rho = \psi\psi^\dagger$ with $\psi \in \mathbb{C}^N$.

It is a remarkable fact that the partial transpose ρ^P of a density matrix ρ need not be positive definite. A density matrix with a positive partial transpose is called a PPT state; thus we define the set of PPT states as $\mathcal{P} = \mathcal{D} \cap \mathcal{D}^P$.

A. The geometry of density matrices

A density matrix $\rho \in \mathcal{D}_N$ lies on the boundary $\partial\mathcal{D}_N$ if it has rank $k < N$. Then it lies in the interior of a face $\mathcal{D}_k(\mathcal{U})$, which is the set of density matrices on a k-dimensional subspace $\mathcal{U} \subset \mathbb{C}^N$. Every face of $\mathcal{D}_N$ is of the type $\mathcal{D}_k(\mathcal{U})$. Thus $\partial\mathcal{D}_N$ consists of faces arranged in the same hierarchical structure as the lattice of subspaces of $\mathbb{C}^N$.

A 2×2 density matrix has the form

$$\rho = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix}, \quad (4)$$

with x,y,z being real and $x^2 + y^2 + z^2 \leqslant 1$. Thus $\mathcal{D}_2$ is a three-dimensional sphere, the Bloch sphere. The boundary states, with $x^2 + y^2 + z^2 = 1$, are the pure states.

The set $\mathcal{D}_3$ of 3×3 density matrices has dimension eight. Its boundary $\partial\mathcal{D}_3$ consists of a seven-dimensional set of rank-two matrices and a four-dimensional set of rank-one matrices, the pure states. Every rank-two state in $\mathcal{D}_3$ lies in the interior of a Bloch sphere $\mathcal{D}_2(\mathcal{U})$.

B. Entanglement and entanglement witnesses

A density matrix $\rho \in \mathcal{D}_N$ is *separable* if it is of the form

$$\rho = \sum_a p_a\, \sigma_a \otimes \tau_a, \quad (5)$$

where $\sigma_a \in \mathcal{D}_m$, $\tau_a \in \mathcal{D}_n$, and $p_a > 0$, $\sum_a p_a = 1$. Its partial transpose is then positive. This fact, that $\mathcal{S} \subset \mathcal{P}$, where $\mathcal{S}$ is the set of separable states, provides an easy test for separability, known as the Peres criterion [10]. A density matrix is *entangled* if it is not separable. The separability problem, how to decide whether a given density matrix is separable or entangled, is difficult because it is difficult to recognize the entangled PPT states.

We define the dual set

$$\mathcal{S}^\circ = \{A \in H_N | \mathrm{Tr}(A\rho) \geqslant 0 \quad \forall\, \rho \in \mathcal{S}\}. \quad (6)$$

The dual of $\mathcal{S}^\circ$ is $\mathcal{S}$; thus ρ is separable if and only if $\mathrm{Tr}(A\rho) \geqslant 0$ for every $A \in \mathcal{S}^\circ$. For this reason we call a matrix $A \in \mathcal{S}^\circ$ an *entanglement witness*. Note that we do not impose here the usual condition on an entanglement witness A that it should have at least one negative eigenvalue, so that $\mathrm{Tr}(A\rho) < 0$ for some $\rho \in \mathcal{D}_N$.

The convex set $\mathcal{S}$ is completely determined by its extremal points, which are the pure product states $\rho = \psi\psi^\dagger = (\phi\phi^\dagger) \otimes (\chi\chi^\dagger)$, with $\psi = \phi \otimes \chi$ and $\phi^\dagger\phi = \chi^\dagger\chi = 1$, for which

$$\mathrm{Tr}(A\rho) = \psi^\dagger A \psi = (\phi \otimes \chi)^\dagger A (\phi \otimes \chi) = \sum_{i,j,k,l} \phi_i^* \chi_j^* A_{ij;kl} \phi_k \chi_l. \quad (7)$$

Thus $A \in \mathcal{S}^\circ$ if and only if the biquadratic form

$$f_A(\phi,\chi) = (\phi \otimes \chi)^\dagger A (\phi \otimes \chi) \quad (8)$$

is nonnegative definite. Since

$$(\phi \otimes \chi)^\dagger A^P (\phi \otimes \chi) = (\phi \otimes \chi^*)^\dagger A (\phi \otimes \chi^*), \quad (9)$$

we see that $A^P \in \mathcal{S}^\circ$ if and only if $A \in \mathcal{S}^\circ$.

C. Maps

A *map*, in the terminology used here, is a linear transformation $\mathbf{M} : H_m \mapsto H_n$ such that $Y = \mathbf{M}X$ when

$$Y_{jl} = \sum_{i,k} M_{jl;ik} X_{ik}. \quad (10)$$

The complex matrix elements $M_{jl;ik}$ defining the map satisfy the relations $M_{jl;ik} = (M_{lj;ki})^*$, so that the map preserves Hermiticity. The relation

$$A_{ij;kl} = M_{jl;ki} \quad (11)$$

defines a one-to-one correspondence between a Hermitian matrix $A \in H_N$ and a map $\mathbf{M} = \mathbf{M}_A$. As we define the correspondence here, it differs slightly from the well-known Choi-Jamiołkowski isomorphism, which associates the matrix A with the map $X \mapsto \mathbf{M}_A(X^T)$.

It is useful to introduce matrices $E_a \in H_m$ and $F_b \in H_n$ that are basis vectors in the two Hilbert spaces. Then the map $\mathbf{M}$ is given by real matrix elements M_{ba} such that

$$\mathbf{M}E_a = \sum_b M_{ba} F_b. \quad (12)$$

We will choose the basis vectors to be orthonormal; then the matrix elements are

$$M_{ba} = \langle F_b, \mathbf{M}E_a \rangle. \quad (13)$$

Furthermore, we will choose the unit matrices as basis vectors, $E_0 = I_m/\sqrt{m}$, $F_0 = I_n/\sqrt{n}$. Then the matrices E_a for $a = 1,2,\ldots,m^2 - 1$ and F_b for $b = 1,2,\ldots,n^2 - 1$ are traceless.

The transposed map $\mathbf{N} = \mathbf{M}^T : H_n \mapsto H_m$ may be defined in a basis-independent way by the condition that

$$\langle \mathbf{M}^T Y, X \rangle = \langle Y, \mathbf{M}X \rangle \quad \text{for all} \quad X \in H_m,\ Y \in H_n. \quad (14)$$

It has matrix elements $N_{ab} = M_{ba}$, or $N_{ik;jl} = M_{lj;ki} = (M_{jl;ik})^*$. Thus transposition of the map $\mathbf{M}$ corresponds to transposition of the real matrix M_{ab} and Hermitian conjugation of the complex matrix $M_{jl;ik}$.

As defined here, maps $\mathbf{M}_A$ and $\mathbf{M}_A^T$ act on $X \in H_m$ and $Y \in H_n$ as follows:

$$\mathbf{M}_A X = \mathrm{Tr}_1[A(X \otimes I_n)], \quad \mathbf{M}_A^T Y = \mathrm{Tr}_2[A(I_m \otimes Y)], \quad (15)$$

where Tr_1 and Tr_2 denote the partial traces. The matrix A may be expanded in terms of the basis vectors $E_a \otimes F_b \in H_N$ as

$$A = \sum_{a,b} M_{ba}\, E_a \otimes F_b. \quad (16)$$

D. Entanglement witnesses and positive maps

By definition, a *positive* map $\mathbf{M} : H_m \mapsto H_n$ transforms positive matrices into positive matrices; that is, if $\rho \geqslant 0$, then $\mathbf{M}\rho \geqslant 0$.

In terms of the map $\mathbf{M}_A$ the biquadratic form introduced in Eq. (8) may be written as

$$f_A(\phi, \chi) = \chi^\dagger \mathbf{M}_A(\phi\phi^\dagger)\, \chi = \phi^\dagger \mathbf{M}_A^T(\chi\chi^\dagger)\, \phi. \quad (17)$$

The condition that $f_A(\phi, \chi) \geqslant 0$ for all $\phi \in \mathbb{C}^m$, $\chi \in \mathbb{C}^n$ means that $\mathbf{M}_A(\phi\phi^\dagger)$ is a positive matrix for every $\phi \in \mathbb{C}^m$. Since the pure states $\phi\phi^\dagger$ are all the extremal points of $\mathcal{D}_m$, this is another way of saying that $\mathbf{M}_A$ is a positive map.

We conclude that A is an entanglement witness if and only if $\mathbf{M}_A$ is a positive map. An equivalent condition is that $\mathbf{M}_A^T$ is a positive map. This correspondence between entanglement witnesses and positive maps places the positive maps in a central position in the theory of quantum entanglement.

E. Completely positive maps

An obvious condition on a physical map $\mathbf{M} : H_m \mapsto H_n$, transforming physical states into physical states, is that it should be positive. A less obvious condition is that it should be *completely positive*, so that every map of the form $\mathbf{I} \otimes \mathbf{M} : H_k \otimes H_m \mapsto H_k \otimes H_n$ is positive, where $\mathbf{I} : H_k \mapsto H_k$ is the identity map.

The most general form of a completely positive map is

$$\mathbf{M}_A X = \sum_a V_a X V_a^\dagger, \quad (18)$$

with $n \times m$ matrices V_a. This form gives the following matrix elements of A:

$$A_{ij;kl} = (A^P)_{il;kj} = \sum_a (V_a)_{jk} (V_a)^*_{li}. \quad (19)$$

We see that the map $\mathbf{M}_A$ is completely positive if and only if A^P is positive. It is well known that A^P may be positive without A being positive because the transposition map $\mathbf{T}X = X^T$ on H_n is positive but not completely positive.

A separable state ρ, as defined in Eq. (5), remains positive under the application of positive maps to its local parts; thus

$$(\mathbf{M} \otimes \mathbf{N})\rho = \sum_a p_a\, (\mathbf{M}\sigma_a) \otimes (\mathbf{N}\tau_a) \quad (20)$$

is a positive matrix when $\mathbf{M} : H_m \mapsto H_k$ and $\mathbf{N} : H_n \mapsto H_l$ are positive maps. In this way positive maps give necessary conditions for separability.

Positive maps can also give sufficient conditions for separability [4]. This follows from the correspondence described above between entanglement witnesses and positive maps. For $\rho \in H_N = H_{mn}$ let us define $\sigma \in H_{m^2}$ as

$$\sigma = \left(\mathbf{T} \otimes \mathbf{M}_A^T\right)\rho, \quad (21)$$

where A is an entanglement witness and $\mathbf{M}_A$ is the corresponding positive map. In index notation we have that

$$\sigma_{ka;ib} = \sum_{j,l} A_{al;bj}\, \rho_{ij;kl}. \quad (22)$$

Clearly, σ is positive if ρ is separable. Conversely, if σ is positive for every entanglement witness A, we have that

$$\mathrm{Tr}(A\rho) = \sum_{i,j,k,l} A_{kl;ij}\, \rho_{ij;kl} = \sum_{k,a,i,b} \delta_{ka} \sigma_{ka;ib} \delta_{ib} \geqslant 0 \quad (23)$$

for every A, implying that ρ is separable.

Note that this proof of sufficiency does not use the full condition that σ should be positive; it uses only the condition that σ should have a nonnegative expectation value in the maximally entangled pure state (Bell state) $\delta_{ik} \in \mathbb{C}^{m^2}$. If we use the full positivity condition, one single positive map can, in principle, reveal the entanglement of many different states for which we would need many different entanglement witnesses. The PPT condition associated with the transposition map is a striking example; it is equivalent to all the conditions provided by a much larger set of entanglement witnesses.

The basic reason for the efficiency of maps in revealing entanglement is nonlinearity: the separability condition that a product $\mathbf{M} \otimes \mathbf{N}$ of positive maps should map ρ into a positive matrix is highly nonlinear in ρ. An entanglement witness A, on the other hand, gives an inequality $\mathrm{Tr}(A\rho) \geqslant 0$ linear in ρ.

F. Decomposable maps

A positive map $\mathbf{M} : H_m \mapsto H_n$ is said to be *decomposable* if it can be written as

$$\mathbf{M} = \mathbf{M}_1 + \mathbf{M}_2 \mathbf{T}_m, \quad (24)$$

where $\mathbf{M}_1$ and $\mathbf{M}_2$ are completely positive and $\mathbf{T}_m$ is the transposition on H_m [8,12].

Since $\mathbf{M}_A \mathbf{T}_m = \mathbf{M}_{A^P}$, it follows that $\mathbf{M}_A$ is decomposable if and only if $A = B + C$, where B and C^P are both positive matrices. This decomposition of A implies for every PPT state ρ that

$$\mathrm{Tr}(A\rho) = \mathrm{Tr}(B\rho) + \mathrm{Tr}(C\rho) = \mathrm{Tr}(B\rho) + \mathrm{Tr}(C^P \rho^P) \geqslant 0. \quad (25)$$

In particular, A is an entanglement witness, and we call it a decomposable witness. It is not very useful as a witness, however, since it cannot reveal the entanglement of an entangled PPT state.

It was shown by Woronowicz that all positive maps $\mathbf{M} : H_m \mapsto H_n$ with $N = mn \leqslant 6$ are decomposable [13]. In these dimensions entangled PPT states therefore do not exist, and the Peres criterion is both necessary and sufficient for separability.

By virtue of the Peres criterion, the development of practically useful separability criteria in dimensions $N \geqslant 8$ is a problem closely related to the understanding of the nondecomposable positive maps.

G. Unital and trace-preserving maps

A map $\mathbf{M} : H_m \mapsto H_n$ is *unital* if it maps the identity $I_m \in H_m$ to the identity $I_n \in H_n$, that is, if

$$\sum_{i,k} M_{jl;ik}\,\delta_{ik} = \sum_{i} M_{jl;ii} = \delta_{jl}. \tag{26}$$

It is *trace preserving* if $\mathrm{Tr}(\mathbf{M}X) = \mathrm{Tr}\,X$ for all $X \in H_m$, that is, if

$$\sum_{j,i,k} M_{jj;ik} X_{ik} = \sum_{i} X_{ii}, \tag{27}$$

which means that

$$\sum_{j} M_{jj;ik} = \delta_{ik}. \tag{28}$$

Thus the map $\mathbf{M}$ is trace preserving if and only if $\mathbf{M}^T$ is unital. A physical map should conserve probability, which means that it should be trace preserving.

In terms of the corresponding matrix $A \in H_N$ the condition for $\mathbf{M}_A$ to be unital is that

$$\mathbf{M}_A I_m = \mathrm{Tr}_1 A = I_n, \tag{29}$$

and the condition for $\mathbf{M}_A$ to be trace preserving is that

$$\mathbf{M}_A^T I_n = \mathrm{Tr}_2 A = I_m. \tag{30}$$

Equation (29) implies that $\mathrm{Tr}\,A = \mathrm{Tr}\,I_n = n$, whereas Eq. (30) implies that $\mathrm{Tr}\,A = \mathrm{Tr}\,I_m = m$. Thus the map may be both unital and trace preserving only if $m = n$. If $m \neq n$, the proper condition is rather that

$$\mathbf{M}_A E_0 = F_0, \quad \mathbf{M}_A^T F_0 = E_0, \tag{31}$$

where $E_0 = I_m/\sqrt{m}$ and $F_0 = I_n/\sqrt{n}$. Since we regard normalization constants as unimportant, we will abuse the language and call the map unital and trace preserving if Eq. (31) holds, even when $m \neq n$.

In terms of the matrix elements $M_{ba} = \langle F_b, \mathbf{M}_A E_a\rangle$ the condition for unitality is that $M_{b0} = \delta_{b0}$, and the condition for trace preservation is that $M_{0a} = \delta_{0a}$. This means that a unital and trace-preserving map $\mathbf{M}_A$ is generated by an entanglement witness of the form

$$A = \frac{I_N}{\sqrt{N}} + \sum_{a=1}^{m^2-1}\sum_{b=1}^{n^2-1} M_{ba}\, E_a \otimes F_b. \tag{32}$$

H. Zeros of entanglement witnesses

Since the conditions defining $A \in H_N$ as an entanglement witness are the infinite set of inequalities

$$f_A(\phi,\chi) \geqslant 0 \quad \text{for all} \quad \phi \in \mathbb{C}^m,\ \chi \in \mathbb{C}^n, \tag{33}$$

it is clear that A is a boundary point of $\mathcal{S}^\circ$ if and only if at least one of these inequalities is an equality. We call the pure product state $\psi_0\psi_0^\dagger$, with $\psi_0 = \phi_0 \otimes \chi_0$ and $\phi_0^\dagger\phi_0 = \chi_0^\dagger\chi_0 = 1$, a zero of A if

$$f_A(\phi_0,\chi_0) = 0. \tag{34}$$

Since a zero is a minimum of the nonnegative function f_A, it follows that the first derivative in every direction at the zero must also vanish. These conditions amount to a set of equalities

$$(\phi_0 \otimes \chi_0)^\dagger A(\phi \otimes \chi_0) = (\phi_0 \otimes \chi_0)^\dagger A(\phi_0 \otimes \chi) = 0 \quad \text{for all} \quad \phi \in \mathbb{C}^m,\ \chi \in \mathbb{C}^n. \tag{35}$$

These equations may be regarded as linear constraints on the matrix A. It should be remembered that each equation is complex and is equivalent to two real equations, except that Eq. (34) is real. A careful counting shows that one zero implies $2(m+n)-3$ real-valued constraints on A.

If the second derivative in every direction is strictly positive, then the zero (ϕ_0,χ_0) is a quadratic minimum, and there are no more constraints on A from this zero. If the second derivative in some direction vanishes, then the third derivative in this direction must also vanish, and the zero is a quartic minimum. This imposes further constraints on A, which we do not detail here.

The distinction between quadratic and quartic zeros is important. Obviously, the quadratic zeros are isolated points. A quartic zero may be an isolated point, but it may also belong to a continuous set of zeros.

An entanglement witness A is extremal if and only if it has so many zeros that all the constraints from all the zeros together determine A uniquely up to a proportionality constant.

I. Extremal positive maps

An extremal entanglement witness A is understood in terms of its zeros. It corresponds to an extremal positive map $\mathbf{M}_A$, and we have to understand what the zeros of the witness imply for the corresponding map.

Let (ϕ_0,χ_0) be an isolated zero of A, and define $Y = \mathbf{M}_A(\phi_0\phi_0^\dagger)$, $X = \mathbf{M}_A^T(\chi_0\chi_0^\dagger)$. We use the identities

$$\chi_0^\dagger Y \chi_0 = \phi_0^\dagger X \phi_0 = (\phi_0 \otimes \chi_0)^\dagger A\,(\phi_0 \otimes \chi_0) = 0. \tag{36}$$

Since Y and X are both positive matrices, it follows that

$$Y\chi_0 = 0, \quad X\phi_0 = 0. \tag{37}$$

We also know that Y and X have no other zero vectors; otherwise, the zero vectors would span subspaces of dimension two or higher, and the zero (ϕ_0,χ_0) would not be isolated. Hence Y has rank $n-1$ and X has rank $m-1$.

To summarize, when A is an extremal entanglement witness, the corresponding maps $\mathbf{M}_A$ and $\mathbf{M}_A^T$ are extremal positive maps. An isolated zero (ϕ_0,χ_0) of A defines a rank-one state $\phi_0\phi_0^\dagger \in \mathcal{D}_m$ mapped by $\mathbf{M}_A$ to a rank-$(n-1)$ state in $\mathcal{D}_n$ and a rank-one state $\chi_0\chi_0^\dagger \in \mathcal{D}_n$ mapped by $\mathbf{M}_A^T$ to a rank-$(m-1)$ state in $\mathcal{D}_m$.

Thus the zero of A defines a point on the boundary of $\mathcal{D}_m$ which is mapped to a point on the boundary of $\mathcal{D}_n$ and a point on the boundary of $\mathcal{D}_n$ which is mapped by the transposed map to a point on the boundary of $\mathcal{D}_m$. We understand that the map $\mathbf{M}_A$ is extremal precisely because the image $\mathbf{M}_A\mathcal{D}_m$ inside $\mathcal{D}_n$ touches the boundary of $\mathcal{D}_n$ in as many points as possible.

J. Rank-one-preserving maps

The identity map $\mathbf{I} : X \mapsto X$ and the transposition map $\mathbf{T} : X \mapsto X^T = X^*$ are two examples of extremal positive maps. They have the special property that they are rank one preservers, mapping pure states to pure states. They are extremal precisely because they map the boundary of $\mathcal{D}$ onto itself. They both preserve volume. In fact, the eigenvalues of $\mathbf{T}$ are ± 1 because $\mathbf{T}$ is its own inverse, $\mathbf{T}^2 = \mathbf{I}$.

More generally, let U be a unitary matrix and define $\mathbf{U} : X \mapsto UXU^\dagger$. Then $\mathbf{U}$ and $\mathbf{UT}$ are both rank one preservers, and they are extremal because they map the boundary of $\mathcal{D}$ onto itself. They also preserve volumes.

These maps and the corresponding entanglement witnesses are fundamentally different from the generic extremal witnesses we find numerically, which have only quadratic zeros, so that the corresponding maps map only a finite number of points from the boundary of $\mathcal{D}$ to the boundary and map all other boundary points to the interior. An extremal map of this generic kind is contractive; it reduces volumes since it maps $\mathcal{D}$ to a subset of itself.

III. TRANSFORMING POSITIVE MAPS TO UNITAL AND TRACE-PRESERVING FORM

We argue now that every positive map $\mathbf{M} = \mathbf{M}_A$, where A is an entanglement witness, may be transformed into a unital and trace-preserving form through a product transformation of the form [9]

$$A \mapsto \widetilde{A} = (U \otimes V)A(U \otimes V)^\dagger, \tag{38}$$

where $U \in \mathrm{GL}(m,\mathbb{C})$ and $V \in \mathrm{GL}(n,\mathbb{C})$. Furthermore, we present here an efficient iteration procedure for doing the transformation numerically.

A product transformation of this kind preserves all the essential characteristics of A. For instance, if A is extremal in $\mathcal{S}^\circ$ and nondecomposable, then so is $\widetilde{A}$, and a zero $\phi \otimes \chi$ of A corresponds to a zero $\widetilde{\phi} \otimes \widetilde{\chi}$ of $\widetilde{A}$, where

$$\widetilde{\phi} = (U^\dagger)^{-1}\phi, \quad \widetilde{\chi} = (V^\dagger)^{-1}\chi. \tag{39}$$

As we saw in Sec. II G the result to be proved is that every entanglement witness A may be transformed to a form as given in Eq. (32). In [14] it was proved that every strictly positive density matrix may be transformed to such a form. But the theorem is actually valid more generally than it is stated there since the proof is valid for every entanglement witness having no zeros, in other words, every witness lying in the interior of $\mathcal{S}^\circ$. Here we want to apply this type of transformation to extremal witnesses, which lie on the boundary $\partial\mathcal{S}^\circ$ and have a maximal number of zeros. What could, in principle, go wrong in the limit when the boundary is approached from the inside of $\mathcal{S}^\circ$ is that the transformation could become singular, but we find in practice that no such problems arise.

In terms of the transformed map $\widetilde{\mathbf{M}} = \mathbf{M}_{\widetilde{A}}$ the conditions to be fulfilled are that $\widetilde{\mathbf{M}} I_m = I_n$ and $\widetilde{\mathbf{M}}^T I_n = I_m$. We will assume here that $m = n$; otherwise, the proper conditions would be that $\widetilde{\mathbf{M}}(I_m/\sqrt{m}) = I_n/\sqrt{n}$ and $\widetilde{\mathbf{M}}^T(I_n/\sqrt{n}) = I_m/\sqrt{m}$.

In index notation Eq. (38) reads as follows:

$$\widetilde{A}_{ij;kl} = \sum_{a,b,c,d} U_{ia} V_{jb}\, A_{ab;cd}\, U^*_{kc} V^*_{ld}. \tag{40}$$

According to Eq. (15), the transformation $Y = \widetilde{\mathbf{M}}X$ then reads as follows:

$$Y_{jl} = \sum_{i,k} \widetilde{A}_{ij;kl}\, X_{ki} = \sum_{i,k,a,b,c,d} V_{jb}(A_{ab;cd}(U^*_{kc}\, X_{ki}\, U_{ia}))V^*_{ld}, \tag{41}$$

or in the index-free notation,

$$Y = \widetilde{\mathbf{M}}X = V(\mathbf{M}(U^\dagger X U))V^\dagger. \tag{42}$$

Similarly, the transformation $X = \widetilde{\mathbf{M}}^T Y$ reads

$$X_{ik} = \sum_{j,l} \widetilde{A}_{ij;kl} Y_{lj} = \sum_{j,l,a,b,c,d} U_{ia}(A_{ab;cd}(V^*_{ld} Y_{lj} V_{jb}))U^*_{kc}, \tag{43}$$

or

$$X = \widetilde{\mathbf{M}}^T Y = U(\mathbf{M}^T(V^\dagger Y V))U^\dagger. \tag{44}$$

The conditions for $\widetilde{\mathbf{M}}$ to be unital and trace preserving are that

$$\begin{aligned} \widetilde{\mathbf{M}} I_m &= V(\mathbf{M}(U^\dagger U))V^\dagger = I_n, \\ \widetilde{\mathbf{M}}^T I_n &= U(\mathbf{M}^T(V^\dagger V))U^\dagger = I_m. \end{aligned} \tag{45}$$

Thus the problem to be solved is to find U and V such that

$$\mathbf{M}(U^\dagger U) = (V^\dagger V)^{-1}, \quad \mathbf{M}^T(V^\dagger V) = (U^\dagger U)^{-1}. \tag{46}$$

Solution by iteration

This problem can be solved in two steps. First, we find positive Hermitian matrices $X = U^\dagger U$ and $Y = V^\dagger V$ solving the equations

$$\mathbf{M}X = Y^{-1}, \quad \mathbf{M}^T Y = X^{-1}. \tag{47}$$

Then we solve the equations

$$U^\dagger U = X, \quad V^\dagger V = Y \tag{48}$$

for U and V. The general solutions are

$$U = U_1 U_2, \quad V = V_1 V_2, \tag{49}$$

where U_1 and V_1 are arbitrary unitary matrices and $U_2 = \sqrt{X}$, $V_2 = \sqrt{Y}$ are the uniquely defined positive Hermitian square roots, which we compute, for example, by diagonalizing X and Y and taking the square roots of the eigenvalues.

Equation (47) makes sense because the matrices $X = U^\dagger U$ and $Y = V^\dagger V$ are strictly positive as long as U and V are nonsingular, and the maps $\mathbf{M}$ and $\mathbf{M}^T$, as well as the inversions $X \mapsto X^{-1}$ and $Y \mapsto Y^{-1}$, transform strictly positive matrices into strictly positive matrices.

The method suggesting itself for solving Eq. (47) is simply to iterate the equations. Given an approximate solution X_k for X, we try to compute a better approximation X_{k+1} by a series of four transformations,

$$\begin{aligned} X_k \mapsto S_k = \mathbf{M}X_k \mapsto Y_k = S_k^{-1} \mapsto T_k \\ = \mathbf{M}^T Y_k \mapsto X_{k+1} = T_k^{-1}. \end{aligned} \tag{50}$$

We start the iterations, for example, with $X_0 = I$.

A sufficient condition for the convergence of X_k to a unique limit X is that each transformation $X_k \mapsto X_{k+1}$ is contractive (except in the direction along X). A small perturbation ΔX_k

of X_k transforms linearly as

$$\Delta X_{k+1} = \mathbf{D}(\Delta X_k), \tag{51}$$

where the linear map $\mathbf{D}$ is the derivative of the nonlinear transformation $X_k \mapsto X_{k+1}$. The transformation is contractive if all eigenvalues of $\mathbf{D}$ (except the special eigenvalue which must be close to 1) are smaller than 1 in absolute value. Now $\mathbf{D}$ is a composition of four linear maps,

$$\mathbf{D} = \mathbf{D}_4\mathbf{M}^T\mathbf{D}_2\mathbf{M}, \tag{52}$$

where $\mathbf{D}_2$ and $\mathbf{D}_4$ are linearizations of the matrix inversions,

$$\mathbf{D}_2(\Delta S_k) = -S_k^{-1}(\Delta S_k)S_k^{-1},$$
$$\mathbf{D}_4(\Delta T_k) = -T_k^{-1}(\Delta T_k)T_k^{-1}. \tag{53}$$

The examples of extremal positive maps $\mathbf{M}$ and $\mathbf{M}^T$ that we have studied numerically are strongly contractive, and we find in practice that this is enough to ensure that $\mathbf{D}$ is contractive, with eigenvalues typically no larger than about 0.5, even though $\mathbf{D}_2$ and $\mathbf{D}_4$ are not contractive.

We have used this iteration scheme on a large number of numerically produced extremal entanglement witnesses [11] and also on many nonextremal witnesses constructed as convex combinations of the extremal ones. Numerically, our attempts, which are in the thousands, always converge, and it also appears that for a given witness A the solution X is unique, independent of the initial guess X_0.

IV. VISUALIZATION OF POSITIVE MAPS

We now specialize to the case $m = n = 3, N = mn = 9$. The set $\mathcal{D} = \mathcal{D}_3$ of normalized density matrices has dimension $3^2 - 1 = 8$. When $m = n$, a linear positive map from H_m to H_n may be transformed to a form in which it is unital and trace preserving.

Given a unital and trace-preserving positive map, $\mathbf{M}: H_3 \mapsto H_3$. It maps $\mathcal{D}$ into $\mathcal{D}$ and the maximally mixed state $I/3$ to itself. We plot two-dimensional planar sections in order to illustrate how the image $\mathbf{M}\mathcal{D}$ lies inside $\mathcal{D}$.

We present such plots here for two different examples of extremal positive maps. The first example corresponds to a randomly chosen generic extremal witness found in a numerical search. Such witnesses have only quadratic zeros. The second example is the Choi-Lam map, which is qualitatively different since the corresponding witness has only quartic zeros.

A. Two-dimensional sections through the set of density matrices

In most of our plots we use three density matrices, $\rho_0, \rho_1, \rho_2 \in \mathcal{D}$, to define a plane $\mathcal{Z} \subset H_3$. We use ρ_0 as the origin in the plane and define coordinate axes

$$B = a(\rho_1 - \rho_0), \quad C = b(\rho_2 - \rho_0) + c(\rho_1 - \rho_0), \tag{54}$$

with real constants a,b,c chosen in some way to be specified later on. We will always choose $a > 0$ and $b > 0$. A matrix $X \in \mathcal{Z}$ is then specified by a coordinate pair (x,y) as

$$X = \rho_0 + xB + yC. \tag{55}$$

Note that $\mathrm{Tr}\,B = \mathrm{Tr}\,C = 0$ because $\mathrm{Tr}\,\rho_i = 1$ for $i = 0,1,2$; hence $\mathrm{Tr}\,X = \mathrm{Tr}\,\rho_0 = 1$.

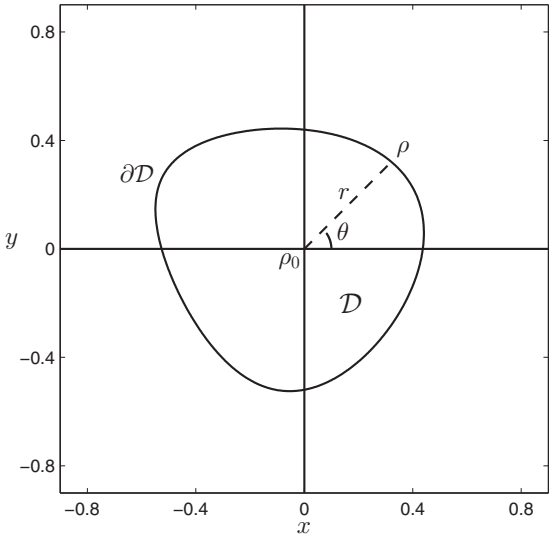


FIG. 1. A two-dimensional section through $\mathcal{D}$, the set of density matrices. The curve represents the boundary $\partial\mathcal{D}$, which consists of the states of less than full rank. For a given angle θ the distance r is calculated numerically. The state $\rho_0 = I/3$ is the origin, and the state ρ on the boundary has rank two.

1. Six types of sections

Figure 1 is an example of a section through the set of density matrices, where we have chosen the maximally mixed state as the origin, that is, $\rho_0 = I/3$. The density matrices ρ_1 and ρ_2 helping to define the section are here chosen at random. By Eq. (54) the x axis goes in the direction from ρ_0 to ρ_1. For this plot we have chosen the constants a,b,c in Eq. (54) such that B and C are orthogonal unit vectors in H_3, that is,

$$\mathrm{Tr}(B^2) = \mathrm{Tr}(C^2) = 1, \quad \mathrm{Tr}(BC) = 0. \tag{56}$$

This means that distances in our plot faithfully represent distances in H_3 as defined by the Hilbert-Schmidt metric. To locate numerically the boundary $\partial\mathcal{D}$ we write $x = r\cos\theta$, $y = r\sin\theta$, and with θ fixed, we determine the largest value of r such that the matrix $X = \rho_0 + xB + yC$ has no negative eigenvalues. The boundary is here a smooth curve consisting entirely of rank-two states.

Figure 2 shows six types of sections of $\mathcal{D}$. In sections A, B, C, and F the origin ρ_0 is the maximally mixed state $I/3$; it is marked by a cross. In sections D and E the cross is the orthogonal projection of the maximally mixed state. All sections, except section E, cut through the interior of $\mathcal{D}$, so that the interior points in the sections are states of the full rank three.

Section A is of the same type as Fig. 1; here ρ_1 and ρ_2 are chosen as random states of rank two, and the boundary is a smooth curve of rank-two states.

In section B we choose ρ_1 to be a pure state, while ρ_2 is a random state of full rank. The pure state is seen in the plot as the single point on the boundary curve where the tangent

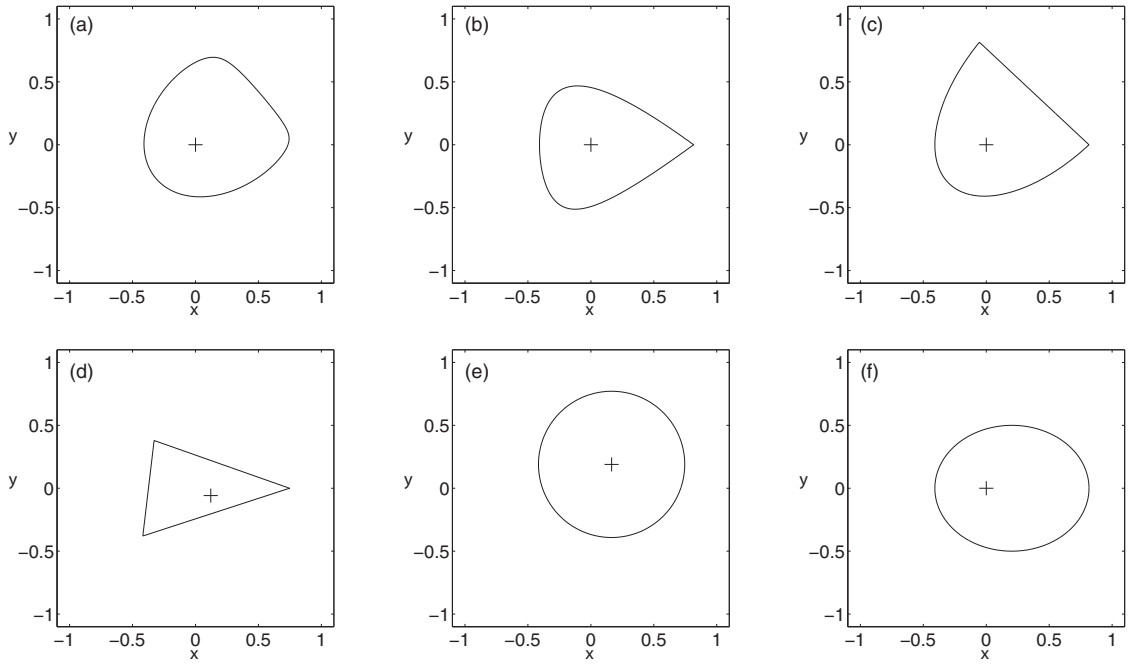


FIG. 2. Six types of two-dimensional sections through $\mathcal{D}$, as explained in the text.

direction is discontinuous. All other boundary points are states of rank two.

In section C both ρ_1 and ρ_2 are pure states; they are joined in the plot by a straight line of rank-two states. The curved part of the boundary also consists of rank-two states.

Section D is a simplex in which the three corners are pure states $\rho_k = \phi_k \phi_k^\dagger$ for $k = 1,2,3$ defined by linearly independent vectors $\phi_k \in \mathbb{C}^3$. The straight lines joining the pure states consist of rank-two states. The origin is chosen as an even mix of the three pure states.

Section E is a two-dimensional section through a three-dimensional Bloch sphere contained in the boundary $\partial\mathcal{D}$. Like section D, it is defined by three pure states, but in this case the three vectors in $\mathbb{C}^3$ are linearly dependent. In this section the interior points have rank two, and the boundary curve is a circle of pure states.

In section F we choose again a pure state $\rho_1 = \phi_1 \phi_1^\dagger$ but a section in which the boundary curve is smooth at this point. The three matrices defining the section are $\rho_0 = I/3$, ρ_1, and $\rho_2 = \rho_1 + D$, where the matrix D is chosen in such a way that $\rho_1 + \epsilon D$ for real ϵ is a pure state to first order in ϵ, that is,

$$(\phi_1 + \epsilon\xi)(\phi_1 + \epsilon\xi)^\dagger = \rho_1 + \epsilon D + O(\epsilon^2) \tag{57}$$

for some $\xi \in \mathbb{C}^3$. We see that the proper choice is

$$D = \phi_1 \xi^\dagger + \xi \phi_1^\dagger. \tag{58}$$

It is of no importance that $\rho_2 \notin \mathcal{D}$. We choose ξ to be orthogonal to ϕ_1 in order to have $\mathrm{Tr}\, D = 0$. The boundary points in this section, apart from ρ_1, are rank-two states.

2. Visualizing maps

The positive map $\mathbf{M}$ maps the plane $\mathcal{Z}$ defined by the states $\rho_i \in \mathcal{D}$ with $i = 0,1,2$ into the plane $\widetilde{\mathcal{Z}} = \mathbf{M}\mathcal{Z}$ defined by the states $\widetilde{\rho}_i = \mathbf{M}\rho_i \in \mathcal{D}$ in such a way that X as given in Eq. (55) is mapped into

$$\widetilde{X} = \mathbf{M}X = \widetilde{\rho}_0 + x\widetilde{B} + y\widetilde{C}, \tag{59}$$

with

$$\widetilde{B} = a(\widetilde{\rho}_1 - \widetilde{\rho}_0), \quad \widetilde{C} = b(\widetilde{\rho}_2 - \widetilde{\rho}_0) + c(\widetilde{\rho}_1 - \widetilde{\rho}_0). \tag{60}$$

It is an important point that $\mathrm{Tr}(\widetilde{B}) = \mathrm{Tr}(\widetilde{C}) = 0$. This follows because we require the map $\mathbf{M}$ to be trace preserving.

The constants a,b,c are the same here as in Eq. (54). We choose them now to have such values that $\widetilde{B}$ and $\widetilde{C}$ are orthogonal unit vectors, that is,

$$\mathrm{Tr}(\widetilde{B}^2) = \mathrm{Tr}(\widetilde{C}^2) = 1, \quad \mathrm{Tr}(\widetilde{B}\widetilde{C}) = 0. \tag{61}$$

The motivation for this choice is that we want our plots of the coordinates (x,y) to represent faithfully the distances in the image plane $\widetilde{\mathcal{Z}}$ of the map $\mathbf{M}$. These values for a,b,c will obviously not, as a rule, be the same values that make B and C orthogonal unit vectors. Thus our (x,y) plot will be a distorted representation of the plane $\mathcal{Z}$. The distortion is of course due to the map $\mathbf{M}$ between the two planes.

To summarize, in one and the same plot the coordinate pair (x,y) represents both the matrix $X \in \mathcal{Z}$, given by Eq. (55), and the matrix $\widetilde{X} = \mathbf{M}X \in \widetilde{\mathcal{Z}}$, given by Eq. (59). Note that our definition of the axes B,C and $\widetilde{B},\widetilde{C}$ is such that $\widetilde{\rho}_1$ always

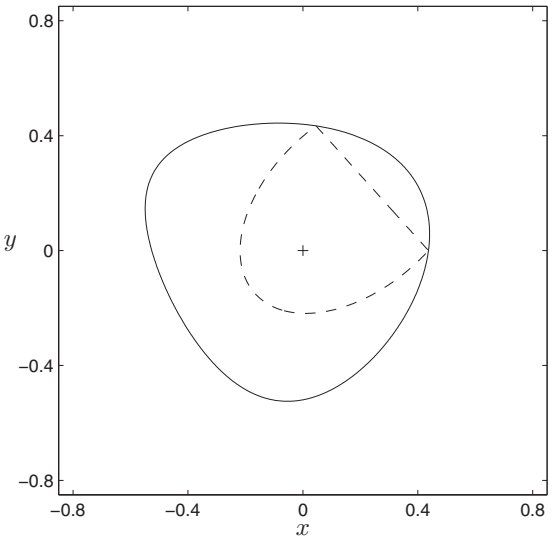


FIG. 3. $\rho_1 = \phi_5\phi_5^\dagger$ and $\rho_2 = \phi_6\phi_6^\dagger$. The origin ρ_0 is the maximally mixed state $I/3$.

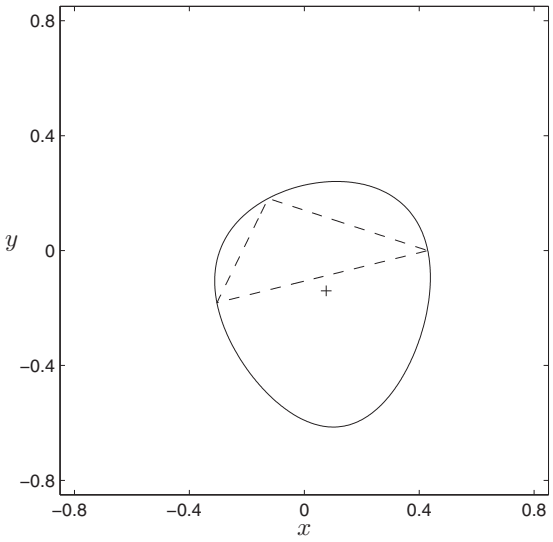


FIG. 4. $\rho_1 = \phi_5\phi_5^\dagger$ and $\rho_2 = \phi_6\phi_6^\dagger$. The origin ρ_0 is an even mix of the pure states ρ_1, ρ_2, and $\rho_3 = \phi_7\phi_7^\dagger$. The triangle is a face of an eight-dimensional simplex defined by the nine zeros of the witness.

has coordinates $y = 0$ and $x > 0$ in the plots, whereas $\widetilde{\rho}_2$ has $y > 0$ but, in general, $x \neq 0$.

We plot the boundary of $\mathcal{D}$ in $\widetilde{\mathcal{Z}}$ by a solid curve and the boundary of $\mathcal{D}$ in $\mathcal{Z}$ by a dashed curve. We calculate these boundaries by the method already described. That is, we write $x = r\cos\theta$, $y = r\sin\theta$. For a fixed value of θ we determine r_1 as the largest value of r such that X has no negative eigenvalue; this defines a point (x_1, y_1) on the boundary of $\mathcal{D}$ in $\mathcal{Z}$, or, equivalently, on the boundary of $\mathbf{M}\mathcal{D}$ in $\widetilde{\mathcal{Z}}$. And we determine r_2 as the largest value of r such that $\widetilde{X}$ has no negative eigenvalue; this defines a point (x_2, y_2) on the boundary of $\mathcal{D}$ in $\widetilde{\mathcal{Z}}$.

B. One example: A generic extremal nondecomposable positive map

In dimensions $m = n = 3$, $N = mn = 9$, we have constructed numerically extremal entanglement witnesses by random searches [11]. By definition, extremal witnesses found in random searches are generic.

The generic extremal witnesses we find have only quadratic zeros. Recall that an extremal witness is uniquely determined by its zeros. Since the number of constraints on the witness from one quadratic zero is $2(m + n) - 3 = 9$ and the number of independent constraints must be $N^2 - 1 = 80$, at least nine zeros are necessary to determine the witness uniquely. Most of our generic extremal witnesses have exactly nine quadratic zeros, giving a total of 81 constraints, of which 80 are then independent.

The plots presented here show the action of one unital and trace-preserving map $\mathbf{M} = \mathbf{M}_A$, where A is the transformed version, as described in Sec. III, of a randomly selected generic extremal entanglement witness from the sample described in [11]. The zeros of the transformed witness A are $\psi_i = \phi_i \otimes \chi_i$, with $i = 1,2,\ldots,9$.

The cross in each plot represents the orthogonal projection of the maximally mixed state $I/3$ on the image plane $\widetilde{\mathcal{Z}} = \mathbf{M}\mathcal{Z}$. Since we use unital maps, if we choose $\rho_0 = I/3$, then

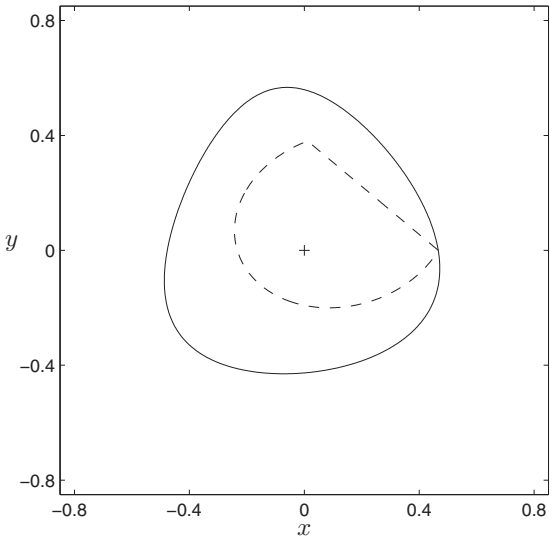


FIG. 5. $\rho_1 = \phi_1\phi_1^\dagger$, ρ_2 is a random pure state, and $\rho_0 = I/3$. Here $\widetilde{\rho}_1 = \mathbf{M}\rho_1$ has rank two and lies on the boundary $\partial\mathcal{D}$, whereas $\widetilde{\rho}_2 = \mathbf{M}\rho_2$ has full rank and lies in the interior of $\mathcal{D}$. The section $\mathcal{Z} \cap \mathcal{D}$ is of type C.

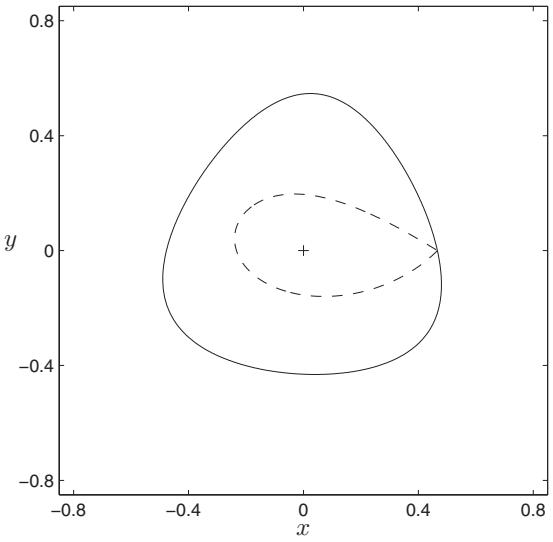


FIG. 6. $\rho_1 = \phi_1\phi_1^\dagger$, ρ_2 is a random state of rank three, and $\rho_0 = I/3$. Again, ρ_1 is mapped to $\partial\mathcal{D}$, while ρ_2 is mapped to the interior of $\mathcal{D}$. The section $\mathcal{Z} \cap \mathcal{D}$ is of type B.

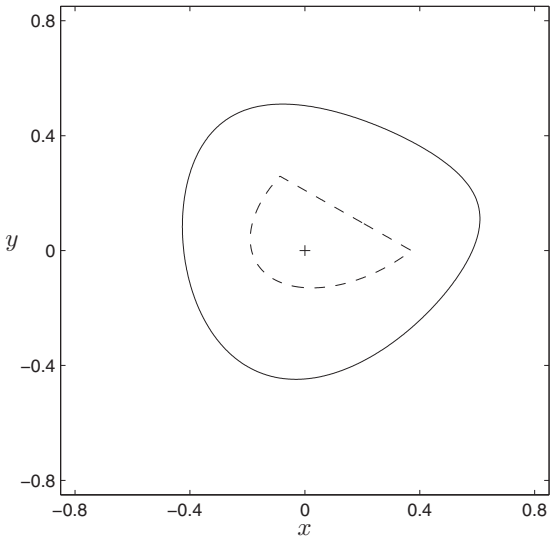


FIG. 8. Here both ρ_1 and ρ_2 are random pure states, and $\rho_0 = I/3$. The section $\mathcal{Z} \cap \mathcal{D}$ is of type C; it is mapped entirely to the interior of $\mathcal{D}$.

$\widetilde{\rho}_0 = \mathbf{M}\rho_0 = \rho_0$, and the cross is at the origin of the plot. The solid curve in each plot represents the boundary $\partial\mathcal{D}$ in the image plane $\widetilde{\mathcal{Z}}$. The dashed curve represents the image under $\mathbf{M}$ of $\partial\mathcal{D}$ in the plane $\mathcal{Z}$.

Figure 3 shows a section with the maximally mixed state at the origin, that is, $\rho_0 = \widetilde{\rho}_0 = I/3$. We have chosen the plane

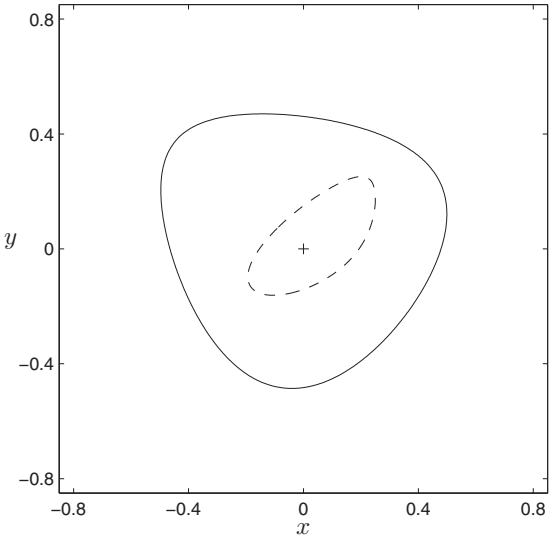


FIG. 7. Both ρ_1 and ρ_2 are random states of rank three, and $\rho_0 = I/3$. Both sections in the plot are of type A.

$\mathcal{Z}$ to go through the pure states $\rho_1 = \phi_5\phi_5^\dagger$ and $\rho_2 = \phi_6\phi_6^\dagger$ corresponding to arbitrarily chosen zeros 5 and 6 of the extremal witness A. The section $\mathcal{Z} \cap \mathcal{D}$, shown by the dashed curve, is of type C shown in Fig. 2. The section $\widetilde{\mathcal{Z}} \cap \mathcal{D}$, shown by the solid curve, is of type A. The pure states ρ_1 and ρ_2 are mapped to rank-two states $\widetilde{\rho}_1$ and $\widetilde{\rho}_2$ on the boundary $\partial\mathcal{D}$.

In Fig. 4 a different origin is used; we have chosen ρ_0 as an even mix of the three pure states $\rho_1 = \phi_5\phi_5^\dagger$, $\rho_2 = \phi_6\phi_6^\dagger$, and $\rho_3 = \phi_7\phi_7^\dagger$, which correspond to zeros of A and are mapped to rank-two states $\widetilde{\rho}_1$, $\widetilde{\rho}_2$, and $\widetilde{\rho}_3$. We observe that the projection of the maximally mixed state $I/3$ is off center. This section $\mathcal{Z} \cap \mathcal{D}$ is of type D shown in Fig. 2, whereas the section $\widetilde{\mathcal{Z}} \cap \mathcal{D}$ is still of type A. Figures 5–9 show various types of sections, as explained in the figure captions.

C. A second example: The Choi-Lam map

The unital and trace-preserving map

$$\mathbf{M} : X \mapsto Y = \frac{1}{2}\begin{pmatrix} X_{11}+X_{33} & -X_{12} & -X_{13} \\ -X_{21} & X_{11}+X_{22} & -X_{23} \\ -X_{31} & -X_{32} & X_{22}+X_{33} \end{pmatrix} \tag{62}$$

was introduced by Choi and Lam in 1977 as the first example of an extremal nondecomposable positive map [7,15]. It has been generalized to a one-parameter family of extremal positive maps, but it is still one of the very few known analytical examples of such maps. It corresponds to the extremal

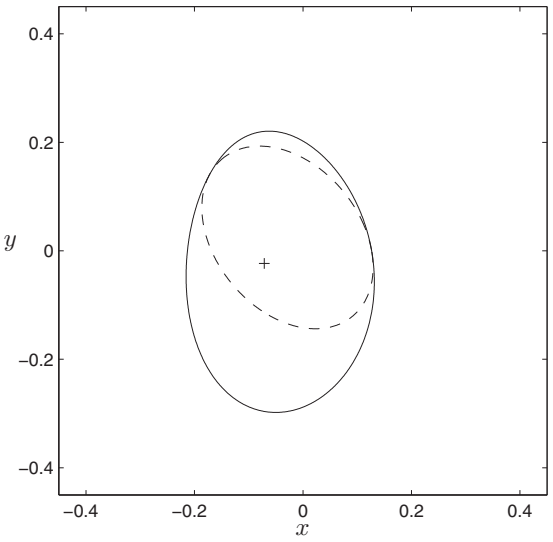


FIG. 9. $\rho_1 = \phi_1\phi_1^\dagger$ and $\rho_2 = \phi_2\phi_2^\dagger$. The origin ρ_0 is an even mix of ρ_1, ρ_2, and $\rho_3 = \xi\xi^\dagger$, where ξ is a linear combination of ϕ_1 and ϕ_2. The section $\mathcal{Z} \cap \mathcal{D}$ is of type E; it cuts through a Bloch sphere in $\partial\mathcal{D}$. Its circular shape is distorted by the map into an ellipse. The pure states ρ_1 and ρ_2 are mapped to rank-two states, on $\partial\mathcal{D}$, whereas all the other pure states on the surface of the Bloch sphere are mapped to rank-three states. The section $\widetilde{\mathcal{Z}} \cap \mathcal{D}$ is of type A; its shape is neither circular nor elliptical.

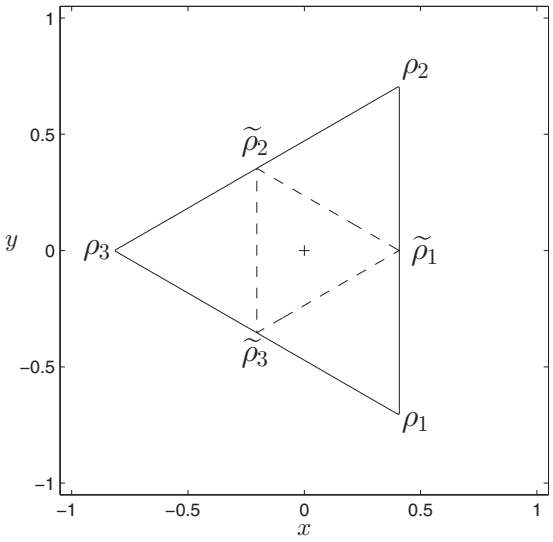


FIG. 10. The Choi-Lam map in the plane of diagonal matrices. $\rho_0 = I/3$, $\rho_1 = e_1e_1^\dagger$, and $\rho_2 = e_2e_2^\dagger$.

entanglement witness W with

$$W^P = \begin{pmatrix} 1 & \cdot & \cdot & \cdot & -1 & \cdot & \cdot & \cdot & -1 \\ \cdot & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & -1 \\ \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ -1 & \cdot & \cdot & \cdot & -1 & \cdot & \cdot & \cdot & 1 \end{pmatrix}, \quad (63)$$

where for clarity the many zero entries are represented as dots. The witness W has three isolated quartic zeros,

$$e_{13} = e_1 \otimes e_3, \quad e_{21} = e_2 \otimes e_1, \quad e_{32} = e_3 \otimes e_2, \quad (64)$$

where e_1, e_2, e_3 are the natural basis vectors in $\mathbb{C}^3$, and a continuum of quartic zeros $\phi \otimes \phi$, where

$$\phi = \phi(\alpha,\beta) = e_1 + e^{i\alpha}\, e_2 + e^{i\beta}\, e_3 \quad (65)$$

and $\alpha, \beta \in \mathbb{R}$.

Equation (62) has a very simple geometrical interpretation. In the subspace of diagonal matrices there is a rotation by 60° about the maximally mixed state, and in the subspace of off-diagonal matrices there is an inversion. There is also an overall contraction by a factor of one-half.

Figure 10 shows the special section through $\mathcal{D}$ containing the diagonal matrices, including the maximally mixed state $\rho_0 = I/3$ and the three pure states,

$$\rho_1 = e_1e_1^\dagger = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \rho_2 = e_2e_2^\dagger = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\rho_3 = e_3e_3^\dagger = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (66)$$

corresponding to the three isolated zeros of the witness W. This section is mapped into itself by the map $\mathbf{M} = \mathbf{M}_W$ since

$$\widetilde{\rho}_1 = (\rho_1 + \rho_2)/2, \quad \widetilde{\rho}_2 = (\rho_2 + \rho_3)/2,$$
$$\widetilde{\rho}_3 = (\rho_3 + \rho_1)/2. \quad (67)$$

Figure 11 shows a section defined by the same pure states ρ_1 and ρ_2 as in Fig. 10, together with a random state of rank three. This section is not mapped into itself.

The continuous set of quartic zeros of the witness W, given in Eq. (65), defines a two-dimensional surface of pure states

$$\rho(\alpha,\beta) = \tfrac{1}{3}\,\phi(\alpha,\beta)[\phi(\alpha,\beta)]^\dagger = \rho_0 + \sigma(\alpha,\beta) \quad (68)$$

mapped by $\mathbf{M}$ to the boundary $\partial\mathcal{D}$. Here $\rho_0 = I/3$, and $\sigma(\alpha,\beta)$ is a completely off-diagonal matrix. We see directly from Eq. (62) that

$$\mathbf{M}\rho(\alpha,\beta) = \rho_0 - \tfrac{1}{2}\,\sigma(\alpha,\beta) = \tfrac{1}{2}\,[3\rho_0 - \rho(\alpha,\beta)]. \quad (69)$$

Thus the straight line through $\mathbf{M}\rho(\alpha,\beta)$ and ρ_0 contains the pure state $\rho(\alpha,\beta)$. This surface of pure states is curved, but we may choose our plane $\mathcal{Z}$ in such a way that it is tangent to the surface. Define, for example,

$$\xi(\alpha) = \frac{\partial}{\partial\alpha}\,\phi(\alpha,\beta) = i e^{i\alpha}\, e_2. \quad (70)$$

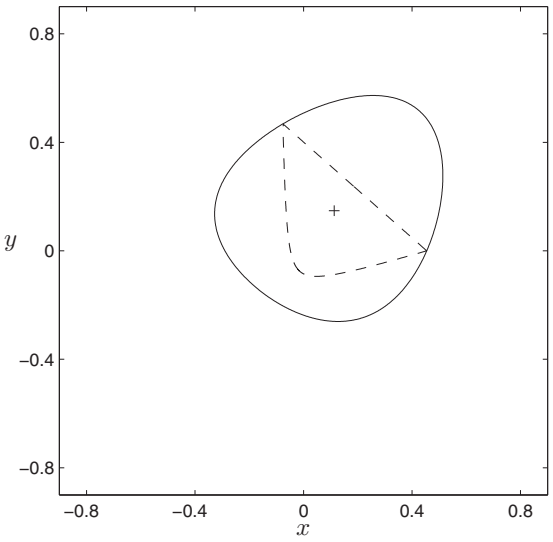


FIG. 11. The Choi-Lam map. Another section with $\rho_1 = e_1 e_1^\dagger$ and $\rho_2 = e_2 e_2^\dagger$. The origin ρ_0 is chosen here as a random state of rank three, so the section $\mathcal{Z} \cap \mathcal{D}$ contains neither $e_3 e_3^\dagger$ nor $I/3$.

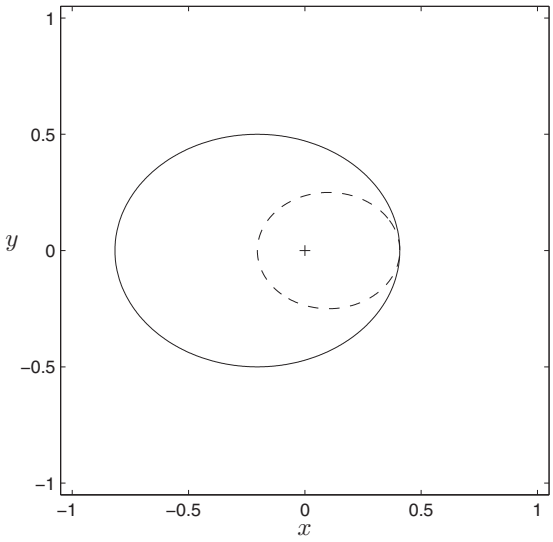


FIG. 12. The Choi-Lam map. $\rho_0 = I/3$, $\rho_1 = \rho(0,0)$, and $\rho_2 = \rho_1 + D(0,0)$, as defined in the text.

Then the matrix

$$D(\alpha,\beta) = \frac{\partial}{\partial\alpha}\,\rho(\alpha,\beta) = \frac{1}{3}\{\xi(\alpha)[\phi(\alpha,\beta)]^\dagger + \phi(\alpha,\beta)[\xi(\alpha)]^\dagger\} \tag{71}$$

is a tangent to the surface such that

$$\rho(\alpha,\beta) + \epsilon D(\alpha,\beta) = \rho(\alpha+\epsilon,\beta) + O(\epsilon^2). \tag{72}$$

Figure 12 shows a section where we have chosen $\rho_0 = I/3$ and

$$\rho_1 = \rho(0,0) = \tfrac{1}{3}\,(e_1+e_2+e_3)(e_1+e_2+e_3)^\dagger\,, \tag{73}$$

$$\rho_2 = \rho_1 + D(0,0) = \rho_1 + \frac{i}{3}\,[e_2(e_1+e_3)^\dagger - (e_1+e_3)e_2^\dagger]. \tag{74}$$

It does not matter that $\rho_2 \notin \mathcal{D}$. The coordinate axes as defined in Eq. (54) are now

$$B = a(\rho_1 - \rho_0) = \frac{a}{3}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix},$$

$$C = b(\rho_2 - \rho_1) = bD(0,0) = \frac{b}{3}\begin{pmatrix} 0 & -i & 0 \\ i & 0 & i \\ 0 & -i & 0 \end{pmatrix}. \tag{75}$$

According to Eq. (62), this section is mapped into itself by a 180° rotation and a scaling by 1/2,

$$B \mapsto \widetilde{B} = -\tfrac{1}{2}\,B, \quad C \mapsto \widetilde{C} = -\tfrac{1}{2}\,C. \tag{76}$$

Both sections $\mathcal{Z} \cap \mathcal{D}$ and $\widetilde{\mathcal{Z}} \cap \mathcal{D}$ shown in Fig. 12 are of type F in our classification. The numerical factors in Eq. (75) making $\widetilde{B}$ and $\widetilde{C}$ orthonormal are $a = \sqrt{6}$, $b = 3$. Figures 13–15 illustrate the Choi-Lam map by other sections, as explained in the figure captions.

D. A third example in 2 × 4 dimensions

Our third example is based on the study of optimal witnesses by Lewenstein *et al.* [16]. As an example they describe how

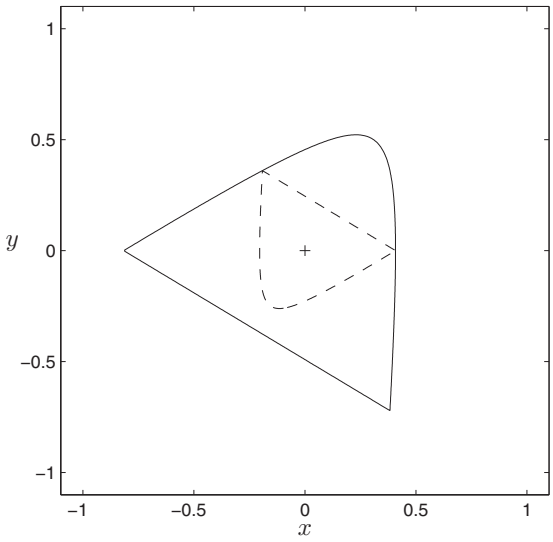


FIG. 13. The Choi-Lam map. We use here $\rho_0 = I/3$ and two randomly chosen pure states $\rho_1 = \phi_1\phi_1^\dagger$, $\rho_2 = \phi_2\phi_2^\dagger$ corresponding to product vectors $\phi_1 \otimes \phi_1^*$, $\phi_2 \otimes \phi_2^*$ from the continuum of quartic zeros.

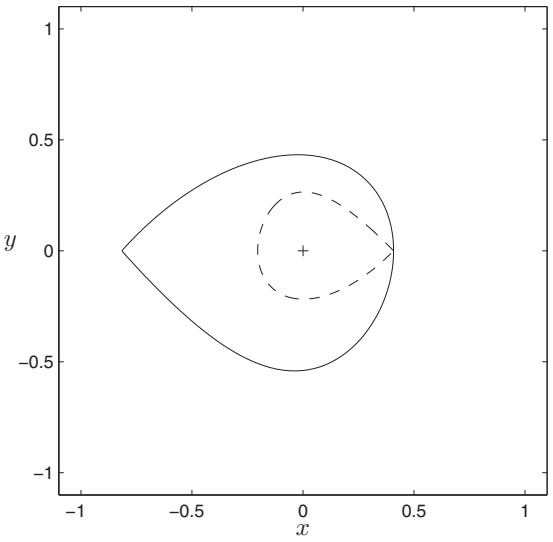


FIG. 14. The Choi-Lam map. $\rho_0 = I/3$, $\rho_1 = e_1 e_1^\dagger$, and ρ_2 is a random state of rank three.

to create optimal witnesses for proving the entanglement of the PPT states in 2×4 dimensions discovered by Horodecki [17]. We have chosen arbitrarily the parameter value $b = 0.6$ and computed numerically the optimal witness as described in [16]. It has exactly eight zeros, all quadratic. It is not extremal but is the center of a two-dimensional circular face of the set of normalized witnesses (normalized to have unit trace). The boundary of this face is a circle of quartic extremal

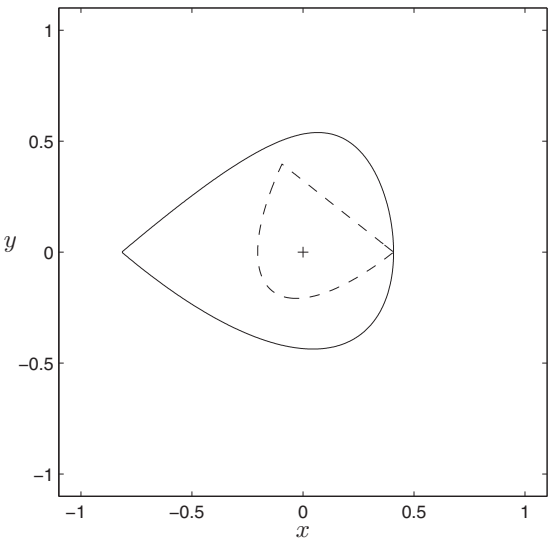


FIG. 15. The Choi-Lam map. $\rho_0 = I/3$, $\rho_1 = e_1 e_1^\dagger$, and ρ_2 is a random pure state.

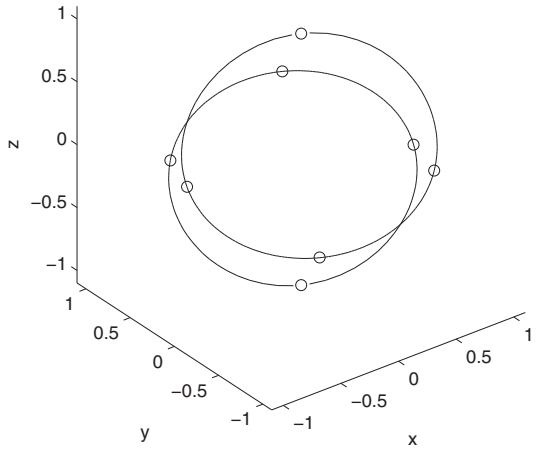


FIG. 16. The two rings of zeros of the extremal witness in dimension 2×4 on the surface of the unit sphere. The eight zeros of the optimal witness are marked by small circles.

witnesses, one of which defines the map described numerically in the Appendix. Note that this map is neither unital nor trace preserving.

This extremal witness has two separate continuous rings of zeros, which are all necessarily quartic since they are not discrete. To a zero $\phi \otimes \chi$ with $\phi^\dagger\phi = \chi^\dagger\chi = 1$ corresponds a density matrix

$$\phi\phi^\dagger = \frac{1}{2}\begin{pmatrix} 1+z & x-iy \\ x+iy & 1-z \end{pmatrix} \tag{77}$$

with $x^2 + y^2 + z^2 = 1$. This rank-one density matrix in $\mathcal{D}_2$ is mapped to a rank-three density matrix in $\mathcal{D}_4$ with χ being an eigenvector of eigenvalue zero. Thus the image of the two rings on the boundary of $\mathcal{D}_2$ is where the image of $\mathcal{D}_2$ touches the inside of the boundary of $\mathcal{D}_4$.

The coordinates x,y,z defining a zero are given by one parameter θ as follows. Let

$$a = 0.1807362587783353, \quad b = 0.047422228589395,$$
$$\theta_0 = 1.121090508802759, \tag{78}$$

$$s = \frac{a\cos(2\theta + \theta_0)}{\cos(\theta - \theta_0)}, \quad t = \frac{-bs \pm \sqrt{1 + s^2 - b^2}}{1 + s^2}, \tag{79}$$

$$x = t\cos\theta, \quad y = t\sin\theta, \quad z = -b - ts. \tag{80}$$

Note that $t \to 0$ and $ts \to -b \pm 1$ as $s \to \pm\infty$.

All the extremal quartic witnesses forming the circular boundary of the face have similar sets of zeros; it is only the parameter θ_0 that varies as we go around the boundary. Any two of them have exactly eight zeros in common, corresponding to the special values $\theta = 0$ and $\theta = \pm\pi/3$, where x,y,z are independent of θ_0, and to $\theta = \theta_0 + \pi/2$, where $x = y = 0$, $z = \pm 1$. Every witness in the interior of the face has exactly these eight zeros.

Figures 16 to 18 show the zeros of the witness defining the map given in the Appendix. Figures 19 to 21 show the same

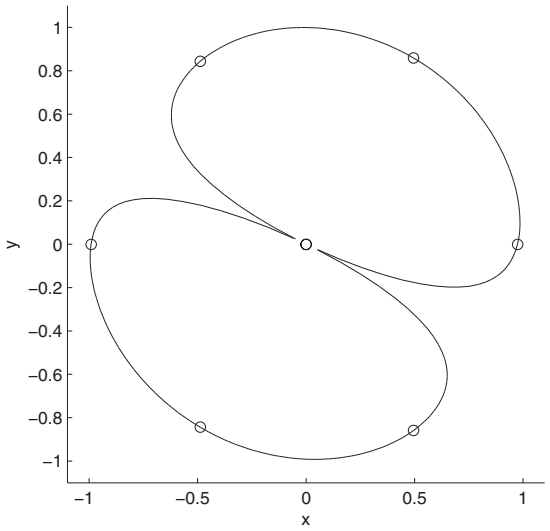


FIG. 17. The xy projection of Fig. 16. Two of the eight special zeros are at the origin $x = y = 0$; the six others are at angles $n\pi/3$ from the x axis with $n = 0,1,2,3,4,5$.

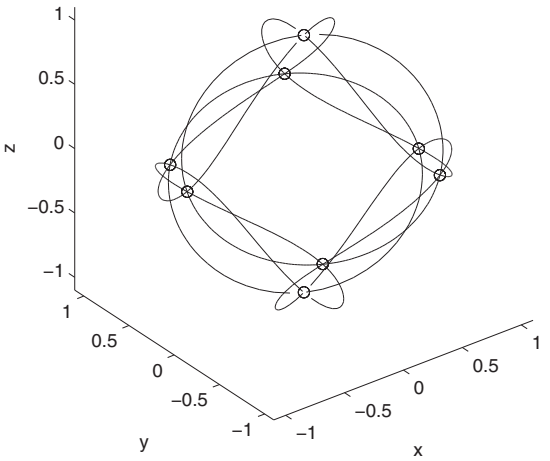


FIG. 19. The same as Fig. 16, but including the zeros of two other extremal witnesses with different values of θ_0.

zeros and in addition the zeros of two other extremal witnesses on the boundary of the same face. The zeros of different such extremal witnesses are found simply by changing the value of θ_0.

V. SUMMARY AND CONCLUSIONS

The purpose of the work presented here has been to gain a more intuitive understanding of the geometry of positive maps, which are related to entanglement witnesses in bipartite quantum systems. We try to visualize the action of an extremal positive map by plotting various two-dimensional sections through the set of density matrices.

For this purpose it is useful to transform the map to some standard form. We argue that any positive map can be transformed into a unital and trace-preserving form through a product transformation of the corresponding entanglement witness. If the witness lies in the interior of $\mathcal{S}^\circ$, which means that it has no zeros, this follows from the proof of a similar result given in [14]. We present an iteration scheme for computing the transformation numerically, and we find in practice in 3×3 dimensions that it works well even for extremal witnesses and other witnesses lying on the boundary

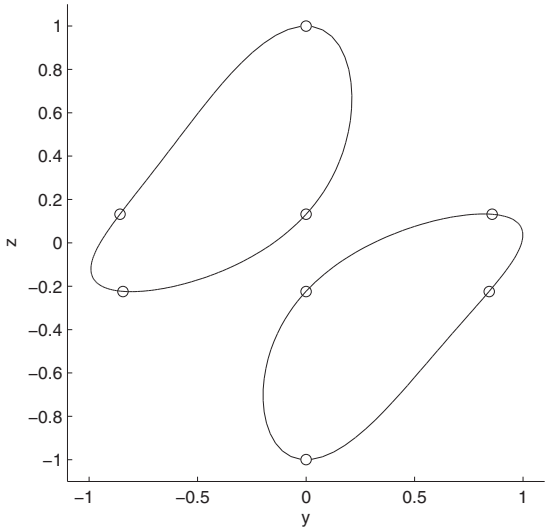


FIG. 18. The yz projection of Fig. 16.

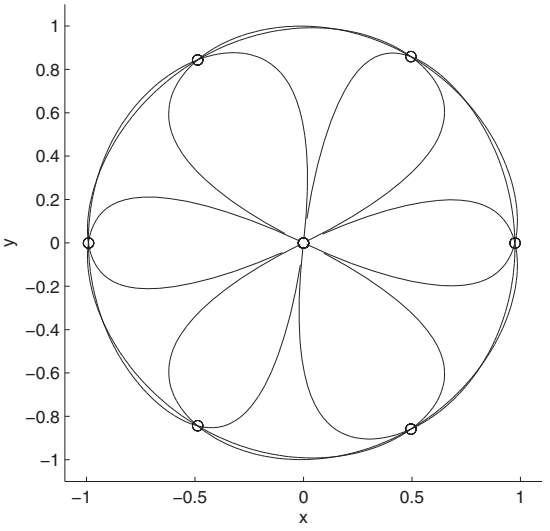


FIG. 20. The xy projection of Fig. 19.

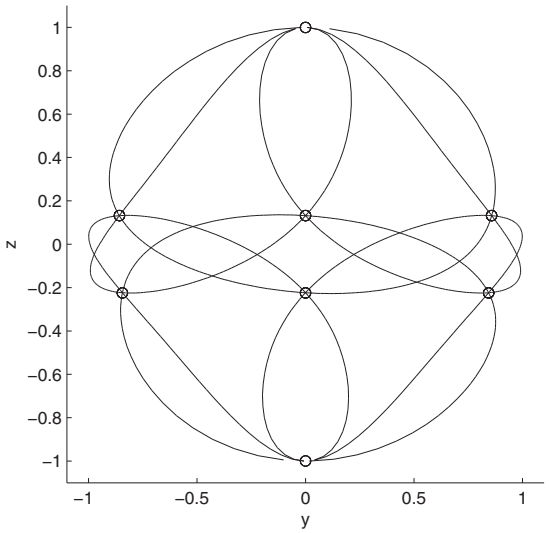


FIG. 21. The yz projection of Fig. 19.

of $\mathcal{S}^\circ$. We find numerically that the unital and trace-preserving form of a positive map is unique up to unitary product transformations.

We present plots related to two different extremal entanglement witnesses in 3×3 dimensions and one in 2×4 dimensions. The first example is a randomly chosen generic extremal entanglement witness with quadratic zeros found in previous numerical searches. We produce the corresponding positive map $\mathbf{M} : H_3 \to H_3$ and then plot two-dimensional sections in order to illustrate how the image $\mathbf{M}\mathcal{D}_3$ lies inside $\mathcal{D}_3$.

We then repeat this scheme for a version of the Choi-Lam map, or entanglement witness, which is extremal but highly nongeneric, having only quartic zeros. It has three isolated quartic zeros and one continuous two-dimensional set of zeros which are necessarily quartic.

The most important feature of the plots is related to the zeros of the two witnesses. A zero defines a pure state in $\mathcal{D}_3$ which is mapped to the boundary of $\mathcal{D}_3$. In particular, a generic extremal witness in 3×3 dimensions, like the one presented here, has nine zeros, defining a simplex in $\mathcal{D}_3$ with nine vertices which is mapped to a simplex in $\mathcal{D}_3$ with nine vertices touching the boundary of $\mathcal{D}_3$ from the inside.

We should emphasize that we have studied here only 3×3 and 2×4 dimensions, which are the simplest nontrivial cases. In higher dimensions the symmetric dimensions $m \times n$ with $m = n > 3$ are clearly the most interesting. The complexity increases much with the dimension because the simplex of pure states in $\mathcal{D}_3$ corresponding to the quadratic zeros of the extremal witness considered here becomes a polytope with a number of vertices larger than m^2 when $m > 3$. For example, with $m = n = 4$, the minimum number of zeros of a quadratic extremal witness is 20, compared to the dimension of $\mathcal{D}_4$, which is 15. With $m = n = 5$ the minimum number of zeros is 37, compared to the dimension of $\mathcal{D}_5$, which is 24 [11].

We believe that the geometrical way of thinking illustrated here may be a fruitful approach when one wants to construct examples of extremal maps and entanglement witnesses. It may be that the increase in complexity with increasing dimension, which is a well-known phenomenon, is easier to handle geometrically than by other methods.

ACKNOWLEDGMENTS

We acknowledge gratefully a research grant from the Norwegian University of Science and Technology (L.O.H.). We thank E. Størmer for his interest in our studies of extremal positive maps and in particular his question about what they look like, which inspired the present paper.

APPENDIX: AN EXTREMAL POSITIVE MAP FROM H_2 TO H_4

We introduce the following 4×4 matrices:

$$B_0 + B_3 = \begin{pmatrix} a_1 & -ia_3 & ia_4 & a_1 \\ ia_3 & a_2 & 0 & ia_3 \\ -ia_4 & 0 & a_2 & -ia_4 \\ a_1 & -ia_3 & ia_4 & a_1 \end{pmatrix}, \tag{A1}$$

with

$$a_1 = 0.0244482760740412, \quad a_2 = 0.2152770862261020,$$
$$a_3 = 0.0114377547217477, \quad a_4 = 0.0500075452822933. \tag{A2}$$

$$B_0 - B_3 = \begin{pmatrix} a_5 & ia_7 & ia_8 & -a_5 \\ -ia_7 & a_6 & 0 & ia_7 \\ -ia_8 & 0 & a_6 & ia_8 \\ -a_5 & -ia_7 & -ia_8 & a_5 \end{pmatrix}, \tag{A3}$$

with

$$a_5 = 0.0644909685779951, \quad a_6 = 0.1957836691218616,$$
$$a_7 = 0.0774551312933996, \quad a_8 = 0.0177155824920755. \tag{A4}$$

$$B_1 = \begin{pmatrix} 0 & -a_9 - ia_{10} & a_{11} - ia_{12} & -ia_{13} \\ -a_9 + ia_{10} & 0 & -a_{14} - ia_{15} & a_{11} + ia_{16} \\ a_{11} + ia_{12} & -a_{14} + ia_{15} & 0 & -a_9 - ia_{17} \\ ia_{13} & a_{11} - ia_{16} & -a_9 + ia_{17} & 0 \end{pmatrix}, \tag{A5}$$

with

$$\begin{aligned} a_9 &= 0.0363521121932822, \quad a_{10} = 0.0276760626964089, \quad a_{11} = 0.0094553411157518, \\ a_{12} &= 0.0293657267910500, \quad a_{13} = 0.0130745578191192, \quad a_{14} = 0.1714859526438769, \\ a_{15} &= 0.0675990471881839, \quad a_{16} = 0.0121590711417975, \quad a_{17} = 0.0384768416753617. \end{aligned} \tag{A6}$$

$$B_2 = \begin{pmatrix} 0 & -a_{11} - ia_{12} & -a_9 + ia_{10} & ia_{18} \\ -a_{11} + ia_{12} & -a_{14} & ia_{19} & a_9 - ia_{17} \\ -a_9 - ia_{10} & -ia_{19} & a_{14} & a_{11} - ia_{16} \\ -ia_{18} & -a_9 + ia_{17} & a_{11} + ia_{16} & 0 \end{pmatrix}, \tag{A7}$$

with

$$a_{18} = 0.0082070224528484, \quad a_{19} = 0.0424325553291989. \tag{A8}$$

The 2×2 matrix

$$A = \frac{1}{2}\begin{pmatrix} u + z & x - iy \\ x + iy & u - z \end{pmatrix} \tag{A9}$$

is positive when $u > 0$ and $u^2 \geqslant x^2 + y^2 + z^2$. In Sec. IV D we discuss the following extremal positive map, which we do not transform to unital and trace-preserving form:

$$\mathbf{M} : A \mapsto B = uB_0 + xB_1 + yB_2 + zB_3. \tag{A10}$$

[1] E. Schrödinger, Discussion of probability relations between separated systems, Proc. Cambridge Philos. Soc. **31**, 555 (1935).

[2] A. Einstein, B. Podolsky, and N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete? Phys. Rev. **47**, 777 (1935).

[3] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[4] M. Horodecki, P. Horodecki, and R. Horodecki, Separability of mixed states: Necessary and sufficient conditions, Phys. Lett. A **223**, 1 (1996).

[5] M. Barbieri, F. De Martini, G. Di Nepi, P. Mataloni, G. M. D'Ariano, and C. Macchiavello, Detection of Entanglement with Polarized Photons: Experimental Realization of an Entanglement Witness, Phys. Rev. Lett. **91**, 227901 (2003).

[6] A. Jamiołkowski, Linear transformations which preserve trace and positive semidefiniteness of operators, Rep. Math. Soc. **3**, 275 (1972).

[7] M. D. Choi, Positive semidefinite biquadratic forms, Linear Algebra Appl. **12**, 95 (1975).

[8] E. Størmer, Positive linear maps of operator algebras, Acta Math. **110**, 233 (1963).

[9] E. Størmer, *Positive Linear Maps of Operator Algebras*, Monographs in Mathematics (Springer, Berlin, 2013).

[10] A. Peres, Separability Criterion for Density Matrices, Phys. Rev. Lett. **77**, 1413 (1996).

[11] L. O. Hansen, A. Hauge, J. Myrheim, and P. Ø. Sollied, Extremal entanglement witnesses, arXiv:1305.2385.

[12] E. Størmer, Decomposable linear maps on C^*-algebras, Proc. Amer. Math. Soc. **86**, 402 (1982).

[13] S. L. Woronowicz, Positive maps of low dimensional matrix algebras, Rep. Math. Phys. **10**, 165 (1976).

[14] J. M. Leinaas, J. Myrheim, and E. Ovrum, Geometrical aspects of entanglement, Phys. Rev. A **74**, 012313 (2006).

[15] M. D. Choi and T. Y. Lam, Extremal positive semidefinite forms, Math. Ann. **231**, 1 (1977).

[16] M. Lewenstein, B. Kraus, J. I. Cirac, and P. Horodecki, Optimization of entanglement witnesses, Phys. Rev. A **62**, 052310 (2000).

[17] P. Horodecki, Separability criterion and inseparable mixed states with positive partial transposition, Phys. Lett. A **232**, 333 (1997).

Paper IV

L.O. Hansen and J. Myrheim

"Nongeneric positive partial transpose states of rank five"

Preprint of paper for submission to Phys. Rev. A

Nongeneric positive partial transpose states of rank five

Leif Ove Hansen and Jan Myrheim
Department of Physics, Norwegian University of Science and Technology, N-7491 Trondheim, Norway

April 7th, 2016

Preprint of paper for submission to
Physical Review A

Abstract. It is known that in 3×3 dimensions, entangled mixed states that are positive under partial transposition (PPT states) must have at least rank four. The classification of these rank four states is believed to be complete, and the next unsolved problem is to understand the extremal PPT states of rank five in 3×3 dimensions. We call two states SL-equivalent (more precisely SL $\times$ SL-equivalent) if they are related by a product transformation. A generic rank five PPT state ρ and its partial transpose ρ^P both have six product vectors in their ranges, and no product vectors in their kernels. Then the three numbers $\{6, 6; 0\}$ are SL-invariants that help us classify the state. We have found no analytical method for constructing generic rank five extremal PPT states, but we have studied numerically a few types of nongeneric states, in particular states with one or more product vectors in the kernels, both extremal and nonextremal. We call a state ρ SL-symmetric if ρ and its partial transpose ρ^P are SL-equivalent, and genuinely SL-symmetric if it is SL-equivalent to a state τ with $\tau = \tau^P$. Genuine SL-symmetry implies SL-symmetry. We have produced a random sample of rank five SL-symmetric states of type $\{6, 6; 0\}$, in which all are extremal and about half are genuinely SL-symmetric. We find an interesting new analytical construction of all rank four extremal PPT states, up to SL-equivalence, where they appear as boundary states on one single five dimensional face on the set of normalized PPT states. The interior of the face consists of rank five states, a simplex of separable states surrounded by entangled PPT states. All these states are real matrices, symmetric under partial transposition.

1. Introduction

Entanglement between subsystems of a composite quantum system is a phenomenon which has no counterpart in classical physics. Entangled quantum states show correlations in measurements which cannot be completely modelled within any local theory, including classical physics. A classical local description of such systems implies several types of so called Bell inequalities [1], which experimentally are shown to be violated [2]. So it is clear that entangled quantum states exhibit a nonlocality that excludes any local theory, though deterministic theories that includes quantum mechanics is still a possibility, but they must be nonlocal.

Pure product states are the only pure quantum states that are not entangled, and they resemble pure classical states in that they have no nonlocal correlations at all. By definition, a mixed quantum state is a statistical ensemble of pure quantum states, and it is represented mathematically by a density matrix.

One single density matrix may represent many different ensembles. A basic postulate is that there is no way to distinguish experimentally between different ensembles represented by the same density matrix.

A mixed quantum state is said to be separable if it can be mixed entirely from pure product states. The entangled mixed states are exactly those that are not separable. While the separability problem for pure states is solved entirely via Schmidt decomposition of state vectors, the problem of how to characterize the set $\mathcal{S}$ of separable mixed states, and to decide whether a mixed state is separable or entangled is known to be a very difficult mathematical problem, and it has been demonstrated that operational procedures are NP-hard [3].

In recent years these problems have been given considerable attention, mainly due to the fact that quantum entanglement have found use in many applications. Many new developments have been made that requires an understanding of entanglement as a resource in quantum communication, quantum cryptography, quantum computers and others. How to prepare, manipulate and detect entangled quantum states have become important in this respect.

Sound operational methods or criteria for entanglement checking only exist for special cases and/or for the low dimensional systems 2×2 and 2×3. The separable states have the property that they remain positive after partial transposition, they are PPT states. The set of PPT states $\mathcal{P}$ is in general larger than the set of separable states $\mathcal{S}$, but the difference between the two sets is small in low dimensions, and in the 2×2 and 2×3 systems $\mathcal{P}=\mathcal{S}$ [4]. Thus the condition of positive partial transpose, known as the Peres separability criterion [5], is completely adequate as long as we for the dimensions of the system have $N=N_AN_B\leq 6$.

For systems with total dimension $N=N_AN_B\geq 8$ entangled PPT states, also known as states with bound entanglement, exist. It is exactly these states that make up the difference between $\mathcal{P}$ and $\mathcal{S}$, and they are in that sense the states that violate the Peres criterion.

Another criterion which is of considerable usefulness is the range criterion. This criterion states that if a quantum state is separable, then its range (or image space) is spanned by a set of product vectors $w_i=u_i\otimes v_i$, and the range of the partial transpose ρ^P is spanned by $\widetilde{w}_i=u_i\otimes v_i^*$. It was shown [6] that this criterion is independent from the PPT criterion, as there are PPT entangled states that violate the range criterion and there are also NPT states satisfying it.

The close relation between PPT states and product vectors has been used to prove the separability of sufficiently low rank PPT states [7], which shows that on an $N_A\times N_B$ system with $N_A\leq N_B$, all states of rank N_B are separable.

For the 3×3 system this means that the entangled PPT states with lowest rank has rank four. Bennet *et.al.* [8, 9] introduced a method for constructing low rank entangled PPT states by using Unextendable Product Bases (UPB). A UPB is defined as a maximal set of orthogonal product vectors which is not a complete basis of the Hilbert space. If one so constructs an orthogonal projection Q and complementary projection $P=I-Q$, then P is an entangled PPT state. The UPB construction is clearly most successful in the case of rank four PPT states in the 3×3 system, where this construction is able to produce all entangled PPT states of rank four [10, 11]. Since the orthogonality is essential to the construction of the PPT states as a projection operator, this UPB strategy fails for rank five states in the 3×3 system.

Due to the lack of such a schedule as the UPB method, the characterization of rank five PPT entangled states in the 3×3 system seems much more challenging than for rank four states. All extremal PPT states of rank four are essentially of one generic type, while for the extremal PPT states of rank five a whole range of nongeneric forms also exist.

Outline of the paper. The contents of the present paper are organized in following manner.

In Section 2 we develop some preliminary linear algebra, which includes density matrices and projection operators.

In Section 3 we extend the description to bipartite composite systems. Product vectors and their importance is discussed, we emphasize the importance of product transformations and discuss PPT states which are equivalent under such transformations. We also briefly discuss PPT states with product vectors in the kernel, and introduce a characterization of such states.

In Section 4 we construct a standard form for PPT states of rank five with four product vectors in the kernel. All these special PPT states of rank $(5,5)$ are nonextremal, but this scheme leads to a new method for constructing extremal PPT states of rank $(4,4)$, and is relevant to our further study of nongeneric PPT states of rank $(5,5)$, presented in Section 7.

In Section 5 we review some features regarding generic subspaces of dimension five in the 3×3 system. We also discuss nongeneric subspaces in the 3×3 system, specifically how to construct pairs of orthogonal subspaces $\mathcal{U}$ and $\mathcal{V}$ with $|\mathcal{U}|=5$ and $|\mathcal{V}|=4$, such that the number of product vectors in $\mathcal{V}$ is nonzero.

In Section 6 a summary of the numerical results obtained during our investigations is presented. This includes data for the generic states with no product vectors in $\operatorname{Ker}\rho$, and several nongeneric cases with up to four product vectors in $\operatorname{Ker}\rho$. We also present some results from our random searches for SL-symmetric states.

In Section 7 we present a collection of nongeneric standard forms for orthogonal subspaces $\mathcal{U}$ and $\mathcal{V}$, with dimensions five and four respectively. The nongeneric feature is that the number of product vectors in $\mathcal{V}$ is nonzero. For the various $\mathcal{U}$ and $\mathcal{V}$, we have produced PPT states of rank $(5,5)$ with $\operatorname{Img}\rho=\mathcal{U}$ and $\operatorname{Ker}\rho=\mathcal{V}$. The number of product vectors in $\operatorname{Img}\rho$ is either six or infinite, while the number of product vectors in $\operatorname{Ker}\rho$ ranges from one to four. The latter is presented as a special case in Section 4.

In Section 8 we present a special five-dimensional subspace of $\mathbb{C}^9$ that contains only two product vectors, and with one product vector in the orthogonal complement space. We construct a set of states found by transformation to standard form of one particular rank $(5,5)$ nonextremal PPT state in this subspace. This nonextremal PPT state was found by complete chance in our searches for SL-symmetric states.

In Section 9 we conclude with a summary of the main results, and also include a brief description of some unsolved issues.

2. Basic linear algebra

We want to review some basic concepts of linear algebra, partly in order to make the paper selfcontained, and partly in order to define our notation.

2.1. **Density matrices.** The natural structure of the set H_N of Hermitian $N \times N$ matrices is that of a real Hilbert space of dimension N^2 with the scalar product

$$(X, Y) = \mathrm{Tr}\,(XY). \tag{1}$$

The set of mixed states, or density matrices, is defined as

$$\mathcal{D} = \mathcal{D}_N = \{\rho \in H_N \,|\, \rho \geq 0,\ \mathrm{Tr}\,\rho = 1\,\}. \tag{2}$$

A density matrix ρ has a spectral representation in terms of a complete set of orthonormal eigenvectors $\psi_i \in \mathbb{C}^N$ with eigenvalues $\lambda_i \geq 0$,

$$\rho = \sum_{i=1}^{N} \lambda_i\, \psi_i {\psi_i}^\dagger, \qquad \sum_{i=1}^{N} \psi_i {\psi_i}^\dagger = I. \tag{3}$$

with ${\psi_i}^\dagger \psi_j = \delta_{ij}$. The spectral representation defines one particular ensemble represented by ρ.

The rank of ρ is the number of eigenvalues $\lambda_i > 0$. The matrices

$$\begin{aligned} P &= \sum_{i,\,\lambda_i > 0} \psi_i {\psi_i}^\dagger, \\ Q = I - P &= \sum_{i,\,\lambda_i = 0} \psi_i {\psi_i}^\dagger, \end{aligned} \tag{4}$$

are Hermitian and project orthogonally onto two complementary orthogonal subspaces of $\mathbb{C}^N$.

P projects onto $\mathrm{Img}\,\rho$, the range of ρ, and Q onto $\mathrm{Ker}\,\rho$, the kernel (or nullspace) of ρ. Some useful relations that follow from this are that $P\rho = \rho P = P\rho P = \rho$ and $Q\rho = \rho Q = Q\rho Q = 0$.

We say that ρ is positive (or positive semidefinite) and write $\rho \geq 0$ when all $\lambda_i \geq 0$. An equivalent condition is that ${\psi}^\dagger \rho \psi \geq 0$ for all $\psi \in \mathbb{C}^N$. It follows from the last inequality and the spectral representation of ρ that ${\psi}^\dagger \rho \psi = 0$ if and only if $\rho\psi = 0$.

The fact that the positivity conditions ${\psi}^\dagger \rho \psi \geq 0$ are linear in ρ implies that $\mathcal{D}$ is a convex set, so that if ρ is a convex combination of $\rho_1, \rho_2 \in \mathcal{D}$,

$$\rho = p\rho_1 + (1-p)\rho, \qquad 0 < p < 1, \tag{5}$$

then $\rho \in \mathcal{D}$. Furthermore, since $\mathrm{Ker}\,\rho = \{\psi \,|\, {\psi}^\dagger \rho \psi = 0\}$ when $\rho \geq 0$, it follows that

$$\mathrm{Ker}\,\rho = \mathrm{Ker}\,\rho_1 \cap \mathrm{Ker}\,\rho_2, \tag{6}$$

independent of p, when ρ is as in (5). Since $\mathrm{Ker}\,\rho$ is independent of p, so is $\mathrm{Img}\,\rho = (\mathrm{Ker}\,\rho)^\perp$.

A compact (closed and bounded) convex set is determined by its extremal points, those points that are not convex combinations of other points in the set. The extremal points of $\mathcal{D}$ are the pure states of the form $\rho = \psi{\psi}^\dagger$ with $\psi \in \mathbb{C}^N$. Thus the spectral representation is an expansion of ρ as a convex combination of N or fewer extremal points in $\mathcal{D}$.

2.2. **Perturbations and extremality in $\mathcal{D}$.** Let ρ be a density matrix, and define the projections P and Q as in Equation (4). Consider a perturbation

$$\rho' = \rho + \epsilon A, \tag{7}$$

where $A \neq 0$ is Hermitian, and $\mathrm{Tr}\,A = 0$ so that $\mathrm{Tr}\,\rho' = \mathrm{Tr}\,\rho$. The real parameter ϵ may be infinitesimal or finite.

We observe that if $\mathrm{Img}\,A \subset \mathrm{Img}\,\rho$, or equivalently if $PAP = A$, then there will be a finite range of values of ϵ, say $\epsilon_1 \leq \epsilon \leq \epsilon_2$ with $\epsilon_1 < 0 < \epsilon_2$, such that $\rho' \in \mathcal{D}$ and $\mathrm{Img}\,\rho = \mathrm{Img}\,\rho'$. This is so because the eigenvectors of ρ with zero eigenvalue will remain eigenvectors of ρ' with zero eigenvalue, and all the positive eigenvalues of ρ will change continuously with ϵ into eigenvalues of ρ'. Since $\mathcal{D}$ is compact, we may choose ϵ_1 and ϵ_2 such that ρ' has at least one negative eigenvalue when either $\epsilon < \epsilon_1$ or $\epsilon > \epsilon_2$. The negative eigenvalue becomes zero at $\epsilon = \epsilon_1$ or $\epsilon = \epsilon_2$, this makes $\mathrm{Ker}\,\rho'$ strictly larger than $\mathrm{Ker}\,\rho$ and $\mathrm{Img}\,\rho'$ strictly smaller than $\mathrm{Img}\,\rho$ in both limits $\epsilon = \epsilon_1$ and $\epsilon = \epsilon_2$.

The other way around, if $\rho' \in \mathcal{D}$ for $\epsilon_1 \leq \epsilon \leq \epsilon_2$ with $\epsilon_1 < 0 < \epsilon_2$, then ρ' is a convex combination of $\rho + \epsilon_1 A$ and $\rho + \epsilon_2 A$ for every ϵ in the open interval $\epsilon_1 < \epsilon < \epsilon_2$. Hence $\mathrm{Img}\,\rho'$ is independent of ϵ in this open interval, implying that $\mathrm{Img}\,A \subseteq \mathrm{Img}\,\rho$ and $PAP = A$.

This gives us three equivalent formulations for the extremality condition of ρ on $\mathcal{D}$. The state ρ is extremal in $\mathcal{D}$ if and only if

(1) There exists no $A \neq 0$ with $\mathrm{Tr}\, A = 0$ and $PAP = A$.

(2) The equation $PAP = A$ for the Hermitean matrix A has $A = \rho$ as its only solution.

(3) There exists no $\rho' \in \mathcal{D}$ with $\rho' \neq \rho$ and $\mathrm{Img}\, \rho' = \mathrm{Img}\, \rho$.

We may replace $PAP = A$ by the weaker condition $QAQ = 0$. By standard perturbation theory the zero eigenvalues of ρ do not change to first order in ϵ. So, with $QAQ = 0$ and ϵ infinitesimal, the perturbation (7) preserves the rank of ρ.

2.3. **Projection operators on H_N.** Using the projections P and Q defined above we can define projection operators $\mathbf{P}, \mathbf{Q}, \mathbf{R}$ on H_N, the real Hilbert space of Hermitean $N \times N$ matrices, as follows,

$$\begin{aligned} \mathbf{P}X &= PXP, \\ \mathbf{Q}X &= QXQ, \\ \mathbf{R}X &= (\mathbf{I} - \mathbf{P} - \mathbf{Q})X. \end{aligned} \tag{8}$$

Here $\mathbf{I}$ is the identity operator on H_N. It is straightforward to verify that these are complementary projections, with $\mathbf{P}^2 = \mathbf{P}$, $\mathbf{Q}^2 = \mathbf{Q}$, $\mathbf{PQ} = \mathbf{QP} = \mathbf{0}$, and so on. They are symmetric with respect to the natural scalar product on H_N, hence they project orthogonally, and relative to an orthonormal basis for H_N they are represented by symmetric matrices.

In an orthonormal basis of $\mathbb{C}^N$ with the first m basis vectors in $\mathrm{Img}\, \rho$ and the last $n = N - m$ basis vectors in $\mathrm{Ker}\, \rho$, an matrix $X \in H_N$ takes the form

$$X = \begin{pmatrix} U & V \\ V^\dagger & W \end{pmatrix}, \tag{9}$$

with $U \in H_m$ and $W \in H_n$. In this basis we have

$$P = \begin{pmatrix} I_m & 0 \\ 0 & 0 \end{pmatrix}, \qquad Q = \begin{pmatrix} 0 & 0 \\ 0 & I_n \end{pmatrix}, \tag{10}$$

and hence

$$\begin{aligned} \mathbf{P}X = \begin{pmatrix} U & 0 \\ 0 & 0 \end{pmatrix}, \qquad \mathbf{Q}X = \begin{pmatrix} 0 & 0 \\ 0 & W \end{pmatrix}, \\ \mathbf{R}X = \begin{pmatrix} 0 & V \\ V^\dagger & 0 \end{pmatrix}. \end{aligned} \tag{11}$$

3. Composite systems

In order to describe entanglement in quantum systems it is necessary to develop the basic theory of tensor product states. We do this for a bipartite system consisting of two subsystems A and B.

3.1. **Product vectors.** If $N = N_A N_B$, the tensor product spaces $\mathbb{C}^N = \mathbb{C}^{N_A} \otimes \mathbb{C}^{N_B}$ (a complex tensor product) and $H_N = H_{N_A} \otimes H_{N_B}$ (a real tensor product) describe a composite quantum system with two subsystems A and B of Hilbert space dimensions N_A and N_B.

A vector $\psi \in \mathbb{C}^N$ then has components $\psi_I = \psi_{ij}$, where

$$\begin{aligned} I &= 1, \ldots, N \\ &\updownarrow \\ ij &= 11, 12, \ldots, 1N_B, 21, 22, \ldots, N_A N_B. \end{aligned} \tag{12}$$

A product vector $\psi = \phi \otimes \chi$ has components $\psi_{ij} = \phi_i \chi_j$. We see that ψ is a product vector if and only if its components satisfy the quadratic equations

$$\psi_{ij}\psi_{kl} - \psi_{il}\psi_{kj} = 0. \tag{13}$$

These equations are not all independent, the number of independent complex equations is

$$\begin{aligned} K &= (N_A - 1)(N_B - 1) \\ &= N - N_A - N_B + 1. \end{aligned} \tag{14}$$

For example, if $\psi_1 \neq 0$ we get a complete set of independent equations by taking $i = j = 1$, $k = 2, \ldots, N_A$, and $l = 2, \ldots, N_B$.

Since the Equations (13) are homogeneous, any solution $\psi \neq 0$ gives rise to a one parameter family of solutions $c\psi$ where $c \in \mathbb{C}$. A vector ψ in a subspace of dimension n has n independent complex components. Since the most general nonzero solution must contain at least one free complex parameter, we conclude that a generic subspace of dimension n will contain nonzero product vectors if and only if

$$n \geq K + 1. \tag{15}$$

The limiting dimension

$$n = K + 1 = N - N_A - N_B + 2, \tag{16}$$

is particularly interesting. In this special case a nonzero solution will contain exactly one free parameter, which has to be a complex normalization constant. Thus up to proportionality there will exist a finite set of product vectors in a generic subspace of

this dimension, in fact the number of product vectors is [12]

$$p = \binom{N_A + N_B - 2}{N_A - 1} = \frac{(N_A + N_B - 2)!}{(N_A - 1)!(N_B - 1)!}. \tag{17}$$

A generic subspace of lower dimension will contain no nonzero product vector, whereas any generic subspace of higher dimension will contain a continuous infinity of different product vectors. By different in this context we mean in the sense that they are not proportional.

These results hold however for *generic* subspaces. It is trivially clear that nongeneric subspaces with low dimensions exist that contain product vectors. In the special case $N_A = N_B = 3$ studied here, the limiting dimension given by Equation (16) is five, and the number of product vectors given by Equation (17) is six. As described in Section 8 we have found one special example of a five dimensional subspace with only two product vectors.

The facts that product vectors always exist in $K + 1$ and higher dimensions, but not always in K and lower dimensions, are special cases of a theorem proved by Parthasarathy for systems composed of any number of subsystems [13]. These results have profound implications for the construction of unextendible product bases.

3.2. **Partial transposition and separability.** The following relation between matrix elements

$$(X^P)_{ij;kl} = X_{il;kj}, \tag{18}$$

defines the partial transpose $X^P = X^{T_B}$ of the matrix X with respect to the second subsystem B. The partial transpose with respect to the first subsystem $X^{T_A} = X^{PT} = X^{T_B T}$ where T denotes total transposition, then simply amounts to complex conjugation. The partial transposition is simply transposition of the individual submatrices of dimension $N_B \times N_B$ in the $N \times N$ matrix X. If $X = Y \otimes Z$ then $X^P = Y \otimes Z^T$.

A density matrix ρ is called separable if it is a convex combination of tensor product pure states,

$$\rho = \sum_k p_k \, w_k w_k^\dagger = \sum_k p_k \, (u_k u_k^\dagger) \otimes (v_k v_k^\dagger), \tag{19}$$

with $w_k = u_k \otimes v_k \in \mathbb{C}^N$, $p_k > 0$, and $\sum_k p_k = 1$. It follows that

$$\rho^P = \sum_k p_k \, (u_k u_k^\dagger) \otimes (v_k v_k^\dagger)^T = \sum_k p_k \, (u_k u_k^\dagger) \otimes (v_k^* v_k^T). \tag{20}$$

We write the set of separable density matrices as $\mathcal{S}$.

The obvious fact that ρ^P is positive when ρ is separable is known as the Peres criterion, it is an easily testable necessary condition for separability. For this reason it is of interest to study the set of PPT or Positive Partial Transpose matrices, defined as

$$\mathcal{P} = \{ \, \rho \in \mathcal{D} \, | \, \rho^P \geq 0 \, \} = \mathcal{D} \cap \mathcal{D}^P. \tag{21}$$

We may call it the Peres set. A well known result is that $\mathcal{P} = \mathcal{S}$ for $N = N_A N_B \leq 6$, whereas $\mathcal{P}$ is strictly larger than $\mathcal{S}$ in higher dimensions [4].

We will classify low rank PPT states by the ranks (m, n) of ρ and ρ^P respectively. Here we study the special case $N_A = N_B = 3$, then the ranks (m, n) and (n, m) are equivalent for the purpose of classification, because of the symmetric roles of the subsystems A and B, and the arbitrariness of choice of which subsystem to partial transpose.

3.3. **Product vectors in the kernel.** Recall that $\psi^\dagger \rho \psi = 0$ if and only if $\rho \psi = 0$ when $\rho \geq 0$, and similarly for ρ^P. The identity

$$(x \otimes y)^\dagger \rho (x \otimes y) = (x \otimes y^*)^\dagger \rho^P (x \otimes y^*) \tag{22}$$

therefore implies, for a PPT state ρ, that $x \otimes y \in \operatorname{Ker} \rho$ if and only if $x \otimes y^* \in \operatorname{Ker} \rho^P$.

Let the number of product vectors in $\operatorname{Img} \rho$, $\operatorname{Img} \rho^P$ and $\operatorname{Ker} \rho$ be respectively n_{img}, $\widetilde{n}_{\text{img}}$, and n_{ker}. Then n_{ker} is also the number of product vectors in $\operatorname{Ker} \rho^P$, and ρ is characterized by $\{n_{\text{img}}, \widetilde{n}_{\text{img}}; n_{\text{ker}}\}$. As we shall see, the generic entangled PPT states of rank $(5, 5)$ are in this way $\{6, 6; 0\}$ states.

We write the product vectors in $\operatorname{Img} \rho$ as

$$w_i = u_i \otimes v_i, \qquad i = 1, \ldots, n_{\text{img}}. \tag{23}$$

And likewise for $\operatorname{Ker} \rho$,

$$z_j = x_j \otimes y_j, \qquad j = 1, \ldots, n_{\text{ker}}. \tag{24}$$

Since the two subspaces are orthogonal, it is necessary that $w_i^\dagger z_j = 0$ for all i, j, hence for every pair i, j we must have either $u_i^\dagger x_j = 0$ or $v_i^\dagger y_j = 0$.

3.4. **The range criterion and edge states.** If ρ is separable, as defined in Equation (19), then $\mathrm{Img}\,\rho$ is spanned by the product vectors $w_k = u_k \otimes v_k$, and $\mathrm{Img}\,\rho^P$ is spanned by the partially conjugated product vectors $\widetilde{w}_k = u_k \otimes v_k^*$. No such relation between product vectors in $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$ need exist if ρ is an entangled PPT state.

The existence of a set of product vectors spanning the range of ρ, such that the partially conjugated product vectors span the range of ρ^P, is seen to be a necessary condition for separability, called the *range criterion* [6]. Since there exist entangled states satisfying the range criterion, the condition is not sufficient.

Please note that the range criterion demands that there should exist at least one such set of product vectors, not that *all* product vectors in the ranges should be related by partial conjugation. We exhibit here, for example, separable states that are convex combinations of five pure product states, such that $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$ contain exactly six product vectors each, but the sixth product vector in $\mathrm{Img}\,\rho^P$ is not the partial conjugate of the sixth product vector in $\mathrm{Img}\,\rho$.

On the same note we define an *edge state* as a PPT state that breaks the range criterion maximally. That is to say that there exists no product vector in the range of ρ with its partial conjugate in the range of ρ^P. Obviously, by the range criterion, all edge states are entangled.

We also see that every extremal entangled PPT state ρ is an edge state. In fact, if $w = u \otimes v \in \mathrm{Img}\,\rho$ and $\widetilde{w} = u \otimes v^* \in \mathrm{Img}\,\rho^P$, then ρ is not extremal, because $(1-\epsilon)\rho + \epsilon\, ww^\dagger$ is a PPT state for both positive and negative ϵ in some finite interval.

3.5. **Extremality in $\mathcal{P}$.** We will now describe the extremality test in $\mathcal{P} = \mathcal{D} \cap \mathcal{D}^P$, which follows directly from the extremality test in $\mathcal{D}$. We want to outline also how to use perturbations with various restrictions in order to calculate the dimensions of and trace numerically surfaces of states in $\mathcal{P}$ of equal ranks. Thus we are interested in perturbations that preserve the ranks (m, n) of ρ and ρ^P simultaneously.

As we did for ρ, we define $\widetilde{\mathbf{P}}$ and $\widetilde{\mathbf{Q}} = \mathbf{I} - \widetilde{\mathbf{P}}$ as the orthogonal projections onto $\mathrm{Img}\,\rho^P$ and $\mathrm{Ker}\,\rho^P$. We then define

$$\begin{aligned} \widetilde{\mathbf{P}}X &= (\widetilde{P}X^P\widetilde{P})^P, \\ \widetilde{\mathbf{Q}}X &= (\widetilde{Q}X^P\widetilde{Q})^P, \\ \widetilde{\mathbf{R}}X &= (\mathbf{I} - \widetilde{\mathbf{P}} - \widetilde{\mathbf{Q}})X. \end{aligned} \tag{25}$$

These are again projections on the real Hilbert space H_N, and like $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ they are symmetric with respect to the natural scalar product on H_N. We use these projection operators on H_N to impose various restrictions on the perturbation matrix A in Equation (7).

Testing for extremality in $\mathcal{P}$. Clearly ρ is extremal in $\mathcal{P}$ if and only if $A = \rho$ is the only simultaneous solution of the two equations $\mathbf{P}A = A$ and $\widetilde{\mathbf{P}}A = A$. Another way to formulate this condition is that there exists no $\rho' \in \mathcal{P}$, $\rho' \neq \rho$, with both $\mathrm{Img}\,\rho' = \mathrm{Img}\,\rho$ and $\mathrm{Img}\,(\rho')^P = \mathrm{Img}\,\rho^P$.

Since $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ are projections, the equations $\mathbf{P}A = A$ and $\widetilde{\mathbf{P}}A = A$ together are equivalent to the single eigenvalue equation

$$(\mathbf{P} + \widetilde{\mathbf{P}})A = 2A. \tag{26}$$

They are also equivalent to either one of the eigenvalue equations

$$\mathbf{P}\widetilde{\mathbf{P}}\mathbf{P}A = A, \qquad \widetilde{\mathbf{P}}\mathbf{P}\widetilde{\mathbf{P}}A = A. \tag{27}$$

Note that the operators $\mathbf{P} + \widetilde{\mathbf{P}}$, $\mathbf{P}\widetilde{\mathbf{P}}\mathbf{P}$, and $\widetilde{\mathbf{P}}\mathbf{P}\widetilde{\mathbf{P}}$ are all real symmetric and positive and therefore have complete sets of nonnegative real eigenvalues and eigenvectors.

When we diagonalize $\mathbf{P} + \widetilde{\mathbf{P}}$ we will always find $A = \rho$ as an eigenvector with eigenvalue 2. If it is the only solution of Equation (26), this proves that ρ is extremal in $\mathcal{P}$. If A is a solution not proportional to ρ, then we may impose the condition $\mathrm{Tr}\,A = 0$ (replace A by $A - (\mathrm{Tr}A)\rho$ if necessary), and we know that there exists a finite range of both positive and negative values of ϵ such that $\rho' = \rho + \epsilon A \in \mathcal{P}$, hence ρ is not extremal in $\mathcal{P}$.

Perturbations preserving the PPT property and ranks. The rank and positivity of ρ is preserved by the perturbation $\rho' = \rho + \epsilon A$ to first order in ϵ, both for $\epsilon > 0$ and $\epsilon < 0$, if and only if $\mathbf{Q}A = 0$. Similarly, the rank and positivity of ρ^P is preserved if and only if $\widetilde{\mathbf{Q}}A = 0$. These two equations together are equivalent to the single eigenvalue equation

$$(\mathbf{Q} + \widetilde{\mathbf{Q}})A = 0. \tag{28}$$

Again $\mathbf{Q} + \widetilde{\mathbf{Q}}$ is real symmetric and has a complete set of real eigenvalues and eigenvectors. The number of linearly independent solutions for A in Equation (28) is then the dimension of the surface of rank (m, n) PPT states going through ρ.

We may want to perturb in such a way that $\mathrm{Img}\,\rho' = \mathrm{Img}\,\rho$, but not necessarily $\mathrm{Img}\,(\rho')^P = \mathrm{Img}\,\rho^P$, that is we only require $\mathrm{Img}\,(\rho')^P$ and $\mathrm{Img}\,\rho^P$ to have the same rank. Then the conditions on A are that $\mathbf{P}A = A$ and $\widetilde{\mathbf{Q}}A = 0$, or equivalently

$$(\mathbf{I} - \mathbf{P} + \widetilde{\mathbf{Q}})A = 0. \tag{29}$$

In this case the number of linearly independent solutions for A is the dimension of the surface of rank (m, n) PPT states with fixed range $\mathrm{Img}\,\rho$ at ρ.

3.6. **Product transformations.** A product transformation of the form

$$\rho \mapsto \rho' = a\,V\rho V^\dagger, \qquad V = V_A \otimes V_B, \tag{30}$$

where $a > 0$ is a normalization factor and $V_A \in \mathrm{SL}(N_A, \mathbb{C})$, $V_B \in \mathrm{SL}(N_B, \mathbb{C})$, preserves positivity, rank, separability and other interesting properties that the density matrix ρ may have. It preserves positivity of the partial transpose because

$$(\rho')^P = a\,\widetilde{V}\rho^P\widetilde{V}^\dagger, \qquad \widetilde{V} = V_A \otimes V_B^*. \tag{31}$$

A transformation of the form of Equation (30) is also sometimes referred to as a *local* SL-transformation.

The range and kernel of ρ and ρ^P transform in the following ways,

$$\begin{aligned} \mathrm{Img}\,\rho' &= V\,\mathrm{Img}\,\rho, \\ \mathrm{Ker}\,\rho' &= (V^\dagger)^{-1}\,\mathrm{Ker}\,\rho, \\ \mathrm{Img}\,(\rho')^P &= \widetilde{V}\,\mathrm{Img}\,\rho^P, \\ \mathrm{Ker}\,(\rho')^P &= (\widetilde{V}^\dagger)^{-1}\,\mathrm{Ker}\,\rho^P. \end{aligned} \tag{32}$$

We say that two density matrices ρ and ρ' related in this way are SL-equivalent. The concept of SL-equivalence is important because it simplifies very much the classification of the low rank PPT states. Since this SL-equivalence is transitive it generates equivalence classes of matrices.

SL-symmetry under partial transposition. We say that the state ρ is SL-symmetric if ρ and ρ^P are SL-equivalent, that is, if

$$\rho^P = a\,V\rho V^\dagger, \qquad V = V_A \otimes V_B. \tag{33}$$

Since SL-transformations of product type $V = V_A \otimes V_B$ preserve the number of product vectors in a subspace, any transformation $\rho \mapsto \rho^P = a\,V\rho V^\dagger$ must transform the set of n_{img} product vectors in the range of ρ to the set of $\widetilde{n}_{\mathrm{img}}$ product vectors in the range of ρ^P, so for SL-symmetric states we always have $n_{\mathrm{img}} = \widetilde{n}_{\mathrm{img}}$. If the product vectors in $\mathrm{Img}\,\rho$ is $w_i = u_i \otimes v_i$ with $i = 1, \ldots, n_{\mathrm{img}}$ and for $\mathrm{Img}\,\rho^P$ we have $\widetilde{w}_i = \widetilde{u}_i \otimes \widetilde{v}_i$ for $i = 1, \ldots, n_{\mathrm{img}}$, then

$$V_A\,u_i = \widetilde{u}_i, \qquad V_B\,v_i = \widetilde{v}_i. \tag{34}$$

Note that we need SL-transformations V_A and V_B that transform *all* vectors u_i and v_i respectively. Since our understanding of the relation between w and $\widetilde{w}$ is quite limited for entangled states, it is difficult to say much in general about what makes some states SL-symmetric and others not. But it is clear that the vectors u_i, v_i which are fixed for a given range of ρ, and the vectors $\widetilde{u}_i$ and $\widetilde{v}_i$ which depend on the specific state ρ (or ρ^P) must have a structure that allows the existence of V_A and V_B to satisfy Equation (34). This structure is very likely nongeneric, so for a generic state ρ no such V_A and V_B will exist.

Genuine SL-symmetry. We say that a state ρ is *genuinely* SL-symmetric if there exists a transformation

$$\rho' = a\,U\rho U^\dagger, \qquad U = U_A \otimes U_B, \tag{35}$$

such that $(\rho')^P = \rho'$. The transformation of ρ implies that

$$(\rho')^P = a\,\widetilde{U}\rho^P\widetilde{U}^\dagger, \tag{36}$$

when we define $\widetilde{U} = U_A \otimes U_B^*$. Then assuming genuine SL-symmetry we get that

$$\widetilde{U}\rho^P\widetilde{U}^\dagger = U\rho U^\dagger \tag{37}$$

and hence

$$\rho^P = V\rho V^\dagger, \tag{38}$$

with $V = \widetilde{U}^{-1}U = I \otimes V_B$ and $V_B = (U_B^*)^{-1}U_B$. This shows that genuine SL-symmetry implies SL-symmetry, and Equation (38) requires that V preserves the trace of ρ, *i.e.* $\mathrm{Tr}\,\rho = \mathrm{Tr}\,\rho^P = \mathrm{Tr}\,(V\rho V^\dagger) = \mathrm{Tr}\,(\rho\,V^\dagger V)$. A sufficient but not necessary condition for this trace preservation is that V_B is unitary.

The relation $V_B = (U_B^*)^{-1}U_B$ implies that V_B has some special properties. One implication is that $V_B^* = V_B^{-1}$, so that V_B is unitary if and only if it is symmetric, $V_B^T = V_B$.

We also infer that $\det V_B = (\det U_B)/(\det U_B)^*$, so that $|\det V_B| = 1$. If we multiply U_B by some phase factor $\mathrm{e}^{\mathrm{i}\alpha}$, then V_B is multiplied by $\mathrm{e}^{2\mathrm{i}\alpha}$, and in this way we may redefine V_B such that $\det V_B = 1$.

We conclude that for the state ρ to be genuinely SL-symmetric it must be SL-symmetric with a transformation of the form given in Equation (38). For the 3×3 system V would have the form

$$V = \begin{pmatrix} V_B & 0 & 0 \\ 0 & V_B & 0 \\ 0 & 0 & V_B \end{pmatrix}, \tag{39}$$

with $V_B \in \mathrm{SL}(3, \mathbb{C})$ and $V_B^* = V_B^{-1}$.

Since the transformation has the form $V = I \otimes V_B$ in the case of genuine SL-symmetry, by Equation (34) the product vectors in $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$ will be related by the transformations $\widetilde{u}_i = u_i$ and $\widetilde{v}_i = V_B v_i$. This is a necessary condition for genuine SL-symmetry which may be tested as soon as we know the product vectors in $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$.

Assume that for a given PPT state ρ we find that ρ and ρ^P are related by a transformation of the form given in Equation (38). Then a further problem to be solved is to find a transformation $U = U_A \otimes U_B$ that demonstrates explicitly the genuine SL-symmetry of ρ. Thus we have to solve the equation $V_B = (U_B^*)^{-1} U_B$ for U_B.

We reason as follows. Assume that λ is an eigenvalue of U_B, then $\mu = \lambda/\lambda^*$ is an eigenvalue of V_B. Hence $|\mu| = 1$, and we may assume that $\lambda = \mathrm{e}^{\mathrm{i}\alpha}$ where α is real. Then λ must be a solution of the equation $\mu = \lambda^2$, suggesting that we may try to take U_B as the matrix square root of V_B. We find in practice that it is possible to choose simply

$$U = I \otimes \sqrt{V_B}. \tag{40}$$

4. The rank $(4, 4)$ extremal PPT states revisited

Rank $(4, 4)$ extremal PPT states in dimension 3×3 are well understood [10, 11, 14, 15]. They are all SL-equivalent to states constructed from unextendible product bases (UPBs). A construction method not using UPBs was discussed in [15], and the structure of a PPT state ρ with at least four product vectors in $\mathrm{Ker}\,\rho$ was derived.

In the present section we will review and expand on the discussion given in [15]. This is relevant for our present study of nongeneric rank $(5, 5)$ PPT states, and it leads to a new construction of the rank $(4, 4)$ extremal PPT states.

Given a PPT state ρ and four product vectors $z_j = x_j \otimes y_j$ in the kernel of ρ, in some definite but arbitrary order. We assume that any three x vectors and any three y vectors are linearly independent. Then we may perform a product transformation as in Equation (30), and subsequent normalizations, so that the vectors take the form

$$x = y = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}, \tag{41}$$

and for the product vectors z_j

$$z = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}. \tag{42}$$

The transformation is unique. In this four dimensional subspace there exist no other product vectors. The real form of the product vectors $z_j \in \mathrm{Ker}\,\rho$ implies that $z_j \in \mathrm{Ker}\,\rho^P$.

It is equally easy to see that there exist exactly six product vectors $w_i = u_i \otimes v_i$ in the orthogonal subspace. In fact, in order to have $(u_i \otimes v_i) \perp (x_j \otimes y_j)$ for all $i = 1, \ldots, 6$ and $j = 1, \ldots, 4$, we must have for each pair i, j that either $u_i \perp x_j$ or $v_i \perp y_j$. Since any three x vectors and any three y vectors are linearly independent, a u vector can be orthogonal to at most two x vectors, and a v vector can be orthogonal to at most two y vectors. This gives the six possibilities for orthogonality listed in the table.

$u_i \otimes v_i$	$u_i \perp x_k, x_l$	$v_i \perp y_m, y_n$
i	k, l	m, n
1	$2, 3$	$1, 4$
2	$1, 3$	$2, 4$
3	$1, 2$	$3, 4$
4	$1, 4$	$2, 3$
5	$2, 4$	$1, 3$
6	$3, 4$	$1, 2$

Table 1. Possibilities for a product vector $u_i \otimes v_i$ to be orthogonal to all four product vectors $x_j \otimes y_j$.

The unique solution is the following list of vectors,

$$u = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 & -1 & 0 \end{pmatrix},$$
$$v = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ -1 & -1 & 0 & 0 & 0 & 1 \end{pmatrix}, \tag{43}$$

and for the product vectors w_i

$$w = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{44}$$

The five dimensional subspace $\mathrm{Img}\,\rho = \mathrm{Span}\,w$ given by (44) defines a face $\mathcal{F} \subset \mathcal{D}$ of dimension $5^2-1 = 24$. Recall that

$$\rho z_j = \rho^P z_j = 0, \qquad j = 1, \ldots, 4. \tag{45}$$

In terms of the face $\mathcal{F}$, Equation (45) means that $\rho \in \mathcal{F}$ and $\rho^P \in \mathcal{F}$, or equivalently, that $\rho \in \mathcal{F} \cap \mathcal{F}^P$. Equation (45) restricts ρ to have the form (47), with real coefficients c_i.

Note that $\rho^P = \rho$. Since $\mathcal{F}$ is a face on $\mathcal{D}$, $\mathcal{F}^P$ is a face on $\mathcal{D}^P$, and the intersection $\mathcal{G} = \mathcal{F} \cap \mathcal{F}^P$ is a face on $\mathcal{P} = \mathcal{D} \cap \mathcal{D}^P$. Equation (47) shows that the face $\mathcal{G}$ has dimension five.

The state ρ defined in (47) is a linear combination, but not necessarily a convex combination, of the six pure product states,

$$\rho = \frac{1}{2} \sum_{i=1}^{6} c_i \, w_i w_i^\dagger. \tag{46}$$

$$\rho = \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & c_1+c_5 & -c_1 & 0 & 0 & 0 & 0 & -c_5 & 0 \\ 0 & -c_1 & c_1+c_4 & 0 & 0 & -c_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & c_2+c_6 & 0 & -c_2 & -c_6 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -c_4 & -c_2 & 0 & c_2+c_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & -c_6 & 0 & 0 & c_3+c_6 & -c_3 & 0 \\ 0 & -c_5 & 0 & 0 & 0 & 0 & -c_3 & c_3+c_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{47}$$

From (47) we observe that the normalization condition for ρ becomes

$$\mathrm{Tr}\,\rho = \sum_{i=1}^{6} c_i = 1. \tag{48}$$

Since the matrices $w_i w_i^\dagger$ are linearly independent, and there are no other pure product states in $\mathrm{Img}\,\rho$, ρ is separable if and only if all the coefficients c_i are nonnegative. However, we will now see that it is possible for ρ to be an entangled PPT state even if one of the coefficients is negative.

An eigenvalue of ρ, and of $\rho^P = \rho$, is a root of the characteristic polynomial

$$\det(\rho - \lambda I) = -\lambda^4 \, f(\lambda), \tag{49}$$

where the function $f(\lambda)$ is a fifth degree polynomial of the form $f(\lambda) = \lambda^5 - d_4\lambda^4 + d_3\lambda^3 - d_2\lambda^2 + d_1\lambda - d_0$.

The constant term in $f(\lambda)$ is

$$d_0 = \frac{3}{16} \sum_{i=1}^{6} \prod_{j \neq i} c_j = \frac{3}{16} \Big(\prod_{j=1}^{6} c_j \Big) \sum_{i=1}^{6} \frac{1}{c_i}. \tag{50}$$

If we start with $c_i > 0$ for $i = 1, \ldots, 6$, then ρ is a rank $(5,5)$ separable state. If we then change the coefficients continuously, ρ will continue to have five positive eigenvalues until we get $d_0 = 0$. Hence $d_0 = 0$ defines the boundary of the set of density matrices, and also of the set of PPT states since $\rho = \rho^P$.

We know that the boundary is not reached before at least one coefficient becomes zero or negative. If two coefficients become zero simultaneously, then $d_0 = 0$ and we have reached a boundary state which is separable. To get negative coefficients while ρ is a rank $(5,5)$ PPT state we have to make one coefficient negative before the others. Let us say,

for example, that $c_1 < 0$, and that we want to make also c_2 negative, while $c_i > 0$ for $i = 3, \ldots, 6$. Then we first have to make $c_2 = 0$, in which case $d_0 = 3\, c_1 c_3 c_4 c_5 c_6/16 < 0$ and we have already crossed the boundary $d_0 = 0$.

In conclusion, the entangled boundary states have $c_i \neq 0$ for $i = 1, \ldots, 6$, and they have one negative and five positive coefficients c_i satisfying

$$\sum_{i=1}^{6} \frac{1}{c_i} = 0. \tag{51}$$

Thus the boundary $d_0 = 0$ consists of two types of states.

(1) Separable states that are convex combinations of up to four of the pure product states $w_i w_i^\dagger$.

(2) Rank $(4, 4)$ entangled PPT states that are linear combinations of all the six pure product states $w_i w_i^\dagger$ with one negative coefficient.

It is well known that rank $(4, 4)$ entangled PPT states are extremal.

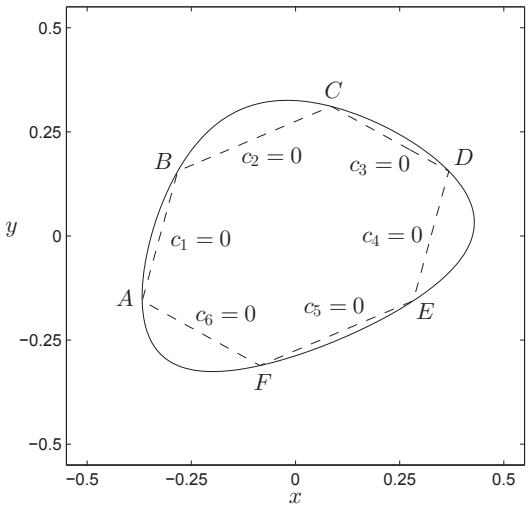


FIGURE 1. A two dimensional cut through the five dimensional face on $\mathcal{P}$ of states given by (47). The outer curve is the common boundary of $\mathcal{D}$ and $\mathcal{P}$, it consists of rank $(4, 4)$ PPT states. The simplex in $\mathcal{S}$ is the hexagon with corners A to F. The region between the two curves is entangled PPT states of rank $(5, 5)$. The coordinates x, y are dimensionless.

Figure 1 shows a two dimensional section through the five dimensional face of $\mathcal{P}$ defined by Equations (46) and (48). The section through $\mathcal{S}$ is the hexagon with corners A to F. On the boundary of the hexagon (dashed), one of the coefficients c_i is zero. The region between the two curves consists of entangled PPT states of rank $(5, 5)$ with exactly one coefficient negative. In Table 2 we have listed the coefficients c_i, multiplied by 12, that define the states A to F by Equation (46). The hexagon is reflection symmetric about two axes.

i	A	B	C	D	E	F
1	0	0	3	6	6	3
2	1	0	0	1	2	2
3	6	3	0	0	3	6
4	2	2	1	0	0	1
5	3	6	6	3	0	0
6	0	1	2	2	1	0

TABLE 2. The coefficients c_i, multiplied by 12, for the states A to F in Figure 1.

The most general rank $(4, 4)$ entangled PPT states. Consider a general rank $(4, 4)$ entangled PPT state ρ in 3×3 dimensions. It is known that any such state is extremal, and has exactly six product vectors in its kernel. We can now see that it is SL-equivalent, in no less than 360 different ways, to such states on the boundary of the five dimensional face of $\mathcal{P}$ that we have described here. The 360 transformations are found in the following way.

Pick any four of the six product vectors in $\operatorname{Ker}\rho$, this can be done in 15 different ways. Order them next in one of the 24 possible ways. Altogether there are $24 \times 15 = 360$ possibilities. There is then a unique product transformation that transforms the four product vectors to the form given in Equation (41). We know that it must transform the state ρ into one of the rank $(4, 4)$ states described by the Equations (46), (48), and (51), since these are the only rank $(4, 4)$ entangled PPT states of this form.

5. Rank $(5, 5)$ PPT states in 3×3 dimensions

Our main purpose with the present study has been to try to understand the rank $(5, 5)$ entangled PPT states in 3×3 dimensions. In particular, we would like to understand better the relation between $\operatorname{Img}\rho$ and $\operatorname{Img}\rho^P$ when ρ is a rank $(5, 5)$ PPT state. A natural question is whether ρ is SL-symmetric, as defined in Equation (33), so that $\operatorname{Img}\rho^P = V \operatorname{Img}\rho$.

The product vectors in $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$ are very useful for answering the question of SL-symmetry, especially when both these spaces have dimension five. By Equation (14), the number of constraints to be satisfied by a product vector is $K = 4$, thus $K+1 = 5$ is precisely the critical dimension at which every subspace contains one or more product vectors, and a generic subspace contains a finite number of product vectors, exactly six in this case.

5.1. **Standard form for generic subspaces.** Any set of five product vectors $w_i = u_i \otimes v_i$ in a generic five dimensional subspace may be transformed by an $\mathrm{SL}\times\mathrm{SL}$ transformation, followed by suitable normalizations, to the standard form [15]

$$u = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & p \\ 0 & 0 & 1 & 1 & q \end{pmatrix}, \qquad v = \begin{pmatrix} 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & r \\ 0 & 0 & 1 & 1 & s \end{pmatrix}, \tag{52}$$

with p, q, r, s as real or complex parameters. By generic we here mean that any three vectors in u and in v are linearly independent. There will also be a sixth product vector which is a linear combination of the above five. The parameters p, q, r, s are determined by the following ratios of determinants,

$$s_1 = -\frac{\det(u_1u_2u_4)\det(u_1u_3u_5)}{\det(u_1u_2u_5)\det(u_1u_3u_4)} = -\frac{p}{q}, \qquad s_2 = -\frac{\det(u_1u_2u_3)\det(u_2u_4u_5)}{\det(u_1u_2u_4)\det(u_2u_3u_5)} = q-1, \tag{53}$$

$$s_3 = \frac{\det(v_1v_2v_3)\det(v_1v_4v_5)}{\det(v_1v_2v_5)\det(v_1v_3v_4)} = \frac{r-s}{s}, \qquad s_4 = \frac{\det(v_1v_3v_5)\det(v_2v_3v_4)}{\det(v_1v_2v_3)\det(v_3v_4v_5)} = \frac{r}{1-r}. \tag{54}$$

All the parameters s_i are invariant under $\mathrm{SL}\times\mathrm{SL}$ transformations. Hence for given vectors u_i and v_i not on standard form these formulas may be used to calculate the values of the parameters p, q, r, s without actually performing the transformation to standard form.

Though only $u_1, \ldots, u_5$ and $v_1, \ldots, v_5$ occur in Equations (53) and (54), since the numbering is arbitrary all six product vectors must be taken into consideration when calculating the invariants. Different permutations of the six product vectors will in general give different values for the invariants.

The standard form in Equation (52), or equivalently the invariants s_i defined in Equations (53) and (54), can be used to check whether ρ and ρ^P are SL-equivalent. We must find the six product vectors in $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$, in some order. Then we either transform these to the standard form in Equation (52), or calculate directly the invariants s_i for both subspaces. For the comparison we must try all the $6! = 720$ permutations of the six product vectors. If the invariants so calculated for $\mathrm{Img}\,\rho$ and $\mathrm{Img}\,\rho^P$ are identical for some permutation, then these subspaces can clearly be transformed by a unique $\mathrm{SL}\times\mathrm{SL}$ transformation to the same standard form in Equation (52) with the same values of p, q, r, s. The transformations of both spaces to a common standard form then define a transformation $V = V_A \otimes V_B$ from $\mathrm{Img}\,\rho$ to $\mathrm{Img}\,\rho^P$. It is then easy to check whether $\rho^P = aV\rho_1V^\dagger$ for some $a > 0$.

Note that the partial transpose of ρ with respect to subsystem A is $(\rho^P)^*$. If ρ is SL-symmetric under this partial transposition, then it means that the invariants of $\mathrm{Img}\,\rho^P$ will be the complex conjugates of the invariants of $\mathrm{Img}\,\rho$.

Separable rank $(5, 5)$ *states.* In a generic five dimensional subspace of $\mathbb{C}^9$ containing six normalized product vectors $w_i = u_i \otimes v_i$, we may construct a five dimensional set of separable states as convex combinations

$$\rho = \sum_{i=1}^{6} c_i\, w_i w_i^\dagger, \tag{55}$$

with $c_i \geq 0$ and $\sum_i c_i = 1$. Hence all the separable states in the subspace are contained in a *simplex* with the six pure product states as vertices. The partial transpose of ρ is

$$\rho^P = \sum_{i=1}^{6} c_i\, \widetilde{w}_i \widetilde{w}_i^\dagger, \tag{56}$$

where $\widetilde{w}_i = u_i \otimes v_i^*$ is the partial conjugate of w_i. The six partially conjugated product vectors will be linearly independent in the generic case, hence the separable states in the interior of the simplex will have rank $(5, 6)$.

On the boundary where one coefficient c_i vanishes, ρ will be a rank $(5, 5)$ PPT state. In this case five of the product vectors in $\mathrm{Img}\,\rho$ and in $\mathrm{Img}\,\rho^P$ are partial conjugates of each other, whereas the sixth

product vectors in the two spaces are related in a more complicated way, unless all vectors v_i are real.

The surface of generic rank $(5,5)$ *PPT states.* It is known that a generic four dimensional subspace is not the range of any entangled rank $(4,4)$ PPT state [10]. It appears however in our numerical investigations that every generic five dimensional subspace is the common range of extremal and hence entangled rank $(5,5)$ PPT states that together form an eight dimensional surface [15]. It is an interesting numerical observation that the dimension of the surface of rank $(5,5)$ PPT states in a generic five dimensional subspace is higher than the dimension of the simplex of separable states.

In order to compute this surface numerically we consider the perturbation

$$\rho' = \rho + \epsilon A, \tag{57}$$

with $\mathrm{Tr}A = 0$ and A satisfying Equation (29). We find eight linearly independent solutions for A in addition to the trivial solution $A = \rho$, meaning that the surface has dimension eight.

Similarily, if we want to find the dimension of the total set of rank $(5,5)$ PPT states we find that Equation (28) has 48 linearly independent nontrivial solutions. The dimensions 8 and 48 are consistent with the fact that the set of five dimensional subspaces has dimension 40 [15].

5.2. **Nongeneric subspaces.** By definition, for a generic set of vectors in $\mathbb{C}^3$ any subset of three vectors will be linearly independent. For a generic rank $(5,5)$ PPT state ρ, $\mathrm{Img}\,\rho$ contains six product vectors $w_i = u_i \otimes v_i$. A nonzero vector $x \in \mathbb{C}^3$ can at most be orthogonal to two vectors u_i, and a nonzero y can at most be orthogonal to two v_i, hence $z = x \otimes y$ can at most be orthogonal to four w_i. Since $\mathrm{Ker}\,\rho = (\mathrm{Img}\,\rho)^\perp$ it is clear that in the generic case it is not possible to have a product vector in $\mathrm{Ker}\,\rho$, as described in Section 3.3. Thus generic states must have $n_{\mathrm{ker}} = 0$.

In order to construct pairs of orthogonal subspaces $\mathcal{U} \subset \mathbb{C}^9$ and $\mathcal{V} = \mathcal{U}^\perp$ with $|\mathcal{U}| = 5$ and $|\mathcal{V}| = 4$, such that $\mathcal{V}$ contains one or more product vectors, we must alter the generic linear dependencies of the u_i and v_i vectors. Instead of the generic condition that any *three* vectors either in u or in v must be linearly independent and span $\mathbb{C}^3$, we introduce the condition that any *four* vectors must span $\mathbb{C}^3$. Then it is possible to have one or more product vectors $z_j = x_j \otimes y_j$ with each x_j orthogonal to u_a, u_b, u_c and y_j orthogonal to v_d, v_e, v_f where $a, b, \ldots, f$ is some permutation of $1, \ldots, 6$.

In the above sense it is the *subspaces* that are interesting, and to a lesser degree the states themselves. A general characterization of two orthogonal subspaces $\mathcal{U}$ and $\mathcal{V}$ with regard to the number of product vectors they contain, is then $\{n_u; n_v\}$.

Since $z_j = x_j \otimes y_j \in \mathrm{Ker}\,\rho$ if and only if $\widetilde{z}_j = x_j \otimes y_j^* \in \mathrm{Ker}\,\rho^P$, it is clear that the kernels of ρ and ρ^P are related when they contain product vectors. In particular, if y_j is real then $z_j = \widetilde{z}_j \in \mathrm{Ker}\,\rho^P$. As long as $n_{\mathrm{ker}} \leq 4$, which is always the case in the examples we have constructed here, we can always choose a standard form where all the vectors x_j and y_j are real.

6. A summary of numerical results

We summarize here the main results of our numerical investigations. For ease of reference we number the cases from I to VII.

6.1. **Generic states.** Case (I). The generic rank $(5,5)$ PPT state is an extremal and hence entangled $\{6,6;0\}$ state. By definition, a generic state is found in a completely random search for rank $(5,5)$ PPT states. The number of generic rank $(5,5)$ PPT entangled states we have generated are in the thousands. They are all non-SL-symmetric, and since they are extremal they are also edge states. The partial transpose ρ^P of a generic rank $(5,5)$ PPT state ρ is again a generic rank $(5,5)$ PPT state.

6.2. **Nongeneric states.** The nongeneric states have from one to four product vectors in the kernel, as discussed in Section 5.2. We have defined standard forms for nongeneric orthogonal subspaces $\mathcal{U}$ and $\mathcal{V}$ with various numbers $\{n_u; n_v\}$ of product vectors. We take the product vectors in $\mathcal{V}$, as many as we need, to have the standard form given in Equation (41). We have then produced a large number of PPT states ρ of rank $(5,5)$ with $\mathrm{Img}\,\rho = \mathcal{U}$ and $\mathrm{Ker}\,\rho = \mathcal{V}$.

In addition to the generic case (I), we have studied the following nongeneric cases:

Case (II), $n_{\mathrm{ker}} = 1$. Then $\mathrm{Ker}\,\rho$ contains z_1 from Equation (41), and $\mathrm{Img}\,\rho$ is defined by Equation (78). The rank $(5,5)$ PPT states found are all extremal and of type $\{6,6;1\}$.

Case (III), $n_{\mathrm{ker}} = 1$. This is the special $\{2;1\}$ subspace presented in Section 8, producing both extremal and nonextremal entangled rank $(5,5)$ PPT

states. The extremal states are of type $\{2,6;1\}$. Some of the nonextremal states are of type $\{2,6;1\}$, others are symmetric under partial transposition and therefore of type $\{2,2;1\}$.

Case (IV), $n_{\ker} = 2$. Then $\operatorname{Ker}\rho$ in the standard form contains z_1, z_2, and $\operatorname{Img}\rho$ is defined by Equation (70). The rank $(5,5)$ PPT states found are all extremal and of type $\{6,6;2\}$.

Case (V), $n_{\ker} = 3$, where $\operatorname{Ker}\rho$ contains z_1, z_2, z_3, and $\operatorname{Img}\rho$ is defined by Equation (64). The rank $(5,5)$ PPT states found are all nonextremal and of type $\{6,6;3\}$.

Case (VI), $n_{\ker} = 4$. This is the case discussed in Section 4, with $\operatorname{Img}\rho$ defined by Equation (43). It gives a new way of constructing the extremal rank $(4,4)$ PPT states.

Case (VII), $n_{\ker} = 2$. $\operatorname{Ker}\rho$ contains z_1, z_2 from Equation (41), whereas $\operatorname{Img}\rho$ is defined by Equation (73) and contains an infinite number of product vectors. The resulting $\{\infty,\infty;2\}$ states include both extremal and nonextremal states, and also some interesting rank $(4,5)$ PPT states. We know that the latter do not exist in generic subspaces [15].

Tables. Using the projection operators and extremality tests outlined in Section 3.5, we can calculate the dimensions of the surfaces defined by Equations (26)–(29) for the various types of states. We use the abbreviation $\mathbf{P}+\widetilde{\mathbf{P}}$ for the operation that preserves the range of both ρ and ρ^P, $\mathbf{Q}+\widetilde{\mathbf{Q}}$ for the preservation of the rank of both ρ and ρ^P, and finally $\mathbf{P}+\widetilde{\mathbf{Q}}$ and $\mathbf{Q}+\widetilde{\mathbf{P}}$ for the two other projections. The dimension given by the trivial solution $A = \rho$ has been subtracted. Note that for many of the states the dimension defined by $\mathbf{P}+\widetilde{\mathbf{P}}$ is zero, indicating that the state is extremal and therefore known with certainty to be entangled.

Presented in Table 3 are data for the various random rank $(5,5)$ PPT states we have produced. In addition to the dimensions defined above, we also indicate whether the states are edge states and if they satisfy the range criterion, both of which are defined in Section 3.4.

All states are non-SL-symmetric and are produced in large numbers, with the exception of the $\{2,6;1\}$ states, where a total of 15 states have been produced.

Case	$\{n_{\text{img}}, \widetilde{n}_{\text{img}}; n_{\ker}\}$	$\mathbf{Q}+\widetilde{\mathbf{Q}}$	$\mathbf{P}+\widetilde{\mathbf{Q}}$	$\mathbf{Q}+\widetilde{\mathbf{P}}$	$\mathbf{P}+\widetilde{\mathbf{P}}$	Edge	Range
I	$\{6,6;0\}$	48	8	8	0	Yes	No
II	$\{6,6;1\}$	49	9	9	0	Yes	No
III	$\{2,6;1\}$	49	9	9	0	Yes	No
IV	$\{6,6;2\}$	50	10	10	0	Yes	No
V	$\{6,6;3\}$	51	11	11	3	No	No
VI	$\{6,6;4\}$	52	12	12	5	No	Yes
VII	$\{\infty,\infty;2\}$	50	12	12	0	Yes	No
	$\{\infty,\infty;2\}$	50	11	11	6	No	No
	$\{\infty,\infty;2\}^a$	50	15	15	9	No	No

TABLE 3. Data for the random rank $(5,5)$ PPT states we have produced based on the standard forms presented in Section 7 and the special $\{2;1\}$ subspace in Section 8. The affix a indicates that these states are PPT states of rank $(4,5)$. All states are non-SL-symmetric. For the dimensions of the various surfaces the trivial solution $A = \rho$ has been subtracted, so the states for which $\mathbf{P}+\widetilde{\mathbf{P}} = 0$, are extremal.

6.3. **SL-symmetric states.** The complete set of PPT states of rank $(5,5)$ has dimension 48, and the set of $\mathrm{SL}(3,\mathbb{C})\times\mathrm{SL}(3,\mathbb{C})$ transformations has dimension $16+16 = 32$, so the set of equivalence classes of rank $(5,5)$ states with respect to $\mathrm{SL}\times\mathrm{SL}$ transformations will have dimension 16 [15]. Hence in a random search for PPT states of rank $(5,5)$ we will never find two states belonging to the same equivalence class. We also find that a generic state ρ does not belong to the same SL-equivalence class as its partial transpose ρ^P. So, SL-symmetric states can be found numerically only by conducting specific searches.

We have randomly produced 50 SL-symmetric PPT states of rank $(5,5)$, which apart from being SL-symmetric, are generic $\{6,6;0\}$ states. These searches are rather special in the sense that we look for transformations of the form (33), but with $a=1$, *i.e.* transformations that are trace preserving. This choice is mainly motivated by the fact that due to Equation (38) we might in this way expect to find states that are genuinely SL-symmetric.

Out of the 50 states produced in this way, about half are genuinely SL-symmetric, as defined by Equation (35). This conclusion is drawn from producing the actual SL $\times$ SL transformations from ρ to ρ^P. Then, by comparing to the form in Equation (39) and constructing the transformation U from (40), one may check for genuine SL-symmetry. We consider two transformations V_1 and V_2 to be identical if they differ only by a phase factor $\mathrm{e}^{\mathrm{i}\alpha}$, *i.e.* $V_1 = \mathrm{e}^{\mathrm{i}\alpha}V_2$. For the cases of genuine SL-symmetry, the number of SL $\times$ SL transformations from ρ to ρ^P is always three, and out of these there is only one that has the form given in Equation (39).

Given three positive matrices $A=[a_{ij}]$, $B=[b_{ij}]$ and $C=[c_{ij}]$, a special class of states proposed by Chruściński and Kossakowski [16] are states of the form

$$\hat{\rho} = \begin{pmatrix} a_{11} & \cdot & \cdot & \cdot & a_{12} & \cdot & \cdot & \cdot & a_{13} \\ \cdot & b_{11} & \cdot & \cdot & \cdot & b_{12} & b_{13} & \cdot & \cdot \\ \cdot & \cdot & c_{11} & c_{12} & \cdot & \cdot & \cdot & c_{13} & \cdot \\ \cdot & \cdot & c_{21} & c_{22} & \cdot & \cdot & \cdot & c_{23} & \cdot \\ a_{21} & \cdot & \cdot & \cdot & a_{22} & \cdot & \cdot & \cdot & a_{23} \\ \cdot & b_{21} & \cdot & \cdot & \cdot & b_{22} & b_{23} & \cdot & \cdot \\ \cdot & b_{23} & \cdot & \cdot & \cdot & b_{32} & b_{33} & \cdot & \cdot \\ \cdot & \cdot & c_{31} & c_{32} & \cdot & \cdot & \cdot & c_{33} & \cdot \\ a_{31} & \cdot & \cdot & \cdot & a_{32} & \cdot & \cdot & \cdot & a_{33} \end{pmatrix}. \tag{58}$$

As long as A, B and C are positive matrices, then so will $\hat{\rho}$. There are additional constraints on the elements a_{ij}, b_{ij} and c_{ij} in order for $\hat{\rho}$ to be a PPT state. Interestingly, all the genuinely SL-symmetric PPT states of rank $(5,5)$ which we have produced, can be transformed by product transformations to the form (58). If we for a genuinely SL-symmetric PPT state ρ of rank $(5,5)$ take the two transformations V_1 and $V_2 \neq V_1$, so that

$$V_1 \rho V_1^\dagger = \rho^P, \qquad V_2 \rho V_2^\dagger = \rho^P, \tag{59}$$

then the transformation $S = {V_1}^{-1} V_2$ will transform from $\mathrm{Img}\,\rho$ to $\mathrm{Img}\,\rho$. Assuming that V_1 is of the form (39), the transformation $S = S_A \otimes S_B$ satisfy $S^3 = I$. If all transformations are nonsingular then S_A and S_B will have the eigenvalue decompositions

$$S_A = \sum_{i=1}^{3} \lambda_i \, g_i g_i^\dagger, \qquad S_B = \sum_{i=1}^{3} \mu_i \, h_i h_i^\dagger, \tag{60}$$

where the eigenvalues λ_i and μ_i are complex roots of the form $\{1, \omega, \omega^2\}$ with $\omega^3 = 1$, and the eigenevectors $\{g_i\}$ and $\{h_i\}$ are orthonormal sets. The transformations

$$\begin{aligned} T_A^{-1} &= T_A^\dagger = [g_1, g_2, g_3], \\ T_B^{-1} &= T_B^\dagger = [h_1, h_2, h_3], \end{aligned} \tag{61}$$

are such that

$$T \rho T^\dagger = \hat{\rho} \qquad \text{with} \qquad T = T_A \otimes T_B. \tag{62}$$

For the SL-symmetric states that are not genuinely SL-symmetric, there is only one SL $\times$ SL transformation from ρ to ρ^P, and this is never of the form given in Equation (39).

Our searches for SL-symmetric states have also produced a small number of states that have several other nongeneric properties, in addition to being SL-symmetric. From the results of these searches we have collected many nongeneric structures, and these have been instrumental to the study of the nongeneric standard forms given in Chapter 7. For instance, the very special subspace discussed in Section 8, which is a $\{2;1\}$ subspace, is collected from a small number of $\{2,2;1\}$ SL-symmetric states found in these random searches.

7. Nongeneric rank $(5,5)$ PPT states

We present several standard forms for nongeneric orthogonal subspaces $\mathcal{U} = \mathrm{Img}\,\rho$ and $\mathcal{V} = \mathrm{Ker}\,\rho$ of dimension five and four. The set of parameters $a_i, b_i, c_i, d_i, e_i, f_i$ is usually assumed to be chosen in a generic, or completely random manner. One may however also make investigations into certain nongeneric choices for these coefficients.

These standard forms can be used to construct nongeneric rank $(5,5)$ PPT states with a range defined by the given standard form, and with $n_{\mathrm{ker}} > 0$. Some of these constructions give states that generically are extremal in $\mathcal{P}$ and therefore entangled, while some return generically only nonextremal states.

The case $n_{\mathrm{ker}} = 3$. Given three product vectors in the kernel of ρ we can make a product transformation so that we get for these $z_i = x_i \otimes y_i$ for $i = 1, 2, 3$

from Equation (41). We choose the orthogonality relations

$$\begin{array}{llll} x_1 \perp u_i & i=2,3,6, & y_3 \perp v_i & i=1,4,5, \\ x_2 \perp u_i & i=1,3,5, & y_2 \perp v_i & i=2,4,6, \\ x_3 \perp u_i & i=1,2,4, & y_1 \perp v_i & i=3,5,6. \end{array} \tag{63}$$

A standard form for the six product vectors $w_i = u_i \otimes v_i$ in the range of ρ is then

$$\begin{aligned} u &= \begin{pmatrix} 1 & 0 & 0 & a_4 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & b_6 \\ 0 & 0 & 1 & 0 & c_5 & 1 \end{pmatrix}, \\ v &= \begin{pmatrix} 0 & d_2 & 1 & 0 & 0 & 1 \\ 1 & 0 & e_3 & 0 & 1 & 0 \\ f_1 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}. \end{aligned} \tag{64}$$

This gives the product vectors

$$w = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 \\ f_1 & 0 & 0 & a_4 & 0 & 0 \\ 0 & d_2 & 0 & 0 & 0 & b_6 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & e_3 & 0 & c_5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{65}$$

Let A be the 6×6 matrix that results from removing the rows of zeros, *i.e.* the rows one, five, and nine from w, then

$$\det A = d_2 e_3 f_1 - a_4 b_6 c_5. \tag{66}$$

Since we want the six product vectors w_i to span a five dimensional space we require that $\det A = 0$. This equation determines the sixth parameter uniquely if we choose random values for five of the parameters $a_4, b_6, c_5, d_2, e_3, f_1$.

We impose the further restrictions that all the six parameters $a_4, \ldots, f_1$ should be nonzero, and that

$$\begin{aligned} \begin{vmatrix} a_4 & 1 & 0 \\ 1 & 0 & b_6 \\ 0 & c_5 & 1 \end{vmatrix} &= -1 - a_4 b_6 c_5 \neq 0, \\ \begin{vmatrix} 0 & d_2 & 1 \\ 1 & 0 & e_3 \\ f_1 & 1 & 0 \end{vmatrix} &= 1 + d_2 e_3 f_1 \neq 0. \end{aligned} \tag{67}$$

Then the three vectors u_2, u_3, u_6 are linearly dependent, and so are u_1, u_3, u_5 and u_1, u_2, u_4, but these are the only sets of three vectors u_i that are linearly dependent. In the same way, the only sets of three linearly dependent vectors v_i are v_1, v_4, v_5, then v_2, v_4, v_6, and v_3, v_5, v_6. These linear dependencies make it possible for the three product vectors z_1, z_2, z_3, and only these, to be orthogonal to all the six product vectors w_i.

One may check that with these choices of parameters there are no more product vectors that are linear combinations of the six vectors w_i.

It should be noted that we have a further freedom of doing diagonal product transformations. In this way we may actually reduce the number of parameters in our standard form from six to one, setting for example $a_4 = c_5 = d_2 = f_1 = 1$. We must then set $b_6 = e_3 \neq \{0, -1\}$ in order to satisfy all the conditions.

A PPT state ρ with $\operatorname{Img}\rho$ spanned by the product vectors $w_1, \ldots, w_6$ and with $z_1, z_2, z_3 \in \operatorname{Ker}\rho$ must also have $z_1, z_2, z_3 \in \operatorname{Ker}\rho^P$. For given w and z these restrictions on ρ reduce our search for such states to a seven dimensional subspace of H_9. Six of the seven dimensions are spanned by the pure product states $w_i w_i^\dagger$, and there is only one dimension left where there is room for entangled PPT states.

In random numerical searches we have produced different sets of product vectors and hundreds of rank $(5,5)$ PPT states. We find that they are all entangled but not extremal. The only extremal PPT states are the six pure product states and a four dimensional surface of rank $(4,4)$ states. The normalized product states $\rho_i = N_i w_i w_i^\dagger$ with normalization factors N_i define a five dimensional simplex, which contains two special equilateral triangles with corners ρ_1, ρ_2, ρ_3 and ρ_4, ρ_5, ρ_6. We never find these separable states in random searches.

A rank $(5,5)$ entangled PPT state arises as a convex combination of an arbitrary rank $(4,4)$ extremal PPT state and a separable state from one of the two special triangles. Fig. 2 shows one three dimensional set of rank $(5,5)$ PPT states bounded by the triangle ρ_4, ρ_5, ρ_6 and a two dimensional surface of extremal rank $(4,4)$ PPT states. This geometry implies that when ρ is a rank $(5,5)$ PPT entangled state there are always three product vectors $w_i = u_i \otimes v_i \in \operatorname{Img}\rho$ such that $\widetilde{w}_i = u_i \otimes v_i^* \in \operatorname{Img}\rho^P$. The two possible sets of three are $i = 1,2,3$ or $i = 4,5,6$. Since there are less than five such product vectors, the range criterion is not fulfilled.

A separable state mixed from n product states ρ_i has rank (n,n) if $n < 6$. It has rank $(5,6)$ and lies in the interior of the simplex if $n = 6$. Rank $(5,6)$ entangled states arise in many ways, for example as convex combinations of one rank $(4,4)$ extremal state

and pure product states from both special triangles, for example ρ_1 and ρ_4. Ranks higher than six are of course impossible, since $\mathrm{Ker}\,\rho^P$ contains the three product vectors z_1, z_2, z_3. Another consequence of the last fact is that $\mathrm{Img}\,\rho^P$ contains six product vectors of the form in Equation (65), usually with other values for the coefficients $a_4, b_6, c_5, d_2, e_3, f_1$.

We conclude that all of the rank $(5,5)$ states produced in these searches are $\{6,6;3\}$ entangled nonextremal states that are neither edge states nor satisfy the range criterion.

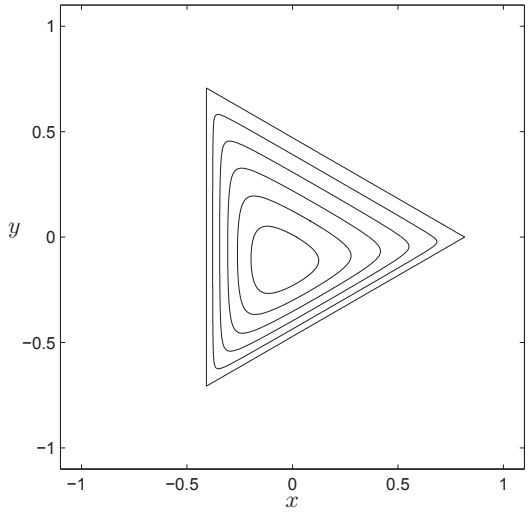


FIGURE 2. A three dimensional set, the interior of which consists of nonextremal entangled rank $(5,5)$ PPT states. The set is bounded by an equilateral triangle at $z = 0$ defined by three pure product states ρ_4, ρ_5, ρ_6, and a two dimensional surface $(z > 0)$ of extremal rank $(4,4)$ PPT states. The curves are equidistant contours at intervals of $\Delta z = 0.07$. The coordinates x, y, z are dimensionless.

The case $n_{\mathrm{ker}} = 2$. Given two product vectors in $\mathrm{Ker}\,\rho$ we can always make a product transformation so that we get $z_i = x_i \otimes y_i$ for $i = 1, 2$ from (41).

We have the freedom of doing further product transformations with

$$\begin{aligned} (V_A^\dagger)^{-1} x_i &= \alpha_i x_i, \\ (V_B^\dagger)^{-1} y_i &= \beta_i y_i, \end{aligned} \tag{68}$$

for $i = 1, 2$ and with $\alpha_i, \beta_i \in \mathbb{C}$. In order to have six product vectors $w_i = u_i \otimes v_i$ orthogonal to z_1, z_2, one possibility is to impose the following orthogonality conditions,

$$\begin{aligned} x_1 &\perp u_i \quad i = 1,2,3, & y_1 &\perp v_i \quad i = 4,5,6, \\ x_2 &\perp u_i \quad i = 3,4,5, & y_2 &\perp v_i \quad i = 1,2,6. \end{aligned} \tag{69}$$

These conditions are satisfied when the six product vectors $w_i = u_i \otimes v_i$ in the range of ρ have the standard form

$$\begin{aligned} u &= \begin{pmatrix} 0 & 0 & 0 & a_4 & a_5 & a_6 \\ b_1 & b_2 & 0 & 0 & 0 & b_6 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \\ v &= \begin{pmatrix} d_1 & d_2 & d_3 & 0 & 0 & 0 \\ 0 & 0 & e_3 & e_4 & e_5 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \end{aligned} \tag{70}$$

with $a_i, b_i, d_i, e_i \in \mathbb{C}$. We then get

$$w = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_4 e_4 & a_5 e_5 & 0 \\ 0 & 0 & 0 & a_4 & a_5 & a_6 \\ b_1 d_1 & b_2 d_2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ b_1 & b_2 & 0 & 0 & 0 & b_6 \\ d_1 & d_2 & d_3 & 0 & 0 & 0 \\ 0 & 0 & e_3 & e_4 & e_5 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{71}$$

We may generate random values of the coefficients a_i, b_i, d_i, e_i defining any five of the six product vectors w_i. Then we always find a sixth product vector, unique up to normalization, which is a linear combination of the five. The complete set of six product vectors then has rank five and automatically takes the form given in Equation (71). It is straightforward to check that z_1, z_2 are the only product vectors orthogonal to all the w_i.

We generate rank $(5,5)$ PPT states ρ with $w_1, \ldots, w_6 \in \mathrm{Img}\,\rho$. With the additional restriction that the two fixed product vectors z_1 and z_2 should be in $\mathrm{Ker}\,\rho^P$, the search can be restricted to an 11 dimensional subspace of the 81 dimensional space H_9 of Hermitian matrices.

With different sets of product vectors generated in this way we have produced numerically, by random searches, hundreds of rank $(5,5)$ PPT states. They are all extremal. Hence they are also edge states, that is, when $u_i \otimes v_i \in \mathrm{Img}\,\rho$ we never have $u_i \otimes v_i^* \in \mathrm{Img}\,\rho^P$. The partial transpose ρ^P of such a state ρ always has the same characteristics, *i.e.* there are six product vectors in $\mathrm{Img}\,\rho^P$ of the form

given in Equation (71), but with different values of the coefficients a_i, b_i, d_i, e_i.

In summary, all the states produced in this way are $\{6, 6; 2\}$ extremal states.

The case $n_{\rm ker} = 2$ with symmetric orthogonality relations. Observe that for the above standard form given in Equation (70), the orthogonality relations in Equation (69) overlap. This has been done intentionally, and the reason will become clear when we now try to impose the more symmetric orthogonality relations

$$\begin{aligned} &x_1 \perp u_i \quad i = 1, 2, 3, \qquad && y_1 \perp v_i \quad i = 4, 5, 6, \\ &x_2 \perp u_i \quad i = 4, 5, 6, \qquad && y_2 \perp v_i \quad i = 1, 2, 3. \end{aligned} \tag{72}$$

This would give six product vectors $w_i = u_i \otimes v_i$ in $\mathrm{Img}\,\rho$ on a standard form

$$\begin{aligned} u &= \begin{pmatrix} 0 & 0 & 0 & a_4 & a_5 & a_6 \\ b_1 & b_2 & b_3 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \\ v &= \begin{pmatrix} d_1 & d_2 & d_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & e_4 & e_5 & e_6 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \end{aligned} \tag{73}$$

where $a_i, b_i, d_i, e_i \in \mathbb{C}$. As compared to Equation (78) it means that we set $b_4 = b_5 = b_6 = e_1 = e_2 = e_3 = 0$. We then get

$$w = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_4e_4 & a_5e_5 & a_6e_6 \\ 0 & 0 & 0 & a_4 & a_5 & a_6 \\ b_1d_1 & b_2d_2 & b_3d_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ b_1 & b_2 & b_3 & 0 & 0 & 0 \\ d_1 & d_2 & d_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & e_4 & e_5 & e_6 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{74}$$

Omitting one of the six product vectors gives $\mathrm{rank}\, w = 5$. However, omitting for example w_6, we can find no sixth product vector of the form w_6 in the five dimensional subspace. On the other hand, there will be infinitely many product vectors that are linear combinations of w_1, w_2, w_3. In fact, u_3 and v_3 must be linear combinations

$$\begin{aligned} u_3 &= \alpha_1 u_1 + \alpha_2 u_2, \\ v_3 &= \beta_1 v_1 + \beta_2 v_2, \end{aligned} \tag{75}$$

and then for any $\gamma_1, \gamma_2 \in \mathbb{C}$ the product vector $w' = u' \otimes v'$ with

$$\begin{aligned} u' &= \gamma_1\alpha_1 u_1 + \gamma_2\alpha_2 u_2, \\ v' &= \gamma_1\beta_1 v_1 + \gamma_2\beta_2 v_2, \end{aligned} \tag{76}$$

will be a linear combination of w_1, w_2, w_3. All these product vectors lie in a three dimensional subspace of a $\mathbb{C}^2 \otimes \mathbb{C}^2$ subspace.

Our search for rank $(5, 5)$ PPT states ρ with $w_1, \ldots, w_5 \in \mathrm{Img}\,\rho$ and $z_1, z_2 \in \mathrm{Ker}\,\rho^P$ can now be restricted to a 13 dimensional subspace of H_9. Using five such randomly generated product vectors $w_1, \ldots, w_5$ to create subspaces of dimension five, we have produced numerically by random searches hundreds of rank $(5, 5)$ PPT states. Most of the states we find are separable, but we also find a very small number of extremal rank $(5, 5)$ PPT states.

We find that 11 of the 13 dimensions in H_9 represent unnormalized separable states, and the entangled PPT states account for the last two of the 13 dimensions. This can be understood as follows. The product vectors w_1, w_2, w_3 span a three dimensional subspace containing infinitely many product vectors. These product vectors generate a set of unnormalized separable states of dimension nine, the same as the complete set of unnormalized density matrices on the three dimensional subspace. In addition we get two more dimensions of separable states from the product vectors w_4 and w_5.

Note that a separable state mixed from more than three product vectors in the $\mathbb{C}^2 \otimes \mathbb{C}^2$ subspace will have rank $(3, 4)$. Hence, if we mix in also one or both of w_4, w_5 we will get separable states of rank $(4, 5)$ or $(5, 6)$, respectively.

The partial transposes of the states we construct numerically have the same characteristics, thus all the states are of type $\{\infty, \infty; 2\}$. It is noteworthy that it is possible for an extremal rank $(5, 5)$ PPT state to have infinitely many product vectors in its range.

The case $n_{\rm ker} = 1$. Given a product vector in the kernel of ρ we may perform a product transformation as in Equation (30) so that the vector after normalization take the form $z_1 = x_1 \otimes y_1$, where again x_1 is taken from Equation (41).

This transformation is not unique, and we have the freedom of doing further transformations with

$$\begin{aligned} (V_A^\dagger)^{-1} x_1 &= \alpha x_1, \\ (V_B^\dagger)^{-1} y_1 &= \beta y_1, \end{aligned} \tag{77}$$

with $\alpha, \beta \in \mathbb{C}$. We assume that the six product vectors $w_i = u_i \otimes v_i$ in the range of ρ have the standard form

$$u = \begin{pmatrix} 0 & 0 & 0 & a_4 & a_5 & a_6 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \qquad v = \begin{pmatrix} d_1 & d_2 & d_3 & 0 & 0 & 0 \\ e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}, \tag{78}$$

where $a_i, b_i, d_i, e_i \in \mathbb{C}$. They are then orthogonal to $z_1 = x_1 \otimes y_1$ because

$$x_1 \perp u_i \quad i = 1, 2, 3, \qquad y_1 \perp v_i \quad i = 4, 5, 6. \tag{79}$$

We then get

$$w = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & a_4e_4 & a_5e_5 & a_6e_6 \\ 0 & 0 & 0 & a_4 & a_5 & a_6 \\ b_1d_1 & b_2d_2 & b_3d_3 & 0 & 0 & 0 \\ b_1e_1 & b_2e_2 & b_3e_3 & b_4e_4 & b_5e_5 & b_6e_6 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \\ d_1 & d_2 & d_3 & 0 & 0 & 0 \\ e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \tag{80}$$

The normalization in Equation (78) presupposes that the third component of every vector is nonzero. There is no loss of generality in this assumption, because of the freedom we have to do product transformations. This freedom could actually be used to reduce the number of parameters in the standard form given in Equation (78).

In our numerical searches we generate, for example, random values for the coefficients a_i, b_i, d_i, e_i for $i = 1, \ldots, 5$. This is then a generic case in which $w_1, \ldots, w_5$ span a five dimensional subspace, and there will be a unique solution for a_6, b_6, e_6 such that w_6 lies in this subspace. There are no other product vectors in the subspace, in the generic case.

Furthermore, there can be no product vector $z' = x' \otimes y'$ orthogonal to all w_i, apart from $z' = z_1$. In fact, x' can at most be orthogonal to three vectors u_i, because any four vectors u_i span $\mathbb{C}^3$. Similarly, y' can at most be orthogonal to three vectors v_i. The only set of three linearly dependent vectors u_i is u_1, u_2, u_3, and the only set of three linearly dependent vectors v_i is v_4, v_5, v_6. The only possibility is therefore that $x' \perp u_i$ with $i = 1, 2, 3$ and $y' \perp v_i$ with $i = 4, 5, 6$, implying that $z' = z_1$.

We generate rank $(5, 5)$ PPT states with this subspace as range. In general, there is a 25 dimensional real vector space of Hermitian matrices operating on a fixed subspace of dimension five. However, when we are searching for a PPT state ρ with $\rho z_1 = 0$, we have the additional restriction that the partially conjugated product vector $\widetilde{z}_1 = x_1 \otimes y_1^* = x_1 \otimes y_1 = z_1$ must be in $\mathrm{Ker}\, \rho^P$. This reduces the 25 dimensions to 17, because the equation $\rho^P z_1 = 0$ gives the four extra complex equations $\rho_{jk} = 0$ with $j = 2, 3$ and $k = 4, 7$. It is straightforward to find rank $(5, 5)$ PPT states in this 17 dimensional space of Hermitian matrices.

In random searches we have produced hundreds of such states for many different sets of product vectors w. All states produced in this manner are extremal. Hence they are edge states, that is, when $u \otimes v \in \mathrm{Img}\, \rho$ we never have $u \otimes v^* \in \mathrm{Img}\, \rho^P$. We find that the partial transpose ρ^P always has the same characteristics as ρ, *i.e.* there are six product vectors in $\mathrm{Img}\, \rho^P$ of the form in Equation (78), but with different values for the coefficients a_i, b_i, d_i, e_i.

So in our classification all these states are $\{6, 6; 1\}$ extremal states.

The case $n_{\mathrm{img}} < 6$. All the above forms have been constructed as to contain (at least) six product vectors. There is however no reason to limit the possible values of n_{img} to this. The maximum dimension of an entangled subspace, a subspace that does not contain any product vectors, is known from [13]. For the 3×3 system the limiting dimension is 4. So any subspace of $\mathbb{C}^9$ of dimension five or higher *must* contain at least one product vector. We shall see in the next section an example of a very special $\{2; 1\}$ subspace of $\mathbb{C}^9$.

For any five dimensional subspace of $\mathbb{C}^9$ to contain less than generic number of six product vectors, the set of equations in (13) must have degenerate solutions. To fully describe extremal rank $(5, 5)$ PPT states according to the $\{n_{\mathrm{img}}, \widetilde{n}_{\mathrm{img}}; n_{\mathrm{ker}}\}$ characteristic, these matters should be further developed.

8. An exceptional subspace of type $\{2; 1\}$

A generic five dimensional subspace of $\mathbb{C}^3 \otimes \mathbb{C}^3$ contains exactly six product vectors. Obviously, a nongeneric five dimensional subspace may well contain infinitely many product vectors, if it contains a whole product space $\mathbb{C}^1 \otimes \mathbb{C}^2$, $\mathbb{C}^2 \otimes \mathbb{C}^1$, or $\mathbb{C}^2 \otimes \mathbb{C}^2$. On the other hand, it may also contain less than six product vectors.

In our random searches for rank $(5,5)$ PPT states with the special symmetry property that the state ρ and its partial transpose ρ^P are SL-equivalent, we have found four such states with $n_{\rm img} = 2$. A further special property of these states is that $n_{\rm ker} = 1$, *i.e.* there is exactly one product vector in $\mathrm{Ker}\,\rho$. The states are not extremal, since three of them lie on a line between a pure product state and a rank $(4,4)$ extremal PPT state, whereas one lies inside a two dimensional region bounded by a closed curve of rank $(4,4)$ extremal PPT states.

Further numerical random searches for rank $(5,5)$ PPT states with the same range as the non-extremal rank $(5,5)$ PPT states described above, but not restricted to SL-symmetric states, reveal extremal states of type $\{2,6;1\}$. It is clear that a subspace with only two product vectors supports just one line of separable states.

We find that the pure product state and the curve of rank $(4,4)$ states are parts of the same geometry, which can be transformed to a standard form as shown in Figure 3. It is an empirical fact that a rank $(4,4)$ state τ_{44} involved in such a construction of a rank $(5,5)$ state ρ can always be transformed to a standard form where the product vectors in $\mathrm{Ker}\,\tau_{44}$ are related to regular icosahedra. We will now describe this standard form in more detail.

8.1. **Rank $(5,5)$ states in a region bounded by rank $(4,4)$ states.** We describe here analytically a set of states found by transformation to standard form of one particular rank $(5,5)$ nonextremal PPT state found numerically in a random search. This particular state lies inside a curve of rank $(4,4)$ states, but the three other states we have found belong to the same geometry. It is remarkable that analytically defined states possessing such very special properties turn up in random searches. Our limited imagination would not have enabled us to deduce their existence.

After transformation to standard form the rank $(5,5)$ state lies inside a circle bounded by extremal rank $(4,4)$ PPT states. The interior of the circle consists entirely of rank $(5,5)$ PPT states, each of which has exactly one product vector in its kernel, this vector is common to all the rank $(4,4)$ and rank $(5,5)$ states. Each of the rank $(4,4)$ states have five additional product vectors in its kernel, these are different for the different states. The rank $(4,4)$ states have no product vectors in their ranges, whereas all the rank $(5,5)$ states have one common range containing exactly two product vectors. All the rank $(4,4)$ and rank $(5,5)$ states are symmetric under partial transposition.

Product vectors from a regular icosahedron. In the standard form we define, all the product vectors in the kernels of all the rank $(4,4)$ states are defined from regular icosahedra, as follows. We define $c_k = \cos(2k\pi/5)$, $s_k = \sin(2k\pi/5)$, thus

$$\begin{aligned} c_1 = c_4 &= \frac{\sqrt{5}-1}{4}, \\ s_1 = -s_4 &= \frac{\sqrt{10+2\sqrt{5}}}{4}, \\ c_2 = c_3 &= -\frac{\sqrt{5}+1}{4}, \\ s_2 = -s_3 &= \frac{\sqrt{10-2\sqrt{5}}}{4}. \end{aligned} \tag{81}$$

Note that $\phi = 2c_1 = 0.61803\ldots$ is the golden mean, defined by the equation $\phi^2 = 1 - \phi$.

We define product vectors $z_k = x_k \otimes y_k$ for $k = 1, 2, \ldots, 6$ with

$$\begin{aligned} x &= \frac{1}{\sqrt{5}} \begin{pmatrix} 2 & 2c_1 & 2c_2 & 2c_3 & 2c_4 & 0 \\ 0 & 2s_1 & 2s_2 & 2s_3 & 2s_4 & 0 \\ 1 & 1 & 1 & 1 & 1 & \sqrt{5} \end{pmatrix}, \\ y &= \frac{1}{\sqrt{5}} \begin{pmatrix} 2c_2 & 2 & 2c_3 & 2c_1 & 2c_4 & 0 \\ 2s_2 & 0 & 2s_3 & 2s_1 & 2s_4 & 0 \\ 1 & 1 & 1 & 1 & 1 & \sqrt{5} \end{pmatrix}. \end{aligned} \tag{82}$$

The 12 vectors $\pm x_k$ are real and are all the corners of a regular icosahedron. The vectors y_k as defined here are the same vectors in a different order. We define x and y in such a way that $x_5 = y_5$ and $x_6 = y_6$. The product vector z_6 is going to play a special role, it will be the one common product vector in the kernels of any two of the rank $(4,4)$ states.

The six product vectors z_k are linearly dependent and define a five dimensional subspace of $\mathbb{C}^9$. The orthogonal projection on this subspace may be written as

$$Q = \frac{5}{6} \sum_{k=1}^{6} z_k z_k^\dagger. \tag{83}$$

Since Q is a projection, with $Q^2 = Q$, and is symmetric under partial transposition, $Q^P = Q$, we conclude that the matrix

$$\tau_{44} = \frac{1}{4}\,(I - Q) \tag{84}$$

is a rank $(4,4)$ PPT state, entangled and extremal. This method for constructing such states is known as

a UPB construction, since any five of the six vectors z_k form an unextendible product basis for the subspace, so called because there is no product vector orthogonal to all of them [8, 9].

Note that Q and τ_{44} are symmetric under partial transposition with respect to subsystem B, in our notation $Q^P = Q$ and $\tau_{44}^P = \tau_{44}$, because the vectors y_k defined in Equation (82) are real. They are also symmetric under partial transposition with respect to subsystem A, in our notation $Q^{PT} = Q$ and $\tau_{44}^{PT} = \tau_{44}$, because the vectors x_k are real.

We now define product vectors $w_k = u_k \otimes v_k$ for $k = 1, 2$ with

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -\mathrm{i} \\ 0 \end{pmatrix} = v_2 = v_1^*, \quad u_2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad (85)$$

Note that $u_2 = x_6 = y_6$ and

$$u_1 = \frac{1}{\sqrt{10}} \sum_{j=1}^{5} \omega^{j-1}\, x_j, \quad (86)$$

with $\omega = \mathrm{e}^{-2\pi\mathrm{i}/5} = c_1 - \mathrm{i}s_1$. We define

$$W = w_1 w_2^\dagger = (u_1 u_2^\dagger) \otimes (v_1 v_2^\dagger), \quad (87)$$

and then the two Hermitian matrices

$$A = W + W^\dagger, \qquad B = \mathrm{i}\,(W - W^\dagger). \quad (88)$$

Since $v_1 = v_2^*$, W is symmetric under partial transposition,

$$\begin{aligned} W &= (u_1 u_2^\dagger) \otimes (v_1 v_1^T) \\ &= (u_1 u_2^\dagger) \otimes (v_1 v_1^T)^T = W^P, \end{aligned} \quad (89)$$

and so are A and B. Define now

$$\rho = \tau_{44} + \alpha A + \beta B, \quad (90)$$

with real parameters α and β. Putting $d_1 = 24\alpha - 2\sqrt{2}c_1$ and $t_1 = 24\beta - 2\sqrt{2}s_1$, the eigenvalues of ρ are four times 0, three times 1/4, and

$$\lambda_\pm = \frac{1}{8} \pm \frac{1}{24}\sqrt{1 + d_1^2 + t_1^2}. \quad (91)$$

Since $\rho^P = \rho$, we get a rank $(4,4)$ PPT state with $\lambda_- = 0$ when

$$\left(24\alpha - 2\sqrt{2}\,c_1\right)^2 + \left(24\beta - 2\sqrt{2}\,s_1\right)^2 = 8. \quad (92)$$

In particular, $\alpha = \beta = 0$ gives the state τ_{44} that we started from. Equation (92) defines a circle of rank $(4,4)$ states, which are in fact extremal PPT states. All the states inside the circle are rank $(5,5)$ non-extremal PPT states. The centre of the circle we call ρ_0, it has $\alpha = \sqrt{2}\,c_1/12$ and $\beta = \sqrt{2}\,s_1/12$.

All the states defined by Equation (90) are symmetric under partial transposition with respect to subsystem B, we have that

$$\begin{aligned} \rho^P &= \tau_{44}^P + \alpha A^P + \beta B^P \\ &= \tau_{44} + \alpha A + \beta B = \rho. \end{aligned} \quad (93)$$

However, because A and B are complex, ρ is complex when $\alpha \neq 0$ or $\beta \neq 0$. Then it is not symmetric under partial transposition with respect to subsystem A, we have then that

$$\rho^{PT} = \rho^T = \rho^* \neq \rho. \quad (94)$$

All the rank $(4,4)$ states on the circle are projections. The rank $(4,4)$ state at an angle γ around the circle from the state τ_{44} is $U\tau_{44}U^\dagger$ where $U = U_A \otimes U_B$ is a a unitary product transformation,

$$\begin{aligned} U_A &= \begin{pmatrix} \cos(\gamma/10) & -\sin(\gamma/10) & 0 \\ \sin(\gamma/10) & \cos(\gamma/10) & 0 \\ 0 & 0 & \mathrm{e}^{-\mathrm{i}\gamma/2} \end{pmatrix}, \\ U_B &= \begin{pmatrix} \cos(\gamma/5) & \sin(\gamma/5) & 0 \\ -\sin(\gamma/5) & \cos(\gamma/5) & 0 \\ 0 & 0 & 1 \end{pmatrix}, \end{aligned} \quad (95)$$

The kernel of $U\tau_{44}U^\dagger$ is defined by the transformed icosahedron vectors $(U_A^\dagger)^{-1}x = U_A x$ and $(U_B^\dagger)^{-1}y = U_B y$. This transformation leaves y_6 invariant and x_6 invariant up to a phase factor. A rotation by $\gamma = 2\pi$ is a cyclic permutation of the first five icosahedron vectors,

$$\begin{aligned} x_1 \mapsto -x_4 \mapsto x_2 \mapsto -x_5 \mapsto x_3 \mapsto -x_1, \\ y_1 \mapsto y_4 \mapsto y_2 \mapsto y_5 \mapsto y_3 \mapsto y_1. \end{aligned} \quad (96)$$

Note that the vectors u_1, u_2, v_1, and v_2 are eigenvectors of the transformation matrices U_A and U_B, for example,

$$U_A u_1 = \mathrm{e}^{\mathrm{i}\gamma/10}\, u_1. \quad (97)$$

Hence the pure states $\rho_k = w_k w_k^\dagger$ are invariant under the transformation $\rho \mapsto U\rho U^\dagger$.

The geometry of this construction is depicted in Figure 3. The four traceless matrices $A, B, \rho_1 - \rho_0$, and $\rho_2 - \rho_0$ define directions in H_9 that are mutually orthogonal. The directions A and B define the plane of the circle centered on ρ_0. Thus, the two states

$\rho_k = w_k {w_k}^\dagger$ with $k = 1, 2$, together with the circle of rank $(4, 4)$ states, define two cones that are different three-dimensional faces of $\mathcal{P}$, the set of PPT states. All the nonextremal states on the surface of a cone are rank $(5, 5)$ states.

A state ρ lying inside a cone may be written as a convex combination

$$\rho = p\rho_{55} + (1-p)\rho_k, \qquad 0 < p < 1, \tag{98}$$

with ρ_{55} inside the circle and $k = 1, 2$. This state also has rank five, since $w_k \in \operatorname{Img} \rho_{55}$. But its partial transpose ρ^P will have rank six, because it is a convex combination of $\rho_{55}^P = \rho_{55}$ and one of the two pure product states

$$\begin{aligned} \rho_1^P &= (u_1 \otimes v_2)(u_1 \otimes v_2)^\dagger, \\ \rho_2^P &= (u_2 \otimes v_1)(u_2 \otimes v_1)^\dagger. \end{aligned} \tag{99}$$

Since $\rho_{55}^{PT} = \rho_{55}^T \neq \rho_{55}$, that is, ρ_{55} is not symmetric under partial transposition with respect to system A, we conclude that neither is ρ. The same conclusion, that $\rho^{PT} \neq \rho$, follows because ρ^{PT} must have rank six, in fact it has the same eigenvalues as ρ^P.

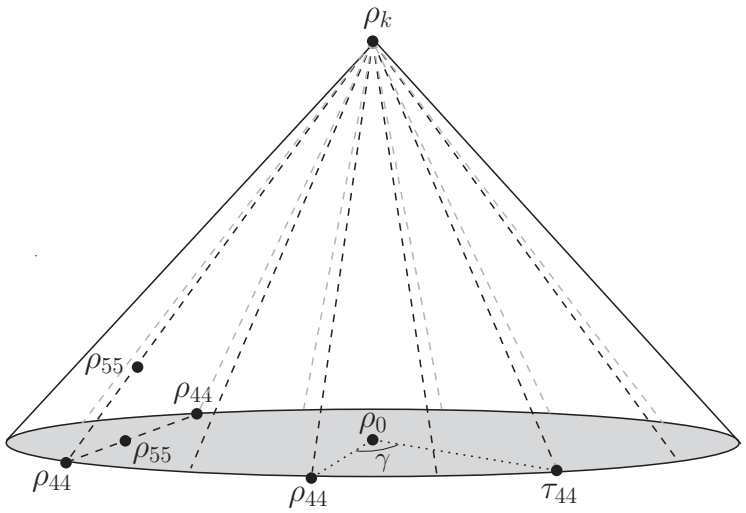


FIGURE 3. A pure product state $\rho_k = w_k {w_k}^\dagger$ with $k = 1$ or $k = 2$, together with the circle of extremal rank $(4, 4)$ PPT states, define a three-dimensional conical face of $\mathcal{P}$. States on the surface that are nonextremal on the cone, are entangled PPT states of rank $(5, 5)$, while the interior of the cone contains entangled PPT states of rank $(5, 6)$. The state ρ_0 is the centre of the circle, and the state τ_{44} is the special state defined in (84).

Since u_2 as defined in Equation (85) is real, it follows that $\rho^{PT} = \rho$ when we define

$$\rho = p\tau_{44} + (1-p)\rho_2, \qquad 0 < p < 1. \tag{100}$$

Thus, a state ρ of this form is a rank $(5, 5)$ state on the surface of a cone, and this state is symmetric under partial transposition with respect to system A. Since ρ^P and ρ^{PT} have the same eigenvalues, we can deduce that ρ and ρ^P have the same eigenvalues, although they are neither SU-equivalent nor SL-equivalent.

If we define

$$\widetilde{\rho} = p\rho_{44} + (1-p)\rho_2, \tag{101}$$

using a different rank $(4, 4)$ state ρ_{44} in place of the special state τ_{44}, then we get a state $\widetilde{\rho}$ that has no

longer the same symmetry as ρ, but is equivalent to ρ by a unitary product transformation U as given in Equation (95). In this case again $\widetilde{\rho}$, $\widetilde{\rho}^P$, and $\widetilde{\rho}^{PT}$ all have the same eigenvalues, in fact, they have the same eigenvalues as ρ, ρ^P, and ρ^{PT}.

If we replace ρ_2 by ρ_1 in Equation (100), then the state ρ we get is not SL-equivalent to any one of ρ^P or ρ^{PT}. All three states ρ, ρ^P, and ρ^{PT} still have the same set of eigenvalues, although the reason is not clear.

Four similar constructions. The construction just described is one of four different ways of extending the four dimensional subspace $\mathrm{Img}\,\tau_{44}$ to a five dimensional subspace which is orthogonal to z_6 and contains exactly two product vectors w_1, w_2. Define

$$a = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad b = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \mathrm{i} \\ 0 \end{pmatrix}. \tag{102}$$

We find that the only possibilities are those given in Table 4. They give rise to four different circles of rank (4,4) extremal PPT states. The specific possibility given by Equation (85) is case 2 in Table 4. Since a is real, it is quite obvious from Table 4 that case 1 and case 2 are related by complex conjugation, and so are also case 3 and case 4.

Case	w_1	w_2
1	$b \otimes b^*$	$a \otimes b$
2	$b^* \otimes b$	$a \otimes b^*$
3	$b \otimes b$	$b^* \otimes a$
4	$b^* \otimes b^*$	$b \otimes a$

TABLE 4. The four ways to extend $\mathrm{Img}\,\tau_{44}$ to a five dimensional subspace orthogonal to z_6, and containing exactly two product vectors w_1, w_2.

In our numerical random searches for SL-symmetric PPT states of rank $(5, 5)$ we have (by pure chance) found four examples of states that are SL-equivalent to states lying on the surface of a cone, as described above. In one example the state lies inside the circle, and is of the form given in Equation (90). This is the case discussed in detail above, corresponding to case 2 in Table 4. In the three other examples the state lies on the side of the cone, and is of the form given in Equation (100). These correspond to the case 3 and case 4 in Table 4, where we have $\rho_2^P = \rho_2$ for the pure product state ρ_2.

9. Summary and outlook

The work presented here is a continuation of previous studies of the entangled PPT states of rank $(5, 5)$ in the 3×3 system, with an emphasis on nongeneric states. We use both numerical and analytical methods, and to some extent we build on ideas from [15]. For dimensions 3×3 it is known that the extremal PPT states of lowest rank, which are not pure product states, are entangled states of rank four. The classification of these rank four states is believed to be complete. The structure of extremal rank five PPT states appear however much more complex.

The equivalence between PPT states under SL $\times$ SL-transformations is a very important concept. The special cases where ρ and its partial transpose ρ^P have this SL-equivalence are SL-symmetric states. We define a state ρ to be *genuinely* SL-symmetric if it is SL-equivalent to a state $\tau = \tau^P$. We then show that genuine SL-symmetry implies SL-symmetry, and that in the case of genuine SL-symmetry at least one SL-transformation from ρ to ρ^P must have a special diagonal block form and in addition be trace preserving. We argue that SL-symmetric states can be found numerically, but only by conducting specific searches. We have randomly produced 50 SL-symmetric PPT states of rank $(5, 5)$, which apart from being SL-symmetric, are generic states. These searches are rather special in the sense that we look for product transformations that are trace preserving. This choice is motivated by the fact that we might in this way expect to find states that are genuinely SL-symmetric. A summary of the results can be found in Section 6.

Generic extremal PPT states of rank $(5, 5)$ have no product vectors in their kernel. How to construct PPT states that have a nonzero number n_{ker} of product vectors in the kernel is discussed in [15], and here these matters are further developed. For the 3×3 system this is essentially a study of orthogonal complementary subspaces $\mathcal{U}, \mathcal{V} \subset \mathbb{C}^9$ of dimension five and four, with certain nongeneric properties as described in Section 5.2. Using these nongeneric properties we have constructed several standard forms for $\mathcal{U}$ and $\mathcal{V}$, with n_{ker} ranging from one to four, and then produced PPT states of rank $(5, 5)$ with $\mathcal{U} = \mathrm{Img}\,\rho$ and $\mathcal{V} = \mathrm{Ker}\,\rho$. For a detailed summary of the results, we refer to Section 6.

For the case $n_{\mathrm{ker}} = 4$ we find an interesting new analytical construction of all rank four extremal PPT states, up to SL-equivalence, where they appear as boundary states on one single five dimensional face

on the set of normalized PPT states. The interior of the face consists of rank five states, a simplex of separable states surrounded by entangled PPT states. All these states are real matrices, symmetric under partial transposition.

Also, a very special subspace, which is a $\{2;1\}$ subspace, is collected from a small number of $\{2,2;1\}$ states found in random searches for SL-symmetric states. We describe analytically a set of states found by transformation to standard form of one particular rank $(5,5)$ nonextremal PPT state in this subspace. After transformation to standard form the rank $(5,5)$ state lies inside a circle bounded by extremal rank $(4,4)$ PPT states. The interior of the circle consists entirely of rank $(5,5)$ PPT states, each of which has exactly one product vector in its kernel, and this product vector is common to all the rank $(4,4)$ and rank $(5,5)$ states. Each of the rank $(4,4)$ states has five additional product vectors in its kernel, these are different for the different states. The rank $(4,4)$ states have no product vectors in their ranges, whereas all the rank $(5,5)$ states have one common range containing exactly two product vectors. All the rank $(4,4)$ and rank $(5,5)$ states are symmetric under partial transposition.

All our nongeneric standard forms have been constructed as to contain (at least) six product vectors. There is however no reason to limit the possible values of n_{img} to this. The maximum dimension of an entangled subspace is known from [13]. For the 3×3 system the limiting dimension is four. So any subspace of $\mathbb{C}^9$ of dimension five or higher *must* contain at least one product vector. An analysis on how to construct subspaces of $\mathbb{C}^9$ with various $n_{\text{img}} < 6$ is given in [11]. To fully describe extremal $(5,5)$ PPT states according to the number of product vectors in the range and kernel, these matters should be further developed.

As for the case of the generic PPT states of rank $(5,5)$, we still do not know how to construct these. An underlying structure that these states might contain, and that would make such a construction scheme possible, has yet to be found. Also, we are even further from a full understanding of higher rank extremal PPT states in 3×3 dimensions, or in higher dimensions.

Acknowledgements

We acknowledge gratefully the research grant from The Norwegian University of Science and Technology (Leif Ove Hansen) and cooperation and help from Børge Irgens especially with development of programs, but also with other insights during the research.

References

[1] J.S. Bell,
On the Einstein Podolsky Rosen paradox.
Physics **1**, 195 (1964).

[2] A. Aspect, J. Dalibard, and G. Roger,
Experimental test of Bell's inequalities using time-varying analyzers.
Phys. Rev. Lett. **49**, 25 (1982).

[3] S. Gharibian,
Strong NP-hardness of the quantum separability problem.
Quant. Inf. and Comp. **10**, 3 (2010).

[4] M. Horodecki, P. Horodecki, and R.Horodecki,
Separability of mixed states: necessary and sufficient conditions.
Phys. Lett. A **223**, 1 (1996).

[5] A. Peres,
Separability criterion for density matrices.
Phys. Rev. Lett **77**, 1413 (1996).

[6] P. Horodecki,
Separability criterion and inseparable mixed states with positive partial transposition.
Phys. Lett. A **232**, 5 (1997).

[7] P. Horodecki, M. Lewenstein, G. Vidal, and I. Cirac,
Operational criterion and constructive checks for the separability of low-rank density matrices.
Phys. Rev. A **62**, 032310 (2000).

[8] C.H. Bennett, D.P. DiVincenzo, T. Mor, P.W. Shor, J.A. Smolin, and B.M. Terhal,
Unextendible Product Bases and Bound Entanglement.
Phys. Rev. Lett. **82**, 5385 (1999).

[9] D.P. DiVincenzo, T.Mor, P.W. Shor, J.A. Smolin, and B.M. Terhal,
Unextendible Product Bases, Uncompletable Product Bases and Bound Entanglement.
Commun. Math. Phys. **238**, 379 (2003).

[10] J.M. Leinaas, J. Myrheim, and P.Ø. Sollid,
Numerical studies of entangled PPT states in composite quantum systems.
Phys. Rev. A **81**, 0062329 (2010).

[11] D. Chen and D.Z. Djokovic,
Description of rank four PPT entangled states of two qutrits.
J. Math. Phys. **52**, 122203 (2011).

[12] R. Hartshorne,
Algebraic Geometry.
Springer, New York (2006).

[13] K.R. Parthasarathy,
On the maximal dimension of a completely entangled subspace for finite level quantum systems,
Proc. Math. Sci. **114**, 464 (2004).

[14] Ł. Skowronek,
Three-by-three bound entanglement with general unextendible product bases.
J. Math. Phys. **52**, 122202 (2011).

[15] L.O. Hansen, A. Hauge, J. Myrheim, and P.Ø. Sollid,
Low-rank positive-partial-transpose states and their relation to product vectors.
Phys. Rev. A **85**, 022309 (2012).

[16] D. Chruściński and A. Kossakowski,
Circulant states with positive partial transpose.
Phys. Rev. A **76**, 032308 (2007).