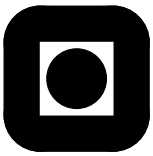
On the Theory of Shannon-Kotel′nikov Mappings in Joint Source-Channel Coding

Pål Anders Floor

A DISSERTATION SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

PHILOSOPHIAE DOCTOR



Department of Electronics and Telecommunications
Norwegian University of Science and Technology

May 2008

Norwegian University of Science and Technology
Faculty of Information Technology, Mathematics, and Electrical Engineering
Department of Electronics and Telecommunications
N-7491 Trondheim
Norway

ISBN 978-82-471-1051-5
ISSN 1503-8181
Doctoral Theses at NTNU, 2008:185

Abstract

In this thesis an approach to joint source-channel coding using direct source to channel mappings is studied. The system studied communicates i.i.d. Gaussian sources on a point-to-point Gaussian memoryless channel with limited feedback (supporting channel state information at most). The mappings, named Shannon-Kotel'nikov (SK) mappings, are memoryless mappings between the source space of dimension M and the channel space of dimension N. Such mappings can be used for error control when $M < N$, called dimension expansion, and for lossy compression when $M > N$, called dimension reduction. The SK-mappings operate on amplitude continuous and time discrete signals (meaning that there is no bits involved) through (piecewise) continuous curves or hyper surfaces in general.

The reason for studying SK-mappings is that they are delay free, robust against varying channel conditions, and have quite good performance at low complexity.

First a theory for determining and categorizing the distortion using SK-mappings for communication is introduced and developed. This theory is further used to show that SK-mappings can reach the information theoretical bound *optimal performance theoretically attainable* (OPTA) when their dimension approach infinity.

One problem is to determine the overall optimal geometry of the SK-mappings. Indications on the overall geometry can be found by studying the codebooks and channel constellations of power constrained channel optimized vector quantizers (PCCOVQ). The PCCOVQ algorithm will find the optimal placing of quantizer representation vectors in the source space and channel symbols in the channel space. A PCCOVQ algorithm giving well performing mappings for the dimension reduction case has been found in the past. In this thesis the PCCOVQ algorithm is modified to give well performing dimension expanding mappings for scalar sources, and 1:2 and 1:3 PCCOVQ examples are given.

Some example SK-mappings are proposed and analyzed. 2:1 and 1:2 PCCOVQ mappings are used as inspiration for making 2:1 and 1:2 SK-

mappings based on the Archimedean spiral. Further 3:1, 4:1, 3:2 and 2:3 SK-mappings are found and analyzed. All example SK-mappings are modeled mathematically using the proposed theory on SK-mappings. These mathematical models are further used to find the optimal coefficients for all the proposed SK-mappings as a function of the channel signal-to-noise ratio (CSNR), making adaptations to varying channel conditions simple.

Preface

This thesis is submitted in partial fulfillment of the requirements for the doctoral degree of *Philosophiae Doctor*, Ph.D. at the Norwegian University of Sciences and Technology (NTNU).

The research was carried out in the period from January 2004 to May 2008 at the Department of Electronics and Telecommunications, at NTNU. The work was founded by the Faculty of Information Technology, Mathematics and Electrical Engineering, via the NFR project "Co-Optimized Ubiquitous Broadband Access Networks" (CUBAN). Professor Tor A. Ramstad has been main supervisor and Associate Professor Lars Lundheim has been co-super visor, both at the Department of Electronics and Telecommunications at NTNU.

Acknowledgements

There are a lot of people who affects my life, and so a long list of people I would like to give an acknowledgement to. But the list is so long that I will only mention the names of people being directly involved with this thesis both when it comes to research and helping out with other problems surrounding it.

First of all I would like to thank my supervisor Professor Tor A. Ramstad. Without his guidance and his large pool of original ideas, this thesis would newer have come into existence. Further he cares for his students and is always available for and interested in discussions (which can go far beyond the specific research field) which is of great inspiration. Cooperation with Tor has been of great value to me. Second, I would like to thank fellow PhD student Fredrik Hekland, who also did a PhD on SK-mappings. Cooperation with Fredrik was of great value and lead into a Journal paper [Hekland, Floor, and Ramstad, 2008]. I would further like to thank Professor of mathematics Yurii Lyubarskii for helping out with some of the mathematics in Chapter 2. Associate Professor Lars Lundheim which has been my co-supervisor has been of great help when it comes to proofreading of both my papers and this thesis. He always has good ideas on how to present difficult concepts in a clear an understandable manner. Visiting PhD student Niklas Wernersson from the Royal Institute of technology (KTH) helped me solve a problem I was stuck with for a long time concerning the PCCOVQ algorithm. Researcher Tor André Myrvoll helped me solve a problem I was stuck with for a long time concerning the approximation distortion lower bound given in Appendix C. He also helped out with some C/C++ programming issues. My office-mate and fellow PhD student Anders Gjendemsjø was of great help when it came to several computer issues and also providing me with some Templates.

I would also like to mention the exceptional environment in the signal processing group at NTNU with all its people. Here I have enjoyed many exiting events both scientifically and socially (where the wine meetings must be mentioned). This is a joyful place to work and an environment

that does not grow on trees. Many of my colleagues has also become good friends of mine.

I would also like thank my family which has always been there for me and helped me out whenever I need it. They have always believed in me and encouraged me to go my own ways.

All my friends play an indispensable role my life making it more colorful in many ways. Without them life would have been dull and I would probably never have been able to do any of the things I do. A special friend has taken my mind to many inspiring places found in everything from everyday life to the great beyond.

Contents

Acronyms

AWGN	additive white Gaussian noise
BPAM	block pulse amplitude modulation
cdf	cumulative distribution function
CSNR	channel signal-to-noise ratio
dB	decibel
FM	frequency modulation
FPM	frequency position modulation
HDA	hybrid digital-analog
HSQLC	hybrid scalar quantizer, linear coder
HVQLC	hybrid vector quantizer, linear coders
i.i.d	independent, identically distributed
JPEG	joint photographic experts group
JSCC	joint source-channel coding
K-PAM	N-fold cartesian products of uniform m-ary PAM with $K = m^N$ points
KKT	Karush-Kuhn-Tucker

MBM	mixed-base modulation
ML	maximum likelihood
MSE	mean-squared error
nD	n-dimensional, $n \in \mathbb{N}$
OPTA	optimal performance theoretically attainable
PAM	pulse-amplitude modulation
PCCOVQ	power-constrained channel-optimized vector quantizer
pdf	probability density function
PM	phase modulation
PPM	pulse position modulation
RD	rate-distortion
SDR	signal-to-distortion ration
SK	Shannon-Kotel'nikov
SNR	signal-to-noise ratio
SQ	scalar quantizer
SSCC	separate source-channel coding
TM	twisted modulation
VQ	vector quantizer

Nomenclature

$\nabla_{x_1 \ldots x_n}$	Gradient over $x_1 \ldots x_n$
$\|\cdot\|$	Euclidian norm
$\|\cdot\|_\infty$	Infinity/Maximum norm
$\|\cdot\|_n$	n norm
$\beta, \gamma, \vartheta, \iota$	Constants used in curve fitting
$\times$	Cartesian product
a	Density of proposed 3:1 and 4:1 mapping
α	Amplification/damping factor
$\mathbf{B}$	Hypothetical inverse of Jacobian matrix J
B_n	Volume of unit n-sphere
$\mathcal{B}(\cdot, \cdot)$	Beta function
$\mathcal{C}$	Codebook for PCCOVQ
$C, \ C(\sigma_z^2)$	Channel capacity
C^1	Set of continuous differentiable functions
$\mathbf{c}_i$	i'th representation vector for PCCOVQ
c_t	Constraining function

$\mathcal{D}$	Domain: $\subseteq \mathbb{R}^M$ or $\subseteq \mathbb{R}^N$
D_t	Total distortion
$\mathbf{d}_r$	Lossless dimension change operator
$d(\cdot,\cdot)$	Distance measure
$d_i(\mathbf{x})$	Distortion associated with representation vector no. i for PCCOVQ
Δ	Distance between two parts of uniform SK-mapping or common scaling of the whole SK-map
δ_n	Deviation where $\delta_n \to 0$ when $n \to \infty$
Δ_1	"Radius" of Helicoid mapping
Δ_2	Distance between two parts of Helicoid mapping
E	Expectation operator
e_i	Error signal for HVQLC
$\text{erf}(\cdot)$	Error function
$\bar{\varepsilon}^2$	Mean square error/Distortion
$\bar{\varepsilon}^2_a$	Approximation distortion
$\bar{\varepsilon}^2_{ch}$	Channel distortion
$\bar{\varepsilon}^2_{th}$	Anomalous/threshold distortion
$\bar{\varepsilon}^2_{wn}$	Weak noise distortion
η	Curve-fitting constant equal to 0.16. Used in approximation to curve-length functions

$f_{\hat{\mathbf{S}}\|x}(\mathbf{S}\|x)$	Likelihood function when using SK-mappings for dimension expansion
$f_x(x)$	Probability density function (source)
G	Metric tensor
g_{ij}	Components of metric tensor
$g_i(\mathbf{x})$	Distortion cost function for PCCOVQ
$\Gamma(\cdot)$	Gamma function
h	General 1D function
$\mathcal{I}$	Set of quantizer indices for PCCOVQ
$\mathbf{I}$	Identity matrix
$\mathbf{i}, \mathbf{j}, \mathbf{k}, \mathbf{l}$	Orthonormal basis vectors
J	Jacobian matrix
J_i	i'th column of Jacobian matrix
k	Spiral type
L	Length of curve
L^p	Space of p-power Lebesgue integrable functions
ℓ	Inverse of φ. Measuring distance from origin to projected point on reducing SK-mapping
λ	Lagrange multiplicator
$\mathcal{L}$	Lagrangian function
M	Source dimension
$\mathcal{M}$	Manifold

$M{:}N$	Mapping between $\mathbb{R}^M$ and $\mathbb{R}^N$
$\mathbb{N}$	Natural numbers
N	Channel dimension
n, $\mathbf{n}$	Noise: scalar, vector
Ω_i	i'th source space partition for PCCOVQ
$\mathcal{P}$	Set of partitions of source space for PCCOVQ
φ	Stretching function (invertible with inverse ℓ)
P_{proj}	Projection matrix
P_N	Power per channel dimension
P, P_{max}	Total system channel power, Maximum allowed power
$\mathbf{p}$	Projection operator
p_r	Probability
$\mathbb{Q}_+$	Positive rational numbers
$\mathbf{q}$	Lossy approximation operator
$\mathbb{R}$	Real numbers
$\mathbf{R}$	Receiver operation (BPAM)
$R(D)$	Rate-distortion function
r	Dimension change factor $= N/M$
ρ	Radial variable in generalized spherical coordinates
ρ_n	Radius of n-sphere
$\mathcal{S}$	SK-mapping (whole set/manifold)

$\mathbf{S}(\cdot)$	Parametric equation of signal hyper surface/curve (SK-mapping)
$\mathbf{s}_0^{'}$	Derivative of sinal curve at x_0: $\mathrm{d}\mathbf{s}(x)/\mathrm{d}x_{\mid x=x_0}$
$\mathbf{s}(\cdot)$	Signal curve
S_i	Component functions of $\mathbf{S}(\cdot)$
$\mathbf{s}_i$	Channel signal vectors for PCCOVQ
$\mathbb{S}^n$	n-sphere
σ_n^2	Noise variance
σ_x^2	Source variance
σ_z^2	Channel signal variance/power
$\mathbf{T}$	Transmitter operation (BPAM)
Θ	Vector with components θ_i, containing all angles in generalized spherical coordinates
θ_i	i'th angle in generalized spherical coordinates
$\mathcal{U}_i$	i'th partition of channel space in nonuniform PCCOVQ
$\mathbf{u}_i$	Equidistant channel signal vectors for uniform PCCOVQ
W_c	Channel bandwidth
W_s	Source bandwidth
x, $\mathbf{x}$	Source signal: scalar, vector
$\hat{x}$, $\hat{x}_{ML}$	ML-estimate for expanding SK-mappings
x_i	Vector component
$\hat{x}_{\pm}$	Detected value when threshold effect has occurred

y, $\mathbf{y}$	Intermediate signal: scalar, vector
z, $\mathbf{z}$	Channel signal: scalar, vector
ζ	Equal to $\eta\Delta$ and used in curve-length approximation

Chapter 1

Introduction

'Information is not knowledge'

-Albert Einstein

Information, or whatever has been labeled *information*, has always been around us in one form or another. Without the exchange of information between humans and between humans and the environment we live in, what would things have been like?

During the last half century a theory on information has emerged. In the aftermath of its development humans have seen an explosion in the access of information all across the planet and also in ways of communicating it. Not only has the possibility to communicate across large distances and orient ourselves at almost every point of the globe become possible, but our understanding of the solar system and far reaches of the universe has increased (or at least we think so) by interpreting the information in the radiation from the universe.

One of the areas where information theory plays an indispensable role is in (tele) communications. One area under strong development is communication of multimedia content like audio, speech, video, pictures etc. The difficulty in communicating multimedia content is its high information content and demands on robustness, yet it is wanted to communicate multimedia content on some of the most difficult media where both power is limited and bandwidth is a scarce resource (like wireless channels).

Information theory also have many applications outside telecommunications. One area is neurobiology, since the human mind is one of the most complex information processing networks we know of. In physics information has become a common term discussed in many areas (where the basic concepts of *entropy* initially came from). Applications also show up in statistics and evolution of molecular codes. The development of quantum

information theory (giving a generalization of classical information theory) probably lies at the frontier of information exchange and processing.

There are still many open problems and undeveloped areas in information theory. The aspects discussed above shed light on the reasons and motives to study and develop the field of information theory and its applications further.

1.1 Information theory

The first development towards a theory of information started out with the term (information) *entropy*, which is a measure of the uncertainty connected with a phenomenon (or a random variable in mathematical terms). The entropy became the measure for the average information content of a source of information. The word entropy comes from physics (the concept dating back to the 1850's), giving a measure of the disorder in a thermodynamical system. Claude E. Shannon borrowed this term and built a theory on information on it [Shannon, 1948]. It is Shannon's concepts that have been developed further to what information theory has become today.

To give bonds on what is ultimately possible for information transfer in any system and the ultimate quality one will get when reducing information, wether it is two units communicating in a point-to-point configuration, or several units communicating in a network, are some of the topics of information theory.

Although not by far every problem has been solved (e.g. the network problem), there are certain cases which have well established bounds on what is ultimately possible. The point to point link and compression of a single source have well established bounds. But even though the bounds are established, and it has been proved that certain coding techniques can be used to reach them, one would still like to find out the best possible way of reaching these bounds for a specific application, something which is basically unknown, and depends on the type of information source in question.

1.1.1 Bounds for coding of sources and capacity of point-to-point links

Figure 1.1 show a typical point-to-point communication system. The transmitter usually consists of both a *source coder*, which is compressing the source information and a *channel coder* protecting the source information from errors.

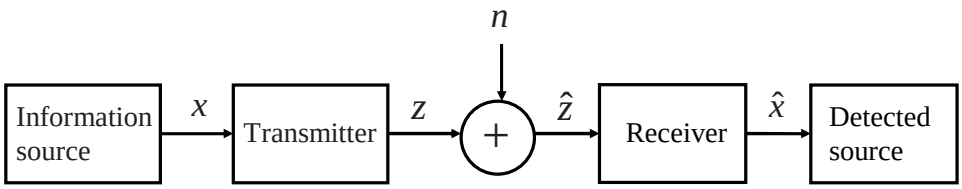


FIGURE 1.1: General point-to-point communication system.

In transmitting a source of information the goal is to represent its information it with the lowest rate possible and with the best possible quality (fidelity). For discrete sources it is possible to do entropy coding, meaning that the rate of the source is reduced without reducing the information content (i.e. lossless compression). For analog sources, however, information must in most (practically all) cases be reduced, meaning that the quality/fidelity of the source is reduced. Given a certain fidelity criterion, it is of interest to represent the reduced source with as low a rate as possible. Shannon derived an absolute lower bound of the rate of a source given a certain distortion, called the *rate-distortion (RD) function* [Shannon, 1959]. In general, it is hard to find a closed form expression for this bound, but for a Gaussian source with bandwidth B, given a mean-squared error (MSE) distortion measure, a closed form expression is shown to be [Cover and Thomas, 1991, pp.344-347]:

$$R\left(D_t\right) = \max\left[W_s \ln\left(\frac{\sigma_x^2}{D_t}\right), 0\right], \tag{1.1}$$

where, W_s is the source bandwidth, σ_x^2 is the source signal power and D_t is the distortion, the ratio σ_x^2/D_t is the signa-to-distortion ratio (SDR), and the rate is in nats per second (using the natural logarithm).

Given that information shall be communicated from transmitter to receiver over some medium (channel) where noise is present (which is always the case in the physical world), there exists an absolute upper limit on the rate that is achievable, guaranteing error free transmission, given a certain Channel signal-to-noise ratio (CSNR), called the *channel capacity* [Shannon, 1948]. Also for channel capacity it is hard to find closed form expressions, but for a channel with additive Gaussian noise, bandwidth W_c, and with an average power constraint (power per channel sample) the capacity is shown to be [Cover and Thomas, 1991, pp.239-242]:

$$C\left(\sigma_z^2\right) = W_c \ln\left(1 + \frac{\sigma_z^2}{\sigma_n^2}\right), \tag{1.2}$$

where σ_z^2 is the average channel power and σ_n^2 is the channel noise power, σ_z^2/σ_n^2 is the CSNR, and the rate is in nats per second (using the natural logarithm).

Both the RD-bound and channel capacity have been proven to be achievable (see e.g. [Cover and Thomas, 1991]), but the bound achieving codes are of infinite length, introducing both infinite complexity and delay into the coders and decoders.

Considering the whole chain from source to receiver the *information transmission theorem* states the following (lossy source coding case) [Berger, 1971, p.282]:

Theorem Let $\epsilon > 0$ and $D \geq 0$ be given. Let $R(D)$ be the RD-function of a time discrete stationary, ergodic source with respect to the single letter fidelity criterion F_d generated by a bounded distortion measure d. Then the source output can be reproduced with fidelity D at the receiving end of any channel of capacity $C > R(D)$. Conversely, fidelity D is unattainable over sufficiently long intervals at the receiving end of any channel of capacity $C < R(D)$. □

(This theorem can be extended to unbounded distortion measures. See [Berger, 1971, p.287]). Looking into this theorem and its proof, it actually states that one can do source and channel coding separately, without any loss (compared to a joint technique), in the case of using achievable codes. This is probably the major reason why almost every communication system today uses separate source and channel coding (SSCC) (also beneficial when it comes to interfaces). But the drawback in the above statements is that it is only guaranteed that nothing is lost using SSCC opposed to joint source-channel coding (JSCC) when the codes in question are optimal, i.e. introducing infinite complexity and delay. However, infinite complexity and delay are impossible in practice, so suboptimal finite complexity source- and channel coders are usually employed. These coders also usually introduces a significant delay into the system, especially if one should close in on the information theoretical bounds. The question is if JSCC will outperform SSCC at finite complexity or at a finite allowed delay.

Analog sources have an infinite information content [Cover and Thomas, 1991, pp.228-229]. Multimedia content like speech, images, audio and video (and for that sake any description of the varieties of nature), are all analog sources of information. To be able to communicate such sources some distortion has to be introduced (except at infinite capacity, which is not realizable in reality), either by the channel noise or by both the channel noise and lossy compression. Distortion is introduced even if the dimensionality of the codes goes to infinity, and so no guarantee for errorfree communication exists for an analog source. Also in showing that the Gaussian channel's capacity is achievable, a countable (countably infinite) number

of separate messages are considered (sphere packing). To communicate an analog source by separate messages, some sort of quantization has had to occur, meaning that some distortion has already been introduced. This distortion, for a given rate, can not be smaller than the distortion given by the rate-distortion function of the source in question. Considering the whole chain from an analog source to the receiver, the only cases where capacity can be said to be truly achieved, without introducing any errors/distortion, would be at infinite capacity (since the information of analog sources is infinite) and by the use of infinite complexity (to reach capacity). Also, in the case where distortion is introduced, infinite complexity is needed to reach the rate-distortion function (e.g. by infinite dimensional vector quantizers).

From this point of view, the rate- distortion function and channel capacity seems like improper measurer in the practical sense, standing alone by themselves, when considering communication of analog sources. A more proper bound describing the information transfer of analog sources is something that gives a certain fidelity at the receiver (signal-to-distortion ratio (SDR)) for a given channel sinal-to-noise ratio (CSNR), in which both the rate-distortion function and channel capacity are included. One such bound is the *optimal performance theoretically attainable* (OPTA) [Berger and Tufts, 1967].

1.1.2 OPTA

OPTA is the result of evaluating the rate-distortion function at channel capacity, and renders the minimum source signal distortion when communicating a source over a channel with a given CSNR [Berger and Tufts, 1967].

The discussion of OPTA is limited to the case of a memoryless Gaussian source and an AWGN channel in this thesis. This serves as the lowest bound for other memoryless sources and channels, as well as e.g. correlated sources and channels with memory. To find OPTA, one equates the source rate and channel capacity: $R = C$. Solving this for the SDR, OPTA is obtained

$$\text{SDR} = \frac{\sigma x^2}{D_t} = \left(1 + \frac{\sigma_z^2}{\sigma_n^2}\right)^{W_c/W_s}. \tag{1.3}$$

The channel/source bandwidth ratio W_c/W_s can in principle take on any real positive number. If $W_c > W_s$, redundant bandwidth is available for communication, and could be used for error control. If $W_c < W_s$, the bandwidth and so the information has to be reduced by some sort of lossy compression before transmission.

Reaching OPTA is possible by introducing infinite complexity source and channel coders, which also introduce infinite delay. Another problem

in combining separate coders is that they are in general non-robust against varying channel conditions. A strong alternative is to use JSCC, where it should be possible to get closer to (and also in some cases reach) OPTA, with both finite complexity and almost no delay. Further, it seems quite simple to make JSCC systems which are robust against varying channel conditions.

1.2 Joint Source-Channel Coding

Considering analog sources (such as video, pictures, sound etc.), JSCC is, for the reasons mentioned in the previous sections, an interesting alternative to separate source and channel coding. Some specific systems illustrate this. It has been proven in [Berger and Tufts, 1967; Gastpar, Rimoldi, and Vetterli, 2003] that for an independent and identically distributed (i.i.d.) source and an additive white Gaussian noise (AWGN) channel, both of the same bandwidth, OPTA is achieved by a direct source-channel mapping. (i.e. a low complexity delay free mapping). Further it has been shown in [Schalkwijk and Bluestein, 1967] that OPTA can be reached also in the bandwidth expansion case (channel bandwidth larger than the source bandwidth), by the use of a system with a noiseless feedback link. Newer results have also showed that even though it is noise on the feedback channel, OPTA can be approximately reached, but now only in a certain range [Ramstad, 2008]. Notice that other cases than expansion with a factor $N \in \mathbb{N}$ has not yet been proved, neither has it been proved that OPTA can be reached in the bandwidth reduction case (source bandwidth larger than channel bandwidth). Probably, also these cases can be reached by finite complexity systems (an approach is suggested later in this thesis to show that OPTA can be reached when $W_c/W_s \in \mathbb{Q}_+$, but then with infinite dimensional systems). Of course, if a well performing feedback channel is not available, the above mentioned systems can not be used, and other types of systems (generally nonlinear) must be considered. An interesting observation is that the above mentioned systems having good performance at low complexity are purely analog.

During the last decades, when most communication systems have become digital, research has been almost uniquely focused on coding and transmission of *digital* sources or in general amplitude- and time discrete sources. In the days of purely analog circuitry, the emphasis was on amplitude- and time continuous sources (often called waveform sources). One field that also got some attention in "the old days" was the communication of amplitude continuous and discrete time sources. All the three mentioned

cases were described and analyzed in detail by Vladimir Aleksandrovich Kotel'nikov in his 1947 PhD dissertation called *The theory of optimum noise immunity* [Kotel'nikov, 1959] (a translation made available in the west in 1959). Since "the takeover" of digital communication it seems that research on communication of amplitude continuous time discrete sources has been almost absent. The reason for this is probably partly the takeover of digital communication and maybe the problems of too high computational complexity for a computer in the 1970's and 80's. The reasons for studying amplitude continuous time discrete systems further are at first their robust character against varying channel conditions, something which digital systems normally lack.

Considering OPTA in (1.3), the bandwidth ratio W_c/W_s can in practice be obtained by e.g. combining M source samples into N channel samples. Assuming Nyquist sampling and an ideal Nyquist channel $W_c/W_s = N/M = r$. Notice that r has to be a positive rational number in this case, i.e. $r \in \mathbb{Q}_+$. Both the source- and channel spaces can be considered Euclidian with dimension M and N respectively. And so the transmitter in Figure 1.1 becomes a mapping between an M dimensional source vector and an N dimensional channel vector (the opposite at the receiver). r is called the "dimension change factor" in the following (or expansion/reduction factor depending on the case under consideration). If $M < N$ the source dimension is expanded onto the channel, meaning that redundant dimensions (e.g. bandwidth) are available, and can be used for noise reduction (error control). If $M > N$ the dimension of the source is reduced, meaning that the information content of the source must be reduced before transmission, and so lossy compression is necessary. In the following, an operation where a source of dimension M is mapped onto a channel of dimension N is called an M:N mapping.

Some known JSCC systems are described in the following sections.

1.2.1 Optimal linear system: BPAM

Having a channel with no feedback BPAM (Block PAM) is the best possible linear solution [Lee and Petersen, 1976]. Assume that an i.i.d. M dimensional Gaussian vector $\mathbf{x}$ should be communicated in an optimal linear way by means of an N dimensional channel vector $\mathbf{z}$. Given that the transmitter is a matrix $\mathbf{T}$ and the receiver is a matrix $\mathbf{R}$, one would like to find the minimum of

$$D_{tot} = E\{(\mathbf{x} - \hat{\mathbf{x}})^2\} \tag{1.4}$$

where

$$\mathbf{z} = \mathbf{T}\mathbf{x} \qquad \hat{\mathbf{x}} = \mathbf{R}\hat{\mathbf{z}} \tag{1.5}$$

and where $\hat{\mathbf{z}} = \mathbf{z} + \mathbf{n}$ and $\mathbf{n}$ is an N dimensional AWGN vector. It can be shown [Lee and Petersen, 1976] that the optimal matrices are

$$\mathbf{T} = \frac{\sigma_z}{\sigma_x}\mathbf{I}_{N\times M} \qquad \mathbf{R} = \frac{\sigma_x\sigma_z}{\sigma_z^2+\sigma_n^2}\mathbf{I}_{M\times N} \tag{1.6}$$

when $M > N$ (dimension reduction), and

$$\mathbf{T} = \sqrt{\frac{N}{M}}\frac{\sigma_x}{\sigma_x}\mathbf{I}_{N\times M} \qquad \mathbf{R} = \sqrt{\frac{M}{N}}\frac{\sigma_x\sigma_z}{\sigma_z^2+\sigma_n^2}\mathbf{I}_{M\times N} \tag{1.7}$$

when $M < N$ (dimension expansion). $\mathbf{I}_{M\times N}$ is an $M \times N$ matrix with ones on the main diagonal and zeros elsewhere. σ_x, σ_z and σ_n is the source-, channel-, and noise standard deviation respectively. Figure 1.2 shows a plot of BPAM and OPTA for some dimension change factors. It is shown

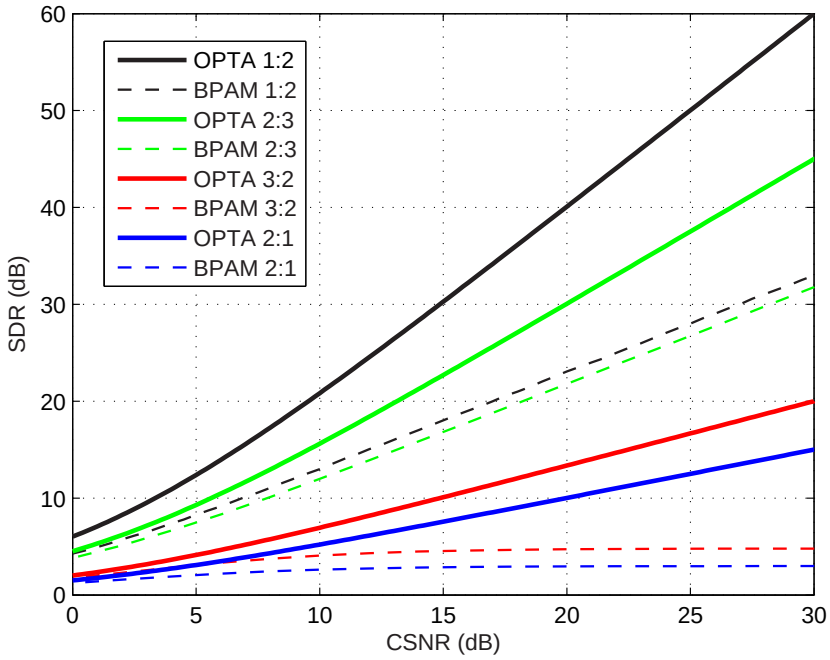


FIGURE 1.2: Some BPAM systems compared to OPTA.

in [Lee and Petersen, 1976] that BPAM reaches OPTA when CSNR $\to -\infty$. Although this case is uninteresting one can see that BPAM are quite close to OPTA for low CSNR (up to approximately 5-10 dB). At high CSNR however, BPAM systems are far away from OPTA (except in the case of $r = 1$

in which BPAM is optimal). To get closer to OPTA it is necessary to turn to nonlinear systems.

1.2.2 Nonlinear Modulation

Some well known nonlinear modulation techniques exists. Of amplitude- and time continuous systems the most well known are Frequency Modulation (FM, which anybody who listen to radio from the 1970's and until today has been involved with, knowing it or not) and Phase Modulation (PM), which are well performing for large bandwidth expansion factors. Among the amplitude continuous and time discrete systems Pulse Position Modulation (PPM) and Frequency Position Modulation (FPM) are well known. Both has been shown to be optimal when the expansion factor N goes to infinity [Wozencraft and Jacobs, 1965, 666-674]. The problem with all of these techniques is that they do not perform well for small bandwidth expansion factors. The reason is that they do not fill the channel space especially well, since ideally, their envelope is constant (it will be explained in chapter 2 why it is important to fill the channel space for small N compared to large N).

Systems built on Kotel'nikovs geometric interpretation [Kotel'nikov, 1959, pp.62-99] using (piecewise) continuous curves (explained further in section 2.1), have shown improved and promising performance for lower bandwidth expansion factors, making them important for applications where bandwidth and power is limited. A special case of these systems is when piecewise continuous line segments are used, such that some sub-channels (of the N-dimensional channel) will have a discrete representation, whereas the other sub-channels have a continuous representation. These are named *Mixed Base Modulation* (MBM). See [McRae, 1971] for some example MBM systems. A common name given to PPM, FPM and Kotel'nikovs 1:N systems is *Twisted Modulation* (TM) [Wozencraft and Jacobs, 1965, pp.611-645].

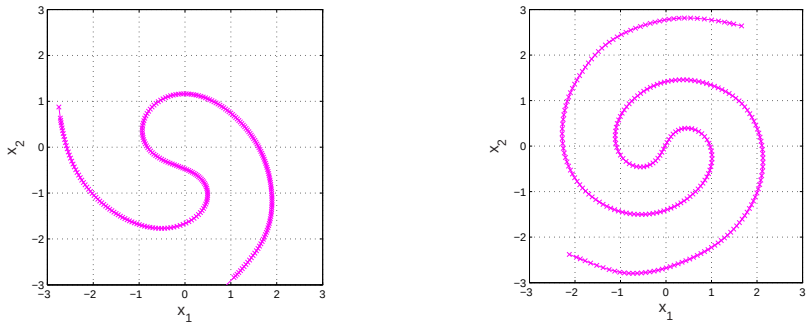
1.2.3 Hybrid Digital Analog systems

More recently developed nonlinear systems are Hybrid Digital Analog systems (HDA) [Coward and Ramstad, 2000; Skoglund, Phamdo, and Alajaji, 2002, 2006]. These schemes are, in principle, built on scalar or vector quantizers and linear coders. Both the centroids (the discrete part, often including channel coders) and the quantization errors (analog part, transmitted with optimal linear systems) are transmitted to perform error control. If no channel coders are included in the discrete part, a HDA system is the same as an MBM system, and so can be included under TM-schemes. A well

performing example of an optimized 1:2 HDA coder, not using traditional channel codes are given in [Coward and Ramstad, 2000] named the *Hybrid Scalar Quantizer Linear Coder* (HSQLC).

1.2.4 PCCOVQ

An example of a system that gives near optimal dimension reduction is *Power Constrained Channel Optimized Vector Quantizers* (PCCOVQ), which is an amplitude- and time discrete system. The PCCOVQ will, however, replicate an amplitude continuous system by using a large number of quantizer representation vectors. Figure 1.3 show a 2:1 PCCOVQ (compression by a factor 2). The representation vectors ($\times$) seem to lay on a continuous curve in the source space (spanned by the components x_1 and x_2). The per-



(a) (b)

FIGURE 1.3: A 2:1 PCCOVQ system. $\times$ is the representation vectors. 1.3(a) Optimized for 12dB CSNR. 1.3(b) Optimized for 20dB CSNR

formance of dimension reducing PCCOVQ system shows close to optimal performance (for all r), in being within approximately 1 dB distance from OPTA [Fuldseth, 1997].

Dimension expanding PCCOVQ mappings have also been developed [Fuldseth, 1997], but none of these had an especially good performance (compared to the MBM and HDA systems mentioned earlier). In Chapter 3, an approach yielding well performing expanding PCCOVQ mappings will be given.

1.2.5 Shannon-Kotel'nikov mappings

Shannon-Kotel'nikov (SK) mappings is an approach on merging source- and channel coders into one delay free (piecewise continuous) mathematical operation, an so is a direct source-channel mapping. The SK-mappings are meant to operate on time discrete and amplitude continuous signals.

The name *Shannon-Kotel'nikov mapping* was developed in two steps. First the name Shannon mapping was chosen in [Ramstad, 2002] to honor Claude E. Shannon, since he presented a geometrical view of the communication problem in his 1949 paper [Shannon, 1949]. However, V. A. Kotel'nikov had already developed a theory for bandwidth expanding modulation in his doctoral dissertation [Kotel'nikov, 1959], dating back to 1947, which implied the same type of structures as Shannon. And so the name Shannon-Kotel'nikov mapping emerged. Kotel'nikov's approach is geometrical, giving good insight into the problem. Such schemes have also been developed further. [Wozencraft and Jacobs, 1965] and [Sakrison, 1968] sums up most of these results.

The SK-mappings have been extended to both M:N dimension expansion [Floor and Ramstad, 2006b] ($M < N$) and reduction [Floor and Ramstad, 2006a] ($M > N$) using hyper surfaces to map between source and channel. These systems have been shown to perform quite well with low complexity and to be quite robust against varying channel conditions [Hekland *et al.*, 2008; Hekland, Øien, and Ramstad, 2005; Floor and Ramstad, 2006c,b,a; Cai and Modestino., 2006]. The mappings, when found, will also be easy to adapt to varying channel conditions, by merely changing a few coefficients in their equations [Hekland *et al.*, 2008]. It is for these reasons of interest to investigate and develop these types of systems further.

Other examples that can be seen as SK-mappings is BPAM, HDA and MBM mentioned in section 1.2.1, 1.2.3 and 1.2.2 respectively. SK-mappings are also closely related to PCCOVQ systems with the exception that PCCOVQ operates on a discrete set of representation vectors. But for a large number of representation vectors, the PCCOVQ will basically be an SK-mapping.

Although no complete theory for these mappings exists, some rules of thumb on how they should be constructed were given in [Lervik, 1996, pp. 102-104]:

1. **Mapping distortion** (only for reducing mappings): Mapping a source space of high dimension to a channel space of lower dimension creates distortion unless noise is absent (the mathematician George Cantor found a way of doing such a mapping in one-by-one manner using *space-filling*

curves [Shannon, 1959], but this technique is impossible to use when noise is added). To minimize the effect of mapping distortion, the SK-mapping must cover the entire source space, such that every source vector have a representation point as close as possible.
2. **Channel signal power**: To minimize the average channel power, source vectors with high probability should be mapped to channel vectors with low amplitude.
3. **Robustness**: To avoid making large errors in the reconstructed vectors, vectors that are close in the channel space should correspond to vectors close in the source space. The opposite, however, is not necessary.

A great deal of analysis on loss factors giving further insight into the construction SK-mappings and also explaining the distance to OPTA for some example SK-mappings are given in [Hekland, Øien, and Ramstad, 2007; Hekland, 2007].

Example on the use of SK- and PCCOVQ mappings

An image coder for flat fading channels built on direct source-channel mappings was presented in [Håkonsen, Ramstad, and Gjendemsjø, 2006] (and a whole thesis on it is given in [Håkonsen, 2007]). The coder shows the potential in using direct source-channel mappings. Its performance is very close to a reference system based on JPEG2000 and adaptive coded modulation. The reference system uses near capacity achieving channel transmission while the proposed system uses simple practical direct source-channel mappings. The reference system is also less robust than the proposed system.

Choice of mappings

The biggest problem with SK-mappings is in general to find them. If M and N are large, they can probably not be found at all (both the SK-maps and the PCCOVQ) due to their complexity. But it was argued in [Coward, 2001; Håkonsen, 2007] that a finite set of well performing low dimensional mappings will together with power allocation do almost as well as an infinite set of M:N mapping. The mappings chosen was

$$\{M{:}0, 4{:}1, 2{:}1, 3{:}2, 1{:}1, 1{:}2\} \tag{1.8}$$

All compressive mappings (except the trivial M:0) are PCCOVQ, while the 1:1 mapping is optimal PAM, and 1:2 is the HSQLC system from [Coward and Ramstad, 2000]. In [Håkonsen and Ramstad, 2006] it was shown that also a 1:3 mapping should be included to make better performance at low

CSNR. It would also seem convenient to have a better resolution between 1:1 and 1:2, like a 2:3 mapping, which should also be possible to find.

1.2.6 Scope of this thesis

Since no general theory on SK-mappings exists, it will be one of the central topics of this thesis. An approach is given so that such systems can be described mathematically. This theory is further used to analyze the asymptotic performance of SK-mappings in the limit of infinite dimensionality.

One of the problems is to determine the overall geometrical structure of the SK-mappings. The only indication one can get on the overall structure will be the PCCOVQ systems. For this reason the PCCOVQ algorithm is modified in chapter 3 to yield better performing expanding PCCOVQ mappings.

Finally, several example mappings are given and analyzed. These will be

$$\{4{:}1, 3{:}1, 2{:}1, 3{:}2, 2{:}3, 1{:}2, 1{:}3\} \tag{1.9}$$

(All of them SK-mappings, except the 1:3 mapping which is a PCCOVQ system). The reason for some of these choices is given in section 1.2.5. Each of these mappings (except the 1:3 PCCOVQ) are given an approximate mathematical model based on the proposed theory on SK-mappings. The model will be compared to actual simulation of each system. All the mathematical models of the proposed SK-mappings are optimized so that their optimal coefficients as a function of the CSNR is found, making them simple to adapt to varying channel conditions.

The theory presented in this thesis will in most cases be for i.i.d. Gaussian vector sources with variance σ_x^2 per dimension and memoryless Gaussian channels with noise variance σ_n^2 per dimension.

1.3 Outline

The reminder of the thesis is arranged as follows
Chapter 2: A theory for SK-mappings is developed. This theoretical model is further used to show that SK-mappings can reach OPTA for all rational dimension change factors $r \in \mathbb{Q}_+$ by letting $M, N \to \infty$.
Chapter 3: The PCCOVQ algorithm is modified to yield better performing 1:N systems. Some problems concerning the algorithm are discussed. Specific examples on 1:2 and 1:3 PCCOVQ mappings are given.
Chapter 4: Example mappings based on Archimedean spirals are intro-

duced. Archimedean spirals can be used for both the 2:1 and 1:2 case. Some mathematical analysis is done, and the systems are optimized. Some mismatch analysis and experiments on the given structures are done.
Chapter 5: Several SK-mappings are introduced. 3:1, 4:1, 3:2 and 2:3 mappings are simulated, and a mathematical model is given. The mathematical models for all the mappings are further optimized. Problems around some more aspects concerning SK-mappings are discussed.
Chapter 6: Conclusions are drawn and some future possible research topics are given.

There are five appendices. Appendix A introduces some necessary results in mathematics. Appendix B derives the minimum weak noise distortion for a dimension expanding SK-mapping. Appendix C derives a uniform spherical distribution, and the approximation distortion lower bound for a dimension reducing SK-mapping. Appendix D derives the channel signal distribution of M:1 SK-mappings. Appendix E derives the 3:2 Helicoid mappings channel distortion.

Chapter 2

Theory on Shannon-Kotel′nikov Mappings

In this chapter a theory for making mathematical models of communication systems using SK-mappings is introduced. This will involve some (simple) concepts from differential geometry. Most concepts are taken from *Riemannian geometry*. The necessary mathematics involved will be described when needed (either in the text or in appendices).

The chapter starts by describing Kotel'nikov's theory on 1:N bandwidth expanding systems [Kotel'nikov, 1959] (expanding the bandwidth of a scalar source by a factor N for transmission), since it is not mentioned in newer literature, and as a reference for further generalization. Next, Kotel'nikov's theory is generalized to include vector sources (of dimension M). This gives an M:N ($M < N$) dimension expanding system. An M:N expanding system performs source error control by mapping the M dimensional vector source onto an N dimensional channel through a parametric hyper surface of dimension M. This mapping will either use N/M times the bandwidth of the source for transmission, or increase the transmission time by a factor N/M. The theory derived is further used to show that M:N expanding systems can close in on OPTA as the dimensionality of the system approaches infinity. Increasing dimensionality means letting the source and channel dimension grow while letting their dimension change factor $r = N/M$ (defined in section 1.2) be constant.

The ideas from the expansion case is further used to develop a theory for M:N dimension reducing ($M > N$) systems. An M:N dimension reducing system performs lossy compression through a projection of an M-

dimensional vector source onto a parametric hyper surface of dimension N. The developed theory is then used to show that also M:N reducing systems close in on OPTA when the dimensionality grows to infinity.

2.1 Kotel'nikovs theory on 1:N bandwidth expanding mappings

Figure 2.1 show a block diagram for the dimension expanding communication system under consideration, which can be used for illustration in this section and section 2.2

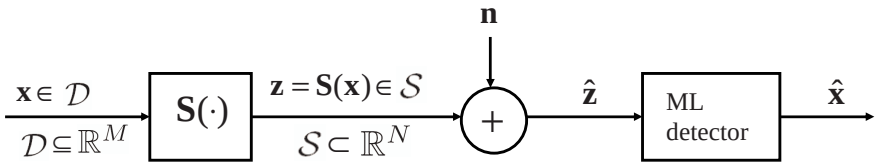


FIGURE 2.1: Block diagram of a general dimension expanding SK-system.

This section describes the necessary aspects (for the sake of this thesis) from part 3 of Kotel'nikovs dissertation [Kotel'nikov, 1959, pp.62-99] in which theory for transmission of amplitude-continuous, discrete-time sources are developed (the theory of this section is presented in a similar way as in [Sakrison, 1968, pp.287-299]).

Consider an amplitude-continuous, time-discrete scalar source $x \in \mathcal{D} \subseteq \mathbb{R}$ (Kotel'nikov refers to this as separate parameter values). The source is communicated using a (parametric) curve in the channel space $x \mapsto \mathbf{s}(x) \in \mathcal{S} \subset \mathbb{R}^N$, called the *signal curve*. Let the noise be denoted $\mathbf{n} \in \mathbb{R}^N$ (i.i.d. and Gaussian). Then the received signal is $\hat{\mathbf{s}}(x) = \mathbf{s}(x) + \mathbf{n}$, with the corresponding likelihood function

$$f_{\hat{\mathbf{s}}|x}(\hat{\mathbf{s}}|x) = \left(\frac{1}{2\pi\sigma_n^2}\right)^{N/2} e^{-\frac{\|\hat{\mathbf{s}}-\mathbf{s}(x)\|^2}{2\sigma_n^2}}. \tag{2.1}$$

The maximum likelihood (ML) estimate is defined by [Therrien, 1992]

$$\hat{x} = \max_{x\in\mathbb{R}} f_{\hat{\mathbf{s}}|\hat{x}}(\hat{\mathbf{s}}|x) \tag{2.2}$$

As the CSNR gets large, the ML estimate approaches that of the optimum estimate (in the mean square sense) [Wozencraft and Jacobs, 1965, pp. 216-219]. (2.1) is maximized by the value x that minimizes the norm

$\|\hat{\mathbf{s}} - \mathbf{s}(x)\|$, and so the ML estimate of x corresponds to the point on the signal curve that is closest to the received vector in Euclidian distance. From this Kotel'nikov reasoned that there are two different contributions to the distortion of the source signal in such systems, *weak noise* which is referred to in the following as *weak noise distortion* and *anomalous errors* which is sometimes referred to as *anomalous distortion*.

When considering weak noise, the signal curve can be approximated in the vicinity of any transmitted signal value x_0 by (linear approximation)

$$\mathbf{s}(x) \approx \mathbf{s}(x_0) + \mathbf{s}_0'(x - x_0), \tag{2.3}$$

assuming that $\mathbf{s}(x) \in C^1$ (continuously differentiable with respect to x). $\mathbf{s}_0' = d\mathbf{s}(x)/dx_{|_{x=x_0}}$. The ML estimate of x can be approximated as the projection onto the tangent line through the value $\mathbf{s}(x_0)$ on the signal curve. Figure 2.2 illustrates this. Given that the value x_0 was transmitted, Ko-

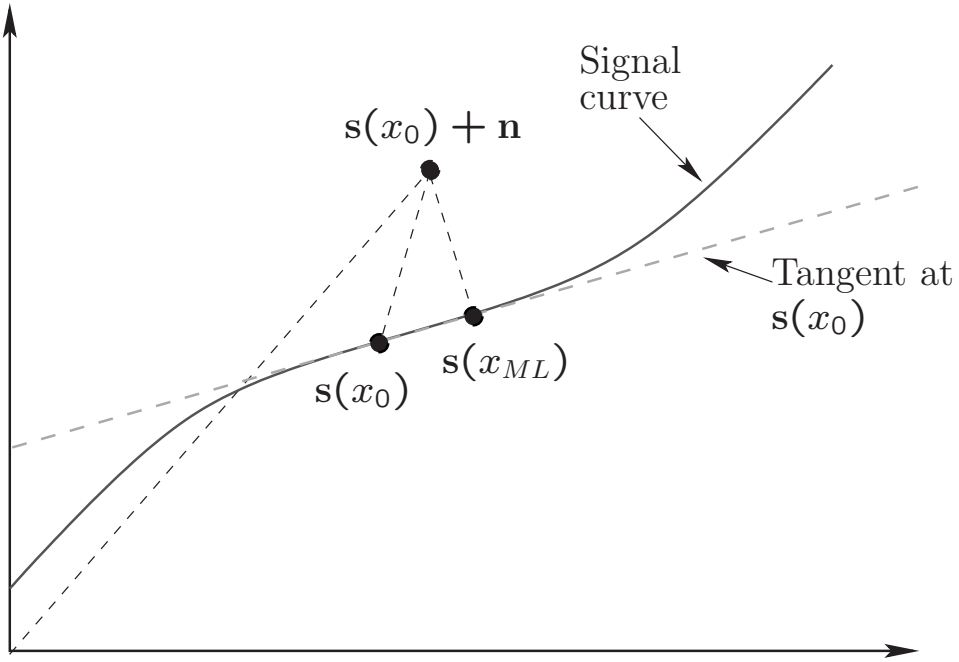


FIGURE 2.2: ML estimate in the weak noise case.

tel'nikov showed that the minimum mean square error (MSE) in the weak noise case is given by

$$\varepsilon_{wn}^2 = E\{(\hat{x}_{\mathrm{ML}} - x)^2 | x = x_0\} = \frac{\sigma_n^2}{\|\mathbf{s}'(x_0)\|^2}. \tag{2.4}$$

σ_n^2 is the noise variance and $\|\mathbf{s}'(x_0)\|$ is the euclidian norm of the curve's velocity (tangent) vector at the parameter value x_0. The weak noise distortion is given by

$$\bar{\varepsilon}_{wn}^2 = E_x\{\varepsilon_{wn}^2(x)\} = \sigma_n^2 \int_{\mathcal{D}} \frac{1}{\|\mathbf{s}\,'(x)\|^2} f_x(x)\mathrm{d}x. \tag{2.5}$$

$\mathcal{D}$ is the source domain and $f_x(x)$ is the probability density function of x. The geometric interpretation of this result is that in order to reduce the noise corruption for a given expansion, without increasing the signal duration or the transmit energy, the signal curve should be made longer (by stretching it like a rubber band, which makes the velocity vectors longer). This should be done without leaving a certain hyper-sphere in order to satisfy a power constraint on the channel. To make the curve as long as possible, it has to be folded/twisted inside this hyper-sphere. I.e. nonlinear systems are needed in order to close inn on OPTA, with the exception of very low CSNR (where linear systems are adequate) or if an ideal feedback channel is available [Schalkwijk and Bluestein, 1967]. This concept is illustrated in Figure 2.3 and 2.4(a). However, the length of the curve cannot be

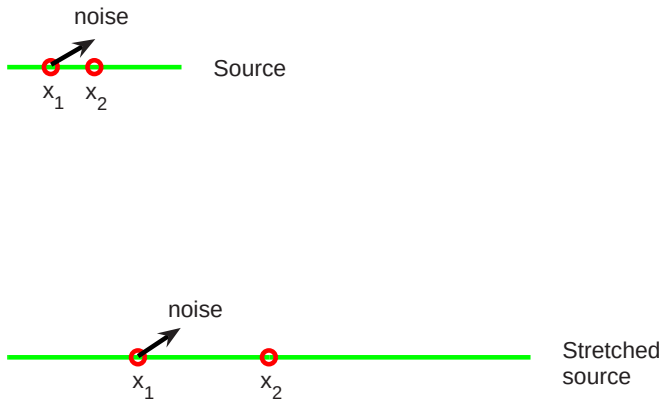


FIGURE 2.3: Kotel'nikovs concept of analog source error control seen intrinsically.

increased beyond a certain length without introducing anomalous errors, also called the *threshold effect* [Shannon, 1949]. These errors are large, since they are the result of the channel noise taking us from one part of the curve to another (and hence rule 3 in section 1.2.5 is violated). The occurrence of these anomalous errors, depends on the relation between the standard

deviation of the channel noise, and the density of the curve. Figure 2.4 illustrates. Due to the severity of the anomalous errors it is wanted to make them occur with as small a probability as possible. Notice that for linear

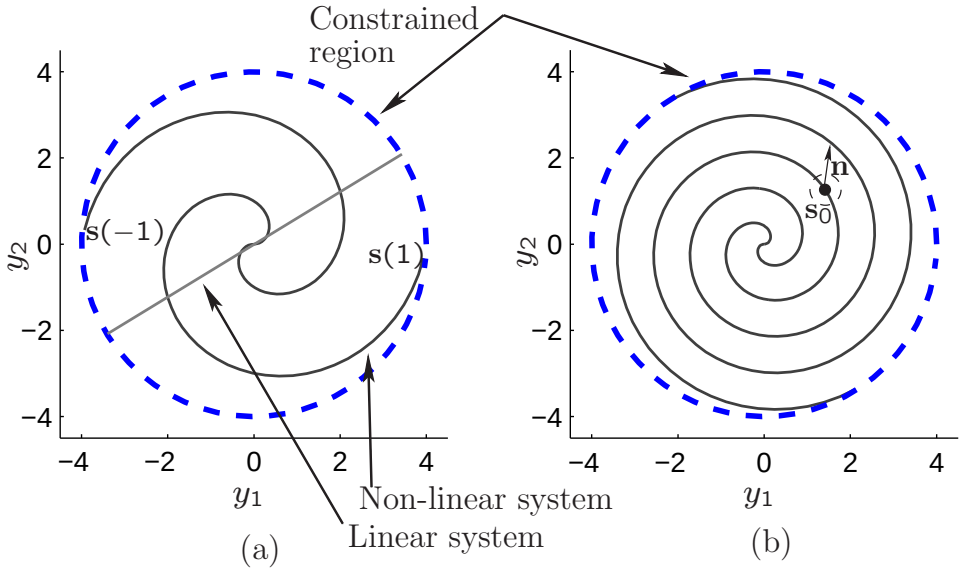


FIGURE 2.4: 1:2 dimension expanding systems. 2.4(a) The straight line illustrates a linear mapping, while the curved line represents a nonlinear mapping. 2.4(b) A nonlinear mapping that has been stretched a significant amount. If stretched too far, different parts of the signal curves come too close. Large values of $\mathbf{n}$ might then take the transmitted point $\mathbf{s}_0$ to the neighboring curve when decoding, resulting in severe distortion.

systems there is no anomalous distortion, i.e. all noise can be considered weak. This also coincides with the way the weak noise distortion is calculated (linear approximation).

There is in addition something to gain in lowering the weak noise distortion further without increasing the probability for the threshold effect. The gain is in the choice of *stretching function*. The stretching function is a bijective function acting on the parameter space, before mapping it onto the parametric curve. It will be denoted $\varphi = \varphi(x)$ in the following. By for instance mapping directly from the parameter x onto the curve, one will in some cases (depending on the structure) get vectors of increasing length for increasing values of x. Then the received distortion will depend on the level of the source amplitudes, and the tangent vectors will be shortest where most of the probability mass of the source resides (which is undesir-

able). Appendix A.3 shows what φ to chose to make the source signal and noise become mutually independent after decoding. φ can be chosen optimally with respect to the pdf of the input signal. How to find the optimum φ for a given pdf, is shown in [Sakrison, 1968, pp.294-297].

2.2 M:N Dimension Expanding SK-mappings.

In this section the theory from section 2.1 is generalized to include vector sources $\mathbf{x} \in \mathcal{D} \subseteq \mathbb{R}^M$. This gives the possibility to consider more general mappings and to exploit dimensionality for a given expansion factor r. Assume that the M source components are i.i.d. and Gaussian. The source will be represented as a "parametric hyper surface" in the channel space $\mathbf{x} \mapsto \mathbf{S}(\mathbf{x}) \in \mathcal{S} \subset \mathbb{R}^N$ which in the following is referred to as the "signal hyper surface" or just $\mathcal{S}$. A general M dimensional hyper surface imbedded in $\mathbb{R}^N$ has the following form

$$\mathbf{S}(\mathbf{x}) = [S_1(\mathbf{x}), S_2(\mathbf{x}), \cdots, S_N(\mathbf{x})] \tag{2.6}$$

where S_i are component functions.

Using $\mathcal{S}$ for communication, the likelihood function of the received signal $\hat{\mathbf{S}} = \mathbf{S}(\mathbf{x}) + \mathbf{n}$ is

$$f_{\hat{\mathbf{S}}|\mathbf{x}}(\hat{\mathbf{S}}|\mathbf{x}) = \left(\frac{1}{2\pi\sigma_n^2}\right)^{N/2} e^{-\frac{\|\hat{\mathbf{S}}-\mathbf{S}(\mathbf{x})\|^2}{2\sigma_n^2}}, \tag{2.7}$$

i.e. the ML estimate of $\mathbf{x}$ corresponds to the point on $\mathcal{S}$ that is closest to the received vector in Euclidian distance.

2.2.1 Weak noise distortion.

In the weak noise case, one can consider the surfaces' tangent space. Figure 2.5 illustrates for the 2:3 case. Assume that each component function of $\mathcal{S}$ is $S_i \in C^1(\mathbb{R}^M)$, $i = 1..N$. The tangent space at a point $\mathbf{x}_0$ is given by (first order Taylor polynomial)

$$\mathbf{S}(\mathbf{x}) \approx \mathbf{S}(\mathbf{x}_0) + J(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0), \tag{2.8}$$

where J is the Jacobian matrix [Munkres, 1991, p.47] (or see Appendix A.1) of $\mathcal{S}$ at $\mathbf{x}_0$. When using an ML detector, the detected value will be

$$\mathbf{S}(\mathbf{x}_{ML}) = \mathbf{S}(\mathbf{x}_0) + P_{proj}\mathbf{n}, \tag{2.9}$$

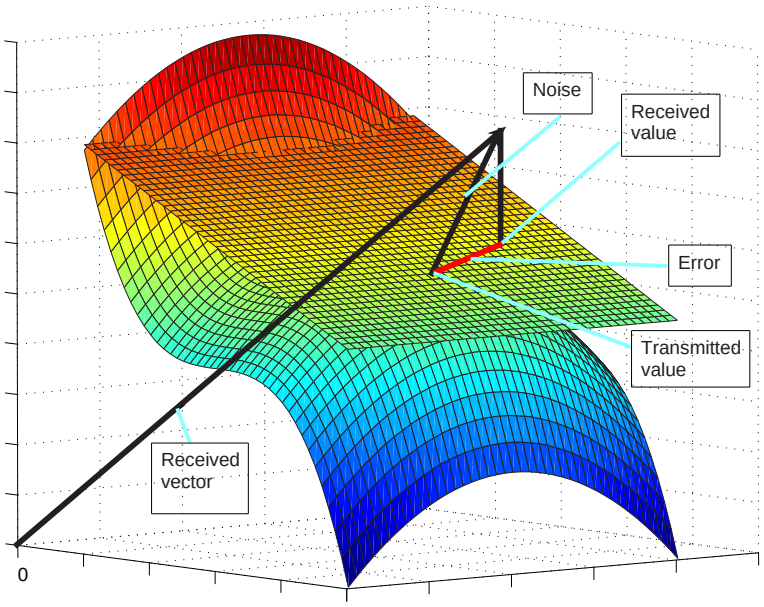


FIGURE 2.5: ML estimate in the weak noise case considering vector sources.

where P_{proj} is a projection matrix given by [Strang, 1986]

$$P_{proj} = J(J^T J)^{-1} J^T = JG^{-1}J^T. \tag{2.10}$$

G is the metric tensor of $\mathcal{S}$ [Spivak, 1999, pp. 301-347] (or see appendix A.1 for a short description). Using (2.8), (2.9) and (2.10) one can easily show that

$$J(\mathbf{x}_{\mathrm{ML}} - \mathbf{x}_0) = JG^{-1}J^T\mathbf{n}. \tag{2.11}$$

By multiplying both sides from the left with the transpose of the Jacobian, then G shows up on both sides. Since G is positive definite, and so also invertible [Strang, 1986]

$$(\mathbf{x}_{\mathrm{ML}} - \mathbf{x}_0) = G^{-1}J^T\mathbf{n}. \tag{2.12}$$

The MSE, given that $\mathbf{x}_0$ was transmitted, is

$$\varepsilon_{wn}^2 = \frac{1}{M}E\{(\mathbf{x}_{\mathrm{ML}} - \mathbf{x}_0)^T(\mathbf{x}_{\mathrm{ML}} - \mathbf{x}_0)\}. \tag{2.13}$$

Inserting (2.12) in (2.13) one can derive that the weak noise distortion given that $\mathbf{x}_0$ was transmitted is (see Appendix B for details)

$$\varepsilon_{wn}^2 = \frac{\sigma_n^2}{M} \sum_{i=1}^{M} \frac{1}{g_{ii}}, \tag{2.14}$$

which is a natural generalization of (2.4) (g_{ii} is the squared norm of the "velocity" vector of $\mathcal{S}$ in the direction of x_i at a point $\mathbf{x}_0$). Considering a Gaussian vector source with i.i.d. components and i.i.d. Gaussian noise on each sub-channel, the above sum is minimized when $g_{ii} = g_{jj}, \forall i, j$ (a spherical shape should be preserved in this case. If $g_{ii} \neq g_{jj}$ a spherical region will be mapped to an elliptical region going from channel to source at the receiver). The weak noise distortion will be given by

$$\bar{\varepsilon}_{wn}^2 = E_{\mathbf{x}}\{\varepsilon_{wn}^2(\mathbf{x})\} = \frac{\sigma_n^2}{M} \iint \cdots \int_{\mathcal{D}} \sum_{i=1}^{M} \frac{1}{g_{ii}(\mathbf{x})} f_{\mathbf{x}}(\mathbf{x}) \mathrm{d}\mathbf{x}. \tag{2.15}$$

This result is a natural generalization of (2.5), and it states that stretching the source space out like a "sheet of rubber" makes the weak noise distortion go down. Notice that although the derivation of (2.15) is done for C^1 functions, one can also use it for piecewise C^1 functions by being cautious during the calculations.

Further one can consider a "shape preserving" mapping, which is to say that every g_{ii} of $\mathcal{S}$ are equal and independent of $\mathbf{x}$ (in another word that the distance between any vectors of the source are equally scaled, not distorted when mapped through $\mathcal{S}$). Then the act of $\mathcal{S}$ can be seen merely as an amplification factor α (from source to channel at the transmitter). In this case (2.15) is reduced to the simple expression

$$\varepsilon_{wn}^2 = \frac{\sigma_n^2}{M} M \frac{1}{\alpha^2} = \frac{\sigma_n^2}{\alpha^2}. \tag{2.16}$$

(2.15) says nothing about any gain from increased dimensionality (this is natural since locally linear systems are considered. In linear systems there is nothing to gain from dimensionality increase). By increasing the dimensionality from e.g. a 1:2 to a 2:4 mapping, and stretching an equal and maximum amount in both parameter directions, gives the same weak noise MSE as in the 1:2 case.

2.2.2 Anomalous distortion and "sphere hardening".

Consider normalized noise vectors in N dimensions $\tilde{\mathbf{n}} = \frac{\mathbf{n}}{\sqrt{N}}$. These vectors have a mean square length equal to σ_n^2. It is shown in [Wozencraft and

Jacobs, 1965, pp. 324-325] that the variance of the squared length goes to zero as $N \to \infty$. So as dimensionality increases, the length of the noise vectors will be more and more localized around the noise standard deviation, and $\lim_{N\to\infty} \|\tilde{\mathbf{n}}\| = \sigma_n$, which will be referred to as the "sphere hardening limit". This is a consequence of the the law of large numbers. The distribution of the length $\rho = \|\tilde{\mathbf{n}}\|$ is given by [Cramér, 1951, p. 237]

$$f_\rho(\rho) = \frac{2(\frac{N}{2})^{\frac{N}{2}}\rho^{N-1}}{\Gamma(\frac{N}{2})\sigma_n^N} e^{-\frac{\frac{N}{2}\rho^2}{\sigma_n^2}} \qquad , N \geq 2, \tag{2.17}$$

where $\Gamma(\cdot)$ is the Gamma function [Bateman, 1953]. Figure 2.6 show (2.17) for some values of N.

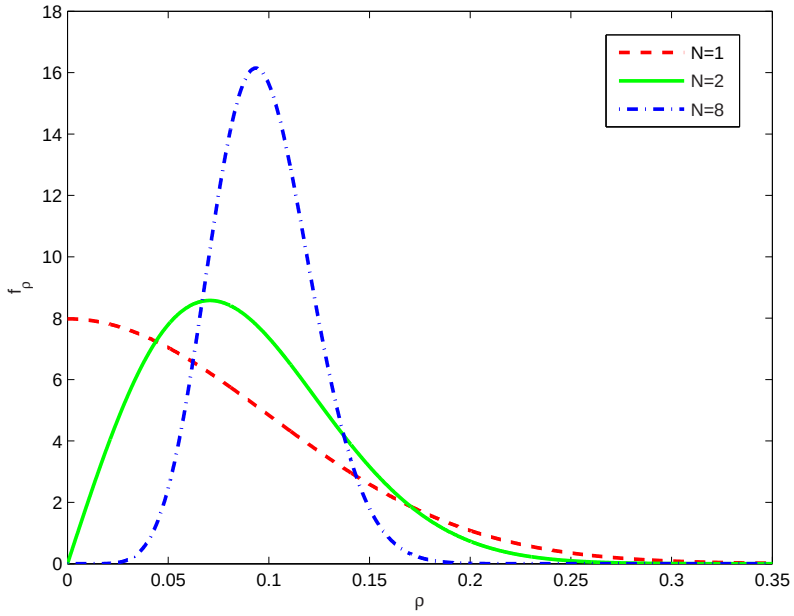


FIGURE 2.6: The pdf for the norm of the normalized noise vectors for some values of N. The noise standard deviation is $\sigma_n = 0.1$.

Considering this effect on SK-mappings, one can benefit from increasing dimensionality in reducing the probability for the anomalous errors, still keeping the same distance between parts/folds of $\mathcal{S}$. In an infinite dimensional system the anomalous distortion can be avoided by letting the distance between folds of $\mathcal{S}$ be at least two times the standard deviation of the noise.

Now assume a map $\mathcal{S}$, where one can stretch the same amount in each parameter direction as was the case for one direction in (2.5). Then according to (2.15), nothing is lost by increasing dimensionality (as long as there are orthogonal basis vectors in the tangent space of $\mathcal{S}$). But the structure can be "packed" more densely in the channel space according to (2.17). This yields additional stretching of the source space, and in principle one should get closer to OPTA by increasing the systems dimensionality. Figure 2.7 illustrates this concept (notice that the intersected 2:4 mapping in the figure is just an example to show the concept, not an actual 2:4 mapping).

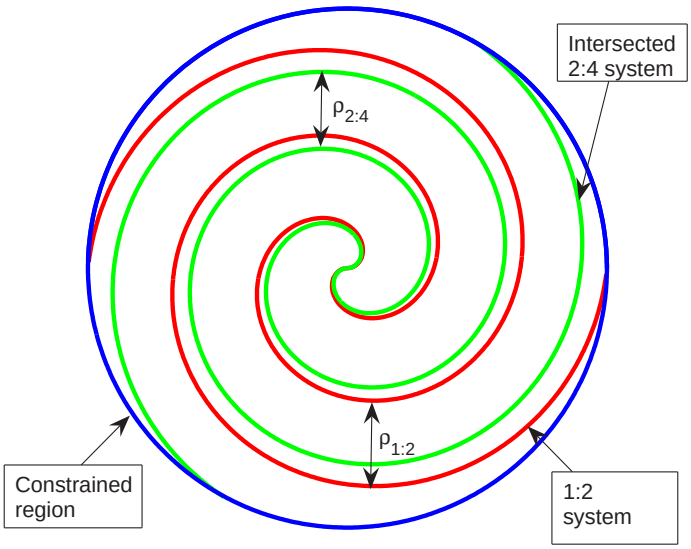


FIGURE 2.7: Illustration on how it might be possible to close in on OPTA by increasing dimensionality. The green line is actually an intersected surface. Since the distance between two parts of the 2:4 system is smaller, the mappings can be made a bit "longer"

2.3 The effect of dimensionality increase on dimension expanding SK-mappings.

In this section one approach is taken to show that it is possible to reach OPTA by using dimension expanding SK-mappings (these results has been

presented in [Floor and Ramstad, 2007]).

Assume that both the channel signal and the noise are normalized with the channel dimension N. Considering a power constrained Gaussian channel, the normalized received vector will lie within an $N-1$ sphere of radius

$$\rho_N = \sqrt{N/N}\sqrt{P_N + \sigma_n^2 + \delta}, \tag{2.18}$$

where P_N is the channel signal power per dimension, and σ_n^2 the noise variance per dimension. By adding the term δ one takes into consideration that ρ_N exceeds $\sqrt{P_N + \sigma_n^2}$ for finite N, so $\delta \to 0$ as $N \to \infty$ (notice that the definition of an N-sphere is $\mathbb{S}^N = \{\mathbf{y} \in \mathbb{R}^{N+1} | d(\mathbf{y}, 0) = \text{constant}\}$ [Spivak, 1999, p.7], where d is the distance from any point $\mathbf{y}$ on $\mathbb{S}^N$ to the origin of $\mathbb{R}^{N+1}$. So the well known *sphere* imbedded in $\mathbb{R}^3$ is denoted $\mathbb{S}^2$, i.e the "2-sphere").

As a starting point the theory done in [Wozencraft and Jacobs, 1965, pp. 666-674] on 1:N mappings is presented, serving as a reference for further generalization.

2.3.1 Asymptotic analysis on 1:N mappings

To be able to establish how long a signal curve can be made (and thereby how small the weak noise distortion can be made) in the limiting case $N \to \infty$, one need to figure out how large a volume of the constrained channel space the signal curve will occupy when anomalous errors should be almost absent (actually avoided in the limit). To figure this out, the signal curve is considered to be the axis of a "hyper cylinder" of dimension $N-1$ and radius ρ_s (which is a function of the noise standard deviation). Arguments in [Wozencraft and Jacobs, 1965, pp. 670-672] suggest that this must be the optimal structure. This cylinder is placed into the given channel hyper sphere. The noise vectors can (conveniently) at every point be decomposed into two statistical independent components: $N-1$ components normal to the curve $\mathbf{n}_a$ (contributing to the anomalous distortion) and one components tangential to the curve $\mathbf{n}_{wn}$ (contributing to the weak noise distortion). For large N one can give an upper bound for the length of the curve. N is assumed to be so large that anomalous errors are almost avoided by letting $\rho_s > \|\mathbf{n}_a\|$ (and $\delta \approx 0$). Let B_N denote the volume contained within an $(N-1)$- sphere of unit radius [Wozencraft and Jacobs,

1965, pp.355-357]

$$B_N = \begin{cases} \frac{\pi^{\frac{N}{2}}}{(\frac{N}{2})!} & , N \text{ even} \\ \frac{2^N \pi^{\frac{N-1}{2}}}{N!}\left(\frac{N-1}{2}\right)! & , N \text{ odd} \end{cases} \tag{2.19}$$

The following inequality must be satisfied in order for the anomalous errors to be absent, and to fulfil the channel power constraint

$$LB_{N-1}\rho_s^{N-1} \leq B_N \rho_N^N \tag{2.20}$$

where L is the length of the signal curve. Since $N-1$ normalized noise components are normal to the curve, $\rho_s = \|\mathbf{n}_a\| = \sqrt{((N-1)/N)\sigma_n^2}$ for very large N. Substituting this and (2.18) in (2.20) gives

$$LB_{N-1}\left(\frac{N-1}{N}\sigma_n^2\right)^{\frac{N-1}{2}} \leq B_N\left(P_N + \sigma_n^2\right)^{\frac{N}{2}}, \tag{2.21}$$

which limits the length of the curve to

$$L \leq \frac{B_N}{B_{N-1}}\sigma_n\left(\frac{1}{1-1/N}\right)^{\frac{N-1}{2}}\left(1+\frac{P_N}{\sigma_n^2}\right)^{\frac{N}{2}}. \tag{2.22}$$

A further elaboration on this is given in [Wozencraft and Jacobs, 1965, pp.673-674].

2.3.2 Asymptotic analysis of M:N expanding SK-mappings

In this section the result from 2.3.1 is generalized to include vector sources to further show that one can reach OPTA by letting the mappings dimensionality grow to infinity. A Gaussian source with limited Euclidian norm, will be distributed on an M-disc or *ball*. This ball is stretched and twisted by $\mathcal{S}$ into the channel hyper-sphere. To figure out how much volume of the channel space the transformed source occupy when the anomalous errors should be almost avoided, a structure that encloses $\mathcal{S}$ needs to be found (like in the 1:N case). This structure must be of the same dimension as the channel sphere, in order to be able to compare their volumes. The structure chosen is $\mathcal{S} \times \mathbb{S}^{N-M-1}$ (which is a natural generalization of the "cylinder" in section 2.3.1). $\mathcal{S}$ can be considered to be a *ball* (or M-disc) with a certain radius ρ_M when considering shape preserving SK-mappings. $\mathbb{S}^{N-M-1}$ is a hyper sphere with radius ρ_{MN} ($\rho_{MN} \geq \|\mathbf{n}_a\|$ at least, to avoid anomalous

errors). If the channel power constraint is to be satisfied and the anomalous errors avoided, the following inequality must be obeyed

$$B_M \rho_M^M B_{N-M} \rho_{MN}^{N-M} \leq B_N \rho_N^N. \tag{2.23}$$

Again, the noise vectors can at each point (of the surface) be decomposed into two statistically independent contributions: M components tangential to the signal hyper surface $\mathbf{n}_{wn}$ (contributing to the weak noise distortion) and $N-M$ components normal to it $\mathbf{n}_a$ (which are the ones causing anomalous errors). Assuming that N is so large that the sphere hardening limit ($\rho_{NM} \to \sqrt{(N-M)/N}\sigma_n$) can be approximately taken into account, (2.23) turns into

$$B_M \rho_M^M B_{N-M} \left(\frac{N-M}{N} \sigma_n^2 \right)^{\frac{N-M}{2}} \leq B_N (P_N + \sigma_n^2)^{\frac{N}{2}}. \tag{2.24}$$

Assuming a shape preserving mapping (each g_{ii} of $\mathcal{S}$ is constant with respect to $\mathbf{x}$ and can be identified by one of the g_{ii}'s) the stretch and thereby the weak noise distortion is determined by the size of the radius ρ_M of $\mathcal{S}$. Solving (2.21) with respect to ρ_M^M gives

$$\rho_M^M \leq \frac{B_N}{B_M B_{N-M}} \left(\frac{1}{1-M/N} \right)^{\frac{N-M}{2}} \frac{1}{\sigma_n^{N-M}} \sigma_n^N \left(1 + \frac{P_N}{\sigma_n^2} \right)^{\frac{N}{2}}, \tag{2.25}$$

and so the following restriction on the radius of $\mathcal{S}$ has to be obeyed

$$\rho_M \leq \sqrt[M]{\tilde{B}} \sigma_n \left(\frac{1}{1-M/N} \right)^{\frac{N-M}{2M}} \left(1 + \frac{P_N}{\sigma_n^2} \right)^{\frac{N}{2M}}. \tag{2.26}$$

For both even and odd N

$$\tilde{B} = \frac{B_N}{B_M B_{N-M}} = \frac{\Gamma\left(\frac{N-M}{2}+1\right)\Gamma\left(\frac{M}{2}+1\right)}{\Gamma\left(\frac{N}{2}+1\right)}. \tag{2.27}$$

This can easily be shown using a similar procedure as equation (C.2)-(C.5) in Appendix C substituting the symbols in question. Putting $M=1$ in (2.26) gives $\rho_M = (L/2)$ compared to (2.22), which shows that (2.26) is a generalization of (2.22). (2.27) can be expressed in terms of the Beta function using the following relation [Bateman, 1953, p. 9]

$$\mathcal{B}(\varrho, \varsigma) = \int_0^1 t^{\varrho-1}(1-t)^{\varsigma-1} \mathrm{d}t = \frac{\Gamma(\varrho)\Gamma(\varsigma)}{\Gamma(\varrho+\varsigma)}, \tag{2.28}$$

and the Functional relation of the Gamma function [Bateman, 1953, p. 3]

$$\Gamma(\varrho+1)=\varrho\Gamma(\varrho). \tag{2.29}$$

Letting $\varrho=(N-M)/2+1$ and $\varsigma=M/2+1$ and using the above relations gives

$$\tilde{B}=\left(\frac{N}{2}+1\right)\mathcal{B}\left(\frac{N-M}{2}+1,\frac{M}{2}+1\right)=\left(\frac{N}{2}+1\right)\mathcal{B}_{(N,M)}. \tag{2.30}$$

Since a shape preserving mapping is considered, the weak noise distortion is given by (2.16), but the decomposition and normalization has to be taken into account. Since M of N components of the normalized noise vectors are the ones contributing to the weak noise distortion

$$\begin{aligned}&\bar{\varepsilon}_{wn}^2=\frac{E\{\|\mathbf{n}_{wn}\|^2\}}{\rho_M^2}=\frac{M\sigma_n^2}{N\rho_M^2}\geq\\&\frac{M}{N}\left(\frac{N}{2}+1\right)^{-\frac{2}{M}}\mathcal{B}_{(N,M)}^{-\frac{2}{M}}\left(\frac{1}{1-M/N}\right)^{-\frac{N-M}{M}}\left(1+\frac{P_N}{\sigma_n^2}\right)^{-\frac{N}{M}}.\end{aligned} \tag{2.31}$$

$\bar{\varepsilon}_{wn}$ can be considered as the total distortion D_t of the M:N system since the anomalous errors can be considered almost absent when $\rho_{MN}\geq\|\mathbf{n}_a\|$ and M,N is close to infinity. Now assume a fixed bandwidth expansion $r=N/M$ (r is the expansion factor from source to channel). Then $M=N/r$, and so

$$D_t=\frac{1}{r}\left(1-\frac{1}{r}\right)^{r-1}\left(\frac{N}{2}+1\right)^{-\frac{2r}{N}}\mathcal{B}_{(N,r)}^{-\frac{2r}{N}}\left(1+\frac{P_N}{\sigma_n^2}\right)^{-r}, \tag{2.32}$$

where

$$\mathcal{B}_{(N,r)}=\int_0^1 t^{\frac{N}{2r}(r-1)}(1-t)^{\frac{N}{2r}}\mathrm{d}t. \tag{2.33}$$

To show that this system can reach the OPTA bound, one need to show that

$$\lim_{N\to\infty}\left(\frac{N}{2}+1\right)^{-\frac{2r}{N}}\mathcal{B}_{(N,r)}^{-\frac{2r}{N}}=r\left(1-\frac{1}{r}\right)^{1-r}. \tag{2.34}$$

Using the product rule for limits [Edwards and Penney, 1998, p.68] gets rid of the first term on the left hand side of (2.34) ($\lim_{N\to\infty}(N/2+1)^{2r/N}=1$). So the problem is reduced to show that

$$\lim_{N\to\infty}\mathcal{B}_{(N,r)}^{-\frac{2r}{N}}=r\left(1-\frac{1}{r}\right)^{1-r}. \tag{2.35}$$

Hölders inequality, described in Appendix A.2, is used for this purpose. Let $f(t) = t^{\frac{N}{2r}(r-1)}(1-t)^{\frac{N}{2r}}$ and $h = 1$ on $I = (0,1)$. Further let $p = \infty$ and $q = 1$. Clearly both f and h are Lebesgue integrable (for $N \geq 0, r \geq 1$), and the norms $\|f\|_\infty$ and $\|h\|_1$ exist. It is easy to see that $\|h\|_1 = 1$. To find $\|f\|_\infty$ the maximum of f must be calculated. Differentiating f with respect to t and then equating to zero, gives

$$t_{max} = 1 - \frac{1}{r}. \tag{2.36}$$

Substituting this in the expression for f gives

$$\mathcal{B}_{(N,r)} \leq \|f\|_\infty = \left(1 - \frac{1}{r}\right)^{\frac{N}{2r}(r-1)} \left(\frac{1}{r}\right)^{\frac{N}{2r}}. \tag{2.37}$$

$\|f\|_\infty$ will dominate more and more over the rest of the contributions of the integral in (2.33) the larger N gets, and when $N \to \infty$ equality in (2.37) is obtained. Raising the right hand side of (2.37) to the power $-2r/N$ gives the wanted result.

The above result does not contain the source variance σ_x^2, so it is valid for unit variance. σ_x can easily be included by letting $\rho_M = \alpha\sigma_x$, where α is an amplification factor (assuming a shape preserving mapping). Solving the new equation with respect to α, and substituting for ρ_M in (2.31), gives the wanted result.

It has to be mentioned that the above result will be valid for the given fixed parameters, meaning that we have a distinct optimal point. If σ_n increase while ρ_{MN} is kept constant the system breaks down rapidly, since the probability for anomalous errors $p_{r_{ae}} \to 1$. If σ_n decrease while ρ_{MN} is kept constant the packing of $\mathcal{S}$ gets non-optimal and the system gets further away from OPTA, but this time in a robust manner (like a linear system, which is apparent from the derivation of the weak noise distortion. The effect can be seen in the example systems in Figure 4.6 and 5.11).

2.3.3 Comments on finite dimensional expanding SK-mappings

At finite dimensionality the anomalous errors must be taken into account, since they will occur with a certain small probability. This is because $\|\tilde{\mathbf{n}}\|$ will get a nonzero variance (if one do not take the fact that the variance of $\|\tilde{\mathbf{n}}\|$ gets larger when N gets smaller into account, the performance will exceed OPTA when N gets smaller). Assuming a nonzero (small) probability for the threshold effect an additional factor must be included in ρ_{MN}

(radius of S^{N-M-1})

$$\rho_{MN} = \sqrt{\frac{N-M}{N}(\sigma_n^2 + \delta_{MN}^2)}. \tag{2.38}$$

Given a certain probability for anomalous errors, δ_{MN} can be found from the cdf of (2.17), substituting $N - M$ for N (since $N - M$ components are normal to $\mathcal{S}$). An additional factor must also be included in ρ_N, since the variance of the received channel vectors also increase

$$\rho_N = \sqrt{P_N + \sigma_n^2 + \delta_N^2}. \tag{2.39}$$

Considering a large CSNR (the anomalous distortion is almost absent at the optimal point for large CSNR), the distortion at the optimal point is approximately given by the weak noise distortion, which in general is

$$\begin{aligned} D_t \approx & \frac{\sigma_n^2}{r}\left(\frac{N}{2}+1\right)^{-\frac{2r}{N}}\left(1-\frac{1}{r}\right)^{r-1} \\ & \mathcal{B}\left(\frac{N}{2}\left(1+\frac{1}{r}\right)+1, \frac{N}{2r}+1\right)\left(\sigma_n^2+\delta_{MN}^2\right)^{(r-1)}\left(P_N+\sigma_n^2+\delta_N^2\right)^{-r} \end{aligned} \tag{2.40}$$

It seems to be more to gain from the sphere hardening effect the larger r is (except when r is so large that the sphere hardening effect must be taken into account, as discussed below). Consider the $r = 2$ versus the $r = 3$ case. In the 1:2 case one will have a curve in a region of the plane, wile in the 1:3 case one will have a curve in a region of space. Given a certain CSNR value, the curve will always be longest in the $r = 3$ case. Assuming that the dimension is increased to a 2:4 ($r = 2$) and 2:6 ($r = 3$) mapping, sphere hardening will take place over the largest "space" in the $r = 3$ case, meaning that there is more to "catch up to" (compared to OPTA) by increasing dimensionality in the $r = 3$ case. This can be stated generally as: A larger r gives a larger $\mathcal{S}$ for a given CSNR, which means that sphere hardening takes place over a larger "space". So one can expect that the gap to OPTA will be larger the larger r is. However, when r gets so large that sphere hardening starts to have an effect, it will again be easier to reach the bounds, since systems with constant envelope (e.g. FPM and PPM, which is shown to be optimal when $r \to \infty$) can be used.

2.4 M:N Dimension Reducing SK-mappings.

Figure 2.8 show a block diagram for the dimension reducing communication system under consideration.

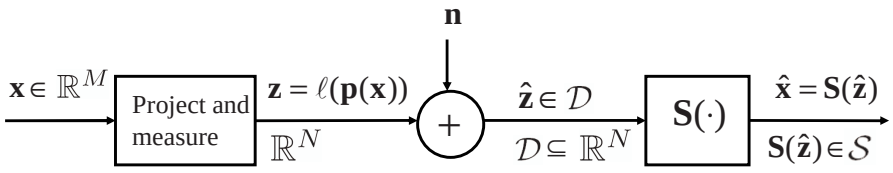


FIGURE 2.8: Block diagram of a general dimension reducing SK-system.

In this section ideas from the expansion case in section 2.2 are used to develop a theory for M:N dimension reducing SK-mappings. To be able to reduce the dimension of the source its information must be reduced (when there is a channel power constraint). This necessitates lossy compression. Compression is done by approximating the source vectors by their projection onto a hyper surface $\mathcal{S}$ which is an N dimensional subset of $\mathbb{R}^M$. This operation is denoted $\mathbf{q}(\mathbf{x}) \in \mathcal{S} \subset \mathbb{R}^M$. The dimension is subsequently changed from M to N by a lossless operator $\mathbf{d}_r : \mathcal{S} \to \mathcal{D} \subseteq \mathbb{R}^N$. The total operation is named the "projection operation", and denoted $\mathbf{p} = \mathbf{d}_r \circ \mathbf{q} : \mathbf{x} \in \mathbb{R}^M \mapsto \mathbf{p}(\mathbf{x}) \in \mathcal{D} \subseteq \mathbb{R}^N$. As for the expansion case $\mathcal{S}$ is called the signal hyper surface (or signal curve in the M:1 case). $\mathcal{S}$ will be a parametric hyper surface with the channel signal $\mathbf{z}$ as its parameters (exchange $\mathbf{x}$ with $\mathbf{z}$ in (2.6)). The point on $\mathcal{S}$, corresponding to $\mathbf{p}(\mathbf{x})$ is given a convenient representation on the channel through an invertible (vector valued) function ℓ. This function determines the way distances are measured from the origin of $\mathcal{S}$ to the given approximated point. The inverse of ℓ is denoted φ (plays a similar role as in the expansion case). The vector $\mathbf{z} = \ell(\mathbf{p}(\mathbf{x}))$ is transmitted over an AWGN channel with noise $\mathbf{n} \in \mathbb{R}^N$. There will be two contributions to the distortion in this system, *approximation distortion* (from the information reducing projection operation), and *channel distortion* (the effect the channel noise has on the signal when mapped through the SK-mapping). These will be described further in the following.

2.4.1 Channel distortion

The received vector $\hat{\mathbf{z}} = \mathbf{z} + \mathbf{n}$ has to be passed through $\mathcal{S}$ to reconstruct $\mathbf{x}$. It is assumed that each component function of $\mathcal{S}$ is $S_i \in C^1(\mathbb{R}^N)$, $i = 1..M$ ($\varphi(\cdot)$ is contained in $\mathbf{S}$ without loss of generality). For a given transmitted channel vector $\mathbf{z_0}$ and small deviations due to the noise $\mathbf{n}$, the received signal $\hat{\mathbf{x}} = \mathbf{S}(\hat{\mathbf{z}})$ can be approximated by (linear approximation)

$$\mathbf{S}(\mathbf{z_0} + \mathbf{n}) \approx \mathbf{S}(\mathbf{z_0}) + J(\mathbf{z_0})\mathbf{n}, \tag{2.41}$$

where $J(\mathbf{z_0})$ is the Jacobian matrix of $\mathcal{S}$ evaluated at $\mathbf{z_0}$. The last term in (2.41) is the distortion due to channel noise. The MSE per source component caused by the channel noise, given that $\mathbf{z}_0$ was transmitted is

$$\begin{aligned}\varepsilon_{ch}^2 &= \frac{1}{M}E\{(J(\mathbf{z}_0)\mathbf{n})^T(J(\mathbf{z}_0)\mathbf{n})\} \\ &= \frac{1}{M}E\Big\{\left(\frac{\partial S_1}{\partial z_1}n_1+\cdots+\frac{\partial S_1}{\partial z_N}n_N\right)^2+\cdots \\ &+\left(\frac{\partial S_M}{\partial z_1}n_1+\cdots+\frac{\partial S_M}{\partial z_N}n_N\right)^2\Big\}.\end{aligned} \tag{2.42}$$

Since the noise on each sub channel is assumed to be independent, $E\{n_i n_j\} = \sigma_n^2\delta_{ij}$. After some rearrangement

$$\begin{aligned}\varepsilon_{ch}^2 &= \frac{\sigma_n^2}{M}\left(\left(\frac{\partial S_1}{\partial z_1}\right)^2+\cdots+\left(\frac{\partial S_M}{\partial z_1}\right)^2+\cdots+\left(\frac{\partial S_1}{\partial z_N}\right)^2+\cdots+\left(\frac{\partial S_M}{\partial z_N}\right)^2\right) \\ &= \frac{\sigma_n^2}{M}(g_{11}+g_{22}+\cdots+g_{NN}) = \frac{\sigma_n^2}{M}\sum_{i=1}^{N}g_{ii}.\end{aligned} \tag{2.43}$$

g_{ii} (partials with respect to z_i), are the diagonal components of the metric tensor of $\mathcal{S}$ (again there is nothing to gain by choosing a nonorthogonal basis for the tangent space of $\mathcal{S}$). The channel distortion will be given by

$$\bar{\varepsilon}_{ch}^2 = E_{\mathbf{z}}\{\varepsilon_{ch}^2(\mathbf{z})\} = \frac{\sigma_n^2}{M}\iint\cdots\int_{\mathcal{D}}\sum_{i=1}^{N}g_{ii}(\mathbf{z})f_{\mathbf{z}}(\mathbf{z})\mathrm{d}\mathbf{z}. \tag{2.44}$$

The base vectors of the tangent space of $\mathcal{S}$ will increase in length when the signal hyper surface is stretched. This means that the more a given surface is stretched out in the source space, the larger is the channel distortion. From this point of view, the surface should be stretched as little as possible. This is the opposite of what was wanted for the expansion case in section 2.2.1.

Considering a shape preserving mapping (as in section 2.2.1) (2.44) is reduced to (for the same reasons as mentioned in section 2.2.1)

$$\bar{\varepsilon}_{ch}^2 = \frac{N\sigma_n^2}{M}\alpha^2 \tag{2.45}$$

where α is the amplification/attenuation from the channel to the source at the receiver.

2.4.2 Approximation distortion

The approximation distortion is the distortion that comes from the lossy projection operation. The size of it is linked to the minimum distance each source vector has to $\mathcal{S}$. For a given dimension reduction, one would like to make the approximation distortion as small as possible. Then $\mathcal{S}$ have to fill the source space as densely as possible (it must be stretched and twisted into a given region). This is in conflict with the requirement of reducing the channel distortion, so there is a tradeoff between the two distortion contributions (as in the expansion case).

Since the approximation distortion is structure dependent, one can not expect to find a general mathematical expression describing it (in a general way). But to be able to do an asymptotic analysis for reducing mappings, some expression valid for large dimensions is needed. By assuming an imbedding of $\mathcal{S}$ where at each point, the distance in the direction of the normal vector (at the point) to another part of $\mathcal{S}$ is constant and equal to Δ (named a *uniform structure*), a similar analysis as for uniform vector quantizers can be used (every centroid having an equal fixed distance Δ between each neighboring centroid). For a uniform vector quantizer in m dimensions (m will be related to M and N later) the distortion can be lower bounded by assuming the decision regions around each centroid to be $(m-1)$-spherical [Gersho, 1979]. Denote the radius of the $(m-1)$-sphere ρ_m. The pdf of the approximation noise will be a *uniform spherical distribution* (a uniform distribution with spherical support, in this case of radius $\Delta/2$) given by (expressed in *generalized spherical coordinates*. See Appendix C)

$$f_{\rho_m,\boldsymbol{\Theta}}(\rho_m,\Theta) = \begin{cases} \frac{m2^{m-1}}{\pi^{\frac{m}{2}}\Delta^m}\Gamma\left(\frac{m}{2}\right), & r \in [0,\Delta/2], \forall\theta_i \\ 0 & \text{elsewhere.} \end{cases} \tag{2.46}$$

The distortion will then be lower bounded by

$$\bar{\varepsilon}_a^2 = E\{\rho_m^2\} = \frac{m}{4(m+2)}\Delta^2. \tag{2.47}$$

The derivations of the distribution in (2.46) and the distortion in (2.47) are given in Appendix C. Due to the fact that the decision regions actually become spherical when $m \to \infty$ using the right construction [Gersho, 1979], (2.47) will be exact when $m \to \infty$ (called *sphere packing*). Notice that this result differs from the well known distortion lower bound derived in [Gersho, 1979]. The reason for this is that Gersho's distortion expression is scaling invariant (independent of the size of the cells containing the centroids), whereas here we want the distortion to depend on the size of the cells so it can be made dependent on the CSNR.

The expression in (2.47) must be modified to take into account a general M-dimensional signal hyper surface. The question is how. Consider first a uniform 2:1 system (uniform spacing between any two parts of it) consisting of concentric circles around the origin in $\mathbb{R}^2$ with a radial distance Δ between them. The $\mathbf{p}$ operator will be 1D and the approximation distortion the same as a scalar quantizer, except that it will be scaled by $1/2$ (divided between two source components). I.e. $\bar{\varepsilon}_a^2 = \Delta^2/24$. Furthermore, consider a uniform 3:2 mapping consisting of concentric spheres around the origin in $\mathbb{R}^3$ with radial distance Δ between them. This is again equivalent to a scalar quantizer, except that the distortion is scaled by $1/3$. Now consider a 3:1 mapping consisting of circles of different radii ($n\Delta$ for $n \in \mathbb{N}$) lying on parallel planes with an equal distance Δ between them (filling out a ball like region in the source space). This time $\mathbf{p}$ is 2D. The approximation distortion will now be equivalent to a 2D uniform vector quantizer except that it will be scaled with $1/3$. The same can be said about a similarly constructed 4:2 mapping consisting of concentric spheres except that the distortion will be scaled by $1/4$. From this it seems apparent that the approximation distortion can be lower bounded by (2.47) by substituting $m = M - N$ and scaling by $1/M$ (distortion per source dimension). And so the lower bound for the approximation distortion will be given by

$$\bar{\varepsilon}_a^2 = \frac{M-N}{4M(M-N+2)}\Delta^2. \tag{2.48}$$

As for for vector quantizers, 2.48 serves as a lower bound, and probably becomes the actual distortion when $M, N \to \infty$ if the right construction (imbedding) is chosen. Notice that in (2.48), Δ must be a function of the CSNR for the expression to be analyzed properly.

2.5 The effect of dimensionality increase in dimension reducing SK-mappings.

In this section it is shown that also M:N dimension reducing systems can close in on OPTA by letting $M, N \to \infty$. To be able to show this, one has to consider sphere hardening (considering the channel signal and noise, and the source signal) and the sphere packing in M dimension (considered in section 2.4.2). As mentioned in section 2.4.2, there is a tradeoff between minimization of the channel distortion (stretching the signal surface as little as possible) and the approximation distortion (stretching the signal surface as much as possible to cover the source space properly). Considering sphere hardening, the length of the source vectors will become

$\rho_M = \|\mathbf{x}\| \to \sqrt{M}\sigma_x$ (non-normalized signals) as $M \to \infty$. This means that there is a smaller space to cover, and so one will get a smaller approximation distortion without increasing the "size" of $\mathcal{S}$ when M, N grows. But as the dimensionality increase $\mathcal{S}$ will also get a smaller spread due to the sphere hardening of the channel signal and noise. It is hard to say if one will gain or loose in total due to these effects. The analysis has to be done in order to figure out what happens.

Asymptotic analysis: To show that it is possible to reach OPTA one has to consider how much the signal surface $\mathcal{S}$ must be stretched in order to cover the source space properly for a given CSNR. This has to be done in a general manner without reference to a specific hyper surface. This can be done by considering volumes. To find out how much space the signal surface will occupy for a certain approximation distortion, it needs to be enclosed in an entity that has the same dimension as the source sphere (to be able to compare the volumes). The entity chosen is $\mathcal{S} \times \mathbb{S}^{M-N-1}$. $\mathcal{S}$ will be a ball-like structure with radius ρ_N considering a shape preserving mapping, while $\mathbb{S}^{M-N-1}$ is a hyper-sphere with radius ρ_{MN}. At all points the distance between two points of $\mathcal{S}$ is kept constant and equal to Δ. This means that $\rho_{MN} = \Delta/2$. The reason why this *uniform covering* is chosen is that (2.48) can be used as an expression for the approximation distortion (M and N large). Notice that in this section (compared to section 2.3.2) it is chosen not to normalize the source and channel signals with respect to either the source- or channel dimension (for simplicity).

To make the approximation distortion as small as possible, $\mathbb{B}^N \times \mathbb{S}^{M-N-1}$ should *cover* the entire source space (for large M and N, and in general the space which has a significant probability associated with it), i.e. the following inequality has to be satisfied

$$B_N \rho_N^N B_{M-N} \rho_{MN}^{M-N} \geq B_M \rho_M^M. \tag{2.49}$$

$\rho_M = \|\mathbf{x}\| = \sqrt{M(\sigma_x^2 + \delta_M)}$ is the radius of the source-space and
$\rho_N = \alpha\sqrt{N(P_N + \sigma_n^2 + \delta_N)}$ is the radius of the channel space, where α is an amplification factor, P_N is the channel power per channel dimension, and σ_n^2 is the noise variance per channel dimension. δ_M and δ_N are included to take the sphere hardening effect into account, so $\delta_M \to 0$ as $M \to \infty$ and $\delta_N \to 0$ as $N \to \infty$. Inserting the above in (2.49) and solving with respect to α gives

$$\alpha \geq \sqrt{\frac{M^{\frac{M}{N}}}{N}} \tilde{B}^{\frac{1}{N}} \left(\frac{\Delta}{2}\right)^{-\frac{M-N}{N}} \sigma_x^{\frac{M}{N}} \sigma_n^{-1} \left(1 + \frac{P_N}{\sigma_n^2}\right)^{-\frac{1}{2}}, \tag{2.50}$$

where

$$\tilde{B} = \frac{B_M}{B_{M-N}B_N} = \frac{\Gamma\left(\frac{M-N}{2}+1\right)\Gamma\left(\frac{N}{2}+1\right)}{\Gamma\left(\frac{M}{2}+1\right)}. \tag{2.51}$$

The last equality in (2.51) can be shown by using the expression for the unit radius hyper spheres given in (2.19) and equations (C.2)-(C.5) in Appendix C. A shape preserving mapping is assumed, and so by inserting (2.50) in (2.45), a general expression for the channel distortion is found

$$\bar{\varepsilon}_{ch}^2 = M^{\frac{M}{N}-1}\tilde{B}^{\frac{2}{N}}\left(\frac{\Delta}{2}\right)^{-2\frac{M-N}{N}}\sigma_x^{2\frac{M}{N}}\left(1+\frac{P_N}{\sigma_n^2}\right)^{-1}. \tag{2.52}$$

Since it is assumed that the approximation distortion and the channel distortion are independent, the total distortion will be given by the sum

$$\begin{aligned} D_t &= \bar{\varepsilon}_a^2 + \bar{\varepsilon}_{ch}^2 \\ &= \frac{M-N}{4M(M-N+2)}\Delta^2 + M^{\frac{M}{N}-1}\tilde{B}^{\frac{2}{N}}\left(\frac{\Delta}{2}\right)^{-2\frac{M-N}{N}}\sigma_x^{2\frac{M}{N}}\left(1+\frac{P_N}{\sigma_n^2}\right)^{-1}. \end{aligned} \tag{2.53}$$

Further the optimal Δ needs to be determined. Differentiating 2.53 with respect to Δ gives

$$\begin{aligned} \frac{\mathrm{d}D_t}{\mathrm{d}\Delta} &= 2\frac{M-N}{4M(M-N+2)}\Delta - \\ &2M^{\frac{M}{N}-1}\frac{M-N}{N}\tilde{B}^{\frac{2}{N}}2^{2\frac{M-N}{N}}\Delta^{-2\frac{M}{N}+1}\sigma_x^{2\frac{M}{N}}\left(1+\frac{P_N}{\sigma_n^2}\right)^{-1}. \end{aligned} \tag{2.54}$$

Equating to zero and solving with respect to Δ gives

$$\begin{aligned} \Delta_{opt} = M^{\frac{M-N}{2M}}&\left(\frac{4M(M-N+2)}{M-N}\right)^{\frac{N}{2M}} \\ &\left(\frac{M-N}{N}\right)^{\frac{N}{2M}}2^{1-\frac{N}{M}}\tilde{B}^{\frac{1}{M}}\sigma_x\left(1+\frac{P_N}{\sigma_n^2}\right)^{-\frac{N}{2M}}. \end{aligned} \tag{2.55}$$

(2.51) can be expressed in terms of the Beta function using (2.28) and (2.29). Letting $\varrho = (M-N)/2+1$ and $\varsigma = N/2+1$ then

$$\tilde{B} = \left(\frac{M}{2}+1\right)\mathcal{B}\left(\frac{M-N}{2}+1, \frac{N}{2}+1\right) = \left(\frac{M}{2}+1\right)\mathcal{B}_{(M,N)}. \tag{2.56}$$

Inserting (2.55) and (2.56) in (2.53) gives

$$\begin{aligned} D_t = &\left(1+\frac{N}{M-N}\right)\left(\frac{M-N}{M-N+2}\right)^{1-\frac{N}{M}} \\ &\left(\frac{M-N}{N}\right)^{\frac{N}{M}}\left(\frac{M}{2}+1\right)^{\frac{2}{M}}\mathcal{B}_{(M,N)}^{\frac{2}{M}}\sigma_x^2\left(1+\frac{P_N}{\sigma_n^2}\right)^{-\frac{N}{M}}. \end{aligned} \tag{2.57}$$

Further assume a fixed dimension reduction $r = N/M$, then $N = Mr$. Substituting this in 2.57 gives

$$\begin{aligned} D_t = &\left(1+\frac{r}{1-r}\right)\left(\frac{1-r}{1-r+2/M}\right)^{1-r} \\ &\left(\frac{1-r}{r}\right)^{r}\left(\frac{M}{2}+1\right)^{\frac{2}{M}}\mathcal{B}_{(M,r)}^{\frac{2}{M}}\sigma_x^2\left(1+\frac{P_N}{\sigma_n^2}\right)^{-r}. \end{aligned} \tag{2.58}$$

Since (using the product rule for limits [Edwards and Penney, 1998, p.68])

$$\lim_{M\to\infty}\left(\frac{M}{2}+1\right)^{\frac{2}{M}} = 1 \text{ and } \lim_{M\to\infty}\left(\frac{1-r}{1-r+2/M}\right)^{1-r} = 1, \tag{2.59}$$

what is left to show is

$$\lim_{M\to\infty}\left(1+\frac{r}{1-r}\right)\left(\frac{1-r}{r}\right)^{r}\mathcal{B}_{(M,r)}^{\frac{2}{M}} = 1, \tag{2.60}$$

where

$$\mathcal{B}_{(M,r)} = \int_0^1 t^{\frac{M}{2}(1-r)}(1-t)^{\frac{Mr}{2}}\,\mathrm{d}t. \tag{2.61}$$

Doing a similar analysis as in section 2.3.2, one will find that

$$\mathcal{B}_{(M,r)} \leq (1-r)^{\frac{M}{2}(1-r)}r^{\frac{Mr}{2}}. \tag{2.62}$$

with equality when $M \to \infty$. Then

$$\begin{aligned} &\left(1+\frac{r}{1-r}\right)\left(\frac{1-r}{r}\right)^{r}\mathcal{B}_{(M,r)}^{\frac{2}{M}} = \\ &\left(1+\frac{r}{1-r}\right)\left(\frac{1-r}{r}\right)^{r}(1-r)^{(1-r)}r^{r} = 1 \end{aligned} \tag{2.63}$$

when $M \to \infty$, which is what we wanted to show.

2.6 Discussion

In this chapter a theory for SK-mappings has been introduced. This theory has further been used to show that the SK-mappings can reach OPTA as the dimensionality grows to infinity.

Fist a theory for dimension expanding mappings was developed. These results show that the overall objective of constructing expanding mappings is to find a structure $\mathcal{S}$ (a signal hyper surface representing the source on the channel), which fill a power constrained region on the channel as properly as possible, by stretching it out as a "sheet of rubber" (minimizing the weak noise distortion). This should be done while the distance between any to folds/parts of $\mathcal{S}$ are kept as large as possible, in order to minimize the effect of the anomalous distortion. This gives a tradeoff between the two distortion contributions, and thus a unique minimum distortion for a given system. This tradeoff seems quite similar to the one in traditional channel coding, where it is desired to place as many codewords as possible into a constrained region (to make the rate as large as possible), and at the same time have as large a distance between them as possible (to minimize the probability of exchanging codewords), as illustrated in [Cover and Thomas, 1991, pp.242-243].

The theory developed was used to show that OPTA can be reached in the limit of infinite dimensionality by using a uniform (fixed distance between any folds of $\mathcal{S}$) and shape preserving (a diagonal metric tensor with constant elements) structure. This, it seems, will not be optimal in the finite dimensional case (especially for low dimensions) which is shown in [Sakrison, 1968, pp. 294-297].

Further a theory for dimension reducing systems was introduced. This theory shows that the overall objective in constructing reducing systems is to choose a structure $\mathcal{S}$ (representing the channel signal in the source space) that cover the source space as properly as possible by keeping the distance between any folds of it as small as possible, in order to minimize the approximation distortion. But on the other hand $\mathcal{S}$ should be stretched out as little as possible to minimize the channel distortion. This gives a tradeoff between the two distortion contributions also for reducing mappings. The tradeoff seems quite similar to the one in lossy source coding, where it is desired to cover the source space as properly as possible (to minimize the distortion) with as few representation vectors as possible (to minimize the rate), as illustrated in [Cover and Thomas, 1991, pp.357-358].

Then it was shown, using the theory developed, that also dimension reducing systems can reach OPTA when the dimensionality grows to infinity, again using uniform shape preserving mappings.

Further research is needed in order to find a method for determining the overall optimal geometrical structure given the source and channel pdf's and their dimensions. Hopefully the theory introduced in this chapter can be extended or modified to some sort of variational calculus problem [Troutman, 1996]. If it is possible to find such (differential) equations, they will probably be solvable in an analytically for low dimensional spaces only.

Chapter 3

Dimension expanding PCCOVQ

In this chapter the PCCOVQ (power-constrained channel-optimized vector quantizer) algorithm from [Fuldseth, 1997] are modified to give well performing bandwidth expanding PCCOVQ mappings for scalar sources (these results has been presented in [Floor, Ramstad, and Wernersson, 2007]).

The reason for studying the PCCOVQ is because it seems like a correctly designed PCCOVQ algorithm will give mappings that perform as close to OPTA as possible for a given source and channel dimensions, due to the high degree of freedom in placing points in space, not being stuck to a specific curve/surface. As mentioned in section 1.2.6, the PCCOVQ will give indications on the overall structure of an SK-mapping. In trying to replicate a PCCOVQ system by a mathematical structure, on would probably be able to find a close to optimal SK-mapping (for a given dimensionality).

As mentioned in section 1.2.4, well performing dimension reducing PCCOVQ has been found [Fuldseth, 1997]. Dimension expanding PCCOVQ has also been studied in [Fuldseth, 1997], but these mappings showed quite poor performance. Fuldseth gave two approaches. One approach, based on general *nonuniform channel signal sets*, would probably give well performing mappings if it was possible to include several hundred points in the channel signal set (representation points in the channel space). Finding these optimal sets numerically requires exhaustive computations, since the objective function is of many variables and irregular with lots of local minima (this problem is discussed in further later in this chapter). The other approach is based on *uniform channel signal sets*: N-fold Cartesian products of uniform m-ary PAM, giving rise to $K = m^N$ points in the channel space, which later will be referred to as K-PAM. This approach is fast, but the de-

grees of freedom is less than in the nonuniform case, since it is stuck to a fixed PAM grid. However, one can get around this problem by using large signal sets (discussed in detail later in this chapter).

This chapter start off by describing the (uniform) PCCOVQ algorithm. The algorithm used is basically the same as the one proposed in [Fuldseth, 1997]. Using different initial conditions and a larger number of centroids and K-PAM symbols are the main differences. If the number of representation vectors and K-PAM symbols is large, the *distortion cost* of using most of them will get too large, meaning that the algorithm automatically chooses a subset of them. With a large enough number of K-PAM symbols (K=8000-11000 symbols are used here), the selected subset will have a regular structure (even though we are stuck to a fixed grid). The inspiration for this was taken from [Wernersson, Karlsson, and Skoglund, 2007]. The problems that shows up when using a PCCOVQ with nonuniform channel signal sets are also shortly addressed. Finally, examples on 1:2 and 1:3 PCCOVQ mappings are given.

3.1 Problem formulation

In this section the PCCOVQ optimization problem from [Fuldseth, 1997] is described, first for the uniform K-PAM case, and then shortly for the nonuniform signal set case. For all cases a Gaussian source $\mathbf{x}$ with zero mean and variance σ_x^2 per component is assumed. Figure 3.1 show the system under consideration.

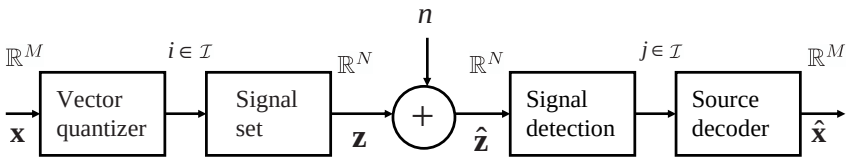


FIGURE 3.1: Block diagram for a PCCOVQ system.

The vector quantizer consists of a codebook $\mathcal{C} \subset \mathbb{R}^M$ of representation vectors $\mathbf{c}_i$, each belonging to a specific partition in a partitioning $\mathcal{P}$ of the source space $\mathbb{R}^M$. Each representation vector is assigned to a specific point $\mathbf{s}_i$ in the channel signal set $\mathcal{S} \subset \mathbb{R}^N$, through an index $i \in \mathcal{I}$, where $\mathcal{I}$ is a set of indices. The received signal set values are detected using an ML detector. The detected signal set values are mapped to specific representation

vectors through an indices $j \in \mathcal{I}$, giving the reconstructed source vectors. Notice that since the emphasis is solely on dimension expansion in this chapter, the symbol $\mathcal{S}$ is used to represent the channel signal set, which will correspond to the signal curve/surface for an expanding SK-mapping.

3.1.1 Uniform channel signal set

Assume that there are K representation vectors in the codebook $\mathcal{C} = \{\mathbf{c}_i\}_{i=0}^{K-1}$. The source space ($\mathbb{R}^M$) is partitioned into K partitions $\mathcal{P} = \{\Omega_i\}_{i=0}^{K-1}$, each containing a specific representation vectors. The vector quantizer approximates each source vector by a representation vector $\mathbf{c}_i$. Each $\mathbf{c}_i$ is represented by an index $i \in \mathcal{I} = \{0, 1, ..., K-1\}$, corresponding to a specific channel symbol $\mathbf{s}_i = \Delta\mathbf{u}_i \in \mathcal{S}$. $\mathcal{S}$ is, in the uniform channel signal set case, a K-PAM alphabet (defined at the start of this chapter). Δ is the minimum distance between points in $\mathcal{S}$, and $\mathbf{u}_i$ are vectors in the K-PAM alphabet with unit distance. On the channel, the transmitted vectors (the $\mathbf{s}_i$'s) are corrupted by an additive noise vector $\mathbf{n}$ generated by a zero mean i.i.d. Gaussian process with variance σ_n^2 per component. At the receiver the detector chooses the index j that corresponds to the point $\mathbf{s}_j$ in $\mathcal{S}$ closest to the received signal $\hat{\mathbf{s}}$, i.e.

$$j = \arg\min_{j \in \mathcal{I}} \|\mathbf{s}_j - \hat{\mathbf{s}}\|^2 \tag{3.1}$$

The average total distortion per source component of this system is

$$D(\mathcal{P}, \mathcal{C}, \Delta) = \frac{1}{M} E\{\|\mathbf{x} - \hat{\mathbf{x}}\|^2\} = \sum_{i=0}^{K-1} \int_{\Omega_i} d_i(\mathbf{x}) f_{\mathbf{x}}(\mathbf{x}) \mathrm{d}\mathbf{x}, \tag{3.2}$$

where $f_{\mathbf{x}}(\mathbf{x})$ is the source pdf, and

$$d_i(\mathbf{x}) = \frac{1}{M} \sum_{j=0}^{K-1} p_{r_\Delta}(j|i) \|\mathbf{x} - \mathbf{c}_j\|^2 \tag{3.3}$$

is the distortion associated with representation vector no.i, where $p_{r_\Delta}(j|i)$ is the probability of receiving index j given that i was transmitted. The transmitted power per channel use is given by

$$P(\mathcal{P}, \Delta) = \frac{\Delta^2}{N} \sum_{i=0}^{K-1} \|\mathbf{u}_i\|^2 \int_{\Omega_i} f_{\mathbf{x}}(\mathbf{x}) \mathrm{d}\mathbf{x} \tag{3.4}$$

The objective is to solve the following constrained optimization problem

$$\min_{\{\mathcal{P}, \Delta, \mathcal{C}\}} [D(\mathcal{P}, \mathcal{C}, \Delta) + \lambda P(\mathcal{P}, \Delta)], \tag{3.5}$$

where λ is a Lagrange multiplier.

3.1.2 Nonuniform Signal set

Considering a nonuniform channel signal set, the equations from the previous section are almost the same, except that $\|\mathbf{u}_i\|$ is replaced by a general $\|\mathbf{s}_i\|$ and $p_{r_\Delta}(j|i)$ is replaced by $p_{r_\mathcal{S}}(j|i)$. $\mathbf{s}_i$ (which corresponds to centroid no. i) can in principle take on any value in the channel space, unlike $\mathbf{u}_i$ which can only take on values in a fixed K-PAM grid. From this it follows that the transition probabilities $p_{r_\mathcal{S}}(j|i)$ will be much harder to calculate than what is the case for the uniform signal set.

3.2 The PCCOVQ Algorithm

As described in [Fuldseth, 1997], the minimum of (3.5) can be found by a *modified generalized Lloyd algorithm*.

3.2.1 Uniform signal set

Select an appropriate initial $\mathcal{C}$ and $\mathcal{S}$
Step 1. Partitioning: Given the codebook $\mathcal{C}$ and the signal set $\mathcal{S}$, the optimal partition $\mathcal{P}$ is given as

$$\Omega_i = \{\mathbf{x} | g_i(\mathbf{x}) \leq g_j(\mathbf{x}), \forall j \in \mathcal{I}\}, \quad i \in \mathcal{I} \tag{3.6}$$

where

$$g_i(\mathbf{x}) = d_i(\mathbf{x}) + \frac{\lambda \Delta^2}{N} \|\mathbf{u}_i\|^2 \tag{3.7}$$

is the *distortion cost function*. $d_i(\mathbf{x})$ is given in (3.3).
Step 2. Update codebook: Given the partition $\mathcal{P}$ and the signal set $\mathcal{S}$, the optimal reconstruction vectors are given by:

$$\mathbf{c}_j = \frac{\sum_{i=0}^{K-1} p_{r_\Delta}(j|i) \int_{\Omega_i} \mathbf{x} f_\mathbf{X}(\mathbf{x}) \mathrm{d}\mathbf{x}}{\sum_{i=0}^{K-1} p_{r_\Delta}(j|i) \int_{\Omega_i} f_\mathbf{X}(\mathbf{x}) \mathrm{d}\mathbf{x}}, \quad j \in \mathcal{I} \tag{3.8}$$

Step 3. Update signal set: For a uniform K-PAM signaling set the only parameter to consider is the minimum distance Δ. Since no known mathematical expression for the optimal Δ given $\mathcal{P}$ and $\mathcal{C}$ exists, it has to be found by a numerical optimization algorithm (gradient search) minimizing (3.5), given $\mathcal{P}$ and $\mathcal{C}$.
Step 4. Repeat or stop: IF the difference between the total distortion at this iteration and the previous is smaller than a certain small convergence criterion ϵ, then STOP! ELSE go to step 1.

3.2.2 Nonuniform signal set

The first two steps are the same as in the uniform case, except that the transition probabilities $p_{r_\Delta}(j|i)$ are replaced by $p_{r_\mathcal{S}}(j|i)$ and $\Delta\mathbf{u}_i$ is replaced by a general $\|\mathbf{s}_i\|$ (that in principle can take on any value in $\mathbb{R}^N$). The transition probabilities are given by

$$p_{r_\mathcal{S}}(j|i) = \int_{U_j} f_\mathbf{n}(\hat{\mathbf{s}} - \mathbf{s}_i)\mathrm{d}\hat{\mathbf{s}}, \quad i,j \in \mathcal{I}, \tag{3.9}$$

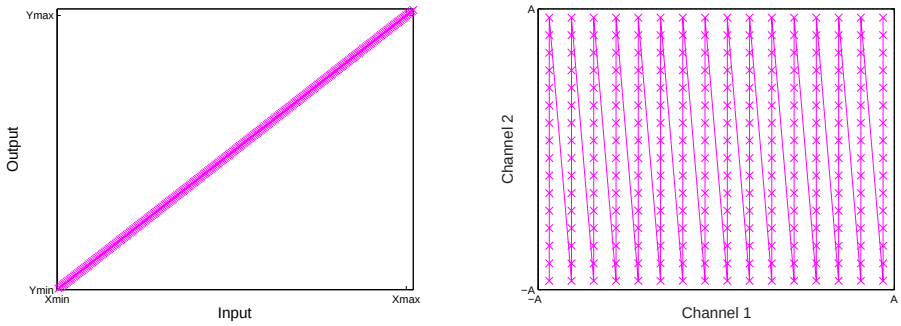
where $\hat{\mathbf{s}}$ is the received signal and $f_\mathbf{n}$ is the noise pdf. $U_j \subset \mathbb{R}^N$ is defined by

$$U_j = \{\hat{\mathbf{s}} | \|\hat{\mathbf{s}} - \mathbf{s}_j\|^2 \leq \|\hat{\mathbf{s}} - \mathbf{s}_k\|^2, \forall k \in \mathcal{I}\}, \quad j \in \mathcal{I}. \tag{3.10}$$

and represents the partitioning of the channel space. The integral in (3.9) has to be calculated at each iteration, which will make the algorithm more complicated than in the uniform case. To find the optimal signal set (since there to our knowledge exists no similar expression as (3.8) for channel signal sets) a numerical optimization algorithm has to be used. Since there are K vectors of dimension N in the signal constellation, the optimization algorithm has to find a solution in KN variables (compared to one variable in the uniform case). Since the objective function is irregular, marred with local minima, the problem gets computationally hard. The chosen optimization algorithm is based on *steepest descent* [Nocedal and Wright, 1999, pp. 21-22]. Although steepest descent is slower than most optimization algorithms it might be the simplest to use in the case of consideration. Due to the irregularity of the objective function, several step sizes has to be tried at each iteration to avoid poor local minima.

3.3 Examples on dimension expanding PCCOVQ

In this section examples on both 1:2 and 1:3 PCCOVQ mappings are given. The emphasis will be on uniform channel signal sets, although a specific example on a nonuniform signal set will be given for the 1:2 case. In [Fuldseth, 1997] *structured projection* of the K-PAM alphabet was used to find the initial codebook. In this thesis it is chosen to use a uniform midrise quantizer as initial condition, which seems to get us closer to the global optimum than structured projection. Each representation vector in the codebook is mapped to a specific K-PAM symbol in the channel signal set. Figure 3.2 shows an example on the correspondence between the initial conditions in the source- and channel space for the 1:2 case. The PCCOVQ algorithm will select a subset of the initial representation vectors and PAM symbols



(a) (b)

FIGURE 3.2: Initial conditions for the 1:2 case using uniform channel signal sets. $\times$ corresponds to representation vectors and PAM-symbols. (a) Source space: uniform midrise quantizer. (b) Channel space: uniform 2-PAM alphabet. The representation vectors with magnitude Ymin is mapped to the PAM symbol (-A,-A) and the representation vectors with magnitude Ymax is mapped to the PAM symbol (A,A). The intermediate representation vectors are mapped according to the line connecting the PAM symbols in 3.2(b). A is chosen to be $4\sigma_x$.

according to the distortion cost function in (3.7) (those who have the least distortion cost associated with them). Using a very large PAM alphabet will guarantee that certain representation vectors and PAM symbols will be left out, and regular structures will emerge.

Notice that after running the uniform PCCOVQ algorithm, representation vectors that were in increasing order of magnitude (like the uniform quantizer in Figure 3.2(a)) before running the algorithm is now highly shuffled and so need to be sorted for the correct structure to appear. By sorting the representation vectors in in increasing order of magnitude the signal set can be found by mapping through the indices corresponding to the representation vectors (the index assignment). Figure 3.3 displays the uniform PCCOVQ concept graphically.

3.3.1 1:2 dimension expanding PCCOVQ

In this section, examples for both uniform and nonuniform signal sets are presented.

Uniform signal set: The optimization is done for several CSNR values. The initial condition is like shown in Figure 3.2 with $A = 4\sigma_x$. Different

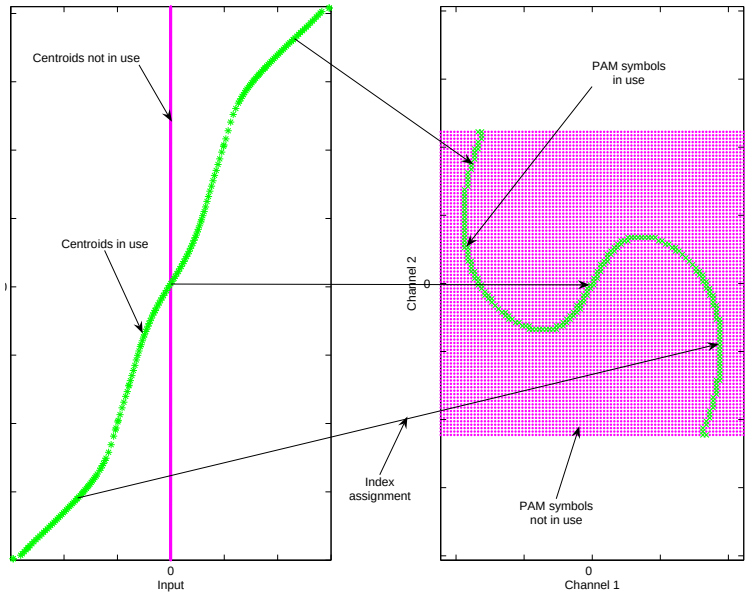


FIGURE 3.3: The concept of the uniform PCCOVQ algorithm illustrated in the 1:2 case. Optimization is done for 5 dB CSNR

structures will emerge depending on the CSNR. The lower the CSNR the simpler the structure. Figure 3.4 shows the channel space structures found for 0, 5, 9 and 13 dB CSNR. At 0 dB CSNR a straight line emerges. This is in accordance with [Lee and Petersen, 1976], where it is shown that linear systems (BPAM) perform close to OPTA when the CSNR is small (see section 1.2.1). For higher CSNR a spiral like structure starts to emerge. At larger CSNR, the structure consists of several curve segments. This seems quite natural since the distortion due to jumps between different parts of the structure caused by the channel noise (the anomalous distortion described in section 2.1) will get smaller (we e.g. get jumps between positive source values instead of jumps from positive to negative values). This will make the system more robust against deteriorating channel conditions. The number of selected symbols will grow with the CSNR (93 points at 0 dB and 706 points at 20 dB).

Nonuniform signal set: In [Fuldseth, 1997], the initial signal set was picked at random from a Gaussian distribution. This initial condition is so far from the optimal, that the algorithm will not evolve (i.e. it will be stuck

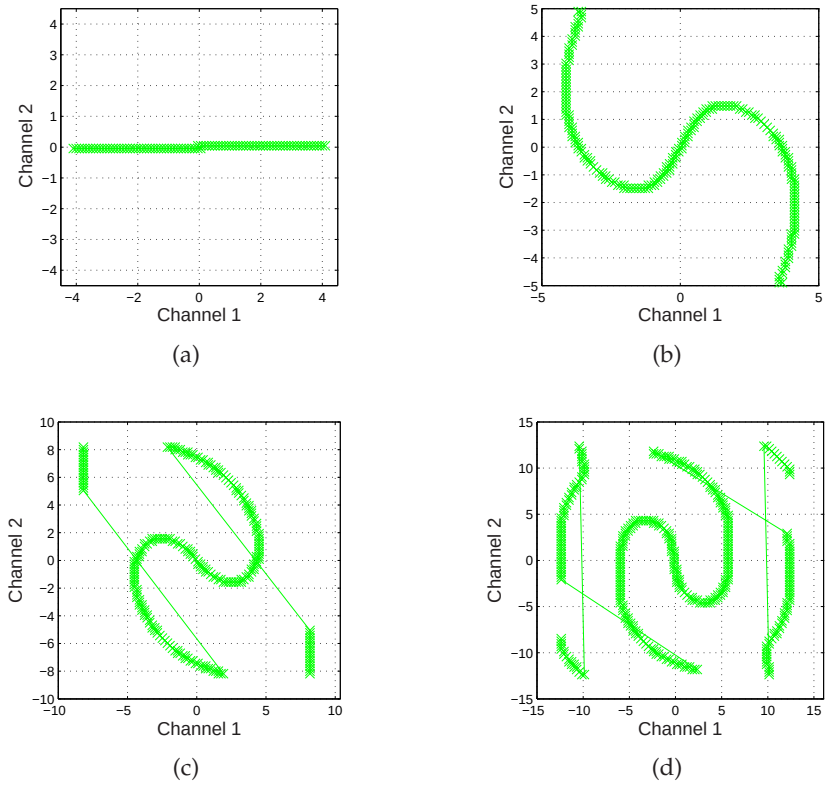
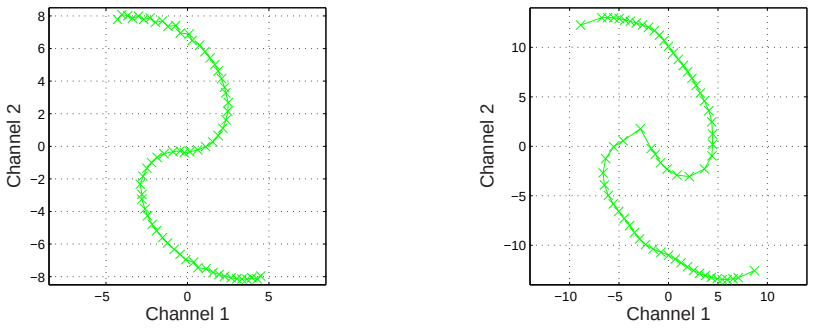


FIGURE 3.4: Channel signal set for 1:2 dimension expanding uniform PCCOVQ. Each $\times$ represent a specific $\mathbf{s}_i$. The lines drawn between the selected $\mathbf{s}_i$'s show the correspondence between neighbors in the one dimensional reconstruction codebook. (a): 0 dB CSNR. (b): 5 dB CSNR. (c): 9 dB CSNR. (d): 13 dB CSNR.

on the same structure for all CSNR values). In our approach the initial channel signal set is chosen to lie on a curve which close to the origin looks like a straight line. Further out it bends back into the channel space in a spiral-like manner (the reason why we let the curve bend is to get a nonzero probability for the threshold effect, or else the algorithm will be stuck on a straight line). All points along this curve are chosen to be equidistant. As in [Fuldseth, 1997], *noisy channel relaxation* is used, meaning that one start by designing the PCCOVQ for very low CSNR, and then in consecutive iterations of the algorithm use the codebook and signal set from the previous CSNR as initial condition for the present. Figure 3.5 shows the channel

signal set for the nonuniform PCCOVQ consisting of 64 points, optimized for 9 dB and 13 dB CSNR. The structures emerging have similarities to the

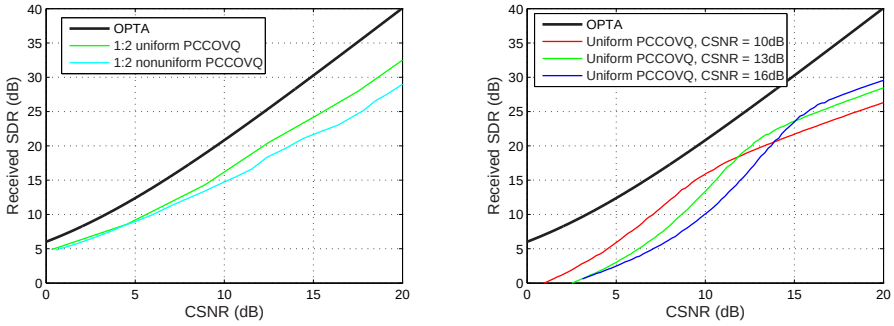


(a) (b)

FIGURE 3.5: Channel signal set for 1:2 bandwidth expanding nonuniform PCCOVQ. Each $\times$ corresponds to a specific $\mathbf{s}_i$. The lines drawn between the selected $\mathbf{s}_i$'s show the correspondence between neighbors in the one dimensional reconstruction codebook. (a): 9 dB CSNR. (b): 13 dB CSNR.

structures in Figure 3.4. The reason why the segmented structures of Figure 3.4 do not show up, might be that a restriction to 64 points in the channel constellation gives rise to different structures, or that the optimization algorithm in question is unable to find piecewise continuous structures, or (most probably) that the algorithm is stuck in a local minimum.

Figure 3.6 shows the performance of the 1:2 PCCOVQ systems. Figure 3.6(a) show an the performance for both uniform and nonuniform PCCOVQ for a range of CSNR values. The curves are interpolations between the optimal points for systems designed for different CSNR. One can observe that the uniform PCCOVQ is performing well for all CSNR in the given range. The nonuniform PCCOVQ is on par with the uniform systems for CSNR below 5 dB CSNR, but start to deteriorate above 5 dB CSNR. The reason why the nonuniform PCCOVQ deteriorates the way it does, is probably because only 64 points are used (compared to using e.g. 235 points for 9 dB CSNR and 368 points for 13 dB CSNR chosen by the algorithm in the uniform case). It is also highly likely that the numerical optimization algorithm get stuck in a local minimum, due to the irregularity and the large number of variables in the objective function. If it was possible to make the nonuniform algorithm find the optimum for several hundred points in the constellation, the nonuniform PCCOVQ would probably be on par with or



(a) (b)

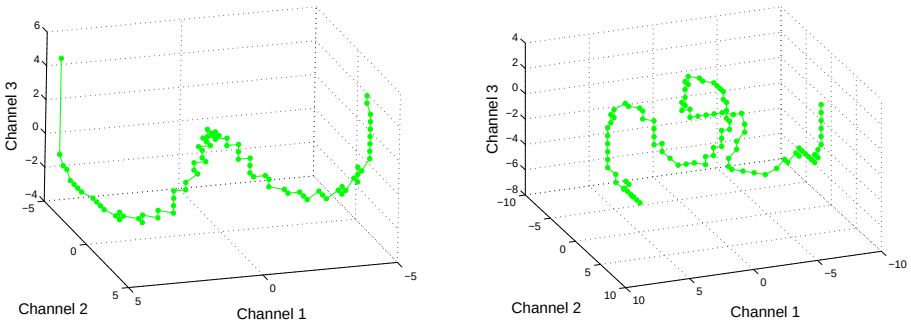
FIGURE 3.6: Performance of 1:2 bandwidth expanding systems. (a): Comparison between the uniform and nonuniform 1:2 PCCOVQ. (b): Robustness plot for some selected 1:2 PCCOVQ mappings.

even outperform the uniform PCCOVQ. Also the uniform PCCOVQ deteriorates above 20 dB CSNR. The reason might be that the algorithm cannot find the global distortion minimum (or at least something close to it). This might improve if one cold implement noisy channel relaxation. The reason why noisy channel relaxation was not used in this case, is that when the algorithm first has left some centroids/PAM symbols out of the problem, we know of no method to take them in again in a correct manner. This could be a subject for further research.

Figure 3.6(b) shows the robustness against varying channel conditions for some selected uniform 1:2 PCCOVQ mappings. One can observe that the PCCOVQ is robust against varying channel conditions.

3.3.2 1:3 dimension expanding PCCOVQ

In this section PCCOVQ mappings for dimension expansion by a factor of 3 are considered. Only uniform signal sets will be considered. The initial condition in the source space is the same as in Figure 3.2(a). The initial condition in the channel space is a generalization of the initial condition shown in Figure 3.2(b). The PAM symbols in Figure 3.2(b) lie in a plane. The initial condition here is several such planes stacked upon each other with distance Δ between them (forming a cube), each connected to its parallel plane at one point. The optimization is done for several CSNR values. Figure 3.7 shows the channel space structures found for 4 dB and 8 dB CSNR. The selected points in the signal set seem to lie on a curve that is twisted around

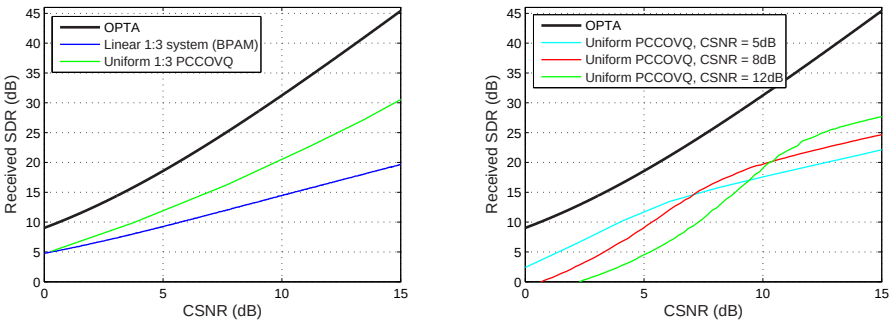


(a) (b)

FIGURE 3.7: Channel signal set for 1:3 bandwidth expanding uniform PCCOVQ. Each • corresponds to a specific $\mathbf{s}_i$. The lines drawn between the selected $\mathbf{s}_i$'s show the correspondence between neighbors in the one dimensional reconstruction codebook. (a) 4 dB CSNR. (b) 8 dB CSNR.

in $\mathbb{R}^3$. At higher CSNR, the 1:3 mapping (like the 1:2 mapping) consists of several curve segments (not shown, since it is hard to visualize it from just one angle in $\mathbb{R}^3$). Also here the number of selected points grows with CSNR.

Figure 3.8 shows the performance of the 1:3 PCCOVQ systems. In Fig-



(a) (b)

FIGURE 3.8: Performance plots. (a) Comparison between 1:3 PCCOVQ and an optimal linear 1:3 system. (b) Robustness plot for some selected 1:3 PCCOVQ mappings.

ure 3.8(a) the optimal performance of each PCCOVQ mapping is compared to 1:3 BPAM. Although the PCCOVQ has a significant gain compared to the linear system, the distance to the optimum bound is quite large. The performance will probably get better if more points are included in the channel signal set during optimization. It might also be that better initial conditions can be chosen. Including noisy channel relaxation might also make the result improve. But as discussed in section 2.3.3, there seem to be more to catch up to (compared to OPTA) by increasing dimensionality the larger r is, so one can not expect to get as close to OPTA as the 1:2 system. In Figure 3.8(b) the robustness against varying channel conditions for the 1:3 PCCOVQ system is shown for some selected mappings. From these results it is evident that the PCCOVQ is robust.

3.4 Discussion.

In this chapter PCCOVQ algorithms for finding 1:N dimension expanding JSCC systems have been introduced.

The PCCOVQ structures showing up in both the 1:2 (Figure 3.4) and 1:3 (Figure 3.7) case seem to be in accordance with the theory on SK-mappings introduced in section 2.1. When the CSNR goes up the noise gets smaller relative to the constrained region on the channel. Then, according to the SK-theory from section 2.1, the signal curve can be made longer. This is what one can observe to happen with the PCCOVQ structures. The larger the CSNR the more PAM symbols are used (ranging from 93 for 0 dB CSNR to 706 for 20 dB CSNR, in the 1:2 case). This results in the source information being spread out over a larger region of the channel space, which again makes it more immune against the channel noise. Further the structures are segmented. It seems like the segmentation is done in such a way that the probability for the threshold effect gets small, and that the noise introduced when it first happens is as small as possible, without significantly increasing the channel power. As seen from Figure 3.4(d), the structure start to get a kind of square like shape. This is probably because the K-PAM alphabet is square. One might expect that the PCCOVQ structure would become more circular if the initial $\mathcal{S}$ was like a disc (and maybe even perform better). Further, it might be possible to make the uniform PCCOVQ algorithm improve by figuring out a way of including noisy channel relaxation.

Using nonuniform channel signal sets in the PCCOVQ should give the best possible result. But the computational complexity gets too large if the necessary amount of variables are taken into the problem. But if it was possible to find an expression similar to (3.8) for signal sets, the problem

would be solved since there would be no need for a numerical optimization algorithm, and the nonuniform PCCOVQ algorithm would speed up significantly. The problem is, compared to (3.8) that such an expression would have to put different weights on crossing different edges in the cells surrounding a specific channel symbol. This could be a subject for further research.

It would be of interest to also develop the algorithm further to general $M{:}N$ expansion. Then one could get an indication on how to construct more general SK-mappings, and also verify if there is any gain from increased dimensionality as predicted in section 2.2.2 and (2.3).

Chapter 4

The Archimedean spiral used as 1:2 and 2:1 SK-mappings

In this chapter examples on 1:2 dimension expanding- and 2:1 dimension reducing SK-mappings are given. This whole chapter is devoted to these examples since the same structure give well performing mappings for both expansion and reduction. Both the 1:2 and 2:1 mapping is realized using parametric curves. One of the most important reasons for using curves (or hyper surfaces in general) compared to e.g. a PCCOVQ system is that once the overall mathematical structure of the curve is found, changing coefficients in its equation is the the only thing needed to adapt to varying channel conditions. Once a specific structure is chosen, one can use the theory from Chapter 2 to calculate its performance.

For both the 1:2 and 2:1 case an approximate mathematical model when using Archimedes' spiral [Zwillinger, 2003] is given. These models are used further to determine the optimal system parameters as functions of the CSNR (these results have been presented in [Hekland *et al.*, 2008]). At the end of the chapter some experiments are performed to determine how robust the proposed systems are against pdf mismatch on both source and channel, and how the given systems can be modified to get closer to OPTA.

Block diagrams showing the systems under considered in this chapter are depicted in Figure 4.1.

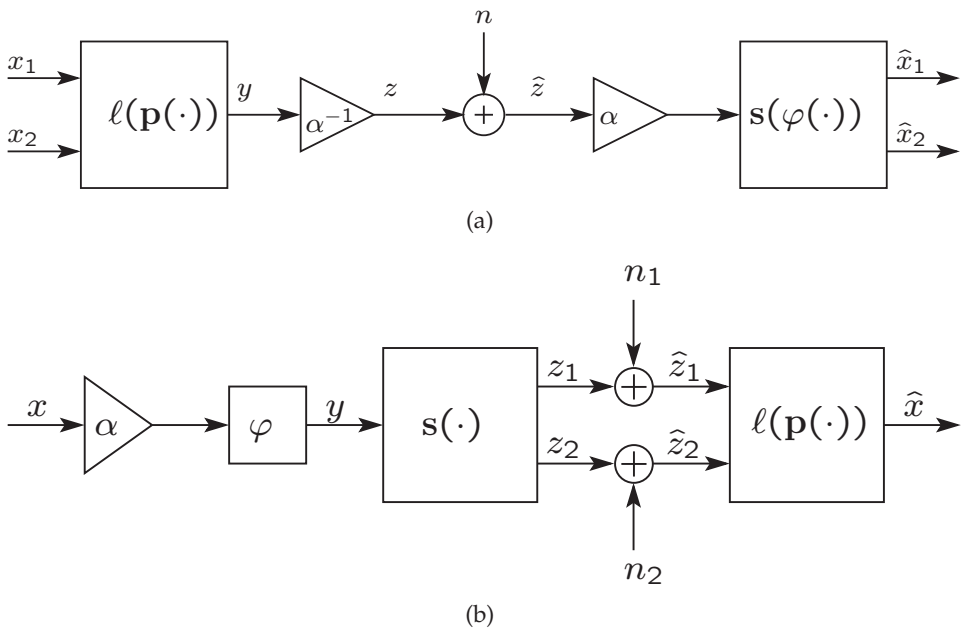


FIGURE 4.1: (a): 2:1 dimension reducing SK-mapping. Two source samples are combined into one channel samples using a non linear map. (b): 1:2 dimension expanding system. One source sample is split onto two channel samples using a non-linear map.

4.1 2:1 dimension reduction with the Archimedes' Spiral

In this section a specific example of a 2:1 dimension reducing SK-mapping for a Gaussian source, and an AWGN channel is given. A factor two dimension reduction, or compression, is achieved by combining two consecutive source samples using a parametric curve in the plane (almost the same analysis was done in [Hekland *et al.*, 2005]. The difference is that here the theory in section 2.4 is used to derive the same results).

As mentioned in section 2.6, no known method for determining the optimal geometry for an SK-mapping exists. Therefore a mapping whose shape resembles the codebook of a 2:1 PCCOVQ is chosen (see Figure 1.3(b) in section 1.2.4). The codebook has a similar shape as a double-intertwined Archimedes' spiral, and so it seems natural to use it as as a 2:1 mapping (in general an Archimedean spiral which will be discussed at the end of this chapter). The double Archimedes' spiral seems to fulfill the requirements for good mappings stated in section 1.2.5. With an average power con-

straint the spiral would cover the source distribution nicely. Furthermore, the most probable source samples would be mapped to small channel amplitudes (if the origin of the channel space corresponds to the "origin" of the spiral), and close neighbors in the channel space will always come from close neighbors in the source space. Thus, the Archimedes' spiral is a suitable choice as a mapping for this specific source and channel pair.

The communication chain considered is shown in Figure 4.1(a). Assume a Gaussian source $\mathbf{x}$ with zero mean. Dimension reduction is performed by transmitting a combination of two source samples x_1 and x_2 as one channel sample y. First the vector $\mathbf{x}$ is projected down on one of the two spirals shown in Figure 4.2. Then the 1D representation of the point on the spiral is mapped through the measuring function ℓ (see section 2.4) onto the channel. The channel signal is corrupted by AWGN. Detection is done by calculating the vector $\hat{\mathbf{x}}$ corresponding to the received value $\hat{z} = z + n$ using the equation of Archimedes' spiral given by (φ contained in $\mathbf{s}$ for simplicity)

$$\mathbf{s}(\hat{z}) = \pm\frac{\Delta}{\pi}\varphi(\alpha\hat{z})\bigg(\cos(\varphi(\alpha\hat{z}))\mathbf{i} + \sin(\varphi(\alpha\hat{z}))\mathbf{j}\bigg), \tag{4.1}$$

where α is a gain factor, φ is a stretching function (see section 2.4), Δ is the (radial) distance between the two spiral arms, $\mathbf{i}$ and $\mathbf{j}$ are orthogonal unit vectors pointing along the positive x_1- and x_2 directions respectively. The equation with positive sign corresponds to positive channel values (solid blue curve in Figure 4.2), while the equation with negative sign corresponds to negative channel values (dashed red curve in Figure 4.2), thus creating a channel representation with no DC components, and the most probable source symbols are represented with the lowest amplitudes. The functions ℓ and φ should be chosen in a convenient manner. The vector projected onto the spiral is first given a 1D representation as the radial distance ρ from the origin out to the given point. Then the $\ell(\cdot)$ operation is performed, mapping from the radius ρ onto the channel. In this case it is chosen to use the square of the radius times a constant

$$y = \ell_{\pm}(\rho) = \pm\zeta\left(\frac{\pi}{\Delta}\right)^2\rho^2 = \pm\zeta\left(\frac{\pi}{\Delta}\right)^2(x_1^2 + x_2^2), \tag{4.2}$$

where '$+$' represents positive channel values and '$-$' represents negative. The rationale behind choosing this representation is that the parameter $\zeta = \eta\Delta = 0.16\Delta$ makes this ℓ-operator an approximation to the length along the spiral. This choice of mapping gives a *unit speed parametrization* (see Appendix A.3), in which the velocity vectors along the curve is of unit norm. This again will make the channel distortion independent of the channel input signal levels (as well as σ_x). The expression in (4.2) is found by

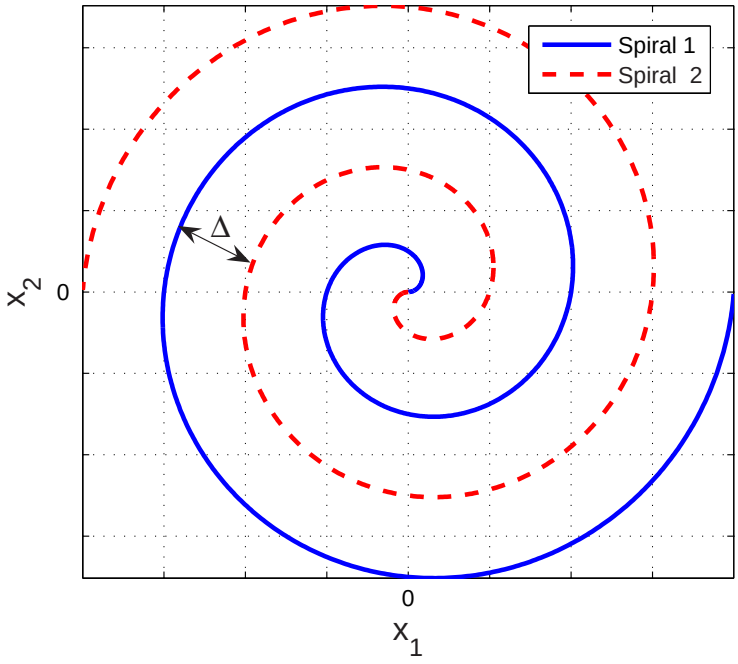


FIGURE 4.2: Archimedes' spiral.

using a nonlinear curve fit to the expression of the true arc length

$$\ell(\rho)_s = \frac{1}{2}\left(\rho\sqrt{1+\left(\frac{\pi}{\Delta}\rho\right)^2} + \frac{\Delta}{\pi}\sinh^{-1}\left(\frac{\pi}{\Delta}\rho\right)\right) \tag{4.3}$$

with the 2'nd degree function in (4.2) as input. The approximation of the true arc length will be quite accurate for large intervals, and therefore a good approximation at high CSNR. Near the origin, this expression will differ from the actual curve-length function. The reason for using this approximative operator instead of using the expression for the true spiral length is that the true arm length (4.3) does not have an analytic inverse. Further, one can find the channel symbol distribution of Y analytically using this approximation something which simplifies the calculation of the resulting channel symbol power. To determine the power of the channel input signal σ_y^2 , it is assumed that the spiral is sufficiently dense to disregard the approximation operation $\mathbf{q}$ in the calculations of it (high CSNR). Then the analysis in Appendix D can be used to find both the channel pdf and its corresponding variance. Considering a 2D source in Appendix D

(substitute $M = 2$) results in a Laplacian distribution with variance

$$\sigma_y^2 = 2\left(2\zeta\frac{\pi^2}{\Delta^2}\sigma_x^2\right)^2, \tag{4.4}$$

where $\zeta = 0.16\Delta$ as in (4.2).

The only parameter to determine in this system is the optimal Δ (which must be found for a given CSNR). To do so a mathematical model is needed, which can be found by using the theory from section 2.4. The distortion after decoding is $D_t = \bar{\varepsilon}_a^2 + \bar{\varepsilon}_{ch}^2$, where $\bar{\varepsilon}_a^2$ is the approximation distortion and $\bar{\varepsilon}_{ch}^2$ is the channel distortion (both defined in section 2.4). The channel input signal y is scaled by a factor $1/\alpha$ in order to satisfy an average power constraint. Thus the channel power is given by

$$P = E\{z^2\} = E\{(y/\alpha)^2\} = \frac{\sigma_y^2}{\alpha^2}, \tag{4.5}$$

and so $\alpha = \sqrt{\sigma_y^2/P}$. In the receiver the channel signal is re-scaled by α. Since the Archimedes' spiral can be seen as a uniform structure (except close to the origin) the approximation distortion is given by (2.48) in section 2.4.2, substituting $M = 2$ and $N = 1$. Further, since a unit speed parametrization is used, the metric component $g_{11} = \|\mathbf{s}'(z)\|^2 = \alpha^2,\ \forall z$ (a shape preserving map) and so the channel distortion is given by (2.45) substituting $M = 2$ and $N = 1$, and so the total distortion is

$$D_t(\Delta) = \bar{\varepsilon}_a^2(\Delta) + \bar{\varepsilon}_{ch}^2(\Delta) = \left(\frac{\Delta^2}{24}\right) + \frac{\sigma_n^2\sigma_y^2}{2P}, \tag{4.6}$$

which corresponds to the total distortion determined in in [Hekland *et al.*, 2005]. The (unconstrained) optimization of this problem is done in [Hekland *et al.*, 2005] and also [Hekland *et al.*, 2008], and the resulting optimal Δ is

$$\Delta_{opt} = 2\pi\sigma_x\sqrt[4]{\frac{6\cdot 0.16^2}{\mathrm{CSNR}}}. \tag{4.7}$$

Further it is also shown that $\bar{\varepsilon}_a^2 = \bar{\varepsilon}_{ch}^2$ at the optimal point. Some further discussion on the proposed 2:1 mapping is given in section 4.3.1.

4.2 1:2 dimension expansion with the Archimedes' Spiral

The communication chain considered in this case is shown in Figure 4.1(b). The source is scaled by α then mapped onto the channel space through the

signal curve $\mathbf{s}$. φ is a stretching function as defined in section 2.1. At the receiver the signal has to be projected back onto the curve (ML-detection) and re-scaled. The different operations in this system are basically the same as in the 2:1 case, except that the projection operation $\mathbf{p}$ does not introduce any approximation distortion.

In chapter 3, specific examples on 1:2 bandwidth expanding PCCOVQ were presented. From Figure 3.4 one can see that these mappings have a segmented spiral like structure. As was discussed, this segmented structure shows up in order to minimize the error done when the threshold effect first occurs. Trying to replicate this structure by a parametrization is possible to some extent, but the equation will get a bit complicated. Therefore it is chosen to use the Archimedes' spiral in (4.1). This spiral will fill the power constrained channel space properly, but the anomalous errors will be more severe than for the PCCOVQ, since when jumps from one spiral arm to another occur, positive source samples will be decoded as negative source samples and vice versa. To reduce the probability of the threshold effect, larger distances between two parts of the curve (Δ) compared to the PCCOVQ is required, something which will limit the total spiral length. As will be demonstrated at the end of this chapter, the effect of the severe anomalous distortion can be reduced by using a nonuniform spiral.

The source will reside along the double spiral in Figure 4.2 (substituting z_1 for x_1 and z_2 for x_2). Positive source values lie on the solid spiral ($\mathbf{s}_1$) and negative values lie on the dashed spiral ($\mathbf{s}_2 = -\mathbf{s}_1$). The only difference from (4.1) is that the parameter will be x instead of $\hat{z}$.

To be able to find the value of Δ that maximizes the performance (SDR) for a given CSNR, the channel power P, the weak noise distortion $\bar{\varepsilon}^2_{wn}$, and the anomalous distortion $\bar{\varepsilon}^2_{th}$ (both defined in section 2.1) needs to be calculated. Similar calculations have been done before in [Thomas, May, and Welti, 1975], using semicircles instead of the Archimedes' spiral. Some of the calculations here will be based on their work.

φ is chosen equal to the inverse curve length approximation

$$\varphi(x) = \pm\sqrt{\frac{x}{\eta\Delta}}. \tag{4.8}$$

The reason for this choice is the same as in Section 4.1 (with velocity vectors of equal length along the whole curve, the signal and noise are mutually independent). This approximation is used in both calculations and simulations. Assume a Gaussian input signal, truncated in the interval $[-1, 1]$ (the curve must have a finite length). By choosing the signal standard deviation small enough (like $\sigma_x = 0.25$), the overload effect will be negligible, and the signal can be considered Gaussian.

Channel power:The channel power is given by (per channel sample)

$$P = \frac{1}{2}E\{\|\mathbf{z}\|^2\} = \frac{1}{2}\int_{-1}^{1}\|\mathbf{s}(x)\|^2 f_x(x)dx = \frac{\alpha\Delta\sigma_x}{\eta\sqrt{2\pi^5}}\left(1 - e^{-\frac{1}{2\sigma_x^2}}\right). \tag{4.9}$$

Weak noise distortion: With a unit speed parametrization, $\|\mathbf{s}'(x_0)\|^2 = \alpha^2$, and so the weak noise distortion will be given by (2.16) (shape preserving mapping)

$$\bar{\varepsilon}_{wn}^2 = \frac{\sigma_n^2}{\alpha^2}. \tag{4.10}$$

Anomalous distortion: This contribution is calculated according to Figure 4.3. Due to symmetry it is sufficient to calculate the anomalous distortion for one of the spiral arms, then multiplying the whole result by 2.

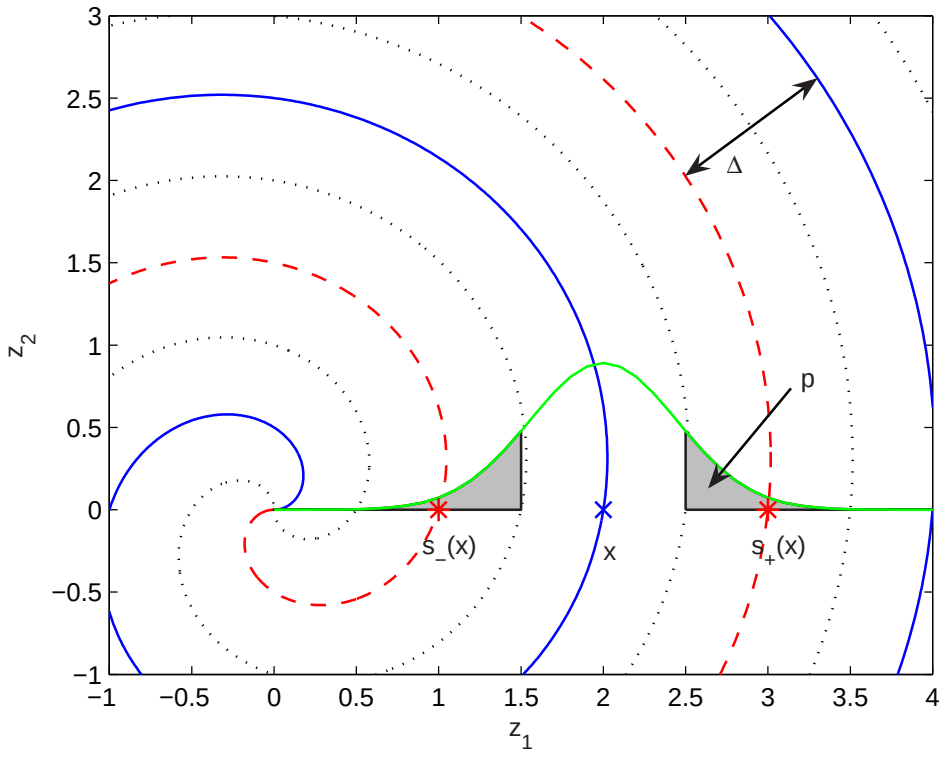


FIGURE 4.3: Calculation of the threshold distortion. The dotted line shows the decision spirals. The probability p for taking e.g. s_+ to be the received value is given by the shaded area outside the decision spiral.

Considering the positive spiral arm, the anomalous distortion can be found from

$$\bar{\varepsilon}_{th}^2 = 2p_r \int_0^1 [(x-\hat{x}_+)^2 + (x-\hat{x}_-)^2] f_x(x)dx, \tag{4.11}$$

where erf(·) denotes the *error function* [Bateman, 1953, p.266], $\hat{x}_\pm$ is the reconstructed values when the threshold effect occur and

$$p_r = \frac{1}{2}\left(1 - \text{erf}\left(\frac{\Delta}{2\sqrt{2}\sigma_n}\right)\right), \tag{4.12}$$

is the probability for the threshold effect. The value one get when the threshold effect first occurs is the one corresponding to the point on the curve in radial distance Δ from the wanted point. To find the two reconstructed values $\hat{x}_\pm$, it is convenient to do the calculations in polar coordinates. Considering transmission on the positive spiral arm the problem is to solve (noise taking us outwards compared to the spiral origin)

$$s_+ = s(x) + \Delta \tag{4.13}$$

$$-\frac{\Delta}{\pi}(\varphi(\alpha\hat{x}_+)) = \frac{\Delta}{\pi}\varphi(\alpha x) + \Delta, \tag{4.14}$$

and so

$$\hat{x}_\pm = -\eta\Delta\left(\sqrt{\frac{x}{\eta\Delta}} \pm \frac{\pi}{\sqrt{\alpha}}\right)^2, \tag{4.15}$$

where the '$-$' sign is added at the end to take into account that the detected x will be negative when the threshold effect occur (when the transmitted x was positive). $\hat{x}_-$ is the closest received value when the noise takes us inwards across the decision boundary to another part of the curve, and $\hat{x}_+$ is the received value the noise takes us outwards. In general, the probability for the threshold effect p_r should be inside the integral in (4.11). But with a uniform structure, p_r will be independent of x except for a little region close to the origin, which gives a negligible contribution. Doing the integration in (4.11) gives

$$\begin{aligned}\bar{\varepsilon}_{th}^2 =& \frac{1}{\sqrt{\pi}\alpha^2}\left(1 - \text{erf}\left(\frac{\Delta}{2\sqrt{2}\sigma_n}\right)\right)\left[\text{erf}\left(\frac{\sqrt{2}}{2\sigma_x}\right)\left(4\sigma_x^2\sqrt{\pi}\alpha^2 + \eta^2\Delta^2\pi^{\frac{9}{2}}\right)\right. \\ &\left. - 4\sqrt{2}\sigma_x\alpha e^{-\frac{1}{2\sigma_x^2}}(\alpha + 2\eta\Delta\pi^2) + 8\sqrt{2}\eta\pi^2\Delta\sigma_x\alpha\right].\end{aligned} \tag{4.16}$$

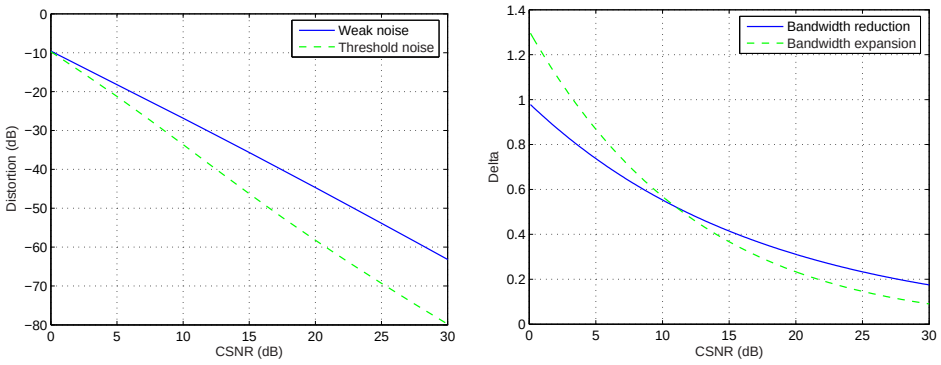
Optimization of the 1:2 mapping: To determine the optimal Δ for a given CSNR, first solve (4.9) with respect to α

$$\alpha = \frac{P\eta\sqrt{2\pi^5}}{\Delta\sigma_x\left(1 - e^{-\frac{1}{(2\sigma_x^2)}}\right)}, \tag{4.17}$$

then substitute this into (4.10) and (4.16) (to reduce the optimization to an unconstrained problem). The anomalous distortion decreases with increasing Δ, whereas the weak noise distortion increases, which implies that there exists a unique global minimum. The following unconstrained optimization problem must be solved

$$\Delta_{opt} = \arg\min_{\Delta} \left[\bar{\varepsilon}^2_{wn}(\Delta) + \bar{\varepsilon}^2_{th}(\Delta)\right]. \tag{4.18}$$

Since the expression for the anomalous distortion (4.16) contains the error function, no analytical expression for Δ_{opt} can be found. Instead a numerical optimization algorithm (Matlab's optimization toolbox) is used. Figure 4.4(a) shows the two distortion contributions at the optimal point as a function of the CSNR. Observe that for large CSNR, the contribution from



(a) (b)

FIGURE 4.4: Results from optimization. (a) The two distortion contributions in a 1:2 bandwidth expanding system, shown at the optimal Δ for a range of CSNR values. The optimization is done for $\sigma_x = 0.25$, and channel power $P = 1$. (b): Comparison of the optimal Δ as a function of CSNR for the 1:2 and 2:1 case. The graphs are made for $\sigma_x = 0.25$ and $\sqrt{P} = 0.25$.

the anomalous distortion is very small compared to the weak noise distortion at the optimum point (this is to be expected since the anomalous distortion is severe). This is in contrast to the 2:1 case, where the two distortion contributions is equal at the optimum point.

It is convenient to find an approximate function for Δ_{opt}. This can be done by using a nonlinear least-squares curve fitting algorithm (Gauss-Newton or Levenberg-Marquardt algorithm [Nocedal and Wright, 1999,

pp.259-270]), for the following function

$$\Delta_{opt} = \sqrt{P}\beta e^{-(\gamma \mathrm{CSNR}_{\mathrm{dB}}^2 + \vartheta \mathrm{CSNR}_{\mathrm{dB}})}. \tag{4.19}$$

The coefficients have to be determined when the source standard deviation is given. With $\sigma_x = 0.25$, the coefficients $\beta = 5.223$, $\gamma = 3 \cdot 10^{-4}$ and $\vartheta = 0.0801$ emerge. This function lies very close to the real Δ_{opt} in the range 0-35dB CSNR. Figure 4.4(b) shows Δ_{opt} as a function of CSNR for both bandwidth compression and expansion when using Archimedes' spirals. The two curves seem to have similar characteristics.

4.3 Results and discussion

In this section simulation results are presented and compared with the theoretical findings from section 4.1 and 4.2. In the simulation $5 \cdot 10^4$ Gaussian distributed test vectors are used for bandwidth reduction. For bandwidth expansion it is necessary to use at least $2 \cdot 10^5$ Gaussian test samples. This is because the threshold effect must occur enough times to give consistent results.

4.3.1 2:1 Bandwidth Reduction

The performance of the system is expressed in terms of the received SDR as a function of the CSNR. By definition CSNR $= P/\sigma_n^2$ and SDR $= \sigma_x^2/D_t$, where $P = \sigma_y^2/\alpha^2$, σ_y^2 is given by (4.4) and D_t is given by (4.6). As one can observe from Figure 4.5 there is a close correspondence between calculations and simulations above 30 dB CSNR. Below 30 dB, however, the results start to differ. For CSNR above the optimum point (marked with a star), one can observe that the curve representing the simulated performance is shifted upwards relative to the calculated performance. This is because the actual approximation distortion near the spiral origin is smaller than $\Delta^2/24$. But still the optimum point from the simulation is further away from OPTA than the calculated results. The reason is that an approximation to the curve length function (4.2) is used. The derivative of the inverse curve-length approximation has a vertical asymptote at the origin, as opposed to the real inverse curve-length derivative. Then the norm of the spiral tangent vectors will get larger close to the origin than what was assumed in the theoretical analysis. This makes the channel distortion larger on average for low CSNR compared to the theoretical analysis, and the simulated optimal points will get further away from OPTA. For a fixed Δ one can also see a difference between calculations and simulations far from the

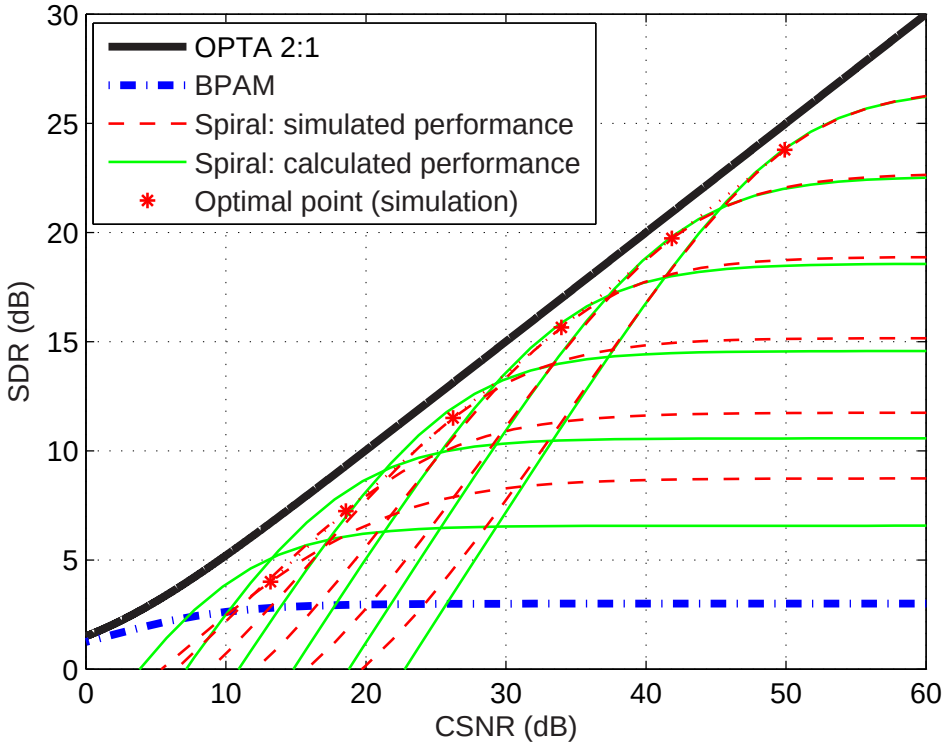


FIGURE 4.5: Performance of a 2:1 dimension reducing system using Archimedes' spiral. Approximate calculations (thin green lines), simulation (dashed red lines) together with BPAM (blue dash-dot) are shown. The different curves correspond to different values of Δ_{opt}. Each curve is made from varying noise variances so that one can see the behavior described in Section 2.4.

optimal point for decreasing CSNR. The reason is that the channel distortion is calculated by using the linear approximation to the signal curve.

4.3.2 1:2 Bandwidth Expansion

The received fidelity is now given by SDR $= \sigma_x^2/(\bar{\varepsilon}_{wn}^2 + \bar{\varepsilon}_{th}^2)$, where $\bar{\varepsilon}_{wn}^2$ and $\bar{\varepsilon}_{th}^2$ are given by (4.10) and (4.16), respectively. Figure 4.6 shows both the calculated and simulated performance of the 1:2 dimension expanding system using Archimedes' spirals, together with an optimal linear system (BPAM). The system is robust in the sense that it is not highly sensitive to incorrect choice of Δ_{opt}, i.e. it shows graceful improvement, and to some extent graceful degradation (especially at low CSNR). The curves optimized

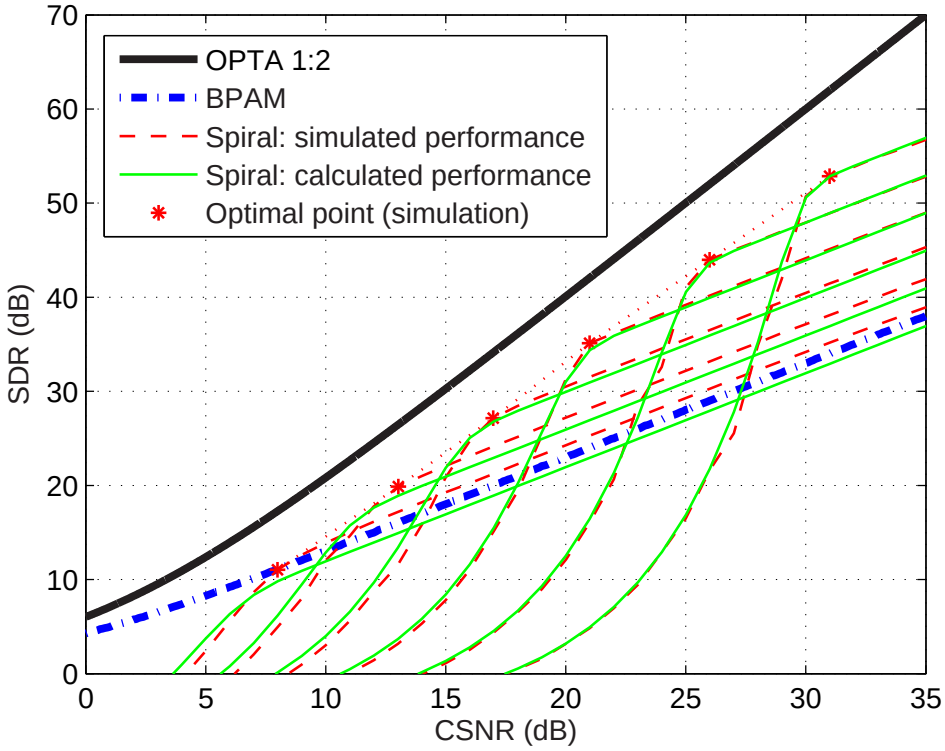


FIGURE 4.6: Performance of a 1:2 dimension expanding system using Archimedes' spiral. Approximate calculations (thin green lines), simulation (dashed red lines) together with BPAM (blue dash-dot) are shown. The different curves correspond to different values of Δ_{opt}. Each curve is made from varying noise variances so that one can observe the behavior described in Section 2.1.

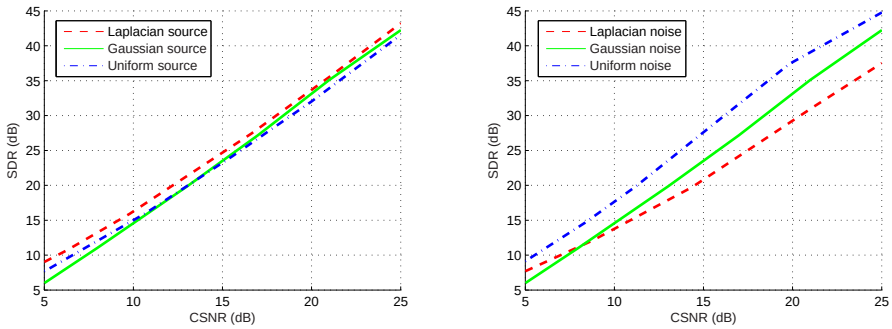
for larger CSNR break down more steeply, although the breakdown is not total, as is the case of traditional channel coders. Notice that there is a considerable gain compared to a linear system except below $8-10$ dB CSNR. Also observe the close correspondence between calculations and simulations above 20 dB CSNR. As in the 2:1 case the difference between the simulated and calculated performance increase when CSNR gets lower. Here the optimal point is closer to OPTA for the simulated case. The explanation for this is similar to the 2:1 case in the previous section, but in this case the longer tangent vectors will reduce the effect of the weak noise distortion, and so the simulated results gets closer to OPTA than the calculated.

4.4 Experiments

In this section the proposed systems are tested against mismatch in both source and channel noise pdf's. Further the proposed systems are modified, making them preform slightly better.

4.4.1 pdf mismatch

To get indications on how robust the example mappings given in section 4.1 and 4.2 are for mismatch in both source- and channel noise distributions, they have been simulated with both Laplacian and uniform distributions for both source and channel noise. The results are shown in Figure 4.7 for the expansion case, and in Figure 4.8 in the compression case. As seen

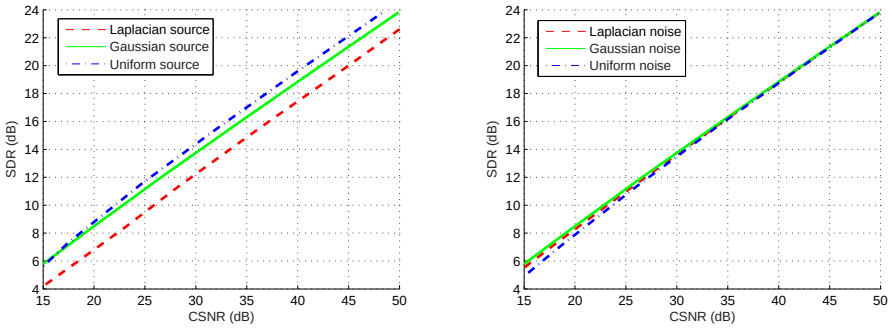


(a) (b)

FIGURE 4.7: (a) Comparison between Gaussian, Laplacian and uniform source in the 1:2 case. (b) Comparison between Gaussian, Laplacian and uniform noise in the 1:2 case.

from figure 4.7(a), the loss/gain is quite insignificant having mismatch in the source pdf in the 1:2 case. The sensitivity towards mismatch in the noise pdf is quite significant though. From figure 4.7(b) one can observe that there is a significant loss in the Laplacian case, whilst a significant gain in the uniform case. The reason for the loss in the Laplacian case is that the Laplacian pdf is "much wider" than the Gaussian pdf. So for a given CSNR, Δ must be larger in the Laplacian case in order to get the same probability for the anomalous errors. Then the curve has to be made shorter to satisfy the channel power constraint, which makes the received SDR go down, since the weak noise distortion gets larger (see equation (2.5)). In the

Uniform noise case the opposite effect occurs. Since Δ can be made smaller for a given CSNR, the received SDR goes up.



(a) (b)

FIGURE 4.8: (a) Comparison between Gaussian, Laplacian and uniform source in the 2:1 case. (b) Comparison between Gaussian, Laplacian and uniform noise in the 2:1 case.

In contrast to the expansion case, the compressive mapping is quite sensitive to source pdf mismatch. From figure 4.8(a) one can observe a quite significant loss in the Laplacian case, whilst some gain in the uniform case. The Laplacian source has a much "wider pdf" than the Gaussian. So given a certain CSNR, the curve must be made longer in the Laplacian case to get the same approximation distortion as in the Gaussian case. A longer curve makes the channel distortion go up (according to equation (2.44)), and so the performance deteriorates. In the uniform case the opposite effect occurs. From figure 4.8(b) one can observe that the sensitivity towards noise pdf mismatch is quite insignificant in the 2:1 case.

From these results it seems like the proposed systems are quite robust against pdf mismatch. The performance seems to gradually decrease or increase with the width of the pdf. The sensitivity towards pdf mismatch is largest in the space where the signal curve resides (source pdf in the compression case and channel noise pdf in the expansion case). The performance is gradually decreasing as the pdf widens, and gradually increasing as the pdf becomes narrower (compared to the Gaussian case).

4.4.2 Archimedean Spiral

In the previous sections only uniform spirals were considered. Using a non-uniform spiral could be beneficial. Indeed, looking at the code-book

for the 2:1 PCCOVQ in Figure 1.3(b) in section 1.2.4, one can see that the spiral is slightly nonuniform. Also in the expansion case, considering the anomalous distortion, it is beneficial that the distance between the spiral arms increase for larger values of the source. This will make the probability for the threshold effect go down for the most severe cases and somewhat compensate for the non-segmented structure chosen (compared to the 1:2 PCCOVQ in Figure 3.4). Archimedean spirals are used in the examples. This spiral is a generalization of Archimedes' spiral, which can be seen from its equation

$$\mathbf{s}(z) = \pm\frac{\Delta}{\pi}(\varphi(\alpha z))^{\frac{1}{k}}\Big(\cos(\varphi(\alpha z))\mathbf{i} + \sin(\varphi(\alpha z))\mathbf{j}\Big). \tag{4.20}$$

k is a factor determining if the spiral is uniform ($k = 1$) or if the distance between the spiral arms increase ($k < 1$) or decrease ($k > 1$) with increasing parameter values.

2:1 **case**: First an Archimedean spiral with $k = 0.9$ is tested (at both transmitter and receiver). φ is again chosen to be the curve length approximation. From Figure 4.9 one can observe that there is a gain compared to the uniform spiral above 30dB CSNR for the $k = 0.9$ system. At 45 dB CSNR, the distance to OPTA is approximately 0.95 dB compared to 1.1 dB. Below 20 dB CSNR however, the nonuniform system with $k = 0.9$ is inferior to the uniform system. Two curves are plotted in addition which outperform the other systems by $0.5 - 1$ dB for low CSNR. Here different spirals are used at the transmitter and receiver, which means that the receiver no longer is the inverse of the transmitter. At the receiver a spiral where the distance between the spiral arms gets smaller the larger the channel signal amplitudes become is used. This means that all received vectors will be slightly moved inwards towards the origin compared to the transmitted vectors. This seems to have the same effect as a Wiener factor [Therrien, 1992] has for a linear coders. The polynomial $\ell(\rho) = \Delta(0.14\rho^2 + 0.55\rho)$ is used as a mapping function (it was found by choosing a 2'nd degree polynomial, and then coarsely tuning the coefficients until the "best" result emerged). The Archimedean spiral systems are also compared to a 2:1 PCCOVQ system using 4096 representation vectors. As one can see at low CSNR, one can get the spiral to perform on par with the PCCOVQ system by using different nonuniform spirals, whilst the spiral (both uniform and nonuniform) actually gets better than the PCCOVQ for high CSNR. The PCCOVQ would probably be on par with the spiral mapping using more representation vectors in the codebook.

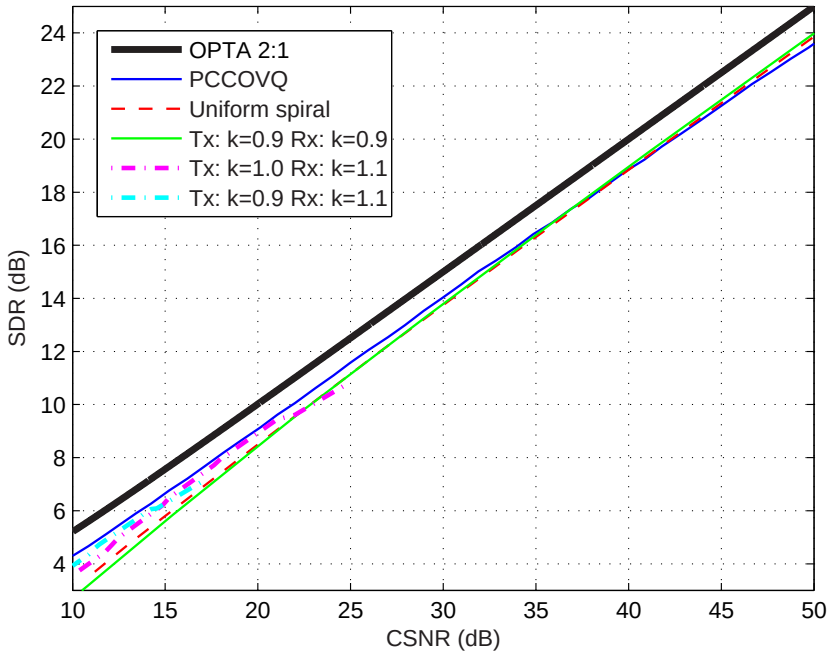


FIGURE 4.9: Comparison between different 2:1 systems, using uniform and nonuniform Archimedean spirals and PCCOVQ. In the plot Tx is denoting the transmitter, while Rx is denoting the receiver.

1:2 case: Here, different Archimedean spirals are tested for dimension expansion. φ is given by (4.8) in all simulations. The simulations in Figure 4.10 shows the comparison between the 1:2 uniform system from Figure 4.6 (dashed red line) and two systems using nonuniform Archimedean spirals together with a 1:2 PCCOVQ system (between 0-20 dB CSNR). Above 15 dB CSNR, the nonuniform system with $k = 0.8$ shows modest gains compared to the uniform system for the optimal choice of Δ, but is more gracefully degrading (compare with Figure 4.6). For CSNR below 15 dB, however, the gain compared to the uniform system is more significant, especially for the $k = 0.7$ system. The PCCOVQ slightly outperforms the spiral systems until about 17 dB CSNR, where it ends up slightly below the spiral systems (probably because the PCCOVQ found by the algorithm at such a high CSNR is a local minimum). In the 1:2 case it seems like one can

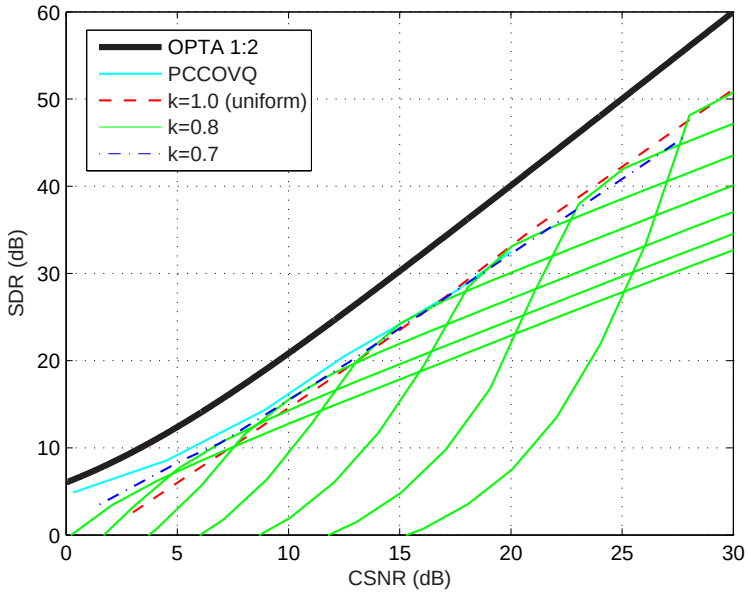


FIGURE 4.10: Comparison between different 1:2 systems, using uniform and non-uniform Archimedean spirals ($k = 0.8$ and $k = 0.7$) and PC-COVQ.

gain something in performance for all CSNR values using an Archimedean spiral with $k < 1$.

Chapter 5

Other mappings and experiments

In this chapter examples on several SK-mappings are given. First three dimension reducing mappings are introduced. The mappings chosen to be analyzed are 3:1 and 4:1 mappings based on parametric curves, and a 3:2 mapping based on a parametric surface. Second an example on a 2:3 dimension expanding mapping is given. This mapping is based on a vector quantizer (VQ) and linear coders. All mappings are modeled mathematically and optimized, yielding the optimal coefficients in their given equations as a function of the CSNR.

The reason why these specific mappings are chosen is given in section 1.2.5. Also, these mappings are simple enough (low enough dimensionality) to determine their approximate mathematical model.

5.1 Dimension reducing mappings

In this section three dimension reducing mappings are presented. First a 3:1 mapping found by generalizing the Archimedean spiral from chapter 4 is given. Then a 4:1 mapping found based on a generalization of the proposed 3:1 mapping. Both mappings are analyzed mathematically. Second a 3:2 mapping is presented and analyzed, illustrating some of the problems showing up going from parametric curves to parametric surfaces (hyper surfaces in general).

Figure 5.1 shows a general block diagram for reducing mappings that will be referred to in the following sections. The M dimensional source vector $\mathbf{x}$ is projected onto a signal hyper surface (or curve) $\mathcal{S}$ and mapped onto an N dimensional channel vector $\mathbf{y}$. ℓ has the same role as in section 4.1.

Each component of the channel vector, y_i, is scaled by $1/\alpha_i$ to satisfy a given power constraint. Then the received vector is re-scaled and mapped through the SK-mapping to reconstruct the transmitted source vector.

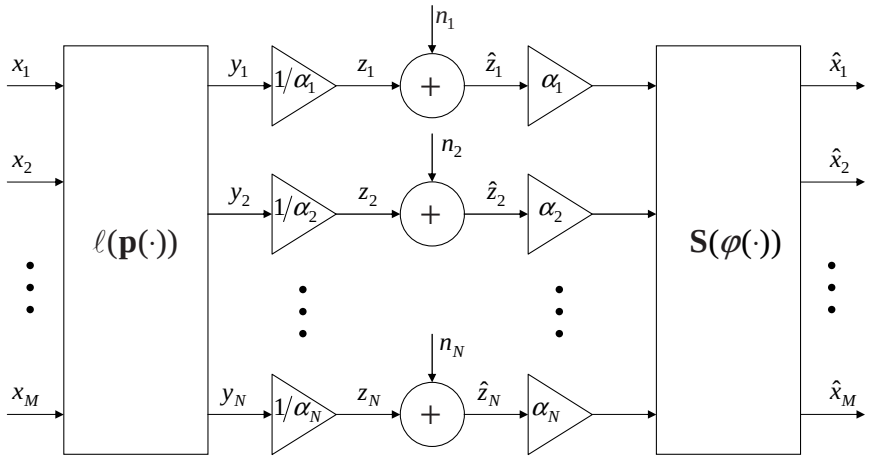


FIGURE 5.1: Block diagram for an M:N dimension reducing system.

5.1.1 3:1 Mapping

The block diagram under consideration in this section is given by Figure 5.1 with $M = 3$ and $N = 1$, and $\alpha_1 = \alpha$.

In section 4.1 it was argued that the spiral structure was chosen based on the codebook of a 2:1 PCCOVQ. Furthermore it was argued that it satisfied the conditions for a good mapping given in section 1.2.5. Since a 2D Gaussian signal has a circular symmetry it is also natural to choose a spiral as a 1D covering.

In the 3:1 case, similar arguments can be given. Since a 3D Gaussian signal has a spherical symmetry, one can naturally imagine a "ball of yarn"-like structure as a 1D cover, where the middle of the "thread", which is at the origin of the source space, corresponds to the origin of the channel space. Then, as for the spiral, the most probable symbols are represented with the lowest power, the points close in the channel space will also be close in the receivers source space guaranteing robustness, and the structure fill a spherical region in the source space in a (relatively) proper manner making the approximation distortion small. The last mentioned point is the most difficult one. One problem is that one has many more

choices of freedom on how to make a curve fill a ball in space then filling a disc in the plane. This makes it more difficult to find equations for well performing systems for a large range of CSNR. As will be seen later a 3:1 PCCOVQ structure (which is also like a ball of yarn) will outperform the proposed system, probably because the PCCOVQ is capable of placing points in space in a close to optimal way for each given CSNR, something which is very hard to do with a given equation.

To find the parametric equations for a ball of yarn like structure, one can start out with a tilted circle in $\mathbb{R}^3$

$$\mathbf{r}(z) = \cos(z)\mathbf{i} + \sin(z)\mathbf{j} + \sin(z)\mathbf{k}. \tag{5.1}$$

A spiral lying on a (tilted) plane shows up by multiplying the circle with the parameter z. Since the functions in the $\mathbf{j}$ and $\mathbf{k}$ directions are linearly dependent, the structure will not "twist" up in the last dimension. By multiplying by two linearly independent functions in these directions (the given functions are orthogonal over periods of 2π), the wanted structure show up. One equation, used for positive channel values, is

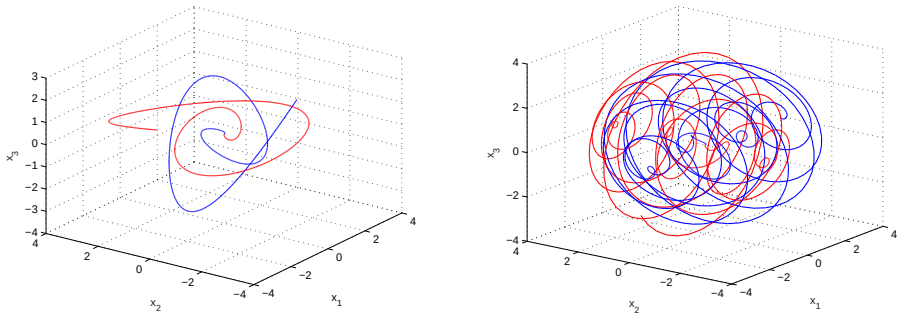
$$\begin{aligned}\mathbf{s}_+(z) =& \frac{\Delta}{\pi}\varphi(\alpha z)\Big(\cos(\varphi(\alpha z))\mathbf{i} + \cos\big(\frac{1}{2\pi}\varphi(\alpha z)^a\big)\sin(\varphi(\alpha z))\mathbf{j} + \\ & \sin\big(\frac{1}{2\pi}\varphi(\alpha z)^a\big)\sin(\varphi(\alpha z))\mathbf{k}\Big).\end{aligned} \tag{5.2}$$

Another configuration is chosen for negative channel values

$$\begin{aligned}\mathbf{s}_-(z) =& -\frac{\Delta}{\pi}\varphi(\alpha z)\Big(\sin\big(\frac{1}{2\pi}\varphi(\alpha z)^a\big)\cos(\varphi(\alpha z))\mathbf{i} + \sin(\varphi(\alpha z))\mathbf{j} + \\ & \cos\big(\frac{1}{2\pi}\varphi(\alpha z)^a\big)\cos(\varphi(\alpha z))\mathbf{k}\Big).\end{aligned} \tag{5.3}$$

Δ has a somewhat different meaning than in the 2:1 case (in that it is not directly giving the distance between two parts of the curve), but can still be seen as a scaling of the whole structure and can therefore be used to adapt to varying channel conditions. φ has the same role as in previous chapters. a is a factor that determines the "density" of the curve in a given region of space (see Figure 5.2). In the following (5.2) and (5.3) together are referred to as "system 1". Figure 5.2 depicts the structure in (5.2) (blue line) and (5.3) (red line).

In the following, an approximate mathematical model for the distortion using (5.2) and (5.3) as a 3:1 mapping assuming $a = 1$ is given. Later in this section different values of a and another structure will be tested and



(a) (b)

FIGURE 5.2: 3:1 dimension reducing mapping, named system 1. This is the structure corresponding to equation 5.2 (blue line) and 5.3 (red line) with two different values of a. (a) $a = 1$. (b) $a = 2.5$.

compared to a 3:1 PCCOVQ system and BPAM. For the same reasons as in the 2:1 case, a unit speed parametrization is chosen (to make the channel input signal and the noise mutually independent, and to simplify the calculations). In seems difficult to find an explicit expression for the curve-length function for the given structure. But by using numerical integration and a nonlinear curve fitting algorithm [Nocedal and Wright, 1999, pp.259-270] (Gauss-Newton or Levenberg-Marquardt), it happily turns out that the same approximation as in equation (4.2) in section 4.1 is good

$$y = \ell_{\pm}(\rho) = \pm\zeta\left(\frac{\pi}{\Delta}\right)^2 (x_1^2 + x_2^2 + x_3^2), \tag{5.4}$$

again with $\zeta = \eta\Delta$ where $\eta = 0.16$.

Channel power: It is shown in Appendix D that the pdf of the channel signal can be approximated (at high CSNR, so the approximation operation can be disregarded during the calculations) by a "double gamma distribution", which has the variance (substituting $M = 3$ in (D.10))

$$\sigma_y^2 = \frac{15(\pi^2\zeta\sigma_x^2)^2}{\Delta^4}. \tag{5.5}$$

And so the channel power will be $P_t = \sigma_y^2/\alpha^2$ (see Figure 5.1).

Channel distortion: (2.44) determines the channel distortion. Using a unit speed parametrization, then for the same reasons as in section 4.1, (2.45)

can be used substituting $M = 3$ and $N = 1$ (shape preserving mapping)

$$\bar{\varepsilon}_{ch}^2 = \frac{\sigma_n^2}{3}\alpha^2 = \frac{\sigma_n^2\sigma_y^2}{3P} = \frac{5(\sigma_n\sigma_x^2\pi^2\eta)^2}{\Delta^2 P} \tag{5.6}$$

Approximation distortion: This contribution can unfortunately not be found analytically due to the more irregular structure on the decision levels (compared to the 2:1 case). The approximation will in this case be a function of both Δ and σ_x. A numerical curve fitting algorithm must be used. Assume that the approximation distortion is on the form

$$\bar{\varepsilon}_a^2 = \beta_{3:1}\Delta^{2-\gamma}. \tag{5.7}$$

Given that $\sigma_x = 1$, the optimal parameters is found to be $\beta_{3:1} = 0.2104$ and $\gamma = 1.09$ by a nonlinear curve fitting algorithm. To simplify the analysis one can say that $\gamma_{3:1} = 1.0$ and find that $\beta =_{3:1} 0.258$ without changing the results significantly (except at lower CSNR).

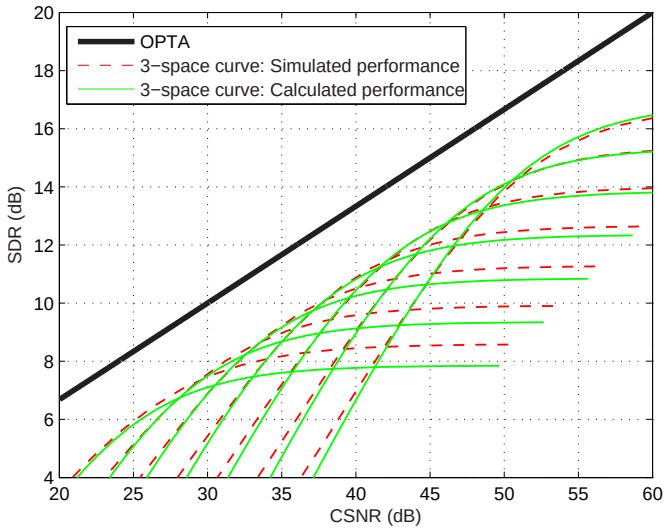
The total distortion is given by

$$D_t = \beta_{3:1}\Delta + \frac{5(\sigma_n\pi^2\eta)^2}{\Delta^2 P}, \quad \sigma_x = 1. \tag{5.8}$$

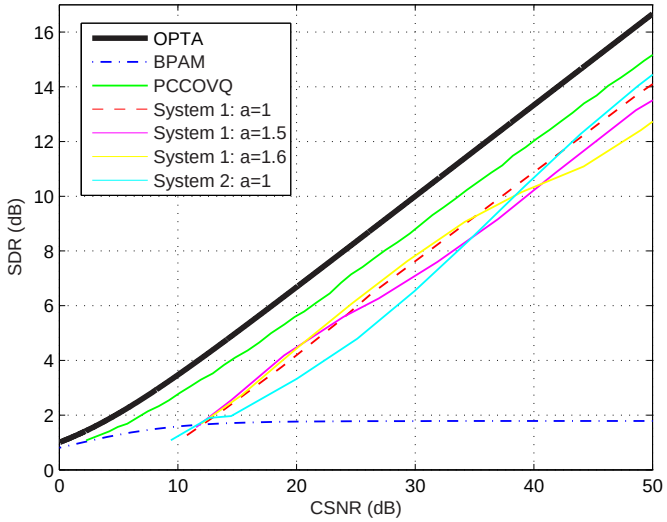
Differentiating (5.8) with respect to Δ, setting the equation equal to zero and solving with respect to Δ gives

$$\Delta_{opt} = \sqrt[3]{\frac{10(\sigma_n\pi^2\eta)^2}{\beta_{3:1}P}} = \sqrt[3]{\frac{10(\eta\pi^2)^2}{\beta_{3:1}\mathrm{CSNR}}}, \quad \sigma_x = 1. \tag{5.9}$$

The comparison between the above theoretical model and simulations of the real system is shown in Figure 5.3(a). As one can see there is quite a good correspondence between the theoretical model and simulations at high CSNR (above 40 dB). At lower CSNR the results start to differ. The reason is partly the same as explained for the 2:1 case in section 4.3.1 (the discrepancy between the derivatives of the curve length function and its approximation close to the origin, and the accuracy in using the curve length approximation for low CSNR) and that the model for the approximation noise was made to fit best at high CSNR (small values of Δ). Considering a specific graph (fixed Δ) at low CSNR it seems apparent that the approximation distortion is deviating the most from from the simulated performance (part of graphs above the optimal point). Fitting the data for the approximation distortion to a more complicated function would probably make the theoretical model coincide better with the simulations at low CSNR.



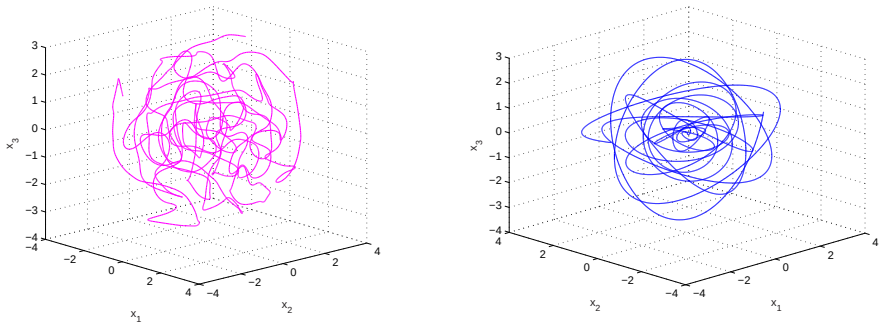
(a)



(b)

FIGURE 5.3: Performance of 3:1 dimension reducing systems. (a) Comparison between the theoretical model and the simulation of system 1. (b) Comparison between different 3:1 systems.

In Figure 5.3(b) several 3:1 systems are compared. One can observe that changing the "density" of system 1 (increasing a) makes it perform slightly better for a certain range of CSNR values, but then much worse for other ranges (compared to $a = 1$). At high CSNR a different structure is tested, called "system 2". The equations of this mapping is the same as (5.2) for positive channel values, and the negative of (5.2) for negative channel values. In all cases an approximate unit speed parametrization is used. Furthermore one can observe that the PCCOVQ system significantly outperforms all the proposed systems. This is probably due to the fact, as mentioned earlier, that the PCCOVQ has a higher degree of freedom in placing points in space, not being stuck to a specific structure.



(a) (b)

FIGURE 5.4: Comparison between PCCOVQ (a) and system 1 (b).

Figure 5.4 shows the comparison between the structures of the PCCOVQ and system 1 with $a = 1$ at 40 dB CSNR. From this figure it is obvious that the PCCOVQ should have a smaller approximation distortion than the parametric curve (covering the space more properly). Unfortunately, the PCCOVQ structure seems too irregular to be modeled by a simple equation, but there might be possible to do a nonlinear curve fitting to the component functions (corresponding to the component functions S_i in a parametrization like (2.6)) of the PCCOVQ showing up when interpolating between the representation vectors for a given CSNR.

The difference between the Archimedes spiral and PCCOVQ in the 2:1 case was not significant. This is probably due to the fact, as mentioned before, that there are not so many choices on how to fill a disc-like region with a curve if one at the same time should obey the three conditions given in section 1.2.5. Filling a ball-like region in space with a curve gives rise

to much more flexibility. The advantages of the proposed SK-mapping(s) over the PCCOVQ is that it is possible to find an approximate mathematical model, which makes it much simpler to adapt to varying channel conditions (i.e. by (5.9) or a more complicated expression).

5.1.2 4:1 mapping

Here an example of a 4:1 mapping using a parametric curve is given. The block diagram under consideration in this section is given by Figure 5.1 with $M = 4$ and $N = 1$, and $\alpha_1 = \alpha$.

Since it is impossible to visualize a structure in 4D, the only thing one can do is to generalize further what was done in section 5.1.1. The following equations are chosen to be analyzed. (5.10) represents positive channel values

$$\begin{aligned}\mathbf{s}_+(z) =&\frac{\Delta}{\pi}\varphi(z)\Big(\sin\big(\frac{1}{3\pi}\varphi(z)^a\big)\cos(\varphi(z))\mathbf{i} + \cos\big(\frac{1}{2\pi}\varphi(z)^a\big)\sin(\varphi(z))\mathbf{j}+\\ &\sin\big(\frac{1}{2\pi}\varphi(z)^a\big)\sin(\varphi(z))\mathbf{k} + \cos\big(\frac{1}{3\pi}\varphi(z)^a\big)\cos(\varphi(z))\mathbf{l}\Big).\end{aligned} \tag{5.10}$$

Negative channel values are represented by

$$\begin{aligned}\mathbf{s}_-(z) =&-\frac{\Delta}{\pi}\varphi(z)\Big(\sin\big(\frac{1}{3\pi}\varphi(z)^a\big)\cos(\varphi(z))\mathbf{i} + \sin\big(\frac{1}{2\pi}\varphi(z)^a\big)\sin(\varphi(z))\mathbf{j}+\\ &\cos\big(\frac{1}{3\pi}\varphi(z)^a\big)\cos(\varphi(z))\mathbf{k} + \cos\big(\frac{1}{2\pi}\varphi(z)^a\big)\sin(\varphi(z))\mathbf{l}\Big).\end{aligned} \tag{5.11}$$

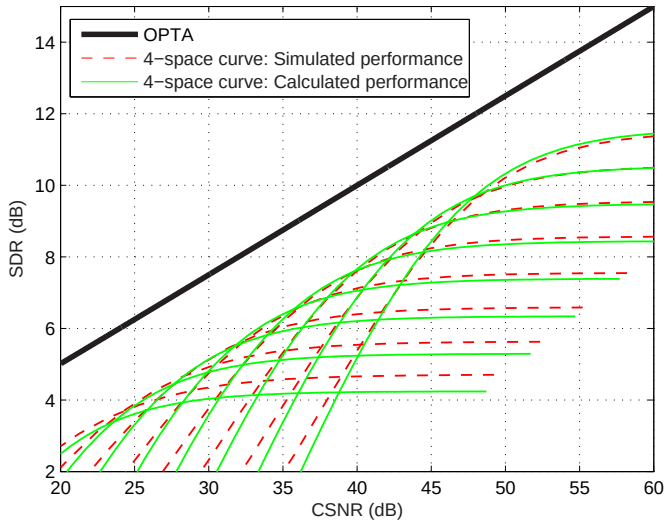
An approximate mathematical model is given for $a = 1$. A unit speed parametrization is chosen. Again, the same curve length approximation as for the 2:1 and 3:1 mapping can be used (with $\eta = 0.16$).

Channel power: The channel pdf and its variance can be found from Appendix D. Substituting $M = 4$ in (D.10) gives the variance of the channel signal

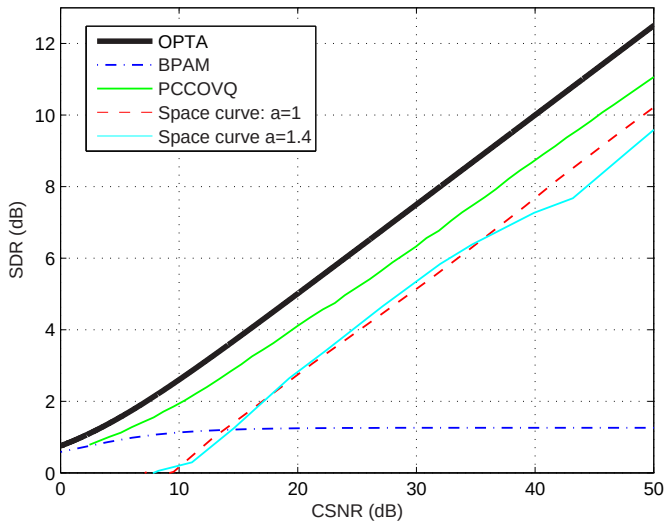
$$\sigma_y^2 = \frac{24(\pi^2\zeta\sigma_x^2)^2}{\Delta^4}. \tag{5.12}$$

Channel distortion: Using a unit speed parametrization, (2.45) determines the channel distortion substituting $M = 4$ and $N = 1$

$$\bar{\varepsilon}_{ch}^2 = \frac{\sigma_n^2}{3}\alpha^2 = \frac{\sigma_n^2\sigma_y^2}{4P} = \frac{6(\sigma_n\sigma_x^2\pi^2\eta)^2}{\Delta^2 P}. \tag{5.13}$$



(a)



(b)

FIGURE 5.5: Performance of 4:1 dimension reducing systems. (a) Comparison between the theoretical model and simulation. (b) Comparison between different 4:1 systems.

Approximation distortion: As for the 3:1 case this contribution can not be found analytically. The approximation will, as for the 3:1 case, be a function of both Δ and σ_x. A similar procedure as in section 5.1.1 gives (for $\sigma_x = 1$)

$$\bar{\varepsilon}_a^2 = \beta_{4:1}\Delta^{\frac{7}{10}}, \tag{5.14}$$

where $\beta = 0.4061$. The total distortion is given by

$$D_t = \beta_{4:1}\Delta^{\frac{7}{10}} + \frac{6(\sigma_n\pi^2\eta)^2}{\Delta^2 P}, \quad \sigma_x = 1. \tag{5.15}$$

Differentiating (5.15) with respect to Δ, equating to zero and solving with respect to Δ gives

$$\Delta_{opt} = \left(\frac{120(\sigma_n\pi^2\eta)^2}{7\beta_{4:1}P}\right)^{\frac{10}{27}} = \left(\frac{120(\eta\pi^2)^2}{7\beta_{4:1}\mathrm{CSNR}}\right)^{\frac{10}{27}}, \quad \sigma_x = 1. \tag{5.16}$$

The comparison between the above theoretical model and simulations of the real system is shown in Figure 5.5(a). As one can see there is quite a good correspondence between the theoretical model and simulations at high CSNR (above 30 dB). At lower CSNR the results starts to differ (for the same reasons as for the 3:1 system).

In Figure 5.5(b) several 4:1 systems are compared. One can observe that increasing a makes the system of equation (5.10) and (5.11) perform slightly better for a certain range of CSNR values, but then much worse for other ranges (in both cases an approximate unit speed parametrization is used). Further one can observe that the PCCOVQ system significantly outperforms all the systems based on fixed curves.

5.1.3 3:2 Mapping.

In this section a surface (2D) will be used to try to cover a ball like region in $\mathbb{R}^3$. Since no known method (at least to the authors knowledge) for finding the equations of a surface which is optimally "wrapped" inside a spherical region of space exists (minimizing the approximation distortion with the "smallest" surface possible), a known surface will be analyzed. The analysis done will reveal some of the difficulties that show up going from one to two parameters in the SK-mapping. This will give some indications on guidelines for finding equations for optimally wrapped surfaces. Further, it might be possible to generalize these concepts to N parameters (dimensions).

The block diagram for the system under consideration is given in Figure 5.1 substituting $M = 3$ and $N = 2$. The structure chosen for $\mathbf{S}$ is a

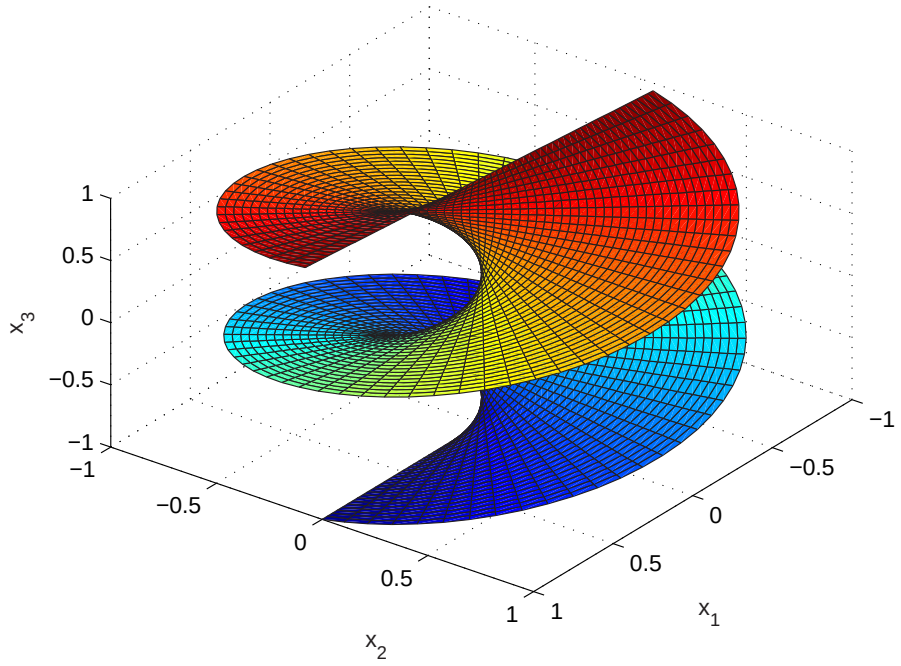


FIGURE 5.6: 3:2 dimension reducing mapping shown in the source space. The structure corresponds to equation 5.17 where $\Delta_1 = \Delta_2 = 1$, $\alpha_1 = \alpha_2 = 1$, and $z_1, z_2 \in [-\pi, \pi]$.

Helicoid, also known as *Archimedes' screw* and is shown in Figure 5.6. The Helicoid has the parametric equation (with the channel signal as parameters)

$$\mathbf{S}(z_1, z_2) = \frac{\Delta_1 \alpha_1}{\pi} z_1 \cos(\alpha_2 z_2)\mathbf{i} + \frac{\Delta_1 \alpha_1}{\pi} z_1 \sin(\alpha_2 z_2)\mathbf{j} + \frac{\Delta_2 \alpha_2}{\pi} z_2 \mathbf{k}. \qquad (5.17)$$

where Δ_2 is corresponding to the distance between two "folds" of the surface (at some distance from the x_3 axis), i.e. it has the somewhat same role as Δ for the 2:1 mapping in section 4.1. Δ_1 is determining the size of the "radius" of the surface in the x_1x_2-plane (see Figure 5.6 from the "top"). To obtain a symmetric channel signal both positive and negative parameter values are used ($(z_1, z_2) = (0, 0)$ corresponds to the origin in $\mathbb{R}^3$). The source vectors $\mathbf{x}$ is projected onto the Helicoid. The point (y_1, y_2) corresponding to the point on the Helicoid represents the two channel signals. The channel signal is further scaled by $1/\alpha_1$ and α_2 to satisfy a given power constraint (giving rise to the pair (z_1, z_2)).

In the following a mathematical model for the distortion using the mapping in (5.17) as a 3:2 mapping will be found. Further, an optimization of the derived mathematical model is done and compared to a 3:2 PCCOVQ system. At the end the difficulties showing up going from one to two variable SK-mappings will be shortly addressed, and some guidelines for finding the correct parametric equations will be given.

Channel distortion: The channel distortion is derived in Appendix E, and is given by

$$\bar{\varepsilon}_{ch}^2 = \frac{\sigma_n^2}{3\pi^2}((\Delta_2\alpha_2)^2 + (\Delta_1\alpha_1)^2(1+\alpha_2^2\sigma_{z_1}^2)). \tag{5.18}$$

From (5.18) one can observe that the channel distortion will increase linearly with the power on channel one ($\sigma_{z_1}^2$). This can also be seen from the metric component g_{22} derived in Appendix E

$$g_{22} = \left(\frac{\Delta_1\alpha_1\alpha_2}{\pi}\right)^2 z_1^2 + \left(\frac{\Delta_2\alpha_2}{\pi}\right)^2. \tag{5.19}$$

One can observe that the length of the velocity vector along the z_2 direction is dependent on z_1. Trying to compensate for the dependence in one direction, will affect the other and vice versa, as will be demonstrated at the end of this section.

Channel power: The channel power can be found by considering equation (5.17) and Figure 5.6. The derivation is done leaving out the scaling factors α_1 and α_2 (can be included at the end), so the statistics of the channel signals y_1 and y_2 are considered. Looking at the Helicoid from above (z_3 direction), one can see a disc in the x_1x_2 plane. The radius of this disc is given by $\rho = (\Delta_1/\pi)y_1$ (the "rotated straight lines"), while y_2 traces out all angles (the "parallel helices"). Thus y_1 is given by the function

$$y_1 = \pm\frac{\pi}{\Delta_1}\sqrt{x_1^2 + x_2^2}. \tag{5.20}$$

According to [Papoulis and Pillai, 2002, p.190], mapping by this function of two Gaussian random variables, gives a Rayleigh distribution. But due to the $\pm$ sign a zero mean "double Rayleigh" distribution shows up in this case

$$f_{y_1}(y_1) = \frac{1}{2}\frac{\Delta_1^2|y_1|}{\pi^2\sigma_x^2}e^{-\frac{\Delta_1^2y_1^2}{2\pi^2\sigma_x^2}}. \tag{5.21}$$

To get the constants right, one has to use the fact that the random variable $h(y) = cy$ has the distribution $f_h(h) = (1/|c|)f_y(h/c)$ [Papoulis and Pillai,

2002, p.131]. Further in [Papoulis and Pillai, 2002, p 148] it is shown that the second moment of the Rayleigh distribution $(y/v^2)e^{-y^2/(2v^2)}$ is $E\{y^2\} = 2v^2$, and so the signal variance on channel one is given by

$$\sigma_{y_1}^2 = 2\left(\frac{\sigma_x \pi}{\Delta_1}\right)^2 \tag{5.22}$$

The other parameter, y_2, will trace out parallel helices with "radius" given by (5.20). Consider first that $y_1 = 0$. Then y_2 will lie on the x_3 axis and have a Gaussian distribution since y_2 and x_3 have a linear relation. Since for this case $x_3 = (\Delta_2/\pi)y_2$ the variance of y_2 will be

$$\sigma_{y_2|y_1=0}^2 = \left(\frac{\pi\sigma_x}{\Delta_2}\right)^2. \tag{5.23}$$

Considering other cases ($y_1 \neq 0$) it seems like the variance (and the pdf) stays almost the same and independent of the value of y_1. Although y_2 will be on a helix, it seems to trace out all of z_3 in the same (linear) way. The little discrepancy from (5.23) seem to be more or less constant and equal to $1/2$ except when Δ_2 is very close to zero (a case which is of no interest), so that

$$\sigma_{y_2}^2 = \left(\frac{\pi\sigma_x}{\Delta_2}\right)^2 + \frac{1}{2} \tag{5.24}$$

Taking the scaling into account, the total channel power become (per channel dimension)

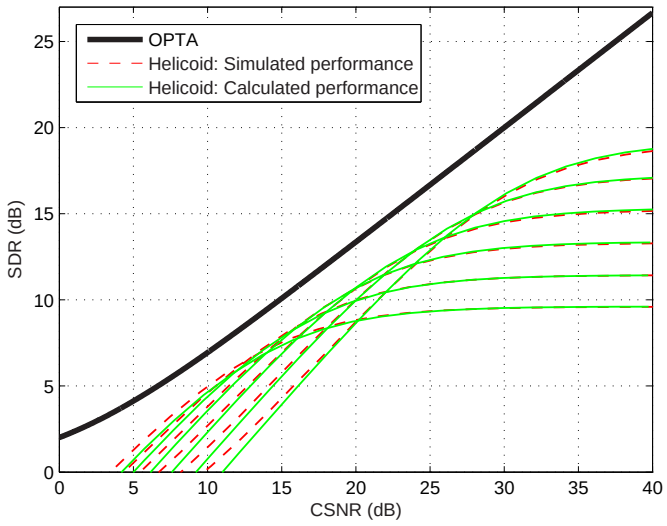
$$P = \frac{1}{2}\left(\frac{\sigma_{y_1}^2}{\alpha_1^2} + \frac{\sigma_{y_2}^2}{\alpha_2^2}\right) = \frac{1}{2}(\sigma_{z_1}^2 + \sigma_{z_2}^2) \tag{5.25}$$

Approximation distortion: As for the other dimension reducing mappings considered in this chapter the approximation distortion seems difficult (if at all possible) to find analytically. So again a numerical model is used. The approximation distortion is nearly independent of Δ_1, and so only Δ_2 determines the size of the approximation distortion. A model that seems to coincide well is (for $\sigma_x = 1$)

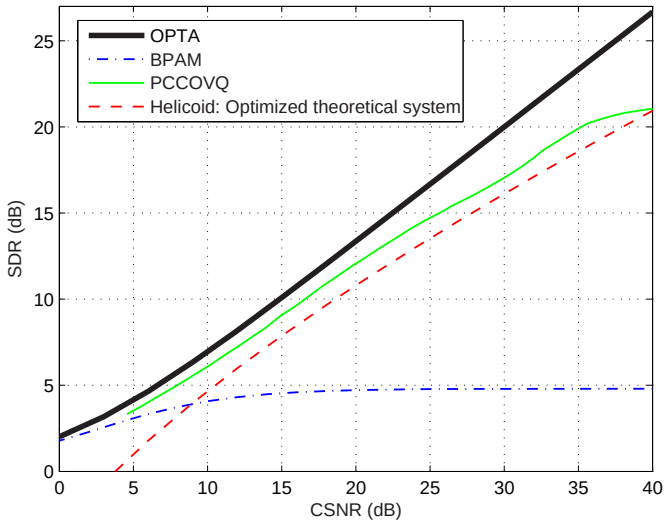
$$\begin{aligned} \bar{\varepsilon}_a^2 &\approx \beta\Delta_2^3 + \gamma\Delta_2^2 + \vartheta\Delta_2 \\ &= -0.0036\Delta_2^3 + 0.024\Delta_2^2 + 0.0056\Delta_2, \quad \Delta_2 \in [0,3],\ \ \sigma_x = 1, \end{aligned} \tag{5.26}$$

found by the previously used nonlinear curve fitting algorithm.

Optimization of the Helicoid system: This procedure is done assuming $\sigma_x = 1$. To find the optimal performance of the Helicoid system, one



(a)



(b)

FIGURE 5.7: Performance of a 3:2 dimension reducing system. (a) Comparison between the theoretical model and a simulation. (b) Comparison between the optimized Helicoid mapping a PCCOVQ and a linear system.

need to optimize it over Δ_1, Δ_2, α_1 and α_2 given a channel power constraint. The Lagrangian of this problem is

$$\mathcal{L}(\Delta_1, \Delta_2, \alpha_1, \alpha_2, \lambda) = \bar{\varepsilon}_a^2(\Delta_2) + \bar{\varepsilon}_{ch}^2(\Delta_1, \Delta_2, \alpha_1, \alpha_2) - \lambda c_t(\Delta_1, \Delta_2, \alpha_1, \alpha_2), \tag{5.27}$$

where the constraint is

$$c_t(\Delta_1, \Delta_2, \alpha_1, \alpha_2) = P_{max} - P(\Delta_1, \Delta_2, \alpha_1, \alpha_2) \geq 0. \tag{5.28}$$

P_{max} is the maximum allowed power per channel and $P(\Delta_1, \Delta_2, \alpha_1, \alpha_2)$ is given by (5.25). The first order necessary conditions for a minimum is given by the the Karush-Kuhn-Tucker (KKT) equations [Nocedal and Wright, 1999, p.328]. One of the criterion that has to be satisfied to be in a KKT point (potential minimum) is

$$\nabla_{\Delta_1 \Delta_2 \alpha_1 \alpha_2} \mathcal{L}(\Delta_1, \Delta_2, \alpha_1, \alpha_2, \lambda) = 0 \tag{5.29}$$

Solving (5.29) analytically is impossible since it contains an equation of high order, having no algebraic solution. Therefore numerical optimization has to be used (optimization toolbox in Matlab). Using nonlinear curve-fitting on the results of the optimization to the following functions

$$\begin{aligned}
\Delta_1(\mathrm{CSNR_{dB}}) = \alpha_1(\mathrm{CSNR_{dB}}) &= \frac{\beta_{\Delta_1}\mathrm{CSNR_{dB}}^2 + \gamma_{\Delta_1}\mathrm{CSNR_{dB}} + \vartheta_{\Delta_1}}{\sqrt[4]{P_{max}}} \\
\Delta_2(\mathrm{CSNR_{dB}}) &= \beta_{\Delta_2} e^{-(\gamma_{\Delta_2}\sqrt{\mathrm{CSNR_{dB}}} + \vartheta_{\Delta_2}\mathrm{CSNR_{dB}})} \\
\alpha_2(\mathrm{CSNR_{dB}}) &= \frac{\beta_{\alpha_2}\mathrm{CSNR_{dB}}^3 + \gamma_{\alpha_2}\mathrm{CSNR_{dB}}^2 + \vartheta_{\alpha_2}\mathrm{CSNR_{dB}} + \iota_{\alpha_2}}{\sqrt{P_{max}}}
\end{aligned} \tag{5.30}$$

gives the coefficients $\beta_{\Delta_1} = 5.253 \cdot 10^{-4}$, $\gamma_{\Delta_1} = 4.733 \cdot 10^{-3}$, $\vartheta_{\Delta_1} = 2.165$, $\beta_{\Delta_2} = 6.559$, $\gamma_{\Delta_2} = 0.1255$, $\vartheta_{\Delta_2} = 5.530 \cdot 10^{-2}$, $\beta_{\alpha_2} = 1.542 \cdot 10^{-4}$, $\gamma_{\alpha_2} = -5.359 \cdot 10^{-3}$, $\vartheta_{\alpha_2} = 0.1381$ and $\iota_{\Delta_2} = 0.3402$. These functions with their given constants are valid for all values P_{max} and $\sigma_x = 1$ in the range $\mathrm{CSNR_{dB}} \in [5, 40]$ dB and coincides well to the actual results. Figure 5.7(a) shows the comparison between the theoretical model given above and a simulation. One can observe that the performance of the theoretical model and the simulation coincide quite well (the functions in (5.30)). Figure 5.7(b) shows the performance of the optimized Helicoid system compared to a PCCOVQ system. Again the PCCOVQ system outperforms the proposed system (the reason why the PCCOVQ system declines is probably that only 4096 centroid are used, which is too small a number for representing a 2D space at high CSNR). The reason is probably the same as discussed earlier in this chapter (section 5.1.1).

Further comments: As mentioned in section 4.1, one can make the channel distortion independent of the channel signal variance by using a unit speed parametrization considering curves. According to appendix A.3. this could be done for any continuous curve. A problem shows up trying to do the same thing for surfaces. Assume that z_1 is mapped through the function $\varphi(z_1)$ before mapping onto the Helicoid. Then g_{22} becomes ($\alpha_1 = \alpha_2 = 1$ for simplicity)

$$g_{22} = \frac{\Delta_2^2 + \Delta_1^2 z_1^2}{\pi^2 \varphi^2(z_1)}. \tag{5.31}$$

Choosing $\varphi(z_1) = \sqrt{\Delta_2^2 + \Delta_1^2 z_1^2}/\pi$ then $g_{22} = 1$. But now g_{11} will become (putting $\Delta_1 = \Delta_2 = 1$ for simplicity)

$$g_{11} = \frac{z_1^4 + (z_1 z_2)^2 + 2z_1^2 + 1}{(1 + z_1^2)^2 \pi^2} \tag{5.32}$$

Figure 5.8 show (5.32). This is worse than what was initially the case (linear

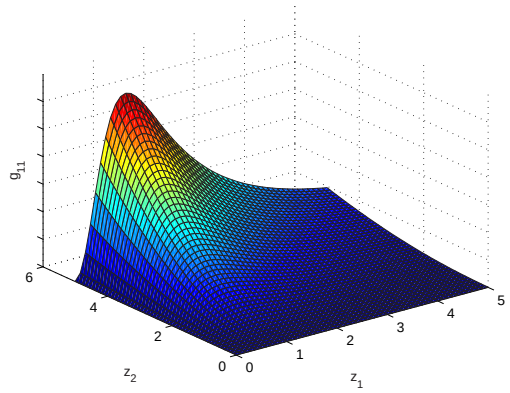


FIGURE 5.8: The first diagonal component of the metric tensor $g_{11}(z_1, z_2)$ for the Helicoid after a unit speed transformation of z_2.

increase of g_{22} as a function of z_1^2 given in (5.19)), since now, if the amplitudes on channel two is relatively large, the noise on channel one will be scaled the most for the most probable channel symbols on channel one, giving a large average distortion after decoding. This interdependence problem will show up in general using parametric surfaces (or for that matter hyper surfaces) for communication. One can get around this by constructing (or for that matter find) a surface which has either a constant metric

tensor for all z_1 and z_2, i.e.

$$G = \begin{bmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{bmatrix} \tag{5.33}$$

(where $\kappa_1 = \kappa_2$ for an i.i.d. channel signal) or where $g_{11}(z_1)$ is a function of z_1 and $g_{22}(z_2)$ is a function of z_2 only. These criteria will also be valid for an N dimensional SK-mapping.

5.2 Expanding 2:3 mapping

In this section a 2:3 mapping is introduced. The chosen mapping is a generalization of the Hybrid Scalar Qunatizer Linear Coder (HSQLC) which was proposed in Coward and Ramstad [2000], and is named *Hybrid Vector Quantizer Linear Coders*, referred to in the following as HVQLC. The system is shown in Figure. 5.9, where i are the VQ indices, e_1 and e_2 the two error components from the quantization and α_1 and α_2 scaling factors to

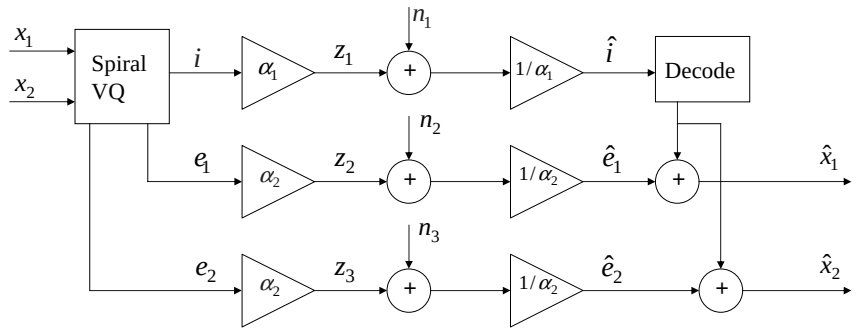
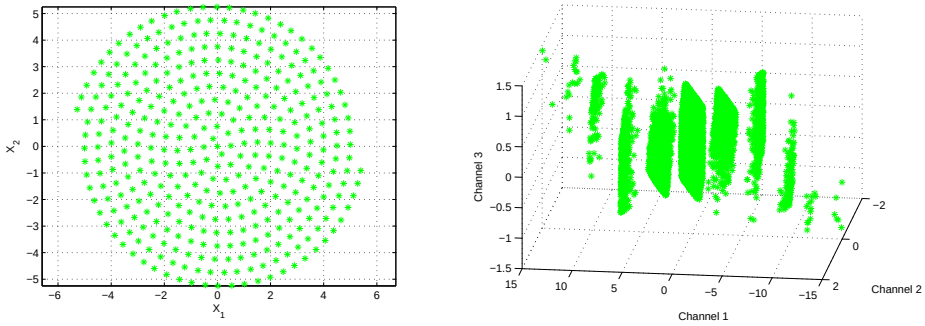


FIGURE 5.9: The HVQLC block diagram.

satisfy a given channel power constraint. To make the VQ code-book more easily adaptable to varying CSNR, the centroids are placed on Archimedes' spirals. A unit speed parametrization (see appendix A.3) is chosen so that all centroids are equidistant. The distance between both the spiral arms and the centroids along the arms are chosen to be equal to Δ, so a uniform VQ structure on a disc shows up. Figure 5.10(a) show the VQ structure. Notice that the cells are somewhat different from an ordinary uniform 2D vector quantizer. The VQ indices are scaled by α_1 and transmitted on channel 1, while the two scaled (by α_2) error components will be transmitted on channel 2 and 3. Figure 5.10(b) shows the simulated structure of the HVQLC mapping on the channel.



(a) (b)

FIGURE 5.10: The HVQLC 2:3 mapping. (a) The Spiral VQ. (b) The channel representation ($\alpha_1 = 1$).

At high CSNR an approximate mathematical model can be made, based on the theory in section 2.2, which further can be used to find the optimal parameters Δ, α_1 and α_2 as a function of the CSNR.

Weak noise distortion: The weak noise distortion can be found from (2.16), since the HVQLC system is a shape preserving mapping ("cutting" the source space is irrelevant to the weak noise distortion). The only contributions to the weak noise distortion comes from the distortion of the two error components, so

$$\bar{\varepsilon}^2_{wn} = \frac{\sigma_n^2}{\alpha_2^2}. \tag{5.34}$$

Anomalous distortion: The anomalous distortion can easily be calculated. Since the VQ indices are scaled by α_1, the distance between each plane in Figure 5.10(b) is α_1 and so the error probability will be the same as for the 1:2 mapping given in (4.12) substituting α_1 for Δ. The error made when the anomalous errors first occur is Δ, since the distance to the nearest neighbor(s) for a given centroid is always Δ. The anomalous distortion is given by

$$\bar{\varepsilon}^2_{th} = \frac{\Delta^2}{2}\left(1 - \operatorname{erf}\left(\frac{\alpha_1}{2\sqrt{2}\sigma_n}\right)\right). \tag{5.35}$$

Channel power: The channel power can be found approximately after making some simplifications. First, since the centroids are placed on the Archimedes' spiral in an equidistant manner, the pdf of the signal on channel one (z_1) will have a "discrete Laplacian like" distribution at high

CSNR. Although the distribution is discrete, its variance can be approximated quite well by the variance of a Laplacian distribution (somewhat the same case as in section 4.1). The power on channel one can therefor be approximated by (4.4) scaled by α_1^2

$$\sigma_{z_1}^2 = 2\alpha_1^2 \left(2\zeta \frac{\pi^2}{\Delta^2}\sigma_x^2\right)^2. \tag{5.36}$$

Also here $\zeta = 0.16\Delta$. Due to the discretization of the spiral, this variance will always be a bit smaller than the real power, but the higher the CSNR the better they coincide. Second, assuming that the spiral is dense (again at high CSNR), each of the error components (e_1 and e_2) are approximately uniformly distributed, limited by $\Delta/2$. So the power on channel two and three can be approximated by (2.47) substituting $m = 1$ and taking the scaling α_2 into account

$$\sigma_{z_2}^2 = \sigma_{z_3}^2 = \frac{\alpha_2^2\Delta^2}{12}, \tag{5.37}$$

which coincides well at high CSNR. The total channel power is

$$P = \frac{2}{3}\left(\left(2\alpha_1\eta\frac{\pi^2}{\Delta}\sigma_x^2\right)^2 + \frac{\alpha_2^2\Delta^2}{12}\right), \tag{5.38}$$

Optimization of the HVQLC To find the optimal performance of the HVQLC, it must be optimized over Δ, α_1 and α_2. The Lagrangian for this problem is

$$\mathcal{L}(\Delta, \alpha_1, \alpha_2, \lambda) = \bar{\varepsilon}_{wn}^2(\alpha_2) + \bar{\varepsilon}_{th}^2(\Delta, \alpha_1) - \lambda c_t(\Delta, \alpha_1, \alpha_2) \tag{5.39}$$

where the constraining function is

$$c_t(\Delta, \alpha_1, \alpha_2) = P_{max} - P(\Delta, \alpha_1, \alpha_2) \geq 0 \tag{5.40}$$

P_{max} is the maximum allowed power per channel and $P(\Delta, \alpha_1, \alpha_2)$ is given in (5.38). As mentioned in section 5.1.3 the KKT conditions [Nocedal and Wright, 1999, p.328] must be satisfied at a potential minimum. One of these conditions are

$$\nabla_{\Delta\alpha_1\alpha_2}\mathcal{L}(\Delta, \alpha_1, \alpha_2, \lambda) = 0 \tag{5.41}$$

To be able to solve this problem analytically, the roots of the above gradient must be found. But since it contains the error function, this can not be done analytically. Therefore numerical optimization is used (optimization toolbox in Matlab), and nonlinear curve-fitting is used to find the optimal Δ, α_1

and α_2 as a function of the CSNR based on the result of the optimization. Doing curve fitting to the following functions

$$\begin{aligned}\Delta(\mathrm{CSNR_{dB}}) &= \sigma_x \beta_\Delta e^{-(\gamma_\Delta \mathrm{CSNR_{dB}^2} + \vartheta_\Delta \mathrm{CSNR_{dB}})} \\ \alpha_1(\mathrm{CSNR_{dB}}) &= \sqrt{P_{max}} \beta_{\alpha_1} e^{-(\gamma_{\alpha_1} \sqrt{\mathrm{CSNR_{dB}}} + \vartheta_{\alpha_1} \mathrm{CSNR_{dB}})} \\ \alpha_2(\mathrm{CSNR_{dB}}) &= \frac{\sqrt{P_{max}}}{\sigma_x} \beta_{\alpha_2} e^{(\gamma_{\alpha_2} \mathrm{CSNR_{dB}^2} + \vartheta_{\alpha_2} \mathrm{CSNR_{dB}})}.\end{aligned} \tag{5.42}$$

gives the following coefficients: $\beta_\Delta = 4.567$, $\gamma_\Delta = 1.352 \cdot 10^{-4}$, $\vartheta_\Delta = 4.293 \cdot 10^{-2}$, $\beta_{\alpha_1} = 2.978$, $\gamma_{\alpha_1} = -0.2796$, $\vartheta_{\alpha_1} = 0.1299$, $\beta_{\alpha_2} = 0.7245$, $\gamma_{\alpha_2} = 9.091 \cdot 10^{-5}$ and $\vartheta_{\alpha_2} = 4.562 \cdot 10^{-2}$. These functions with their given constants are valid for all values of both P_{max} and σ_x in the range $\mathrm{CSNR_{dB}} \in [5, 60]$ dB and coincide well with the actual functions. The performance (robustness plot) of the optimized HVQLC system is shown in Figure 5.11. The functions given in (5.42) are used to find the optimal parameters for both the calculated and simulated system. The theoretical model

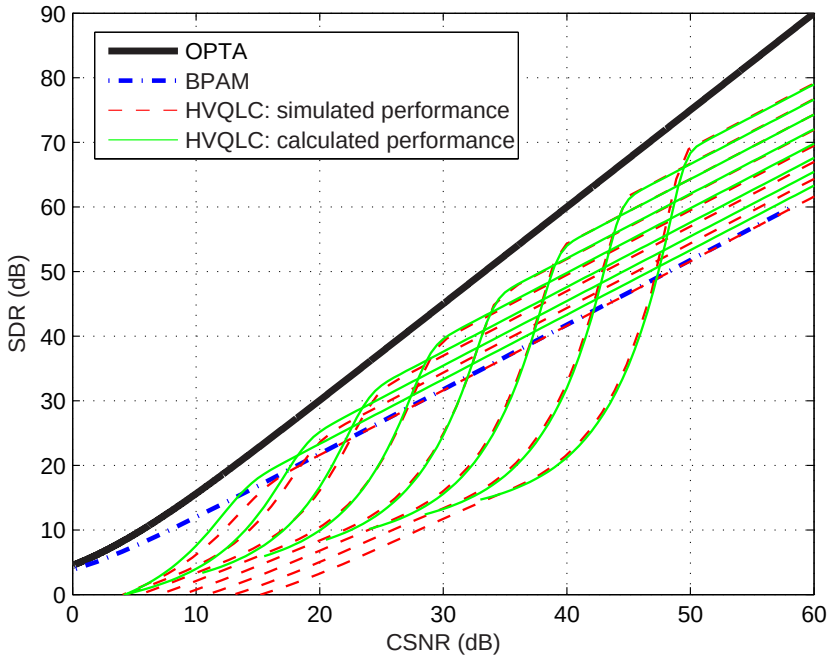


FIGURE 5.11: Comparison between simulated and calculated performance of the HVQLC system and BPAM.

coincides very well with simulations at high CSNR, wile there is a significant difference at low CSNR, something that was expected. The reason for the difference between the theoretical and simulated results in the horizontal direction (CSNR) is probably that the power determined by the theoretical model is smaller than the real power (noticeably so at low CSNR). The reason for the difference in the vertical direction (SDR) is that it is assumed that the two error components are equal and uniformly distributed. This does not coincide with reality if most of the probability mass of the source is close to the origin of the spiral. At low CSNR, therefore, it will probably be suboptimal to have equal scaling of the two error components. Further, at low CSNR, a Wiener like scaling factor at the receiver is more beneficial than just the inverse of the scaling at the transmitter (but by introducing such a factor, the mathematical analysis become significantly more difficult). For fixed parameters (a given curve) one can also observe a discrepancy between the calculated and simulated performance far from the optimal point for decreasing CSNR. The reason for this is that in calculating the anomalous distortion only mapping to the nearest neighboring planes was considered.

One can observe that the HVQLC system has quite good performance above 20dB CSNR, where it is in the range of 5dB from OPTA. Below 15dB however, BPAM will outperform the proposed system (simulated), and so, it is no point in using the HVQLC below this CSNR (BPAM is anyhow quite close to OPTA below 15dB CSNR).

5.3 Discussion

In this chapter several SK-mappings has been introduced. The proposed reducing mappings do not perform as well as the PCCOVQ system, meaning that there is something more to gain when constructing them. The main reason for the loss using curves (3:1 and 4:1 case) seems to a be non-optimal source-space filling. It might also be possible to gain something in these cases by finding the optimal φ (no known method for the dimension reduction case exists). In the 3:2 case the problem seems to be the dependence on both channel parameters in the metric components of the parametric surface. It might be that all of these problems can be avoided if one could discover a differential equation determining the overall optimal structure (like a Variational calculus problem). Further, it might be possible that some curve fitting or polynomial approximation can be used to replicate the component functions (in each dimension) of the PCCOVQ showing up when interpolating between the representation vectors/PAM

symbols for a given CSNR.

Also a 2:3 HVQLC mapping was tested. This SK-mapping seem to have quite good performance for high CSNR (no other known structure to compare it with except for BPAM exists to the authors knowledge). At CSNR below 15 dB however, the HVQLC performs worse than BPAM. The HVQLC will perform better for lower CSNR by introducing some Wiener factor like scaling at the receiver for the two error components, but the mathematical analysis will become more difficult (since the theory introduced assumes invertibility). To figure out the best possible performance for the HVQLC system, the optimization that was done for the HSQLC system in [Coward, 2001] should be generalized to the 2:3 case.

Chapter 6

Conclusions, discussion and Future research

'Experience teaches only the teachable'

-Aldous Huxley

In this dissertation both theory and systems for compression and error control of analog sources of information (amplitude continuous time discrete) have been introduced and investigated. The case studied is Gaussian memoryless sources communicated over a point-to-point link with reduced possibilities for feedback (supports CSI at most). The systems studied are called Shannon-Kotel'nikov (SK) mappings. The name SK-mapping comes from the people who first suggested the use of such systems. Shannon and Kotel'nikov independently suggested schemes using (piecewise) continuous curves for communicating analog sources. Shannon (and for that sake Kotel'nikov although the bounds in information theory were still to be discovered at the time he proposed these schemes) argued that such schemes might be the way to close in on what is optimal. Some of the advantages of such systems are their robustness against varying channel conditions, that they are delay free, and have seemingly good performance at low complexity.

In this thesis the concept of SK-mappings, developed by V. A. Kotel'nikov (for bandwidth expansion of scalar sources), has been generalized and extended to vector sources (using hyper surfaces) to be able to look at more general dimension expanding mappings, and also make a theory on dimension reducing mappings. A mathematical framework for for calculating the signal distortion using such mappings has been developed. The presented theory has further been used to show that SK-mappings have the potential

to reach OPTA by letting their dimensionality approach infinity. To reach OPTA it seems like a "shape preserving uniform mapping" will do (for both expansion and compression), something which is usually suboptimal for finite dimensional systems. The theory presented does, unfortunately not, yield the overall optimal geometry of the SK-mapping. Maybe the presented theory can be extended to some variational calculus problem giving as solutions the optimal overall geometry.

The PCCOVQ algorithm probably gives close to optimal geometrical structures in a numerical manner for a given dimension change factor, and source and channel noise pdf's. Well performing PCCOVQ mappings has been found in the past [Fuldseth, 1997]. Some small modifications on the PCOOVQ algorithm, done in this thesis, yield well performing dimension expanding mappings for scalar sources as well (small expansion factors). Examples on 1:2 and 1:3 PCCOVQ mappings are given. The complexity in finding higher dimensional 1:N expanding PCCOVQs and also expansion for vector sources, increases dramatically with the dimension due to the number of representation vectors/PAM symbols involved. For the uniform PCCOVQ it might be possible to include a smaller number of representation vectors/PAM symbols if some sort of method for doing noisy channel relaxation could be found. This would probably also make it possible to make the example systems mentioned above perform better at high CSNR ($> 15 - 20$dB). For the nonuniform PCCOVQ (probably giving the best possible result) it would be beneficial to find (if possible) a mathematical expression giving the optimal placing of the channel symbols given the partitioning of the source space and the VQ coodebook. This would significantly speed up the algorithm, since one can avoid the use of numerical optimization algorithms (which is slow for the objective function in question).

Replicating the 2:1 and 1:2 PCCOVQ structure by Archimedean spirals, gives well performing SK-mappings for both cases (preform on par with the PCCOVQ). The 2:1 mapping is further used as a basis for generalization to 3:1 and 4:1 mappings. These mappings have a reasonable good performance: $1.5 - 2$ dB from OPTA. But they are outperformed by the PCCOVQ by approximately 1 dB. The reason for this is probably that the PCCOVQ will place the representation vectors/PAM symbols in the best possible manner in the spaces in question. The PCCOCQ mappings seem to be too irregular to be replicated by one simple equation, but it might be possible that some curve fitting or polynomial approximation can be used to replicate the component functions (in each dimension) showing up when interpolating between the representation values/PAM symbols for a given CSNR. Although the PCCOVQ perform better than the SK-mappings, the

advantage of using the proposed SK-mappings over PCCOVQ is that the same structure can be used for all CSNR by merely changing coefficients in their equations. Further a 3:2 system built on "Archimedes screw" (Helicoid) was tested. The performance of it was reasonably good for CSNR < 15 dB. This example also shed light on some of the problems showing up going from 1D to 2D structures. At last a 2:3 SK-mapping was given, based on a generalization of the HSQLC system to vector quantizers, named HVQLC. The HVQLC has good performance for high CSNR (> 15 dB): around 5 dB from OPTA (unfortunately no other reference system than BPAM exists). It is also possible to make a simple approximate mathematical model of it using the proposed theory of SK-mappings.

For all SK-mapping proposed in this thesis, approximate mathematical models were given. Further, an optimization was done for all proposed SK-mappings, determining the optimal coefficients in their equations as a function of the CSNR, making them simple to adapt to varying channel conditions.

It seem generally hard to find high dimensional SK-mappings (and PCCOVQ mappings for that matter) due to the increased complexity of the problem. But once a structure is found, however, these systems perform quite well and are simple to adapt to varying channel conditions.

6.1 Contributions of this thesis

- A general theory for categorizing and calculating distortion using SK-mappings are derived.
- It is shown that SK-mappings has the potential to reach OPTA.
- The PCCOVQ algorithm is modified to yield well performing dimension expansion for scalar sources.
- Concrete examples on 1:2 expanding and 2:1 reducing mappings are given, using the Archimedean spiral for both cases.
- Some experiments on the proposed systems based on the Archimedean spiral is given, and their robustness against pdf mismatch is analyzed and discussed.
- Several example systems are developed: 3:1, 4:1 and 3:2 dimension reducing systems and a 2:3 dimension expanding system is found and analyzed.

- All proposed systems are given a (approximate) mathematical model using the proposed theory on SK-mappings. The mathematical models were optimized to yield the optimal coefficients for each mapping as a function of CSNR.
- Analysis yielding the channel signal statistics for each example mapping is given.

6.2 Ideas to future research

- Extending (if possible) or finding a new theory giving the overall optimal geometry for SK-mappings given the dimension change factor and the source and channel pdf's. Maybe some variational calculus problem could be formulated.
- The PCCOVQ algorithm should be generalized to M:N dimension expanding mappings.
- Dimension expanding PCCOVQ: Noisy channel relaxation should be incorporated in the uniform PCCOVQ algorithm. In the nonuniform PCCOVQ a mathematical relation giving the optimal placing of the channel symbols given the source partitioning and the codebook should be found (if possible) to significantly speed up the algorithm.
- Find SK-mappings for other cases than Gaussian. The PCCOVQ algorithm could also in these cases be used as a reference to determine the optimal geometry.
- Figure out how sensitive the SK-mappings are to some practical aspects like non ideal carrier recovery, phase drift etc.
- Find general efficient ways to project vectors onto the SK-mappings, and analyze their computational complexity in general.
- Find out how sensitive SK-mappings are to different noise phenomenon, eventually include such cases in the theory.
- Find a way to compare SK-mappings to traditional SSCC schemes, to figure what one will gain/loose in using such systems
- Do curve fitting/find polynomials that can be used to replicate the "component functions" of the PCCOVQ.

Appendix A

Some necessary mathematical results

A.1 The Metric Tensor

See [Spivak, 1999, pp.301-347] or [Dalarsson and Dalarsson, 2005, pp. 43-53] for definition and more involved details.

Consider an imbedding of an M-manifold $\mathcal{M}$ given by the parametric equation

$$\mathbf{S}(\mathbf{x}) = [S_1(\mathbf{x}), S_2(\mathbf{x}), \cdots, S_N(\mathbf{x})] \tag{A.1}$$

where S_i are component functions. The metric tensor for a smooth imbedding of $\mathcal{M}$ in $\mathbb{R}^N$ ($M \leq N$) is given by:

$$G = J^T J = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1M} \\ g_{21} & g_{22} & \cdots & g_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ g_{M1} & g_{M2} & \cdots & g_{MM}, \end{bmatrix} \tag{A.2}$$

where J is the Jacobian [Munkres, 1991, p.47] of $\mathcal{M}$, given by

$$J = \begin{bmatrix} \frac{\partial s_1}{\partial x_1} & \frac{\partial s_2}{\partial x_1} & \cdots & \frac{\partial s_N}{\partial x_1} \\ \frac{\partial s_1}{\partial x_2} & \frac{\partial s_2}{\partial x_2} & \cdots & \frac{\partial s_N}{\partial x_2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial s_1}{\partial x_M} & \frac{\partial s_2}{\partial x_M} & \cdots & \frac{\partial s_N}{\partial x_M} \end{bmatrix}^T \tag{A.3}$$

The metric tensor G is symmetric and positive definite [Callahan, 2000, pp. 208-209]. g_{ii} can be interpreted as the squared length of the tangent vector

in the direction of parameter x_i, where x_i is the i'th parameter in a parametric description of $\mathcal{M}$. All "cross terms" g_{ij}, are the inner product of the tangent vectors in the direction of x_i and x_j.

A.2 Hölders inequality

The following Lemma is taken from [Gasquet and Witomski, 1999]

Lemma Assume that $f \in L^p(I)$ ($|f|^p$ is Lebesgue integrable on the interval $I \in \mathbb{R}$) and $h \in L^q(I)$, where $\frac{1}{p} + \frac{1}{q} = 1$, then

$$\int_I |f(t)h(t)|\mathrm{d}t \leq \|f\|_p \|h\|_q \tag{A.4}$$

□

Proof See[Gasquet and Witomski, 1999, p. 135-136] □

A.3 Unit speed parametrization.

Assume a parametrization of a curve $\mathbf{s}(\varphi(x))$. For every continuous parameter curve there exists a special φ that makes all tangent vectors along the curve unit vectors.

Let $\mathbf{s} : x \in [a, b] \subseteq \mathbb{R} \rightarrow \mathbf{s}(x) \in \mathbb{R}^N$ be a parametrization for the curve C. Let $\ell(x)$ denote the curve length function of $\mathbf{s}$ given by

$$\ell(x) = \int_a^x \|\mathbf{s}'(q)\|\mathrm{d}q \tag{A.5}$$

and φ denote its inverse.

Theorem Let $\mathbf{y}(\ell)$ be a curve length parametrization of C. Then $\mathbf{y}(\ell)$ and $\mathbf{s}(\varphi(x))$ will have the same image, and $\|\mathbf{y}'(\ell)\| = \|\mathbf{s}'(\varphi(\ell))\| \equiv 1, \quad \forall \ell$. □

Proof See[Callahan, 2000, pp. 115-116]. □

Appendix B

Derivation of the weak noise distortion

Here it is shown that

$$\frac{1}{M}E\{(G^{-1}J^T\mathbf{n})^T(G^{-1}J^T\mathbf{n})\} = \frac{\sigma_n^2}{M}\sum_{i=1}^{M}\frac{1}{g_{ii}}, \tag{B.1}$$

gives the smallest possible weak noise distortion. G and J are described in Appendix A.1. To simplify the analysis (so the general matrix multiplications can be avoided) the N-dimensional noise vector $\mathbf{n}$ is replaced, without loss of generality, by its M dimensional projection $\mathbf{n}_P$, which will also be Gaussian (since P_{proj} is a linear transformation [Strang, 1986, p.117]). Let $J = J(\mathbf{x}_0)$. Further it is assume that an hypothetical inverse $\mathbf{B} = J^{-1}$ exists (which is also the case when the analysis is restricted to the M dimensional tangent space). Let $\mathbf{S}_t$ denote the tangent hyper plane of $\mathbf{S}$ at $\mathbf{x}_0$. Further, let the inverse of $\mathbf{S}$ be denoted $\mathbf{S}^{-1}$. Considering weak noise, the linear approximation of $\mathbf{S}^{-1}$ can be considered. Taking all the above into account, the received vector will be given by

$$\hat{\mathbf{x}} = \mathbf{S}^{-1}(\mathbf{S}_t(\mathbf{x}_0) + \mathbf{n}_P) \approx \mathbf{S}^{-1}(\mathbf{S}_t(\mathbf{x}_0)) + \mathbf{B}\mathbf{n}_P = \mathbf{x}_0 + \mathbf{B}\mathbf{n}_P, \tag{B.2}$$

and so, the MSE given that $\mathbf{x}_0$ was transmitted is

$$\varepsilon_{wn}^2 = \frac{1}{M}E\{\mathbf{n}_P^T\mathbf{B}^T\mathbf{B}\mathbf{n}_P\} = \frac{1}{M}\sum_{i=1}^{M}\sum_{j=1}^{M}\mathbf{b}_i^T\mathbf{b}_j E\{n_i n_j\}, \tag{B.3}$$

where $\mathbf{b}_i$ is column vector no. i in $\mathbf{B}$. Since the noise is considered i.i.d. each component of $\mathbf{n}_P$ is independent, so $E\{n_i n_j\} = \sigma_n^2\delta_{ij}$ and (B.3) is reduced to

$$\varepsilon_{wn}^2 = \frac{1}{M} E\{\mathbf{n}_P^T \mathbf{B}^T \mathbf{B} \mathbf{n}_P\} = \frac{\sigma_n^2}{M} \sum_{i=1}^{M} \mathbf{b}_i^T \mathbf{b}_i = \frac{\sigma_n^2}{M} \sum_{i=1}^{M} \|\mathbf{b}_i\|^2 \tag{B.4}$$

Since $\mathbf{B} = J^{-1}$, and it is well known that a matrix with orthogonal columns has an inverse (if it exists), the above result tells us that there is nothing to gain by choosing a nonorthogonal basis in the tangent space of $\mathcal{S}$, so the basis can always be chosen orthogonal. This simplifies both the analysis and the system itself. Making the Jacobian orthogonal will make the metric tensor diagonal. Therefore G^{-1} is also diagonal with diagonal elements $1/g_{ii}$. Now consider B.1 again. Using the fact that G^{-2} is diagonal, $E\{n_i n_j\} = \sigma_n^2 \delta_{ij}$ and with (B.4) in mind one can easily derive the following

$$\begin{aligned} \bar{\varepsilon}_{wn}^2 &= \frac{1}{M} E\{(G^{-1} J^T \mathbf{n})^T (G^{-1} J^T \mathbf{n})\} \\ &= \frac{1}{M} E\{(J^T \mathbf{n})^T G^{-2} (J^T \mathbf{n})\} = \frac{\sigma_n^2}{M} \sum_{i=1}^{M} \frac{1}{g_{ii}^2} \|J_i\|^2 \\ &= \frac{\sigma_n^2}{M} \sum_{i=1}^{M} \frac{1}{g_{ii}}, \end{aligned} \tag{B.5}$$

where J_i is column vector no. i of J and $\|J_i\|^2 \equiv g_{ii}$.

Appendix C

Derivation of the uniform spherical distribution and the approximation distortion lower bound

A uniform spherical distribution can be found by integrating a constant over a spherical region of $\mathbb{R}^m$ and equating it to one. It is most convenient to do the integration in *generalized spherical coordinates* [Richter, 2007]

$$\int_0^{\pi} \cdots \int_0^{\pi} \int_0^{2\pi} \int_0^{\Delta/2} \kappa \rho^{m-1} \prod_{k=1}^{m-1} \sin(\theta_k)^{m-1-k} \mathrm{d}\rho \mathrm{d}\theta_k = 1. \tag{C.1}$$

This integral equals the volume of an m-sphere with radius $\Delta/2$ scaled by the constant κ

$$\mathcal{V}_{\int} = \kappa \begin{cases} \frac{\pi^{\frac{m}{2}}}{\left(\frac{m}{2}\right)!} \left(\frac{\Delta}{2}\right)^m & , m \text{ even} \\ \frac{2^m \pi^{\frac{m-1}{2}} \left(\frac{m-1}{2}\right)!}{m!} \left(\frac{\Delta}{2}\right)^m & , m \text{ odd} \end{cases}$$

Using the relation $\Gamma(n+1) = n!$ [Bateman, 1953, p.3] for the Gamma function, then for even m

$$\mathcal{V}_{even} = \frac{\pi^{\frac{m}{2}} \Delta^m}{2^m \frac{m}{2} \left(\frac{m}{2} - 1\right)!} = \frac{\pi^{\frac{m}{2}} \Delta^m}{2^{m-1} m \Gamma\left(\frac{m}{2}\right)} \tag{C.2}$$

For odd m, things get more involved. *Legendre's duplication formula* [Bateman, 1953, p.5] must be used

$$\Gamma(2n) = (2\pi)^{-\frac{1}{2}} 2^{2n-\frac{1}{2}} \Gamma(n)\Gamma\left(n+\frac{1}{2}\right) \tag{C.3}$$

then

$$\frac{\Gamma(n)}{\Gamma(2n)} = \frac{(2\pi)^{\frac{1}{2}}}{2^{2n-\frac{1}{2}}\Gamma\left(n+\frac{1}{2}\right)} \tag{C.4}$$

and so

$$\begin{aligned} \mathcal{V}_{odd} &= \frac{2^m \pi^{\frac{m-1}{2}}}{m(m-1)(m-2)!}\left(\frac{m-1}{2}\right)\left(\frac{m-3}{2}\right)! = \frac{2^m \pi^{\frac{m-1}{2}}}{2m}\frac{\Gamma\left(\frac{m-1}{2}\right)}{\Gamma(m-1)} \\ &= \frac{2^m \pi^{\frac{m-1}{2}}}{2m}\frac{1}{2^{(m-1)-\frac{1}{2}}\Gamma\left(\frac{m-1}{2}+\frac{(2\pi)^{\frac{1}{2}}}{2}\right)} = \frac{\pi^{\frac{m}{2}}\Delta^m}{2^{m-1}m\Gamma\left(\frac{m}{2}\right)} \end{aligned} \tag{C.5}$$

This implies that the pdf of the approximation distortion (uniform spherical distribution) is given by

$$f_{\rho,\Theta}(\rho,\Theta) = \begin{cases} \frac{m2^{m-1}}{\pi^{\frac{m}{2}}\Delta^m}\Gamma\left(\frac{m}{2}\right) & ,\rho \in [0,\Delta/2], \forall\theta_i \\ 0 & \text{elsewhere} \end{cases} \tag{C.6}$$

Assuming a uniform source and equal distance between each neighbouring cell and one centroid at the origin, the approximation distortion can be found by

$$\bar{\varepsilon}_a^2 = \int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}\int_0^{\Delta/2} \rho^2 f_{\rho,\Theta}(\rho,\Theta)\rho^{m-1}\prod_{k=1}^{m-1}\sin(\theta_k)^{m-1-k}\mathrm{d}\rho\mathrm{d}\theta_k \tag{C.7}$$

The innermost integral in (C.1) is

$$\kappa\int_0^{\Delta/2}\rho^{m-1}\mathrm{d}r = \frac{1}{m}\left(\frac{\Delta}{2}\right)^m \tag{C.8}$$

and so (the contributions from the other integrals will cancel out with $f_{\rho,\Theta}$)

$$\begin{aligned} \bar{\varepsilon}_a^2 &= \frac{1}{m}\left(\frac{\Delta}{2}\right)^m \int_0^{\Delta/2}\rho^2 f_{\rho,\Theta}(\rho,\Theta)\rho^{m-1}\mathrm{d}\rho = \frac{1}{m}\left(\frac{\Delta}{2}\right)^m \frac{1}{m+2}\left(\frac{\Delta}{2}\right)^{m+2} \\ &= \frac{m}{4(m+2)}\Delta^2 \end{aligned} \tag{C.9}$$

Appendix D

Channel symbol distributions for M:1 mappings

In this appendix the channel symbol distribution that shows up when using M:1 mappings with unit speed parametrization (proposed in this thesis) is derived. Since all the compressive mappings based on parametric curves found in this thesis seems to have a curve-length function that can be approximated by the same parabola, a common procedure can be used to find their channel pdf and variance.

Assuming that the given parametric curve is so dense (small Δ) that one can disregard the approximation operation in the calculations, the whole M:1 mapping operation can be seen as a mapping $h : \mathbb{R}^M \to \mathbb{R}$. All the parametric curves used for dimension reduction in this thesis seem to have a curve-length function that can be approximated by the second degree function $\ell(\rho) = \pm\zeta(\pi/\Delta)^2\rho^2$. The domain of consideration is a ball in $\mathbb{R}^M$. The channel signal y is given by

$$Y = h(X_1, X_2, \cdots, X_M) = \pm\zeta\frac{\pi^2}{\Delta^2}(X_1^2 + X_2^2 + \cdots + X_M^2). \tag{D.1}$$

The resulting cumulative distribution of y can be found by a generalization of [Papoulis and Pillai, 2002, pp.180-181]

$$\begin{aligned} F_y(y) &= p_r\{Y \leq y\} = p_r\{(X_1, \cdots, X_M) \in \mathcal{D}_Y^+ \cup \mathcal{D}_Y^-\} \\ &= \iint \cdots \int_{\mathcal{D}_Y^+ \cup \mathcal{D}_Y^-} f_{X_1 \cdots X_M}(x_1, \cdots, x_M)\mathrm{d}x_1 \cdots \mathrm{d}x_M, \end{aligned} \tag{D.2}$$

where

$$f_{X_1 \cdots X_M}(x_1, \cdots, x_M) = \frac{1}{(2\pi)^{\frac{M}{2}}\sigma_x^M} e^{-\frac{x_1^2 + \cdots + x_M^2}{2\sigma_x^2}} \tag{D.3}$$

is the joint Gaussian distribution and

$$\begin{aligned}\mathcal{D}_Y^+ &= \left\{(x_1,\cdots,x_M)\big|x_1^2+\cdots+x_M^2 \le \frac{\Delta^2 y}{\zeta\pi^2}, y \ge 0\right\},\\ \mathcal{D}_Y^- &= \left\{(x_1,\cdots,x_2)\big|x_1^2+\cdots+x_M^2 \ge -\frac{\Delta^2 y}{\zeta\pi^2}, y < 0\right\}.\end{aligned} \tag{D.4}$$

The pdf can be found by differentiating (D.2) with respect to y. The simplest is to consider only the positive domain in (D.4) then use the fact that the channel pdf will be symmetric about the origin. Further, since the domain in consideration is spherically symmetric, it is convenient to do the integration in generalized spherical coordinates [Richter, 2007]

$$f_y(y) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}y}\int_0^{\pi}\cdots\int_0^{\pi}\int_0^{2\pi}\int_0^{\frac{\Delta\sqrt{y}}{\pi\sqrt{\zeta}}} f_\rho(\rho)\rho^{M-1}\prod_{k=1}^{M-1}\sin(\theta_k)^{M-1-k}\mathrm{d}\rho\mathrm{d}\theta_k \tag{D.5}$$

From Appendix C one can observe that the integral over all θ_i is

$$I_\Theta = \frac{\pi^{\frac{M}{2}}}{2\Gamma\left(\frac{M}{2}\right)}. \tag{D.6}$$

Further

$$\frac{\mathrm{d}}{\mathrm{d}y}\int_0^{C_1\sqrt{y}}\rho^{M-1}e^{-\frac{\rho^2}{C_2}} = \frac{C_1^M}{2}y^{\frac{M}{2}-1}e^{-\frac{C_1^2 y}{C_2}}, \tag{D.7}$$

where $C_1 = \Delta/(\pi\sqrt{\zeta})$ and $C_2 = 2\sigma_x^2$. Using this and multiplying (D.6) and (D.7) and further using absolute value (to make the pdf symmetric around the origin) gives

$$f_y(y) = \frac{\Delta^M|y|^{\frac{M}{2}-1}}{2(\sqrt{2\zeta}\pi\sigma_x)^M\Gamma\left(\frac{M}{2}\right)}e^{-\frac{\Delta^2|y|}{2\pi^2\zeta\sigma_x^2}} \tag{D.8}$$

According to [Papoulis and Pillai, 2002, p.87] the gamma distribution has the general form

$$f_\gamma(x) = \frac{1}{\Gamma(c)b^c}x^{c-1}e^{-\frac{x}{b}}u(x). \tag{D.9}$$

This means that the M:1 channel symbol distribution in (D.8) is a "double gamma distribution" where $c = M/2$ and $b = (2\pi^2\zeta\sigma_x^2)/\Delta^2$. Further in [Papoulis and Pillai, 2002, p.154], it is shown that the second moment of of (D.9) is given by $\mathrm{E}\{x^2\} = c(c+1)b^2$, and so since (D.8) is a double

gamma distribution with zero mean (second moment equal to second central moment when the mean is zero), the variance of (D.8) is

$$\sigma_y^2 = \frac{M(M+2)(\pi^2\zeta\sigma_x^2)^2}{\Delta^4}. \tag{D.10}$$

(D.10) gives the channel power (before scaling) in the M:1 case using uniform structures with unit speed parametrization.

Appendix E

Derivation of the Helicoid's channel distortion

In this appendix the channel distortion using the Helicoid as a 3:2 SK-mapping is derived. The equation of the Helicoid is given by

$$\mathbf{S}(z_1, z_2) = \frac{\Delta_1 \alpha_1}{\pi} z_1 \cos(\alpha_2 z_2)\mathbf{i} + \frac{\Delta_1 \alpha_1}{\pi} z_1 \sin(\alpha_2 z_2)\mathbf{j} + \frac{\Delta_2 \alpha_2}{\pi} z_2 \mathbf{k}. \tag{E.1}$$

To determine the channel distortion one will need the diagonal components of the metric tensor of $\mathbf{S}$

$$\begin{aligned} g_{11} &= \left\| \frac{\partial \mathbf{S}}{\partial z_1} \right\|^2 = \left\| \frac{\Delta_1 \alpha_1}{\pi} \cos(\alpha_2 z_2)\mathbf{i} + \frac{\Delta_1 \alpha_1}{\pi} \sin(\alpha_2 z_2)\mathbf{j} + 0\mathbf{k} \right\|^2 \\ &= \left(\frac{\Delta_1 \alpha_1}{\pi} \right)^2, \end{aligned} \tag{E.2}$$

and

$$\begin{aligned} g_{22} &= \left\| \frac{\partial \mathbf{S}}{\partial z_2} \right\|^2 \\ &= \left\| -\frac{\Delta_1 \alpha_1}{\pi} z_1 \sin(\alpha_2 z_2)\alpha_2 \mathbf{i} + \frac{\Delta_1 \alpha_1}{\pi} z_1 \cos(\alpha_2 z_2)\alpha_2 \mathbf{j} + \frac{\Delta_2 \alpha_2}{\pi} \mathbf{k} \right\|^2 \\ &= \left(\frac{\Delta_1 \alpha_1 \alpha_2}{\pi} \right)^2 z_1^2 + \left(\frac{\Delta_2 \alpha_2}{\pi} \right)^2. \end{aligned} \tag{E.3}$$

The off diagonal elements can easily be shown to be 0 (using the derivatives in (E.2) and (E.3))

$$\begin{aligned} g_{12} = g_{21} &= \frac{\partial \mathbf{S}}{\partial z_1}\frac{\partial \mathbf{S}}{\partial z_2} = \\ &- \left(\frac{\Delta_1\alpha_1}{\pi}\right)^2 \alpha_2 z_1 \cos(\alpha_2 z_2)\sin(\alpha_2 z_2) + \left(\frac{\Delta_1\alpha_1}{\pi}\right)^2 \alpha_2 z_1 \sin(\alpha_2 z_2)\cos(\alpha_2 z_2) \\ &= 0. \end{aligned} \tag{E.4}$$

Using (2.44) with $M = 3$ and $N = 2$ and the two metric components in (E.2) and (E.3), one can calculate the channel distortion

$$\begin{aligned} &\bar{\varepsilon}_{ch}^2 = \\ &\frac{\sigma_n^2}{3}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\left(\frac{\Delta_1\alpha_1}{\pi}\right)^2 + \left(\frac{\Delta_1\alpha_1\alpha_2}{\pi}\right)^2 z_1^2 + \left(\frac{\Delta_2\alpha_2}{\pi}\right)^2\right] f_{z_1,z_2}(z_1,z_2)\mathrm{d}z_1\mathrm{d}z_2 \\ &= \frac{\sigma_n^2}{3}\int_{-\infty}^{\infty}\left[\left(\frac{\Delta_1\alpha_1}{\pi}\right)^2 + \left(\frac{\Delta_1\alpha_1\alpha_2}{\pi}\right)^2 z_1^2 + \left(\frac{\Delta_2\alpha_2}{\pi}\right)^2\right]\int_{-\infty}^{\infty} f_{z_1,z_2}(z_1,z_2)\mathrm{d}z_2\mathrm{d}z_1 \\ &= \frac{\sigma_n^2}{3\pi^2}\left((\Delta_1\alpha_1)^2 + (\Delta_2\alpha_2)^2 + (\Delta_1\alpha_1\alpha_2)^2\int_{-\infty}^{\infty} z_1^2 f_{z_1}(z_1)\mathrm{d}z_1\right) \\ &= \frac{\sigma_n^2}{3\pi^2}((\Delta_2\alpha_2)^2 + (\Delta_1\alpha_1)^2(1+\alpha_2^2\sigma_{z_1}^2)). \end{aligned} \tag{E.5}$$

References

Bateman, H. (1953).
Higher Transcendental Functions, Volume One.
McGraw-Hill book company, Inc.

Berger, T. (1971).
Rate Distortion Theory: A Mathematical Basis for Data Compression.
Englewood Cliffs, New Jersey: Prentice-Hall, Inc.

Berger, T. and D. W. Tufts (1967, April).
Optimum pulse amplitude modulation part I: Transmitter-receiver design and bounds from information theory.
IEEE Trans. Information Theory IT-13(2), 196–208.

Cai, X. and J. W. Modestino. (2006, March).
Bandwidth expansion shannon mapping for analog error-control coding.
In *40th Annual Conference on Information Sciences and Systems*. IEEE.

Callahan, J. J. (2000).
The Geometry of Spacetime: An Introduction to Special and General Relativity.
New York: Springer-Verlag, Inc.

Cover, T. M. and J. A. Thomas (1991).
Elements of Information Theory.
New York: Wiley.

Coward, H. (2001).
Joint Source-Channel Coding: Development of Methods and Utilization in Image Communications.
Ph. D. thesis, NTNU Trondheim.

Coward, H. and T. A. Ramstad (2000, June).
Quantizer optimization in hybrid digital-analog transmission of analog source signals.

In *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Proc. (ICASSP)*, Volume 5, Istanbul, Turkey, pp. 2637–2640. IEEE.

Cramér, H. (1951).
Mathematical Methods of Statistics (First (reprint) ed.).
Princeton University Press.

Dalarsson, M. and N. Dalarsson (2005).
Tensors, Relativity and Cosmology.
Elsevier academic press, Inc.

Edwards, C. H. and D. E. Penney (1998).
Calculus with analytic geometry.
Prentice-Hall, Inc.

Floor, P. A. and T. A. Ramstad (2006a, June).
Dimension reducing mappings in joint source-channel coding.
In *Nordic Signal Processing Symposium*, Reykjavik, Iceland. IEEE.

Floor, P. A. and T. A. Ramstad (2006b, July).
Noise analysis for dimension expanding mappings in source-channel coding.
In *7th Workshop on Signal Processing Advances in Wireless Communications*, Cannes, France. IEEE.

Floor, P. A. and T. A. Ramstad (2006c, March).
Noise immunity for 1:N and M:1 nonlinear mappings for source-channel coding.
In *Data Compression Conference, 2006. DCC 2006. Proceedings*.

Floor, P. A. and T. A. Ramstad (2007, September).
Optimality of dimension expanding shannon-kotel'nikov mappings.
In *Information Theory Workshop*, Tahoe City, CA, USA. IEEE.

Floor, P. A., T. A. Ramstad, and N. Wernersson (2007, October).
Power constrained channel optimized vector quantizers used for bandwidth expansion.
In *International Symposium on Wireless Communication Systems*, Trondheim, Norway. IEEE.

Fuldseth, A. (1997).
Robust subband video compression for noisy channels with multilevel signaling.
Ph. D. thesis, Norwegian University of Science and Engineering (NTNU).

Gasquet, C. and P. Witomski (1999).
Fourier Analysis and Applications (First ed.).
Springer-Verlag New York, Inc.

Gastpar, M., B. Rimoldi, and M. Vetterli (2003, May).
To code, or not to code: Lossy source-channel communication revisited.
IEEE Trans. Information Theory 49(5), 1147–1158.

Gersho, A. (1979, July).
Asymptotically optimal block quantization.
IEEE Trans. Information Theory 25(4), 373–380.

Håkonsen, G. H. and T. A. Ramstad (2006, June).
On losses of performance in a joint sourcechannel coder.
In *Proc. Norwegian Signal Processing Symposium and Workshop (NORSIG/IEEE)*, Reykjavik, Iceland. IEEE.

Hekland, F. (2007).
On the Design and Analysis of Shannon-Kotel'nikov Mappings for Joint Source-Channel Coding.
Ph. D. thesis, Norwegian University of Science and Engineering (NTNU).

Hekland, F., P. A. Floor, and T. A. Ramstad (2008).
Shannon-kotel'nikov mappings in joint source-channel coding.
IEEE Transactions on Communications (to appear).

Hekland, F., G. E. Øien, and T. A. Ramstad (2007).
Quantifying performance losses in source-channel coding.
In *Proc. European wireless*, Paris, France. IEEE.

Hekland, F., G. E. Øien, and T. A. Ramstad (2005, March).
Using 2:1 Shannon mapping for joint source-channel coding.
In *Proc. Data Compression Conference*, Snowbird, Utah, pp. 223–232. IEEE: IEEE Computer Society Press.

Håkonsen, G. H. (2007).
Joint Source-Channel Coding for Image Transmission over Flat Fading Channels.
Ph. D. thesis, Norwegian University of Science and Engineering (NTNU).

Håkonsen, G. H., T. A. Ramstad, and A. Gjendemsjø (2006).
Image transmission with adaptive power and rate allocation over flat fading channel using joint source channel coding.
In *International Conference on wireless Information Networks and Systems*, Setubal, Portugal. IEEE.

Kotel'nikov, V. A. (1959).
The Theory of Optimum Noise Immunity.
New York: McGraw-Hill Book Company, Inc.

Lee, K.-H. and D. P. Petersen (1976, December).
Optimal linear coding for vector channels.
IEEE Trans. Commun. COM-24(12), 1283–1290.

Lervik, J. M. (1996).
Subband Image Communication over Digital Transparent and Analog Waveform Channels.
Ph. D. thesis, NTNU.

McRae, D. (1971, August).
Performance evaluation of a new modulation technique.
IEEE Trans. Information Theory 19(4), 431–445.

Munkres, J. R. (1991).
Analysis on Manifolds.
Westview Press.

Nocedal, J. and S. J. Wright (1999).
Numerical optimization.
New York: Springer-Verlag.

Papoulis, A. and S. U. Pillai (2002).
Probability, Random Variables and Stochastic Processes (Fourth ed.).
New York: McGraw-Hill higher education, Inc.

Ramstad, T. A. (2002).
Shannon mappings for robust communication.
Telektronikk 98(1), 114–128.

Ramstad, T. A. (2008).
Simple and reliable low power image communication based on dpcm and multiple refinements throuhg feedback quantifying performance losses in source-channel coding.
In *3rd International Symposium on Communications, Control and Signal Processing*, Malta. IEEE.

Richter, W. D. (2007).
Generalized spherical and simplicial coordinates.
Journal of Mathematical Analysis and Applications 336, 1187–1202.

Sakrison, D. J. (1968).
Communication Theory: Transmission of Waveforms and Digital Information.
New York: John Wiley & Sons, Inc.

Schalkwijk, J. and L. Bluestein (1967).
Transmission of analog waveforms through channels with feedback.
IEEE Trans. Information Theory 13, 617–619.

Shannon, C. E. (1948).
A mathematical theory of communication.
The Bell System technical journal 27, 379–423.

Shannon, C. E. (1949, January).
Communication in the presence of noise.
Proc. IRE 37, 10–21.

Shannon, C. E. (1959).
Coding theorems for a discrete source with a fidelity criterion.
IRE Nat. Conv. Rec 7, 142–163.

Skoglund, M., N. Phamdo, and F. Alajaji (2002, March).
Design and performance of VQ-based hybrid digital-analog joint source-channel codes.
IEEE Trans. Information Theory 48(3), 708–720.

Skoglund, M., N. Phamdo, and F. Alajaji (2006, August).
Hybrid digital-analog source-channel coding for bandwidth compression/expansion.
IEEE Trans. Information Theory 52(8), 3757–3763.

Spivak, M. (1999).
A Comprehensive Introduction to Differential Geometry (Third ed.).
Publish or Perish, Houston Texas, Inc.

Strang, G. (1986).
Linear Algebra and its Applications (Third ed.).
Thomson Learning, Inc.

Therrien, C. (1992).
Discrete Random Signals and Statistical Signal Processing.
Prentice Hall.

Thomas, C., C. May, and G. Welti (1975, June).
Hybrid amplitude-and-phase modulation for analog data transmission.
IEEE Trans. Commun. 23(6), 634–645.

Troutman, J. L. (1996).
Variational Calculus and Optimal Control.
Springer-Verlag.

Wernersson, N., J. Karlsson, and M. Skoglund (2007, June).
Distributed scalar quantizers for gaussian channels.
In *International Symposium on Information Theory*, Nice, France. IEEE.

Wozencraft, J. M. and I. M. Jacobs (1965).
Principles of Communication Engineering.
New York: John Wiley & Sons, Inc.

Zwillinger, D. (2003).
CRC Standard Mathematical Tables and Formulae (31st. ed.).
Boca Raton, Fla.: CRC Press.