ISBN 978-82-326-1586-5 (printed ver.)
ISBN 978-82-326-1587-2 (electronic ver.)
ISSN 1503-8181

Reza Parseh
Doctoral thesis

Doctoral theses at NTNU, 2016:124

Reza Parseh

Estimation over MIMO Fading Channels: Outage and Diversity Analysis

Doctoral theses at NTNU, 2016:124

NTNU
Norwegian University of Science and Technology
Thesis for the Degree of Philosophiae Doctor
Faculty of Information Technology, Mathematics and Electrical Engineering
Department of Electronics and Telecommunications

NTNU
Norwegian University of Science and Technology

NTNU
Norwegian University of Science and Technology

Reza Parseh

Estimation over MIMO Fading Channels: Outage and Diversity Analysis

Thesis for the Degree of Philosophiae Doctor

Trondheim, May 2016

Norwegian University of Science and Technology
Faculty of Information Technology,
Mathematics and Electrical Engineering
Department of Electronics and Telecommunications

NTNU
Norwegian University of
Science and Technology

NTNU
Norwegian University of Science and Technology

Thesis for the Degree of Philosophiae Doctor

Faculty of Information Technology, Mathematics and Electrical Engineering
Department of Electronics and Telecommunications

ISBN 978-82-326-1586-5 (printed ver.)
ISBN 978-82-326-1587-2 (electronic ver.)
ISSN 1503-8181

Doctoral theses at NTNU, 2016:124

Printed by NTNU Grafisk senter

Abstract

In this thesis, estimation of signals over fading channels for analog uncoded transmission is considered. In communication settings with tight delay requirements, e.g. in real-time control over wireless fading channels and vehicle-to-vehicle communication, the use of efficient and therefore long channel codes for reliability is not possible. Without channel codes, one needs to seek out alternative techniques. One such technique is to send uncompressed discrete-time source samples directly over the channel and then estimate the source signals from the channel outputs at the receiver. It is then required that the estimation quality is assessed with respect to source and system parameters, such that suitable system parameters can be selected for a given setting.

This work begins by considering scalar Gauss-Markov sources and communication over scalar Rayleigh fading channels. For estimation, the optimal minimum mean square filter, i.e. the Kalman filter, is used in order to estimate the signal at the receiver. In order to evaluate the performance of the Kalman filter, the estimation error outage probability is selected as a measure of quality. For random fading channels, the instantaneous estimation error variance for the Kalman filter is also random. Principally, the outage probability criterion measures the probability that the instantaneous estimation error variance of the Kalman filter exceeds a certain threshold. This measure is most meaningful in settings where delay is of concern. The presented results in this thesis include characterization of the estimation outage probability, derivation of the upper and lower bounds for a certain range of outage thresholds, and characterization of the behavior of the outage probability in the high signal-to-noise ratio regime.

Next, the channel model is extended to include multiple receivers in order to obtain a diversity gain and improve the estimation quality. Due to lack of coding, diversity is a very a suitable way to obtain extra reliability when needed. For this setting too, upper and lower bounds are obtained for the outage probability. It is then shown that in the high signal-to-noise ratio regime, the estimation error outage probability decreases inversely polynomially with the signal-to-noise-ratio to the power of the number of receivers.

Afterwards, the source is extended to be a vector of arbitrary dimension and the channel to be a general MIMO Rayleigh fading channel. A joint Kalman filter and space-time coding scheme is proposed to allow for transmission of sources with any dimensions over the channel. The space-time codes are incorporated in order to parallelize the channels and obtain full diversity for the estimation outage probability.

Finally, two special scenarios are considered. The first is when the source dimension grows large. In this case, the source may be considered as a collection of users in a sensor network who transmit their measurements over a general MIMO fading channel without compression or other extra processing. The performance is analyzed in terms of the high signal-to-noise-ratio behavior of the average mean squared error. The second scenario, considers a case when source samples are not correlated in time, and are sent over parallel MIMO channels. The estimation quality is then improved by transmission over a larger number of channels than the source dimension.

Preface

This thesis is submitted to the Norwegian University of Science and Technology (NTNU) for partial fulfillment of the requirements for the degree of *philosophiae doctor* (PhD). My main advisor has been Associate Prof. Kimmo Kansanen at the Department of Electronics and Telecommunications. I have also had two co-advisors, Prof. Tor A. Ramstad and Prof. Tor Arne Johansen, from the Department of Electronics and Telecommunications, and the Department of Engineering Cybernetics, respectively.

Besides the research activities, the PhD work also included compulsory course studies corresponding to one semester of full-time studies, and thirteen months of teaching assistantship duties. This work was conducted in the period of Sep. 2011 – Oct. 2015. In the period of Jan. 2014 – May 2014, I was a visiting researcher at the EURECOM Institute in Sophia-Antipolis, Nice, France, and worked under the supervision of Prof. Dirk Slock.

The three-year period of PhD research was funded by the Faculty of Information Technology, Mathematics and Electrical Engineering at NTNU. The thirteen months of teaching assistantship was funded by the Department of Electronics and Telecommunications at NTNU.

Acknowledgements

I want to give special thanks to my main advisor Associate Prof. Kimmo Kansanen. Kimmo has been an excellent mentor and advisor. We have had innumerable fruitful discussions during this period, and many of the ideas of this work were shaped during those discussions. With his love for science, direct and honest comments, and a sense of humor which he kept at all times, in ups and downs of research, he taught me better ways of doing research, creative thinking, and decision making. As a PhD student, I have been very lucky to have had him as my advisor, and this work could not have shaped without his help, supervision and support. It has been an honor to have worked with him in these four years.

Many thanks go to Prof. Tor Ramstad, for all the discussions and the support during these years, including the proofreading of this thesis. Tor has always welcomed scientific discussions with a genuine interest, and has been a very good mentor to me, both in research and in life. My gratitude goes also to Prof. Dirk Slock at the EURECOM Institute, where I spent five months as a visiting researcher. His knowledge and talent in the field of signal processing and wireless communications is exemplary. He has also been very supportive and helpful during my stay at EURECOM, which has made that period one the most memorable times of my life. Special thanks go to the respected committee members, Prof. Krishna Narayanan, Prof. Daniel Quevedo, and Dr. Nadezda Sokolova, for their critical review of this work, whose valuable comments helped me to improve the quality of this work.

I also want to thank my former colleagues at the Department of Electronics and Telecommunications for the nice and friendly working environment I have had during this period. Special thanks go to Ms. Kirsten Ekseth, whose remarkable support I have enjoyed in the last four years. She made my life and work in the new country much easier, and I am very happy to have met and worked with her. Many thanks go also to former and current PhD students in the signal processing group, namely Babak, John, and Mohammad, whose company and consultations I have gladly enjoyed in this period.

Last, but not least, I humbly thank my mom and dad, Jarieh and Jamshid, who unconditionally supported my education despite all the problems and challenges they faced. Many thanks go also to my loving girlfriend Synnøve, whose love and support has made this work considerably easier for me.

When I started this work four years ago, I was not sure what I was going to do exactly, and what I was going to end up with. These four years have been rich with experiences, ups and downs, and significant changes in my life. I owe a lot to the many individuals I have met during this period, who contributed to shaping my work and life. It was mostly chance that I made the decision to move to Norway to do this PhD, and I am happy to have made that choice. The work-life balance, independence, trust, the friendly environment and all the help and support I have enjoyed during this period is remarkable, and I am very grateful for that.

Contents

List of Figures

Abbreviations

AR Auto-Regressive.

DFT Discrete Fourier Transform.

EKF Extended Kalman Filter.

EnKF Ensemble Kalman Filter.

EOP Estimation Error Outage Probability.

GM Gauss-Markov.

iDFT Inverse Discrete Fourier Transform.

IEV Instantaneous Estimation Error Variance.

ITS Intelligent Transport Systems.

MIMO Multi-Input Multi-Output.

MMSE Minimum Mean Square Error.

MSE Mean Squared Error.

NCS Network Control Systems.

RRE Random Riccati Equation.

SIMO Single-Input Multi-Output.

SISO Single-Input Single-Output.

SNR Signal to Noise Ratio.

UKF Unscented Kalman Filter.

V2V Vehicle to Vehicle.

WSN Wireless Sensor Networks.

List of Symbols

A State transition matrix.

C_u Process noise covariance matrix.

C_v Channel noise covariance matrix.

G Coding gain.

G_k Kalman gain.

H Fading channel.

K Source signal dimension.

M Estimation error (co)-variance.

M_{th} Estimation error variance threshold.

N Received signal dimension.

N_c Number of channel uses.

P Prediction error (co)-variance.

P_{out} Outage probability.

P_x Source average power.

$T()$ Prefiltering transform.

$\Gamma()$ Incomplete Gamma function.

$\hat{\mathbf{x}}$ Estimated signal.

$\mathbf{u}$ Process noise.

$\mathbf{v}$ Channel noise.

$\mathbf{x}$ Source signal.

$\mathbf{x}_T$ Input to the channel.

$\mathbf{x}_f$ Prefiltered source signal.

$\mathbf{y}$ Received signal at the estimator.

$\mathbf{y}_R$ Received signal at the channel output.

$\mathcal{C}$ Set of complex numbers.

$\mathcal{N}()$ Normal distribution.

$\mathcal{O}()$ Big O asymptotic notation.

$\mathcal{R}$ Set of real numbers.

$\mathcal{U}()$ Unit step function.

$\overline{\text{MSE}}$ Average MSE.

ρ Correlation coefficient for scalar ARGM processes.

σ_u^2 Process noise variance.

σ_v^2 Channel noise variance.

Im() Imaginary part of a complex number.

Pr() Probability.

Re() Real part of a complex number.

exp() Exponential function.

tr() Matrix trace.

d_{ord} Diversity order.

d_{th} Distortion threshold.

n Discrete time index.

$o()$ Small o asymptotic notation.

Chapter 1

Introduction

Low or zero delay transmission of measurements of a dynamic system to a remote controller/observer is required in applications such as network control systems (NCS), wireless sensor networks (WSN), intelligent transport systems such as vehicle-to-vehicle (V2V) communication, and generally for remote real-time signal processing. Due to tight delay conditions in these cases, high-performance block-wise channel codes, which incur unacceptable delays, are not permitted.

It is possible to send the measurements directly over the channel without quantization, compression or coding, using analog uncoded transmission. One then performs estimation on the channel outputs at the receiver in order to recover the transmitted signals. Transmission of discrete-time continuous-valued signals over channels is then an attractive alternative to the state-of-the-art digital communication schemes due to its simplicity and zero-delay property. This scheme is also attractive for application in WSN, where the sensors have very limited processing ability, and thus simple but reliable transmission schemes are needed. Due to that, sophisticated signal processing operations are rather performed at a fusion center. In addition, such sensors are often battery-powered and therefore power-efficient communication schemes need to be considered as well.

As the transmitted signals are continuous-valued and are not quantized, estimation rather than detection is performed at the receiver side. A relevant measure for estimation quality assessment should also be sought out. Then, it is important that the performance of the suggested transmission scheme is carefully analyzed with respect to that measure. The analysis of the signal estimation quality is necessary to ensure satisfactory performance for the possible practical implementations.

One prevalent measure of quality of estimation is the mean squared error (MSE), i.e. the mean of the squared error between the original signal and the estimated one. While the error is a random process due to the presence of noise, the MSE is a deterministic value. However, when there are other sources of randomness in the system, e.g. when the channel is random, the MSE (average with respect to noises only) becomes random itself. The random MSE may converge in distribution, given that certain convergence criteria is sat-

isfied by the system parameters and the channel. Based on that, deterministic measures of quality can be obtained from the distribution of the MSE. The simplest of such measures is the average MSE, where the average is now computed over all the randomness in the system, including the random channel. The average MSE measure is applicable when an overall assessment of the estimation quality is required. A more elaborate and complete measure is the estimation error outage probability. Estimation error outage probability expresses the probability that the MSE exceeds a threshold. While the outage is obtained from the complementary cumulative distribution function (complementary cdf) of the random MSE, the average MSE measure is the mean of its probability density function (pdf). For ergodic processes, this probability can also be interpreted as *how often* the random MSE exceeds that threshold. This measure is most insightful when delay is of concern, e.g. an outage value is needed for each channel realization, and when there is a need and a meaning for comparing the estimation performance against a threshold value, e.g. in order to ensure a certain level of statistical reliability in mission critical applications. The outage probability measure is also more informative than the average MSE, as the average MSE can be obtained from the outage probability function if needed, while the other way around is not possible. Yet, it might be easier and less complicated to get a closed-form expression for the MSE. Thus, the outage probability measure is utilized mainly, but the average MSE measure is also considered when outage probability calculation becomes tedious. Also, as the notion of random MSE is already clarified, the term *estimation error outage probability* is used for the outage probability of the random MSE, because MSE traditionally has a deterministic value and this might lead to confusion or misunderstanding. Alternatively, the term distortion outage probability is used when vector sources and normalized instantaneous MSE are considered.

In the aforementioned applications, the transmitted signals are quite often correlated. If the signals follow an auto-regressive (AR) model, which many natural signals do at least approximately, the use of Kalman (-like) estimation algorithms is optimal at the receiver. The optimality is for the minimum mean square error (MMSE) criterion and under certain conditions for the channel and the source. If e.g. the signals follow a linear Gauss-Markov (GM) model, the optimal estimator at the receiver is the Kalman filter. Even if the Kalman filter is not optimal, it is still the best linear estimator under certain conditions.

If a state-space model is not at hand, one could consider a Wiener filter as the estimator in order to recover the signal from the noisy observations. Being a forerunner to the Kalman filter, one could consider it a special case of the Kalman filter as well. The Wiener filter can also be used instead of the Kalman filter when there is an uncertainty in the state-space model. In some cases, the model uncertainty is so significant that the extra gain obtained by using the state-space model is overshadowed by such inaccuracy. It might then be more beneficial to just use the Wiener filter instead.

The major difference between our work and most classical Kalman and Wiener filtering/estimation settings is the existence of a random and time-varying channel. With a random channel, the instantaneous MSE of the filter and the filter itself are then random. Nevertheless, it can be shown that these random variables converge in distribution. This

type of statistical convergence allows for the calculation of the outage probabilities and the average MSE's. Obtaining the estimation or prediction errors' statistical properties then paves the way for analyzing the performance of the filters.

In this work, the behavior of the estimation error outage probability and the mean estimation error are studied in two different signal to noise ratio (SNR) regimes, namely the finite SNR regime and the asymptotic infinite SNR regime, also known in the literature as the high SNR regime. Ideally, one would like to find the exact value for the outage probabilities in the finite SNR regime. However, as we will see in the following chapters, this task seems to be quite daunting even for the simplest cases. We then resort to finding upper and lower bounds for the outage probabilities whenever possible, and can even in some cases show tightness of bounds in certain regimes. Although it is mostly analysis which is performed in this work, rather than design, these bounds may be applied in system design, where certain outage probabilities are required to guarantee a certain level of estimation quality.

The other main approach for characterizing the estimation quality in this thesis is a diversity analysis in the high SNR regime. The results are often a simple characterization of the high SNR behavior of the estimation error outage probability for the different schemes which are considered here. In some settings, it is the case that the estimation error outage probability decreases inversely polynomially with a certain power of SNR in the high SNR regime. This polynomial behavior in the high SNR regime can then be characterized using two parameters, namely the order of the polynomial and the leading multiplier. To be consistent with the related concepts in the digital communication domain, the names diversity order and coding gain are used for the aforementioned order and leading multiplier (the coding gain is calculated relative to a benchmark value), respectively. The diversity order and the coding gain can together fully characterize the asymptotic performance of the estimation error outage probability.

In the next section, we review the related previous works in order to better understand the relevance and implications of the current thesis.

1.1 Previous Related Works

This section begins by reviewing the challenges in some of the main application areas of analog source transmission, and the problems this transmission scheme might potentially help solve or mitigate. Later on, the signal model, the channel model, and the estimation algorithm and estimation quality assessment methods are introduced, and a review of the previous work related to these categories is provided.

1.1.1 Application Areas for Analog Source Transmission

The literature for network communication and control is diverse and rich. One of the main challenges concerning NCS with communication links in between, is stability. One of the pioneering works which has provided an overview of major limitations of the communication networks which affect the stability of NCS is [122]. In that regard, one can consider networked induced delay, multiple packet transmission and packet drops. The delay occurs due to exchange of information between sensors, actuators and controller within the network and can be a serious issue. In extreme cases, such delays may lead to instability. With multiple packet transmission (vs. single packet transmission), the sensor and actuator data may be sent via separate network packets and may arrive at different times, which may degrade the performance. Packet drops also occur occasionally due to link failures, congestion, and message collisions in networks and can be another cause for performance degradation and instability. For that matter, it can be argued that by using the proposed analog uncoded transmission scheme, the link failure issue may be avoided or mitigated. Instead of complete link failure, one can achieve a rather smooth and graceful performance degradation compared to abrupt failures occurring in NCS which utilize digital schemes.

In a collection of works [4], several problems regarding stability, performance and reliability in NCS are discussed and the current state of the art for those problems are presented. Most importantly, the issues with integration of the wireless transmission media into NCS, such as delay and network communication challenges, are mentioned and real time NCS are discussed. Several of the articles presented in [4] are based on practical problems and projects, which makes it an interesting starting point on the NCS research problems.

Considering a point to point (P2P) link in a network, it is necessary to reduce the link failure probability in order to prevent instability. As [61] duly states, Shannon-type capacity achieving codes may not be used in NCS, because those codes are sufficiently long and incur an intolerable delay. Causality, the main characteristic of control applications, is not considered in the Shannon-sense error performance either. For causal systems, especially in the control context, issues such as stability have the utmost importance. This has lead to the development of the notion of anytime capacity in [84], initially for scalar systems. In the anytime sense, successful communication is not only defined by its diminishing decoding error probability, but rather by a stability criterion. The anytime capacity is then given by the log of the largest unstable system gain or equivalently the spectral radius. Anytime capacity has been extended to the case of vector systems in [83], which provides the theoretical benchmark for practical development, e.g. by anytime coding.

Recent developments on the topic of anytime codes include the existence of universal (not requiring the channel transition probabilities) anytime codes presented in [21]. Reliable transmission of real-valued information over binary erasure channels (see [15, Chapter 7] for the definition) is studied in [14], where two anytime coding strategies are presented, one with high complexity and exponential convergence rate and the other one with lower

complexity, but sub-exponential convergence rate. Although interesting, they have the limitation of being only applicable in erasure channels. While very interesting in nature and promising for the future works, anytime codes are still under development and thus are not considered further in this work.

Wireless sensor networks are also an interesting application area for the schemes considered in this thesis. The literature on WSN is also abundant as it engages several communication and networking challenges. One could list the main functionalities of the sensor networks as sensing and detecting relevant phenomena and first (simple) stages of processing [60]. The main application areas are in machine, ship, animal, vehicle, environmental and medical monitoring. Although the communications and networking constraints for WSN are somewhat different from those of NCS, there are issues such as delay which are relevant for both cases. While some WSN applications may tolerate delay, applications which involve real-time monitoring could benefit from our proposed scheme. Another issue in WSN is power management, because in many applications, the sensors are battery-powered and their computational operations are required to be energy-efficient [80]. As a result, a simple communication scheme, which does not involve compression and coding, but can still reliably transmit the information content of the signal is very useful. The need for power efficiency and low data rates is also mentioned in [77], one of the pioneering works on WSN. There is also an extensive amount of information related about WSN challenges and the involved design techniques in the recent work of [1]. The joint-source channel coding approach is also mentioned in [1, Ch. 4.3.2] and estimation subject to power allocation is discussed. For another review work on WSN, the interested reader is also referred to [95]. In short, from the P2P communication-link point of view for WSN, a power-saving, cheap, and (almost) real-time/delay-free communication scheme is required.

Another concrete example for a potential application area is intelligent transport systems (ITS). The requirements for the communication links among the vehicles on the road or between the vehicle and the roadside traffic station for traffic control purposes, are also in many ways similar to the requirements for NCS and WSN, especially the delay issues. At high speeds and on the highways, real-time communication is a must for ITS, especially if autonomous cars are to be introduced to the traffic system. While swarms of cars collaborating and communicating in a network incurs network issues similar to those of WSN, the real-time requirements may be similar to those of NCS. According to [101], one important type of message for V2V communication is the transmission of position information from one vehicle to the other. Given the correlated nature of position information, vehicles' high speed and their relative short distances (in contrast to e.g. cellular communications), it seems that analog communication is well suited for such purposes. Channel models for V2V communication using the dedicated short range communication band is considered in [97] and it is recommended that a Nakagami [33, Sec. 3.2.2] fading channel model be used, whose parameters need to be determined using empirical data. Network structure and certain cooperative techniques were also considered in [9,53,117]. As the technology and structure for V2V communication is still under development, there is a prospect that the proposed schemes in this thesis may be used as part of the future

protocols and standards.

1.1.2 The Signal and Estimator Model

One common characteristic of the signals observed and transmitted in the three aforementioned real-time communication settings is that they are correlated in time. For vector signals, the signals in each dimension will be correlated as well. The correlation is a result of the signals being recorded from natural phenomena. The state of a system at each sample time is related to its past. As an example, a sensor network monitoring temperature and humidity in a forest bed usually records correlated signals as temperature and humidity signals do not usually change abruptly and are often correlated, given that the sampling times are short enough. The position of a car on a highway at the current time is a function of its position at the previous time instant and its current speed. Furthermore, different position dimensions (in a 3D space) are also correlated and a function of the traffic situation. One then needs a signal model which is general enough such that it takes into account the correlation in time and across dimensions, but not so complicated that it makes the analysis too tedious or unrealistic. One such model is the first order ARGM signal model.

The use of GM models for natural signal modeling is widespread and has a long history. A first order ARGM model states that the current system state/signal sample is only a function of the previous system state/signal sample and some randomness. To name a few of the application of ARGM signals in the analysis of the dynamic system, we can mention the pioneer work [113], which is a survey of design methods for failure detection in dynamic systems, [81] which discusses maximum likelihood estimators for dynamic systems, and [62] which discusses optimal adaptive estimators for stochastic processes. For the application of ARGM models in wireless sensor networks, we could mention [124], which considers a GM mobility model for localization in sensor networks. In [3], the placement of nodes in a sensor network is modeled as GM random fields. In the context of target tracking for sensor networks, [54] uses Gauss-Markov mobility model for the mobile node movements in the surveillant field and presents a new distributed localization scheme based on that assumption.

For the several applications mentioned so far and various others, an ARGM model is quite suitable for modeling the dynamic systems and their observed signals. If the signals are transmitted over a communication channel, they are prone to the destructive effects of the communication channel, such as noise, fading, interference, packet losses, network problems, etc. In order to recover the transmitted signal from the channel outputs, an estimation algorithm needs to be used. One such algorithm which incorporates the dynamic model of the system is the Kalman filter. Given certain conditions (so-called conditionally Gaussian systems), the Kalman filter is optimal in the sense that it minimizes the MSE. For some other less stringent conditions, the Kalman filter is still optimal, if only linear filters are to be used, i.e. the Kalman filter is the optimal linear filter. Linear filters/estimators have simple designs and low computational complexity, and are therefore

interesting in many applications, though they might not be optimal in the general sense, which may imply non-linear filtering.

After its introduction in [50], the Kalman filter and similar Kalman-like algorithms have been applied in many engineering applications for purposes, including but not limited to estimation, filtering, smoothing, and prediction. The Kalman filter is optimal for linear systems with known parameters. If these two conditions are violated, the optimality might also be violated. In addition, some of the filtering operations are very costly for large dimensions. In the following, a review of some of the major contributions to the theory and practice of the Kalman filter regarding the aforementioned issues is presented.

For nonlinear dynamical systems, the development of the optimal (sequential) filter requires a complete characterization of the conditional pdf of the source symbols, given the observations, at each step [56]. This might necessitate acquiring knowledge of an (potentially) infinite number of parameters. This challenge has been addressed in several ways. To deal with the pdf calculation problem, some suboptimal approximations have been proposed [34], [47], [55], [63], [94]. However, a more popular approach is the linearization of the system model using Taylor series expansion, and then using a Kalman filter for the linearized model. The corresponding filter has been named the extended Kalman filter (EKF), and has been the de facto estimation and tracking algorithm for many years after its introduction. A good description of the fundamental theory of the extended Kalman filter can be found in [47, ch. 8]. Several other extensions to the original Kalman filter including colored and correlated noises are also treated in [47]. Although a significant improvement, EKF suffers from some drawbacks such as complications with implementation, tuning and reliability. Another approach first proposed in [24] is called the ensemble Kalman filter (EnKF). The EnKF uses Monte-Carlo simulation methods in order to track the propagation of mean and covariance of a Gaussian random variable through a nonlinear system, in order to approximate the posterior probabilities needed in the Kalman filter. The EnKF then uses an ensemble (or even more than one ensemble [44]) of points for tracking the mean and the covariance at each step. For large data sizes, the EnKF can help reduce the computational complexity significantly. Another method introduced in [48] is called the unscented Kalman filter (UKF), which uses carefully selected and fewer data points, made possible by a special transformation called the unscented transformation. The points are also chosen using a deterministic algorithm, rather than being selected at random. While the EKF is based on linearization of the underlying nonlinear system, the EnKF and UKF are based on numerically estimating the underlying (posterior) conditional probability distribution functions. The UKF was further extended to parameter estimation (compared to state estimation) and for machine learning problems in [108]. Some numerical improvements in terms of computational complexity and stability for parameter estimation were also achieved in [106].

Another constraint on the optimality of the Kalman filter, as briefly mentioned previously, is that it is only valid for known system parameters. In reality, the system parameters are not always known and must be estimated. This has lead to the analysis of the effects of erroneous parameter estimation in Kalman filtering design and implementation. In fact,

the treatment of modeling errors in Kalman filtering started shortly after the introduction of the original Kalman filter. A good review of the early results around the effects of modeling error can be found in [Ch. 8] [47], which is mentioned here again, due to its close connection to part of the work in Chapter 3. In [7], the modeling error for the measurement error covariance matrix is considered. In [25], the modeling error in the state transition matrix and the process noise covariance matrix are considered, and in [67], the effect of error in the initial state is studied. A more complete analysis is presented in [39], where the effect of modeling errors in the covariance of the initial state vector, the covariance of the stochastic inputs to the system, and the covariance of the uncorrelated measurement noise is studied. There, a recursive equation for the actual covariance matrix of the estimation error with modeling errors was also derived. The modeling error in process and observation noise covariance matrices, as well as biases, were also studied in [78]. One of the most complete modeling error analyses can be found in [35], where all the possible modeling error types, except the errors in the bias term, were considered. However, as it was discovered in [Ch. 8] [47], some of the derivations of [35] appear to be incorrect. The corrected equations for the actual estimation and prediction error covariance matrices, when all the system parameters may be prone to modeling error (including the biases), can thus be found in [Ch. 8] [47].

One of the ways to mitigate the modeling error and uncertainty in the system model problems in Kalman filtering is robust designs. There exist several methods for robust state estimator design, but the details are beyond the scope of this work. However, a good review on different robust state estimation techniques can be found in [85]. According to [85], robust Kalman filtering may be performed via three different types of algorithms, namely the H_∞ filtering method, the set-valued method, and the guaranteed cost method. Briefly, the H_∞ methods (see [87]) are based on bounding the induced L_2-norm (equivalently the largest singular value) of the total mapping operator of the filter (from inputs to the error). The set-valued approach (see [8]) models the uncertainties as belonging to a particular set, e.g. satisfying a certain energy constraint, to achieve robustness. In the guaranteed cost design approach (see [116]), the target is to devise a class of estimators which can guarantee an upper bound for the steady-state variance of the state estimation error for the whole range of uncertainties in the model. In addition, a new robust approach is also developed in [85]. There, the objective is to minimize the maximum residual norm (estimation cost function) over the whole range of uncertainties, whereas in the standard Kalman filter, the objective is only to minimize the residual norm. The resulting algorithm has the benefit of not requiring to verify any existence conditions, which is problem for robust H_∞ and set-valued methods. In addition, this approach yields stable filters, while having the benefit of being applicable to a general class of parametric uncertainties.

The Kalman filter and its successors are able to minimize/reduce MSE, when there is an available state-space model for correlation between consecutive source samples. This however, might not always be the case. The Wiener-Kolmogorov theory, which was developed years prior to the Kalman filter and in separate works by Wiener and Kolmogorov [93], considers in turn only wide sense stationary signals. In order to perform filtering through the Wiener-Kolmogorov theory, or simply Wiener filtering, one needs to

have the autocorrelation and cross-correlation functions of the original and noisy versions of the signal. The major difference between the Wiener and Kalman filters is the existence of the state-space model for the signals in the Kalman formulation, which allows for very efficient recursive updates of the estimates for the current signal sample based on previous signal estimates, rather than the whole signal history. Kalman filtering also allows for non-stationary signal models. Although replaced by the Kalman-like filters in applications where a state-space model is available, the Wiener filter still has its application in various fields, including but not limited to channel estimation [41], wavelet de-noising [32], filtering of cyclo-stationary signals [28], speech enhancement [64], and image de-noising [92]. While Wiener filter has quite significant importance in signal estimation, the focus of this work is on estimation of the ARGM signals, and therefore the suitable estimation technique is the Kalman filter. The Wiener filter is however considered in Ch. 6 for estimation of band-limited signals where no such state-space model is available.

1.1.3 Analog Source Transmission Over Fading Channels

With the signal structure and the corresponding estimation techniques reviewed, we turn to the issue of source transmission. The main objective is to find schemes, suitable for real-time communication, which incur zero delay and have low complexity. One such scheme as previously mentioned, is analog uncoded transmission. To be more specific, we refer to schemes which transmit continuous-valued but discrete-time signals, and which do not utilize conventional source and channel coding techniques. Conventional digital schemes usually require buffering, which incurs delay, and therefore are not suitable for real-time applications. It is, however, allowed to utilize transforms, power allocation, and similar operations which do not incur delay and have low computational complexity. In the following, some of the recent works which have considered the analog uncoded transmission schemes for various applications are reviewed. Note that the use of the term *analog* is to distinguish between the transmission of uncoded continuous-valued samples, which is considered in this thesis, and general uncoded transmission of bits in digital communication.

A significant result for analog uncoded transmission is optimality for transmission of Gaussian signals over Gaussian channels and single source broadcast channels [30] and for some simple Gaussian sensor networks [29]. The optimality of analog uncoded transmission was also proven for large Gaussian relay networks in [31]. Furthermore, it is shown that for transmission of independent and identically distributed (i.i.d.) bivariate Gaussian signals over multiple access channels, the uncoded transmission is optimal if the channel SNR is below a certain threshold [58]. This was also shown to be true even if the competitive block coding schemes enjoy a feedback channel as well [59]. In addition, transmission of a memoryless bivariate Gaussian source over an average-power-constrained one-to-two Gaussian broadcast channel is studied in [11]. There, it is shown that below a certain SNR threshold, transmission of a linear combination of the source

components (still considered uncoded) achieves the optimal power-distortion trade-off.

While the number of cases where analog uncoded transmission is optimal, is limited, it is still an interesting scheme due to its simplicity and zero-delay property. It is therefore considered in several other applications, either on its own or in combination with digital transmission. In [5], analog transmission of sensor data in a Gaussian network is considered, when the transmissions are subject to noise and fading. There, the optimal power allocation scheme for minimizing the total transmission power subject to some distortion constraint for the estimates is obtained. Analog source transmission for sensor networks have also been considered in other works, such as [107] and [91], for estimation over fading channels. A hybrid digital-analog scheme is considered in [26] for joint source-channel coding of a bivariate Gaussian source sent over Gaussian multiple access channel. Analog transmission is also considered in [88] in the context of systematic lossy source channel coding, as means of providing the decoder with side information. In [12], transmission of analog i.i.d. sources over multi-input multi-output (MIMO) block fading channels is considered and a new hybrid digital-analog communication scheme using space-time codes is presented. In [65] too, hybrid digital-analog source-channel codes are introduced for broadcasting and robust communication. Hybrid digital-analog schemes are also considered in [37] for join source-channel coding over MIMO block-fading channels. A Hybrid digital-analog scheme for transmission of analog sources over additive white Gaussian noise channels is also considered in [16], where the main objective is quantizer optimization. Another hybrid digital-analog coding scheme is introduced in [27] for transmission of a Gaussian source over a Gaussian channel, and which generalizes the work of [11].

Although analog transmission avoids bit and packet errors, packet drops and outage in digital communication, it is still prone to noise, fading and interference. In this work, the focus is only on the effects of fading. In line with similar other works within the context of analog uncoded transmission over fading channels, e.g. [37], [75], [12], [105], and [19], Rayleigh fading channels are mainly considered. Rayleigh fading is a good model for propagation and scattering effects in urban environments and in general for non-line-of-sight communication. It is also a simple model which allows for easier analysis and therefor more insightful results. For that reason, much of the early and fundamental analysis for fading channels in digital communication consider the Rayleigh fading model as well [102].

Besides the fading channel type in terms of distribution, the fading channel's behavior in time and frequency also plays an important role for system development and analysis. While the fading channel is in principle a random multiplicative distortion, it can be considered (approximately) non-changing for a certain period of time, known as the coherence time (for mathematical definitions of different fading models, concepts, and parameters, please see [102]). In that regard, the fading can be fast or slow. Fast fading is the case when the coherence time of the channel is smaller than one symbol duration. Otherwise, we deal with slow fading. If the channel is the same for a block of source transmissions, then it is conveniently called a block fading model. Many of the works

studying Rayleigh fading, especially in digital communication, consider the block fading model, as it best suits high rate cellular transmission (and naturally very short symbol durations). The block fading model is also considered for analog source transmission over fading channels [37], [75], [12], [105], and [19]. However, the block size in this model is reduced for lower symbol rates, which might in fact be the case in settings such as wireless sensor networks.

From a spectral point of view, the channel can be frequency selective or non-frequency selective (flat). If the source bandwidth is smaller than coherence bandwidth of the channel (inversely proportional to the root mean square of the delay spread), then the fading is flat [102]. The input-output relationship for the signals sent over flat fading channels is simpler, compared to the frequency selective case, as inter-symbol interference (ISI) might be avoided. However, this assumption limits the source bandwidth to the coherence bandwidth of the channel. The flat fading model is also used considerably for analog source transmission over fading channels [37], [75], [12], [105], and [19].

1.1.4 Kalman Filtering with Random Parameters

Due to the randomness of the fading channel, the estimator e.g. the Kalman or Wiener filter is random and does not necessarily converge to a constant value. The instantaneous estimation error covariance matrix used to qualify the quality of estimation is random as well. Kalman filtering with random systems parameters has been studied from different aspects and for different settings due to its wide application spectrum. In the following, some of those works are reviewed.

In [90], Kalman filtering over packet delaying networks with random delays is considered and the filter performance is characterized using a probabilistic approach. In [89], a binary sensor power scheduling algorithm for Kalman filtering over packet dropping networks is presented. The algorithm minimizes the expected terminal estimation error covariance at the remote estimator. In [82], the performance of the Kalman filter is studied when the measurements are only intermittently available and a method is presented to determine the boundedness of the expected value of the estimation error covariance matrix. In [118], stability of Kalman filtering over a network subject to random packet losses is studied. There, the packet losses are modeled as a time-homogeneous ergodic Markov process. Then for second-order systems, necessary and sufficient conditions for stability of the mean estimation error covariance matrices are derived. In [79], stochastic stability of centralized Kalman filtering for linear time-varying systems with wireless sensors and for transmission over fading channels is considered. There, a power control strategy is proposed to mitigate the effects of the packet drops. Also, the stabilizing power control policies are formulated such that they minimize the total sensor power budget. In [96], joint rate and power control algorithms for wireless sensor networks are presented, which consider state-space models and also account for congestion in the network. In [125], power efficient algorithms for state estimation in sensor networks with non-ideal channels are addressed. There, a dimensionality reduction approach is presented which minimizes

the MSE of Kalman filtered state estimates formed at the fusion center. Kalman filtering for nonlinear systems can also lead to estimation and prediction error covariance matrices which are random as in [38] and [22], even though system parameters are deterministic.

In the Kalman filter, the estimation error covariance matrix is a function of the channel randomness and the filter memory and is related to another matrix, the prediction error covariance, through a simple matrix transformation. The prediction error covariance matrix also propagates through a Riccati equation, studied extensively in the literature. With random channel matrices, the prediction error covariance matrices then constitute the well-known stochastic process referred to as the random Riccati equation (RRE) [114]. While the target here is the analysis of the estimation error quality, other works, especially in the control literature, have focused on the prediction error covariance matrix, because it is used directly in the controller design and that the Riccati formulation makes the analysis more feasible, in general. The results on the prediction error covariance matrix analysis may, if required, be extended to the estimation error covariance matrix usually by means of simple algebraic and probabilistic manipulations. In the following, some of the works which have considered the RRE are reviewed.

In [110], stability of the RRE is studied and it is shown that under mild assumptions on the random observability Gramian matrix, it is both L_r and exponentially stable[1]. In [115], the peak covariance stability of the RRE resulting from Kalman filtering with random observation losses is studied. Boundedness of the covariance matrix in the usual sense is also considered in the same work. In [66], an adaptive filtering scheme based on the Riccati equation is proposed for state estimation in network control systems subject to delays, packet drops and missing measurements. In [51], it was shown that a sequence of random covariance matrices converges in probability when observations are sent over a packet erasure channel where the erasure event is a Bernoulli i.i.d. process. The stationary distributions for infinitely large random matrices were studied in [104] and [105] for two classes of random Riccati and Lyapunov equations. There, a Stieltjes transform approach was used to obtain the moments of the stationary eigenvalue distribution of the prediction error covariance matrix for Lyapunov and Riccati equations. However, in order to obtain the intended results, the state transition matrix is either considered equal to unity or random, which are not of very practical use. While [105] succeeded in obtaining closed-form values for moments of the Lyapunov equation, it only presents a set of interconnected Stieltjes transform equations for two of the Riccati equations it considers. It also mentions that a stable numerical algorithm was yet to be developed. However, the eigenvalue distribution of the prediction error covariance matrix for the single-input multi-output case is obtained, given that the system dimensions grow asymptotically large.

[1]The supremum of the expected value of the r-th norm is finite, where the norm is defined as the maximum of the singular values.

1.1.5 Quality Assessment Criteria and Related Issues

In order to measure the quality of estimation at the receiver for analog communication as well as source-channel coding, two main criteria have been considered before, namely end-to-end average distortion and distortion outage probability. The behavior of these two criteria in different regimes of SNR, number of sensors, channels, etc have also been under the spotlight. The end-to-end average distortion measure and the corresponding distortion exponent figure of merit were first studied in [57] in the context of source transmission over a pair of slow fading channels. The distortion exponent shows how fast the average distortion decays with SNR and is similar to our notion of diversity gain for outage probability, which will be reviewed shortly after. The main result of [57] is that for the case of two non-ergodic (slow fading) channels, multiple-description encoding with joint source-channel decoding outperforms other schemes. Note that [57] considers blocks of i.i.d. source samples. The behavior of the distortion exponent was further considered in [12] for a hybrid digital-analog scheme and its relation with channel-to-source bandwidth ratio further outlined. The results imply that in general, a higher source-to-channel bandwidth ratio leads to lower distortion exponent and thus higher average distortion. In [36], a joint source–channel coding for cooperative relay systems is considered and the optimal distortion exponent behavior in the high SNR regime is presented for full-duplex system and for all bandwidth ratios. Also, a tight upper bound for small and large bandwidth ratios is provided for half-duplex relay channels. In [42] and using an automatic repeat request scheme, a new framework is developed for optimizing the trade-off between diversity, multiplexing, and delay in MIMO systems, and which minimizes the end-to-end distortion. For AR Gauss-Markov models, the end-to-end average distortion, corresponding to the mean of the covariance matrix for the Kalman filter, was considered in [19], and two converging sequences were introduced which act as upper and lower bounds for the mean of the error covariance.

In the spirit of outage analysis for fading channels in digital communication, we utilize estimation error outage as a criterion for estimation performance assessment. Estimation outage event is defined as the event when the estimation error exceeds a certain threshold. The outage measures are most insightful when delay is of concern and when the MSE is random. From a practical viewpoint, this measure could be used as a design parameter for a control or monitoring system which observes the process. A similar property, namely distortion outage was proposed in [75] for MIMO block fading channels from an information theoretical point of view. There, a transmitter-informed lower bound and a separate source-channel coding scheme are studied. It is shown that the bounds achieve the same diversity order and the same outage probability for Gaussian sources and also for sources with discrete alphabets. In addition, the trade-off between the distortion outage diversity order and the source-to-channel-bandwidth-ratio is derived. In this thesis however, the interest is rather in analyzing the performance of practical estimators, especially the Kalman and Wiener filters with respect to their outage behavior. Our approach is thus constructive, rather than information theoretic. Also note that due to the Gauss-Markov nature of the source, the source blocks in this work are mainly correlated and not independent.

One way to obtain the outage values is by first obtaining the asymptotic pdf of the eigenvalues of the estimation error covariance matrix. In that regard, efforts have been made to characterize the pdf of the instantaneous random covariance matrices resulting from Kalman filtering over packet drop networks, as well as with analog transmission over fading channels. It was shown in [51] that the sequence of random covariance matrices converges in probability when observations are sent over a packet erasure channel where the erasure event is a Bernoulli i.i.d. process. Again in [104] and [105], the system model is such that it can simply be adapted to continuous-valued source transmission over fading channels. There, the equations regarding the moments of the stationary eigenvalue distributions for infinitely large random matrices were found. However, analyzing smaller dimensions is still an open research topic and also the distributions themselves rather than their moments remain unsolved for more general cases.

Considering the distortion outage probability as a measure of quality in this work, the focus is especially on its high SNR behavior. The aim is to show that similar diversity results hold for estimation error outage probability in Kalman filtering over fading channels compared to diversity results for error and outage probability in digital transmission. The outage measure for estimation has also been considered before in [17, 75, 109]. The diversity results provide simple and effective means of quantifying the estimation reliability. In [17], estimation outage and estimation diversity are considered in the context of distributed sensing, where several sensors observe an i.i.d. process and transmit their measurements over parallel fading channels. There, it is shown that using a certain power allocation scheme, a diversity order equal to the number of sensors can be achieved. A similar system model is considered in [109], where the focus is on distortion outage minimization. One noteworthy result in [109] is that full channel state information at the transmitter can increase the diversity order by a factor of the logarithm of the number of sensor nodes, compared to [17]. The work of [75] should be mentioned again regarding the bounds and diversity analysis which are presented for MIMO-block fading channels.

In order to provide the means for spatial diversity for general MIMO channels, we have turned to space-time codes and have employed methods similar to those of digital communication for analog source transmission. The idea of orthogonal space-time coding for digital communication was first introduced in [2], allowing transmit diversity gain in addition to receive diversity gain by use of multiple transmit antennae. General complex block orthogonal space-time codes were introduced in [98], where it was shown that they are able to achieve the full diversity of general MIMO channels. There exist, however, many space-time coding schemes other than the (complex) orthogonal space-time codes, of which a few will be reviewed in the following. For digital communication, there exist a well-known diversity-rate trade-off, as a fundamental limit of the MIMO channels [123]. While block orthogonal space-time codes are able to achieve full diversity, quasi-orthogonal space-time block codes, which were introduced in [45], allow for full-rate transmission and partial diversity. Space-time trellis codes were introduced in [99], which provide a coding gain in addition to the diversity gain. The coding gain performance was also later improved in [6] and [46]. There is also the family of non-orthogonal space-time codes originally introduced in [100], which provide high rates and good cod-

ing performances. As for analog communication, there are fewer works available. However, one can still refer to [13], which uses complex block orthogonal space-time codes for channel estimation, but mainly mentions half-rate codes introduced in [98]. In this work however, that concept is extended to any general complex block orthogonal code.

Another interesting issue in Kalman filtering over fading channels is the issue of dimensionality, especially for uncoded transmission. There are processes for which the dimension is high and space-time coding might be too costly, e.g. in simple sensor networks. Also, a set of nodes in a sensor network may sometimes be modeled as a high dimensional source. It is very interesting to be able to have a performance metric on the estimation quality of the whole process/system for analog uncoded transmission. In such cases, tools from large system analysis and random matrix theory have gotten much attention in recent years [103]. While the application areas for digital communication are abundant and can be reviewed in [103, Ch. 3], there is not much available work on analog communication. It is however possible to mention [104] and [105] again, which have models that resemble uncoded transmission over fading channels. One of the targets of this thesis is to show the application and importance of tools from large system analysis for estimation of outputs from fading channels.

In the next section, a generic system model is provided, which is utilized in the rest of this thesis. The system model will be somewhat adapted for the different subproblems in each chapter, but the main blocks and their tasks will remain the same.

1.2 The Generic System Model

In this section and first, the operations of different parts of the system model (as depicted in Figure 1.1) are described. Then, the mathematical framework which we will deal with in the rest of this thesis is presented. It should however be noted that the mathematical formulas for this generic system model will be pruned and adapted to the different subproblems which are considered in each of the following chapters. Finally, the practical implications of these choices are discussed, and their benefits as well as possible shortcomings and limitations are considered.

The source $\mathbf{x}(n)$ is a vector signal whose elements may be correlated with each other and also in time. It is assumed that $\mathbf{x}(n)$ is a random vector of dimension K, which is considered to be drawn from a circular-symmetric (proper) complex Gaussian distribution. In addition, it is assumed that the source follows a first order ARGM model. The only exception is Ch. 6, in which Gaussian band-limited sources are considered.

The pre-filtering module $T(\cdot)$ represents the transformation on the source signal before transmission, and the result is called $\mathbf{x}_f(n)$. In this thesis, the notion of pre-filtering is applied in a more general sense, as the pre-filter represents power allocation, spectrum shaping and channel parallelization, but not lossy compression such as quantization. While power allocation and spectrum shaping are performed by linear operations, the pre-

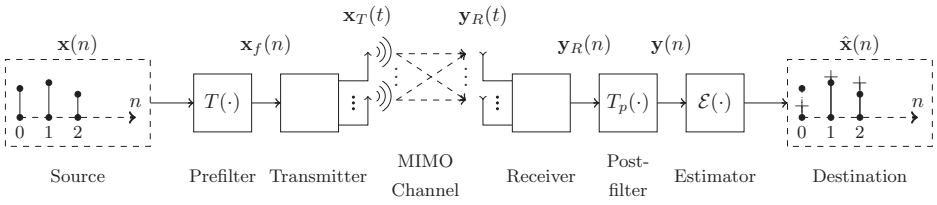

Figure 1.1: Generic system model for the suggested analog source transmission scheme

filter which is used to parallelize the MIMO channel in Chapter 4 is non-linear. It should be noted that using $T(\cdot)$ for spectrum shaping, which is discussed further in Chapter 6, might incur some delay. The filtered source $\mathbf{x}_f(n)$ is then made available to the transmitter in order to be sent over the channel.

The transmitter module performs the necessary operations in order to send the discrete-time continuous-valued transformed signal over the channel as the continuous-time waveform $\mathbf{x}_T(t)$. The operations include, but might not be limited to, pulse shaping and up-conversion for transmission over bandpass channels. In this work, no particular modulation type is chosen, as there exist several options, each with different design considerations. The possible choices are however discussed later on in this section. Regardless of the modulation scheme, the receiver, knowing the modulation type and the particular channel tap (the channel is assumed known), performs sampling, matched filtering with the overall transfer function of transmitted pulses and the channel, and finally extracts the now distorted information content.

The input to the channel is $\mathbf{x}_T(t)$, and the output at the receiver is $\mathbf{y}_R(t)$. The transformation from $\mathbf{x}_T(t)$ to $\mathbf{y}_R(t)$ is potentially prone to distorting phenomena such as noise, large scale and small scale fading, ISI and interference from other users. In this work, the focus is on the effects of small scale fading only. In that regard, it is assumed throughout this work to have non-frequency selective (flat), Rayleigh block fading channels, a common assumption in the analysis of many wireless communication systems. In addition, it is assumed throughout this thesis that the channel is known at the receiver, but not the transmitter.

The next module is the post-filter $T_p(\cdot)$, which is responsible to revert the pre-filtering operation. The result after this stage may be called $\mathbf{y}(n)$. The operation of this module may be considered as part of the estimator for certain operations, such as spectrum shaping, which only involves a linear transform. For space-time decoding however, the decoding operation is performed before the actual estimation procedure.

The last module in the system model is the estimator. The task of the estimator is to recover the transmitted signal $\mathbf{x}(n)$ from $\mathbf{y}(n)$. The output of this module is an estimated version of $\mathbf{x}(n)$, namely $\hat{\mathbf{x}}(n)$. The estimation is performed such that the MSE for the given random channel instance is minimized. With the Gaussian assumption, it is possible to show that the filter which minimizes the instantaneous MSE is linear, i.e. a linear

MMSE estimator will be used. For AR signals, the optimal linear MMSE filter is the Kalman filter. If there is no state-space model or if there is no correlation, a Wiener filter may be used. Both cases are addressed in this thesis. The Kalman filter is the filter of choice though Chapters 2 – 5, whereas the Wiener filter is considered in Chapter 6.

With the channel being random, the MSE is also a random variable for each channel use. The main focus of this thesis is then analyzing the instantaneous random MSE in accordance with the fading channel and design and analysis of diversity schemes in order to maximize the estimation quality.

In the following, the mathematical relations between the different variables in the system model are presented.

The most generic model which is considered for the source $\mathbf{x}(n)$ is

$$\mathbf{x}(n) = A\mathbf{x}(n-1) + \mathbf{u}(n). \tag{1.1}$$

In (1.1), $A \in \mathcal{R}^{K\times K}$ is the state-transition matrix, where $\mathcal{R}$ represents the space of real numbers. In accordance with $\mathbf{x}(n)$, the vector $\mathbf{u}(n)$ is a complex Gaussian random vector of dimension K, with zero mean and covariance matrix C_u.

The pre-filtering operation is denoted by

$$\mathbf{x}_f(n) = T(\mathbf{x}(n)), \tag{1.2}$$

and the fading channel effects by

$$\begin{aligned}\mathbf{y}_R(n) &= H(n)\mathbf{x}_f(n) + \mathbf{v}(n)\\ &= H(n)T(\mathbf{x}(n)) + \mathbf{v}(n).\end{aligned} \tag{1.3}$$

In (1.3), $H(n)$ is an $N \times K$ matrix representing the effect of a MIMO flat fading channel and drawn from a known pdf. Note that the conversion of $\mathbf{x}_f(n)$ to continuous-time signals at the transmitter and sampling of continuous-time received signals at the receiver also take place as an intermediate stage. Note that for (1.3) to be correct, Nyquist channels are required. In most of the following chapters, Rayleigh fading channels are considered. In that case, the elements of $H(n)$ are i.i.d. complex Gaussian random variables. In addition, the channel values over different blocks are assumed i.i.d. as well. To simplify the analysis, it is assumed as well that the channel is known at the receiver. This assumption is rather strong and can be costly to achieve in practice, e.g. it might require considerable overhead for channel estimation. However, this is a simplifying assumption which allows for development of new theory and thus provides a good starting point for understanding the performance of estimation over fading channels.

The receiver noise is also modeled by $\mathbf{v}(n)$, which is an $N \times 1$ vector of zero mean complex Gaussian random variables with the covariance matrix equal to C_v.

The instantaneous random MSE between the signal and its estimated version is defined as

$$M(n) = E\left(\|\mathbf{x}(n) - \hat{\mathbf{x}}(n)\|^2 | H(n)\right), \tag{1.4}$$

where the operator $E(\cdot)$ represents the statistical expected value (for each channel realization $H(n)$), and is taken with respect to the process and channel noises. The estimated version is then found such that the MSE is minimized, i.e.

$$\hat{\mathbf{x}}(n) = \underset{\mathbf{w}\in\mathcal{C}^K}{\operatorname{argmin}} E(\|\mathbf{x}(n) - \mathbf{w}\|^2), \tag{1.5}$$

which for the the linear MMSE filter simplifies to

$$\hat{\mathbf{x}}(n) = \mathbf{g}_k(n)\mathbf{y}(n), \tag{1.6}$$

where $\mathbf{g}_k(n)$ represents a linear filtering operation. The details of the linear estimation/filtering operation, namely the Kalman and Wiener filtering, and the filter forms are presented in their corresponding chapters.

In order to measure the estimation quality, we consider mainly the notion of estimation error outage probability defined as

$$P_{\text{out}}(M_{\text{th}}) = \Pr(M(n) \geqslant M_{\text{th}}) \tag{1.7}$$

for scalar sources, i.e. $K = 1$, and distortion outage probability defined as

$$P_{\text{out}}(d_{\text{th}}) = \Pr(d(n) \geqslant d_{\text{th}}) \tag{1.8}$$

for $K > 1$, with $d(n) = \dfrac{1}{K}\text{tr}(M(n))$, and $\text{tr}(\cdot)$ denoting the matrix trace operation. Note that for $K = 1$, both definitions are equal. The notation $d(n)$ is used for the normalized instantaneous MSE, namely the distortion, and the parameters M_{th} and d_{th} in (1.7) and (1.8) are arbitrary threshold values, corresponding to the quality of estimation. An outage probability of 10^{-2} e.g. when $M_{\text{th}} = 1$, shows that random MSE $M(n)$ is smaller than 1 with the probability of 0.99, as well as 99% of the time, when $M(n)$ is ergodic (in our settings, $M(n)$ is always ergodic). Also note that the outage probability $P_{\text{out}}(\cdot)$ is always a function of the threshold parameters M_{th} and d_{th}. However, it is mostly referred to as P_{out} for simplicity.

The SNR at the receiver can be obtained as

$$\text{SNR} = \frac{E(\|H(n)\|^2\|T(\mathbf{x}(n))\|^2)}{E(\|\mathbf{v}(n)\|^2)}. \tag{1.9}$$

As already mentioned in Section 1.1, the asymptotic high SNR behavior of P_{out} is of interest, as well as its finite SNR performance. When the outage probability has a diversity order of d_{ord} in the high SNR regime, it has the following form

$$P_{\text{out}} \approx (G \cdot \text{SNR})^{-d_{\text{ord}}} + o\left(\text{SNR}^{-d_{\text{ord}}}\right). \tag{1.10}$$

The parameters d_{ord} and G are referred to as the diversity order and the coding gain of the outage probability, respectively. This is parallel to the names used in digital communication for outage probabilities of error functions [111]. While the diversity order is the

slope of the outage probability function vs. SNR in the log-log scale, G, when given in dB, expresses the shift of the outage probability curve in SNR relative to the benchmark curve of $\text{SNR}^{-d_{\text{ord}}}$. Although the diversity order is a good rule of thumb for evaluating the reliability of estimation at high SNR, the coding gain provides a more complete characterization of performance, which allows for comparison of systems which have the same diversity order, but different coding gains.

Another reliability measure is the average MSE (or average distortion), defined as

$$\overline{\text{MSE}} = E\left(\frac{1}{K}\text{tr}\left(M(n)\right)\right). \tag{1.11}$$

The average MSE measure is an overall reliability measure which does not involve any threshold. The average MSE measure contains thus less information about the quality of estimation than the outage probability. However, this measure is considered in Chapter 5, where obtaining the outage probabilities becomes too complicated.

After presenting the details for the different parts of the system model and their mathematical representations, we now discuss the benefits as well as the shortcomings and challenges of the system model.

Considering the whole system model, the most significant benefit of analog uncoded transmission is its simplicity and low computational complexity, especially at the transmitter side, which makes it suitable for applications such as WSN. In addition, it is possible to have reliable delay-free communication under the settings discussed in this work, which makes it a promising scheme for emerging applications such as V2V. However, it should be noted that there are also some shortcomings associated with analog transmission compared to state-of-the-art digital schemes. It is e.g. much easier to arrange multiple access to the shared spectrum with digital schemes. In addition, analog communication is more prone to eavesdropping than digital communication, as there exist sophisticated cryptography techniques, which are only applicable for digital data. Consequently, while the analog communication scheme might be well-suited for certain application areas, it might be a poor choice for other applications and therefore should be selected carefully.

Regarding the source, the Gaussian assumption allows us to claim some optimality and convergence results in the next chapters. It is also commonly used in the literature for several reasons. To begin with, many natural phenomena can be modeled as Gaussian processes. Furthermore, the Gaussian pdf has only two independent parameters, mean and covariance, which in practice may be estimated very effectively. In addition, the Gaussian pdf has the maximum entropy for zero mean and given variance, thus providing an upper bound for the performance, e.g. when the underlying distributions are unknown. The complex signal model is used in order to make the system model as general as possible, yet all the results can easily be interpreted and adapted for real signals due to the circular symmetric assumption. Although the system model can in practice be extended to incorporate higher order AR processes, only the first-order AR processes are studied in order to simplify the analysis. The choice of real space (and not complex) for the elements of A (in (1.1)) is to avoid further complications regarding the optimal filters, as we

may then incorporate the linear Kalman filter. For a complex A and when optimality is of concern, a widely linear Kalman filter as introduced in [20], should be used. In case a complex-valued A is used, many of the algorithms and filters can still be applied without considerable modification, and the performance results are at worst an upper bound to the case if widely linear filters had been used.

Regarding the transmitter structure, it should be noted that the information bearing signal $\mathbf{x}_f(n)$ can be sent over the channel in different ways, e.g. as the amplitude, phase, frequency or width of the modulation pulses. Each method then leads to different practical design considerations and consequently different challenges. If the signal values are to be embedded as amplitude information, the result is a simple modified pulse amplitude modulation scheme. There, $\mathbf{x}_f(n)$ is the amplitude of the pulses which are to be sent over the channel. Note that as a Gaussian distribution (with theoretically unlimited support) is assumed for the source, this could lead to peak to average power problems in the transmitter if very high values are to be transmitted. Yet, the probability of very high amplitudes decreases exponentially with the square of the amplitude for Gaussian signals and application specific cut-off regions for the source amplitude can be selected. This may then lead to overload errors. However, if high enough cut-off threshold is used, the effect on the performance is negligible. The receiver's operation would then be limited to recovering the amplitude from the channel output wave-form $\mathbf{y}_R(t)$ in order to obtain the discrete time received vector $\mathbf{y}_R(n)$. The information content can also be transmitted in the signal phase, which does not have the peak to average power ratio problem at the transmitter. Nevertheless, it has the limitation that the mapping from Gaussian signal values to a phase limited to the interval $[0, 2\pi)$ is theoretically lossy. This problem might be overcome if the information is mapped to the frequency of the pulses instead. This is however limited by the available bandwidth. The information content may also be mapped to the width of the pulses, constituting a pulse width modulation (PPM) scheme. Another option is to map the information to the position of some known pulses, constituting a modified pulse position modulation (PPM) for analog communication. In general, PPM is more power efficient compared to PAM, but requires more bandwidth in order to deliver the same performance.

In this work, the main focus is a point-to-point communication scheme. As a result, the effects of interference from other possible users are not studied. It should however be noted that for non-cellular communication, bandwidth requirements, e.g. in sensor networks, are usually not as stringent as in cellular communications. As a result, interference might be mitigated by allocating different frequency bands to different users. If a feedback channel is allowed, the carrier frequencies can be decided during some training phase. If not, the carrier frequencies need to be pre-defined for the transmitters. Disregarding the inter-user interference, one can focus on the effects of fading only. For simplicity, only small scale fading is considered. If the channel is flat as assumed in this thesis, then there is no inter-symbol interference, and the analysis can be simplified even more. In practice, this assumption limits the allowed source bandwidth up to the fading channel's coherence bandwidth. This should be considered while selecting the modulation type. In addition, it may limit the application to sources which have low bandwidth. In this work, the i.i.d.

block fading model is considered, i.e. the channel is constant over a block of source symbols and that different blocks are independent and identically distributed (i.i.d.). For most of the work, only a block size of one is considered in order to allow for the development of new theory, as this assumption is a sufficient (but not necessary) condition for some of the theoretical analysis. It is true that this assumption might be somewhat limiting in practice. However, such limitation may be overcome, if needed, at the cost of more hardware complexity[2]. At worst, the i.i.d. assumption for source block lengths of one, may limit the transmission's (symbol) rate. Finally, assuming known channel at the receiver requires that the channel taps should be estimated with estimation algorithms that can provide very high accuracy. As we have mentioned the possibility of the transmitters using multiple frequency bands or adapting their carrier frequencies to avoid inter-user interference, we have inherently assumed that a feedback channel from the receiver (may be a control center) to the transmitting node might exist. If so, this allows for full or partial channel knowledge at the transmitter. However, the channel at the transmitter side is not assumed known, because that would require ongoing channel estimation at the nodes, which is costly and prohibitive, given the assumed simple structure for the transmitter. It is true nevertheless, that full or partial channel information at the transmitter may in general improve the performance.

Given the generic system model, there are several subproblems which are studied in this thesis and for which new or improved theory and solutions are presented. The next section provides an overview of the structure of this thesis and the major contributions of this work.

1.3 Thesis Structure and Contributions

In Chapter 2, we study the case when measurements of a scalar Gauss-Markov process are sent over a scalar fading channel (the SISO model) with i.i.d. channel realizations. This model best suits e.g. low-cost sensor networks with processing at the fusion center. The major contributions of this chapter are as follows.

- It is shown that for general i.i.d. fading channels, the first order pdf for the instantaneous MSE may be obtained through a recursive integral equation.
- For the case of Rayleigh fading channel, upper and lower bounds for the outage probability are provided. It is then shown that the bounds are tight in the high SNR regime.
- It is shown that the outage probability decreases linearly with inverse of the channel SNR in the high SNR regime.

The material in Chapter 2 is based on the following publication ([70])

[2]This can e.g. be done by switching the carrier frequencies or using multiple antennas in order to obtain independent channels.

(I) Reza Parseh and Kimmo Kansanen. On estimation error outage for scalar Gauss–Markov signals sent over fading channels. *IEEE Transactions on Signal Processing*, 62(23):6225–6234, Dec. 2014.

In Chapter 3, the focus is on the case where first-order scalar Gauss-Markov signals are sent over d parallel i.i.d. Rayleigh fading channels (the SIMO model). The major contributions of this chapter are as follows.

- A closed-form integral equation for the pdf and outage of the instantaneous estimation error variance is obtained for a certain range of thresholds.
- It is shown that in the high SNR regime, the estimation error outage probability has a diversity order of d.
- Upper and lower bounds for the outage probability are developed, which provide a more complete characterization of the performance, especially in the medium SNR regime.
- The analysis of the diversity order is improved by considering the effect of imperfect system parameter knowledge. It is shown that under certain conditions, the diversity order is not changed by using a wrong or mismatched parameter in the Kalman filter.

The material in Chapter 3 is based on the following publication ([71])

(II) Reza Parseh and Kimmo Kansanen. Diversity effects in Kalman filtering over Rayleigh fading channels. *IEEE Transactions on Signal Processing*, 63(23):6329 – 6342, Dec. 2015.

In Chapter 4, the general case of transmission of a vector signal over a general fading channel (the MIMO model) is considered. The major contributions of this chapter are as follows.

- Complex block orthogonal space-time codes are employed and adapted to be used in the Kalman filtering setting by performing the decoding operation in separate real and imaginary parts of the signal.
- The joint space-time and Kalman filtering scheme is analyzed in terms of the high SNR behavior of the distortion outage probability, and it is shown that this scheme can achieve the maximum diversity order for transmission over a $N \times K$ MIMO fading channel, i.e. KN.
- Upper and lower bounds for the outage probability are provided, which can be used for a wide range of system parameters.
- Upper and lower bounds are derived for the coding gain of the outage probability, which provide a more complete high SNR analysis.

The material in Chapter 4 is based on the following publication ([72])

(III) Reza Parseh, Kimmo Kansanen, and Dirk Slock. Distortion outage analysis for joint space-time coding and Kalman filtering. *Submitted to IEEE Transactions on Signal Processing*, 2016.

In Chapter 5, we consider the case when an infinite number of Gauss-Markov signals send their measurements over a fading channel using analog transmission (large system regime). The Kalman filter is used at the receiver. The criterion for estimation quality assessment in this chapter is the average MSE. The major contributions of this chapter are as follows.

- Via making a high SNR assumption, the RRE is approximated with a new equation which leads to the development of an approximation for the eigenvalue distribution of the estimation error covariance matrix.
- The approximated pdf is used to further approximate the first moment of the eigenvalue distribution, corresponding to the mean estimation MSE.
- The high SNR performance of the mean estimation MSE is derived as a function of SNR and the number of transmit and receive signal dimensions.

The material in Chapter 5 is based on the following publication ([73])

(IV) Reza Parseh, Dirk Slock, and Kimmo Kansanen. Mean estimation MSE for Kalman filtering of large dimensional sources sent over fading channels. In *IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 499–503, June 2014.

Finally in Chapter 6, transmission of a band-limited signal over a MIMO fading channel is considered and the estimation error outage is incorporated as a measure for quality assessment. Diversity order analysis is then performed in the high SNR regime. It is assumed that each K source samples can be transmitted over N separate channels. There, it is shown that for parallel channels, if a linear MMSE filter, i.e. Wiener filter, is used at the receiver, the estimation error outage probability vanishes with the $(N-K+1)$'st power of SNR for the high SNR regime, and thus achieves a diversity order of $N-K+1$.

The material in Chapter 6 is based on the following publication ([74])

(V) Reza Parseh, Dirk Slock, and Kimmo Kansanen. Oversampling diversity for uncoded transmission of bandlimited sources over parallel fading channels. In *IEEE International Conference on Communications (ICC)*, pages 4653 – 4658, 2015.

In addition and as part of the research, the following papers were published within the topic of estimation and coding over channels and specifically for smart grid applications, and which are mentioned here for future reference ([69], [119], [120]).

(I) Reza Parseh, Santiago Sanchez Acevedo, Kimmo Kansanen, Marta Molinas, and Tor A Ramstad. Real-time compression of measurements in distribution grids. In *IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, pages 223–228, 2012.

(II) Mehdi K. Zadeh, Reza Parseh, Marta Molinas, and Kimmo Kansanen. Bifurcation in PWM converter-based systems with wireless communication-based current controller. In *4th IEEE/PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, pages 1–5, 2013.

(III) Mehdi K. Zadeh, Reza Parseh, Marta Molinas, and Kimmo Kansanen. Modeling and simulation of wireless communication based robust controller for multiconverter systems. In *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 738–743, 2013.

Chapter 2

Scalar Signals and a Single Receiver

In this chapter, transmission of a scalar ARGM signal over a scalar fading channel ($N = K = 1$) is considered. The main reason to select a scalar model for the source and channel as a starting point for the whole work, is that it significantly simplifies the problem. This allows us to provide insightful results for the outage probability. Vector sources and multiple antennas will be considered in the following chapters. The fading channel is modeled as i.i.d. distributed random variables with known realization at the receiver. As we will see later in the chapter, the i.i.d. assumption is a sufficient condition for convergence of the steady-state pdf of the outage event. In the absence of a necessary condition, the i.i.d. assumption is kept for convenience. The optimal estimator at the receiver is the Kalman filter. The first part of this chapter deals with characterizing the stationary probability density function of the random MSE of the Kalman filter. There, it is shown that the first order pdf may be obtained through a recursive integral equation. Furthermore and for the particular case of the i.i.d. Rayleigh fading channels, upper and lower bounds for the outage probability are derived. It is also shown that the bounds are tight in the high SNR regime, and that the outage probability decreases linearly with the inverse of the average SNR. This suggests a diversity order of one for the outage probability. In addition, the coding gain is obtained for the outage probability and shown to be a function of the source parameters and the threshold.

2.1 Updated System Model and Problem Definition

We update the system mode for the $N = K = 1$ case as follows

$$\begin{aligned} x(n) &= \rho x(n-1) + u(n),\ n \geq 1,\ x(0) \sim \mathcal{CN}(0, M(0)) \\ y(n) &= h(n)x(n) + v(n). \end{aligned} \tag{2.1}$$

In (2.1), $u(n)$ and $v(n)$ are scalar white circularly symmetric complex Gaussian random variables with variances $\sigma_u^2 > 0$ and $\sigma_v^2 > 0$, respectively. Consider $h(n)$ to be i.i.d.

samples of a random variable (starting from Sec. 2.2.2, the channel will be considered as Rayleigh fading). It is also assumed that $h(n)$ cannot be equal to zero for all n (non-existent channel is not included). It is assumed that perfect knowledge of the random channel $h(n)$ is available at the receiver and $h(n)$ are also independent of $u(n)$ and $v(n)$. Note that ρ is a system parameter which imposes correlation on $x(n)$ and replaces the matrix A in the system model of Chapter 1. In addition, the transform $T(\cdot)$ only accounts for power allocation (constant for given SNR), and is combined within the definition of $h(n)$ for simplicity. For further development in Sec. 2.2.1, it is required that $h(1) \neq 0$ and $\rho \neq 0$. The objective at the receiver is optimal estimation of the signal $x(n)$, given the channel outputs.

Given the previous assumptions, and with $h(n)$ independent of $u(n)$ and $v(n)$, the optimal MMSE estimator of $x(n)$ based on the observations $y(n)$ is the scalar Kalman filter with the following formulation adapted from [52]

$$\hat{x}(n|n-1) = \rho\hat{x}(n-1|n-1) \tag{2.2}$$
$$P(n) = \rho^2 M(n-1) + \sigma_u^2 \tag{2.3}$$
$$g_k(n) = P(n)h^*(n)\left(\sigma_v^2 + |h(n)|^2 P(n)\right)^{-1} \tag{2.4}$$
$$\hat{x}(n|n) = \hat{x}(n|n-1) + g_k(n)(y(n) - h(n)\hat{x}(n|n-1)) \tag{2.5}$$
$$M(n) = \left(1 - g_k(n)h(n)\right) P(n). \tag{2.6}$$

Concisely stated, eq. (2.2) is the prediction of the current state based on the previous estimated state (*a priori estimate*) using the system model and eq. (2.3) is the instantaneous prediction error variance (with respect to noise) $P(n)$. Equation (2.4) is the corresponding Kalman gain equation and eq. (2.5) is the correction equation based on the Kalman gain update (*a posteriori estimate*). Finally eq. (2.6) provides us with the instantaneous estimation error variance (IEV) denoted by $M(n)$.

It is straightforward to show that both $P(n)$ and $M(n)$ may be written recursively in terms of their previous values and current value of $h(n)$, where one is a deterministic function of the other for deterministic system parameters. The statistical properties of the one may then be acquired using the statistical properties of the other. The recursion for $P(n)$ is a well-defined and studied RRE. However, the focus of this work is on the instantaneous MSE, and the interest is mainly in obtaining $M(n)$.

The recursion for $M(n)$ is obtained as follows

$$\begin{aligned} M(n) &= \left(1 - g_k(n)h(n)\right) P(n) \\ &= \left(1 - \frac{P(n)|h(n)|^2}{\sigma_v^2 + |h(n)|^2 P(n)}\right) P(n) \\ &= P(n) - \frac{P^2(n)|h(n)|^2}{\sigma_v^2 + |h(n)|^2 P(n)} \\ &= \frac{P(n)}{1 + |h(n)|^2 P(n)/\sigma_v^2}, \end{aligned} \tag{2.7}$$

which can be simplified to

$$M(n) = \frac{\rho^2 M(n-1) + \sigma_u^2}{1 + \gamma(n)\left(\rho^2 M(n-1) + \sigma_u^2\right)}, \tag{2.8}$$

by setting $P(n) = \rho^2 M(n-1) + \sigma_u^2$ and taking $\gamma(n) = |h(n)|^2/\sigma_v^2$.

In this setting, the estimation error outage probability (EOP) for time n is equal to

$$P_{\text{out}}^n(M_{\text{th}}) = \Pr(M(n) \geq M_{\text{th}}). \tag{2.9}$$

The asymptotic EOP is then equal to

$$P_{\text{out}}(M_{\text{th}}) = \lim_{n\to\infty} P_{\text{out}}^n(M_{\text{th}}) = \lim_{n\to\infty} \Pr(M(n) \geq M_{\text{th}}). \tag{2.10}$$

Clearly $P_{\text{out}}^n(M_{\text{th}}) = 1 - F_{M(n)}(M_{\text{th}})$ and $P_{\text{out}}(M_{\text{th}}) = 1 - F_M(M_{\text{th}})$, where $F_{M(n)}(M)$ $\big(F_M(M)\big)$ is the (steady state) cdf of $M(n)$. It is easy to verify that the source power for $|\rho| < 1$ is equal to $\sigma_x^2 = P_x = \dfrac{\sigma_u^2}{1-\rho^2}$ and the SNR at the receiver is consequently equal to $\text{SNR} = \dfrac{P_x E(|h(n)|^2)}{\sigma_v^2}$. In addition to the outage values, we are also interested in the high SNR (asymptotic) behavior of the outage probability.

2.2 Statistical Properties of the IEV

In this section, the steady-state probability density function of the IEV is studied. Because $F_M(M)$ and $f_M(M)$ are related with a linear operation (derivative), one can begin to study $f_M(M)$. After that, the EOP will readily be obtained with one integration operation.

2.2.1 Steady-state Pdf of The IEV

Given (2.8), it is easy to verify that for any arbitrary positive real number M, $M(n) \leqslant M$ leads to

$$\gamma(n) \geqslant \frac{1}{M} - \frac{1}{\rho^2 M(n-1) + \sigma_u^2}.$$

Also, we have that $\gamma(n) \geqslant 0$ and $0 < M(n) < M_{\text{max}}$, where M_{max}, the upper limit for $M(n)$ is obtained from

$$M_{\text{max}} = \begin{cases} \infty, & |\rho| \geqslant 1 \\ P_x, & |\rho| < 1. \end{cases} \tag{2.11}$$

$M_{\max}$ is effectively the estimation error variance for the worst channel, i.e. $\gamma(n) = 0$ with probability 1. In that case, the best estimator is the average mean, i.e. $\hat{x}(n) = E(x(n)) = 0$ and therefore the estimation error variance is equal to $\sigma_x^2 = P_x$. Also note that the case $|\rho| \geqslant 1$ is of little practical importance in our case, because for a divergent signal, continuous-amplitude uncoded transmission would not be practical, as the transmission would require infinite SNR. It is however included in the definition of $M_{\max}$ to show that the analysis does not depend on the value of ρ.

Given the above limits and conditions for $\gamma(n)$ and $M(n)$ and according to [68], it is possible to get the cdf of $M(n)$, i.e. $F_{M(n)}(M)$ as follows.

$$F_{M(n)}(M) = \int_0^{M_{\max}} \int_{\frac{1}{M} - \frac{1}{\rho^2 M(n-1) + \sigma_u^2}}^{\infty} f_{\gamma(n), M(n-1)}\big(\gamma(n), M(n-1)\big)\, \mathrm{d}\gamma(n)\, \mathrm{d}M(n-1), \quad (2.12)$$

where $f_{\gamma(n),M(n-1)}\big(\gamma(n), M(n-1)\big)$ is the joint pdf of $\gamma(n)$ and $M(n-1)$. The pdf for $M(n)$ is then obtained by simply applying $f_{M(n)}(M) = \frac{\partial}{\partial M} F_{M(n)}(M)$. That leads to

$$\begin{aligned} & f_{M(n)}(M) \\ &= \int_0^{M_{\max}} \frac{\partial}{\partial M} \int_{\frac{1}{M} - \frac{1}{\rho^2 M(n-1) + \sigma_u^2}}^{\infty} f_{\gamma(n), M(n-1)}\big(\gamma(n), M(n-1)\big)\, \mathrm{d}\gamma(n)\, \mathrm{d}M(n-1) \\ &= \int_0^{M_{\max}} \frac{1}{M^2} f_{\gamma(n), M(n-1)}\left(\frac{1}{M} - \frac{1}{\rho^2 M(n-1) + \sigma_u^2}, M(n-1)\right) \mathrm{d}M(n-1), \quad (2.13) \end{aligned}$$

or with some change of notation,

$$f_{M(n)}(M) = \int_0^{M_{\max}} \frac{1}{M^2} f_{\gamma(n), M(n-1)}\left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2}, m\right) \mathrm{d}m. \quad (2.14)$$

Now if we assume that $\gamma(n)$ is independent of $M(n-1)$ $\big(\gamma(n) \perp\!\!\!\perp M(n-1)\big)$, we may rewrite (2.14) as

$$f_{M(n)}(M) = \frac{1}{M^2} \int_0^{M_{\max}} f_{\gamma(n)}\left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2}\right) f_{M(n-1)}(m)\, \mathrm{d}m. \quad (2.15)$$

Note that as we have assumed i.i.d. channels, then we have that $\gamma(n) \perp\!\!\!\perp \gamma(i)$ for $i < n$. It can simply be assumed that $\gamma(i) \perp\!\!\!\perp M(0)$ ($M(0)$ is a constant). As a result, we obtain that $\gamma(n) \perp\!\!\!\perp M(n-1)$ because $M(n-1)$ is a function of $M(0)$ and $\gamma(1), \gamma(2), \cdots, \gamma(n-1)$ only. Thus i.i.d. channel assumption is a sufficient condition to get the main result in (2.15). As i.i.d. channels are assumed, all $\gamma(n)$ have the same pdf, i.e. $f_{\gamma(n)}(\gamma(n)) = f_\gamma(\gamma(n))$. Therefore, it is possible to rewrite (2.15) to obtain

$$f_{M(n)}(M) = \frac{1}{M^2} \int_0^{M_{\max}} f_\gamma\left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2}\right) f_{M(n-1)}(m)\, \mathrm{d}m. \quad (2.16)$$

Equation (2.16) may be used to find $f_{M(n)}(M)$ for any n by starting from $f_{M(0)}(M) = \delta(M - M(0))$ and iterating over n. However, the main objective is outage analysis and for that purpose, the steady-state distribution is needed. In the following, Theorem 1 is presented, which proves the existence of a steady-state distribution for $M(n)$, namely $f_M(M)$ and outlines how it can be obtained.

Theorem 1. *The random process $M(n)$ converges in law and has a steady-state distribution, namely $f_M(M)$ which satisfies the following equality*

$$f_M(M) = \frac{1}{M^2} \int_0^{M_{\max}} f_\gamma \left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2} \right) f_M(m) \, \mathrm{d}m. \tag{2.17}$$

Proof. In order to prove Theorem 1, we refer to [10] which considers Kalman filtering with random coefficients. In [10], a general vector state-space Kalman filter comprises the system model, which can be shown to include our system model as well. There and based on a contraction property of the Kalman filter, it is proven that the sequence of estimation error covariance matrices converges in law to a stationary process [10, Theorem 2.4], given that some ergodicity conditions are met. In the following, it is shown in Lemma 1 that those conditions are also met in our system model. As a result, the equivalent random variable in our case, i.e. $M(n)$ also converges in distribution and therefore $f_M(M)$ in fact exists. Then, it is shown in Lemma 2 that $f_M(M)$ can in fact be obtained from (2.17). □

Lemma 1. *The instantaneous estimation error variance $M(n)$ converges in distribution.*

Proof. According to [10, Theorem 2.4], three conditions are required for convergence of the estimation error covariance matrix of the Kalman filter with stochastic system parameters. Firstly, a hypothesis $\mathcal{H}$ (introduced in [10, Section 2]) should be satisfied. Secondly, it is required that the system is weakly observable and weakly controllable as defined in [10, Definition 2.1] (also see Remark 1). Thirdly, certain (random) system parameters, specified later, should be integrable.

For the first condition, it is mentioned in [10, Section 2] that a conditionally Gaussian system satisfies hypothesis $\mathcal{H}$. In our system, $u(n)$ and $v(n)$ are i.i.d. Gaussian random variables and also independent of $h(n)$. Therefore, our system is also conditionally Gaussian and satisfies the condition.

For the second condition, it must be shown that the system is weakly controllable and weakly observable. Using the definition for weak controllability, it must then be shown that

$$\Pr\left(\sigma_u^2 + \rho^2\sigma_u^2 + \rho^4\sigma_u^2 + \ldots + \rho^{2n}\sigma_u^2 \neq 0\right) > 0, \tag{2.18}$$

which obviously holds as long as $\sigma_u^2 > 0$. For weak controllability, we must show that

$$\Pr\left(\rho^2\gamma(1) + \rho^4\gamma(2) + \ldots + \rho^{2n}\gamma(n) \neq 0\right) > 0. \tag{2.19}$$

It is possible to show that (2.19) holds for all channel distributions apart from the non-existent channel ($h(n) = 0$ for all n). Therefore, the second condition for convergence is also met.

For the third condition to hold, it should be such that the (random) variables $\log\log^+(\rho)$, $\log\log^+(\rho^{-1})$, $\log\log^+(\sigma_u^2)$ and $\log\log^+(h(1))$ are integrable, where

$$\log^+(x) = \max\left(\log(x), 0\right), \tag{2.20}$$

i.e. they have a well-defined expectation value (see e.g. [112] Chapter 13 for a definition of integrable random variables). Obviously, $\rho \neq 0$ and $\sigma_u^2 > 0$ are deterministic variables. Therefore, they are integrable. $\log\log^+(h(1))$, is also integrable, given that $h(n)$ is defined as in Sec. 2.1. As a result, our system model satisfies all the prerequisites of Theorem 2.4 in [10]. The consequence of the aforementioned theorem is that $M(n)$ converges in distribution (law) and therefore $f_M(M)$ exists. □

Remark 1. The exact definitions for weak controllability and weak observability are provided in [10, Definition 2.1] and for a vector system model, similar to the generic system model in Chapter 1 of this thesis. In order to define weak controllability and weak observability, the following probabilities are defined first.

$$\epsilon_o = \Pr\left(\text{Det}\left(A^T\|H(1)\|^2A + (A^T)^2\|H(2)\|^2A^2 + \ldots + (A^T)^n\|H(n)\|^2A^n\right) > 0\right)$$
$$\epsilon_c = \Pr\left(\text{Det}\left(C_u + AC_uA^T + A^2C_u(A^T)^2 + \ldots + A^nC_u(A^T)^n\right) > 0\right),$$

where Det(.) represents the determinant of a matrix. For the scalar system model, the above probabilities simplify to the ones given in 2.18 and 2.19.

For weak observability, it must hold that ϵ_o is non-zero. For weak controllability, it must hold that ϵ_c is non-zero. Weak controllability and weak observability are concepts which are similar to the widespread concepts of observability and controllability for deterministic systems. However, weak controllability and weak observability are probabilistic definitions and thus are "weaker" conditions than those for deterministic systems, i.e. it is in general easier to satisfy them.

Lemma 2. *The steady-state distribution of $M(n)$ can be obtained from*

$$f_M(M) = \frac{1}{M^2}\int_0^{M_{\max}} f_\gamma\left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2}\right) f_M(m)\,\mathrm{d}m. \tag{2.21}$$

Proof. We use the fact that $M(n)$ is a Markov process. This is due to the fact that $M(n)$ is determined by only $M(n-1)$ and $\gamma(n)$. It was also shown in Lemma 1 that $M(n)$ converges in distribution. From the theory of Markov processes in [40] and utilizing the relationship between $f_{M(n-1)}(M)$ and $f_{M(n)}(M)$ in (2.16), it can be verified that $M(n)$ has a transition probability measure (function) equal to $\frac{1}{M^2}f_\gamma\left(\frac{1}{M} - \frac{1}{\rho^2 m+\sigma_u^2}\right)$. One can then refer to Theorem 2.3.5 (ii) in [40] and conclude that

$$f_M(M) = \frac{1}{M^2}\int_0^{M_{\max}} f_\gamma\left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2}\right) f_M(m)\,\mathrm{d}m, \tag{2.22}$$

as stated in (2.17). □

After stating Theorem 1, (2.17) will be utilized for the rest of the analysis in this chapter in order to characterize $f_M(M)$. To be more specific with (2.17), we use the fact that $\gamma(n) \geqslant 0$. That necessitates that the argument of the function $f_\gamma(\cdot)$ should always be positive. Clearly, if $M \leqslant \sigma_u^2$, the term $\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2}$ is always positive. However for $M > \sigma_u^2$, the integral should be taken over the range of γ where $\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2} \geqslant 0$, i.e. for $m \geqslant \frac{M - \sigma_u^2}{\rho^2}$. With this background, we can finally provide the following lemma that describes the asymptotic pdf of $M(n)$, i.e. $f_M(M)$ in terms of itself integrated with a kernel which is a function of $\gamma(n)$. Solving this equation leads to $f_M(M)$ and with one integration to P_{out}, which is the target.

Lemma 3. *Asymptotic pdf of $M(n)$, i.e. $f_M(M)$ can be obtained from the following equation*

$$f_M(M) = \begin{cases} \frac{1}{M^2} \int_0^{M_{\max}} f_\gamma \left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m, & M \leqslant \sigma_u^2 \\ \frac{1}{M^2} \int_{\frac{M-\sigma_u^2}{\rho^2}}^{M_{\max}} f_\gamma \left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m, & M > \sigma_u^2. \end{cases} \tag{2.23}$$

The solution is general and is explicitly given in terms of the pdf of $\gamma(n)$ and other system parameters. Though (2.23) can be solved numerically if needed, the general closed-form solution does not seem to be readily attainable. In the following, the focus is on the important case of Rayleigh fading channels where $f_\gamma(\gamma) = \lambda \mathrm{e}^{-\lambda\gamma} \mathcal{U}(\gamma)$ $\big(\mathcal{U}(\cdot)$ is the unit step function$\big)$. Note that with this definition, $\lambda = 1/E(\gamma(n)) = \sigma_v^2 / E(|h(n)|^2)$, i.e. stronger channels yield smaller values for λ and vice versa. Also that the average SNR at the receiver is equal to $\text{SNR} = \dfrac{\mathrm{P}_x}{\lambda}$.

2.2.2 Outage Probability for Rayleigh Fading Channels

We can rewrite (2.23) given that channel is i.i.d. Rayleigh fading. Using the Rayleigh fading assumption leads to

$$f_M(M) = \frac{\lambda}{M^2} \exp(\frac{-\lambda}{M}) \begin{cases} \int_0^{M_{\max}} \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m, & M \leqslant \sigma_u^2 \\ \int_{\frac{M-\sigma_u^2}{\rho^2}}^{M_{\max}} \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m, & M > \sigma_u^2, \end{cases} \tag{2.24}$$

which in order to get more insight and with some algebraic manipulation, can also be written as

$$f_M(M) = \frac{\lambda}{M^2} \exp(\frac{-\lambda}{M}) \begin{cases} \kappa, & M \leqslant \sigma_u^2 \\ \kappa - \int_0^{\frac{M-\sigma_u^2}{\rho^2}} \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m, & M > \sigma_u^2, \end{cases} \tag{2.25}$$

where

$$\kappa = \int_0^{M_{\max}} \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m. \tag{2.26}$$

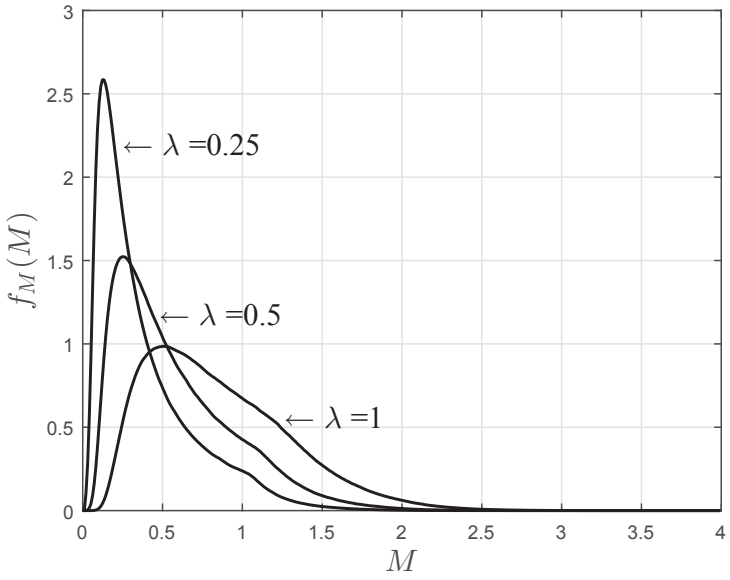

Figure 2.1: Asymptotic pdf of $M(n)$ for $\sigma_u^2 = \sigma_v^2 = 1$, $\rho = 0.95$, $\lambda = 1, 0.5, 0.25$.

Though in general κ depends on the pdf itself, (2.25) is insightful in the sense that it shows the exact shape of the pdf for the first part where $M \leqslant \sigma_u^2$. Typical shapes of such pdf's which are obtained through Monte-Carlo simulations are depicted in Fig. 2.1 to further highlight the points mentioned. For Fig. 2.1, it is assumed that $\sigma_u^2 = \sigma_v^2 = 1$, $\lambda = 1, 0.5, 0.25$ and $\rho = 0.95$. This leads to $\mathrm{P}_x = 10.11$ dB and average SNR values equal to 10.11, 13.11, and 16.11 dB, respectively. Note that the pdf support is theoretically bounded in this case at point $M_{\max} = \sigma_u^2/(1-\rho^2) \approx 10.26$ (not shown in the figure due to its insignificance). Also note that the break point, $M = \sigma_u^2$ where the pdf changes shape is quite visible in Fig. 2.1.

The break point $M = \sigma_u^2$ corresponds to the steady-state variance of the signal when there is no correlation ($\rho = 0$), whereas the point $M_{\max} = \sigma_u^2/(1-\rho^2)$ corresponds to the upper limit for the support of $f_M(M)$ (maximum value for the IEV) for the worst channel ($\gamma(n) = 0$) when no information gets passed the channel and the estimator is equal to $\hat{x}(n) = E(x(n)) = 0$. It is quite visible and theoretically verifiable that the pdf tail vanishes rapidly after the break point if the SNR increases. Also that the higher the threshold, the lower the outage probability would be. As a result, getting bounds on the first part helps with understanding the pdf's behavior better and at the same time get approximate values and bounds for $P_{\text{out}}(M_{\text{th}})$. Using (2.25) and (2.26), it is possible to find upper and lower bounds for κ, approximations for the pdf and upper and lower bounds for the outage for the first part ($M \leqslant \sigma_u^2$). Another insight from (2.26) is that the pdf shape is independent of whether the system is stable ($\rho < 1$) or unstable ($\rho \geqslant 1$), though the value of κ

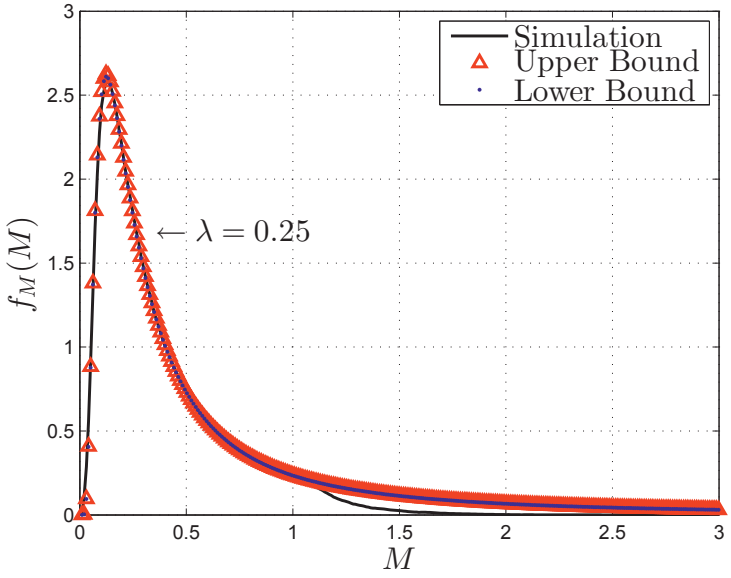

Figure 2.2: Asymptotic pdf of $M(n)$ and its approximates using upper and lower bounds for κ given that $\sigma_u^2 = \sigma_v^2 = 1$, $\rho = 0.95$, $\lambda = 0.25$ (SNR = 16.11 dB).

depends on ρ. Though the pdf is given by the equation $f_M(M) = \frac{\kappa\lambda}{M^2}\exp(\frac{-\lambda}{M})\,(M \leqslant \sigma_u^2)$, the exact value of κ depends on the whole pdf and cannot be known without solving (2.25). However, it is possible to obtain the following bounds for κ, namely $\kappa_l < \kappa < \kappa_u$, which later on are used to characterize two functions $P_{\text{out}}^l(M_{\text{th}})$ and $P_{\text{out}}^u(M_{\text{th}})$ for which $P_{\text{out}}^l(M_{\text{th}}) < P_{\text{out}}(M_{\text{th}}) < P_{\text{out}}^u(M_{\text{th}})$ for all $M \leqslant \sigma_u^2$.

Lemma 4. *For all $M \leqslant \sigma_u^2$, we have $\kappa_l < \kappa < \kappa_u$, where*

$$\kappa_u = \frac{1}{\left(a_\kappa \exp\left(\frac{-\lambda}{\sigma_u^2(1+\rho^2)}\right) + \exp(-\frac{\lambda}{\sigma_u^2})\right)} \tag{2.27}$$

$$\kappa_l = \frac{1}{\left(a_\kappa \exp(\frac{-\lambda}{\rho^2 M_{\max} + \sigma_u^2}) + \exp(-\frac{\lambda}{\sigma_u^2})\right)}, \tag{2.28}$$

where it was defined that

$$a_\kappa = 1 - \int_0^{\sigma_u^2} \exp(\frac{\lambda}{\rho^2 m + \sigma_u^2})(\frac{\lambda}{m^2})\exp(\frac{-\lambda}{m})\,\mathrm{d}m. \tag{2.29}$$

Proof. See Appendix A. □

Note that for stable systems, $M_{\max} = \mathrm{P}_x = \frac{\sigma_u^2}{1-\rho^2}$ and not surprisingly $\rho^2 M_{\max} + \sigma_u^2 = M_{\max}$, which is the steady-state equation for $\sigma_x^2 = \mathrm{P}_x = M_{\max}$. Following that, we obtain

$$\kappa_l^b = \frac{1}{\left(a_\kappa \exp(\frac{-\lambda}{M_{\max}}) + \exp(-\frac{\lambda}{\sigma_u^2})\right)} \tag{2.30}$$

For unstable systems we have $M_{\max} \to \infty$, and as a result

$$\kappa_l^\infty = \frac{1}{\left(a_\kappa + \exp(-\frac{\lambda}{\sigma_u^2})\right)}. \tag{2.31}$$

To show how useful the bounds for κ are, the simulated pdf along with two approximates using the bounds for κ are plotted in Fig. 2.2, given that $\sigma_u^2 = \sigma_v^2 = 1$, $\lambda = 0.25$ (SNR = 16.11 dB), and $\rho = 0.95$.

With Lemma 4 at hand, we are now ready to present the bounds for $P_{\text{out}}(M_{\text{th}})$. It is then shown that the bounds are tight for the high SNR regime, i.e. $\lambda \to 0$. This is discussed in the next section.

2.3 Upper and Lower Bounds for Outage Probability

In this section, we first get upper and lower bounds for $P_{\text{out}}(M_{\text{th}})$. It is then shown that the bounds are tight in the high SNR regime. In addition, it is shown that for a given non-zero M_{th}, EOP decreases with inverse of the average SNR, i.e. the outage probability has a diversity order of one. The coding gain can also be calculated as a result.

As defined before, $P_{\text{out}}(M_{\text{th}})$ is given by

$$P_{\text{out}}(M_{\text{th}}) = \int_{M_{\text{th}}}^{M_{\max}} f_M(M)\,\mathrm{d}M. \tag{2.32}$$

For $M \leqslant \sigma_u^2$, we get

$$P_{\text{out}}(M_{\text{th}}) = \int_{M_{\text{th}}}^{M_{\max}} \frac{\kappa\lambda}{M^2}\exp(\frac{-\lambda}{M})\,\mathrm{d}M = 1 - \kappa\exp(\frac{-\lambda}{M_{\text{th}}}). \tag{2.33}$$

As shown in the previous section, $\kappa_l < \kappa < \kappa_u$. As a result, we get

$$1 - \kappa_u\exp(\frac{-\lambda}{M_{\text{th}}}) < P_{\text{out}}(M_{\text{th}}) < 1 - \kappa_l\exp(\frac{-\lambda}{M_{\text{th}}}), \tag{2.34}$$

which gives us an upper bound and a lower bound for $P_{\text{out}}(M_{\text{th}})$. Figure 2.3 depicts the outage probability and the bounds for the case when $\sigma_u^2 = \sigma_v^2 = 1$, $\rho = 0.95$, and SNR = 8, 10, 15, 20 dB ($\lambda \approx 1.626, 1.026, 0.324, 0.103$), and for $M \leqslant \sigma_u^2$. As seen in Fig. 2.3, a smaller λ yields a smaller outage probability. From Fig. 2.3, one can observe that while

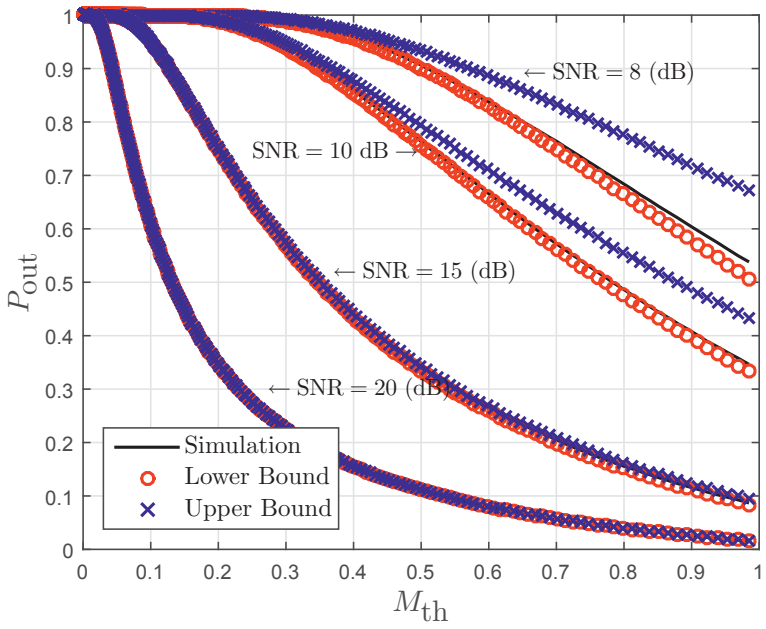

Figure 2.3: $P_{\text{out}}(M_{\text{th}})$ and its upper and lower bounds for $\sigma_u^2 = \sigma_v^2 = 1$, $\rho = 0.95$, SNR = 8, 10, 15, 20 dB ($\lambda \approx 1.626, 1.026, 0.324, 0.103$).

the lower bound has a very good visible accuracy for a wide range of SNR values, the upper bound is not as good for lower SNR values. However, the accuracy of the bounds improve while increasing the SNR. This suggests that the bounds might be tight in the asymptotic high SNR regime i.e. when $\lambda \to 0$. In the following lemma, it is shown that in fact, this is the case.

Lemma 5. *The upper and lower bounds for $P_{\text{out}}(M_{\text{th}})$ are tight for $\lambda \to 0$.*

Proof. we first show that when $\lambda \to 0$, we have that

$$\lim_{\lambda \to 0} \kappa = \lim_{\lambda \to 0} \kappa_u = \lim_{\lambda \to 0} \kappa_l = 1. \tag{2.35}$$

We have for the upper bound that

$$\kappa_u = \frac{1}{\left(a_\kappa \exp\left(\frac{-\lambda}{\sigma_u^2(1+\rho^2)} \right) + \exp(-\frac{\lambda}{\sigma_u^2}) \right)}. \tag{2.36}$$

For finite σ_u^2, the condition $\lambda \to 0$ can be extended to $\lambda/\sigma_u^2 \to 0$. This assumption is made to simplify the calculations. Assume $\lambda = \alpha\sigma_u^2$, then

$$\kappa_u = \frac{1}{\left(a_\kappa(\alpha) \exp\left(\frac{-\alpha}{(1+\rho^2)} \right) + \exp(-\alpha) \right)}. \tag{2.37}$$

For finite σ_u^2, the condition $\lambda \to 0$ will be equal to $\alpha \to 0$. We can then see that

$$\lim_{\alpha \to 0} \kappa_u(\alpha) = \frac{1}{1 + \lim_{\alpha \to 0} a_\kappa(\alpha)}. \tag{2.38}$$

Next, it is shown that $\lim_{\alpha \to 0} a_\kappa(\alpha) = 0$. We have

$$\begin{aligned} a_\kappa(\alpha) &= 1 - \int_0^{\sigma_u^2} \exp(\frac{\lambda}{\rho^2 m + \sigma_u^2}) \left(\frac{\lambda}{m^2}\right) \exp(\frac{-\lambda}{m}) \, \mathrm{d}m \\ &= 1 - \int_0^1 \exp(\frac{\alpha}{1 + \rho^2 v}) \left(\frac{\alpha}{v^2}\right) \exp(\frac{-\alpha}{v}) \, \mathrm{d}v, \end{aligned} \tag{2.39}$$

where the change of variable $v = \frac{m}{\sigma_u^2}$ was made. Now take $a_\kappa(\alpha) = 1 - I(\alpha)$, where

$$\begin{aligned} I(\alpha) &= \int_0^1 \exp(\frac{\alpha}{1 + \rho^2 v}) \left(\frac{\alpha}{v^2}\right) \exp(\frac{-\alpha}{v}) \, \mathrm{d}v \\ &= \exp(\frac{\alpha}{1 + \rho^2 v}) \exp(\frac{-\alpha}{v}) \Big|_0^1 - \int_0^1 \exp(\frac{-\alpha}{v}) \exp(\frac{\alpha}{1 + \rho^2 v}) \left(\frac{-\alpha \rho^2}{(1 + \rho^2 v)^2}\right) \mathrm{d}v \\ &= \exp(\frac{\alpha}{1 + \rho^2}) \exp(-\alpha) + \rho^2 \alpha \int_0^1 \exp(\frac{-\alpha}{v}) \exp(\frac{\alpha}{1 + \rho^2 v}) \frac{1}{(1 + \rho^2 v)^2} \, \mathrm{d}v. \end{aligned} \tag{2.40}$$

Now, because all of the functions $\exp(\frac{-\alpha}{v}), \exp(\frac{\alpha}{1+\rho^2 v}), \frac{1}{(1+\rho^2 v)^2}$ are finite for $v \in [0, 1]$, then the integral term in (2.40) is also finite. As a result, $\lim_{\alpha \to 0} I(\alpha) = 1$ and $\lim_{\alpha \to 0} \kappa_u(\alpha) = 1$. Similar results also hold for $\kappa_l^\infty, \kappa_l^b$.

For the limiting behavior for κ, the result of Appendix C is used for the Taylor series expansion of κ at $\lambda = 0$. Although Appendix C discusses the more general case of multiple receivers rather than a single one, the case of κ for a single receiver can be considered as a special case there and is therefore omitted from here in order to avoid repetition.

From there, we have that the Taylor series expansion for κ at $\lambda = 0$ can be obtained from

$$\begin{aligned} \kappa &= \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} \frac{1}{(\sigma_u^2)^l} \\ &= 1 + \frac{\lambda}{\sigma_u^2} + \mathcal{O}(\lambda^2) \end{aligned} \tag{2.41}$$

As a consequence, we can verify from (2.41) that $\lim_{\lambda \to 0} \kappa = 1$, as claimed before.

This proves the lemma, because then the outage probability and the bounds will have the same values as $\kappa, \kappa_u, \kappa_l$ converge to the same value. □

It is also interesting to see how fast κ converges to 1 for small values of λ and for which values of λ, the upper and lower bound are approximately equal. This is depicted in Fig.

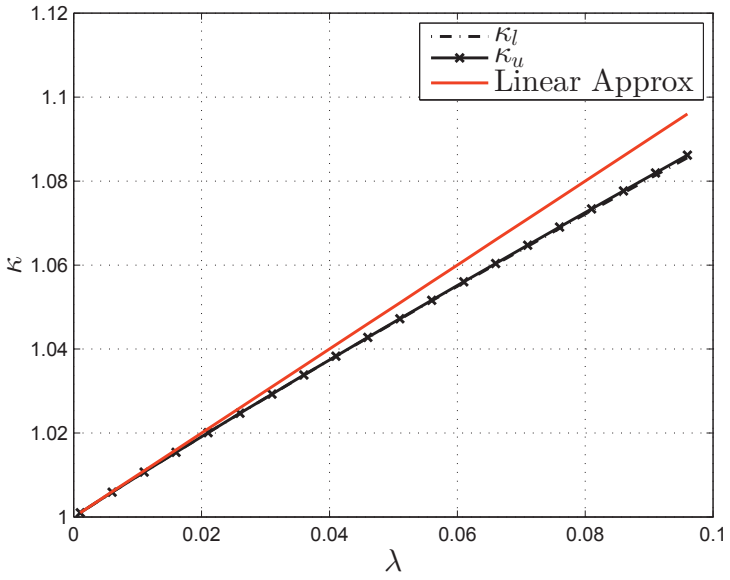

Figure 2.4: κ_l, κ_u as a function of λ for $\sigma_u^2 = \sigma_v^2 = 1$, $\rho = 0.95$.

2.4. Quite visibly, for values of λ close to 0.01 (SNR close to 30dB) the upper and lower bounds for κ are very close. Due to the fact that the bounds for κ are tight, the bounds for $P_{\text{out}}(M_{\text{th}})$ are also tight. Even for the range of medium SNR depicted in Fig. 2.3, it is quite visible that the upper and lower bounds for the outage probability are quite close to the one obtained from the simulation and that increasing the SNR improves their accuracy. However, the bounds, especially the upper bound lose their accuracy in the low SNR regime. This necessitates taking extra precautions if the bounds are to be used in applications prone to low SNR's. It is also quite visible from Fig. 2.4 that the linear approximation obtained from the Taylor series expansion of κ as a function of λ is quite accurate.

At this point, the asymptotic behavior of the outage probability in the high SNR regime can be presented. This is expressed in lemma 6.

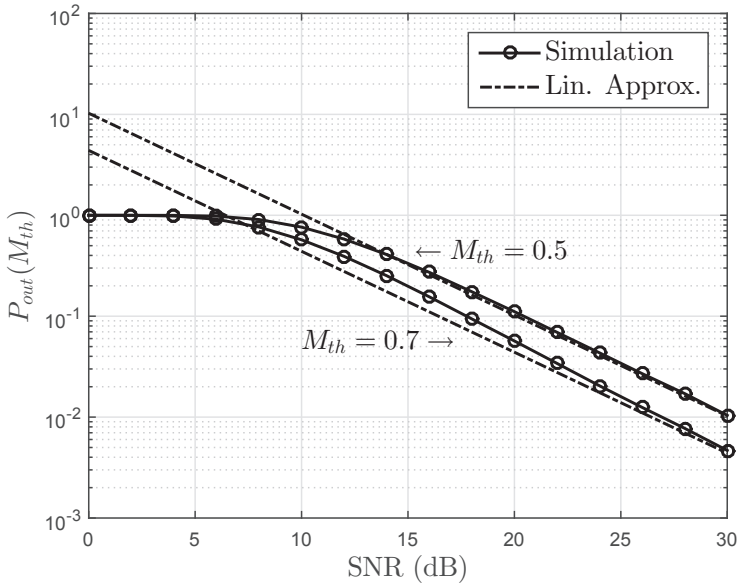

Figure 2.5: P_{out} as a function of λ for $\sigma_u^2 = \sigma_v^2 = 1$, $\rho = 0.95$ and its linear approximation

Lemma 6. *For the high SNR regime, $P_{\text{out}}(M_{\text{th}})$ decreases inversely with the SNR.*

Proof. We can use the Taylor series expansion of κ around $\lambda = 0$ from (2.41) to approximate the outage probability for the high SNR regime. Using the Taylor series expansion for κ and $\exp(\frac{-\lambda}{M_{\text{th}}})$, we obtain

$$\begin{aligned} P_{\text{out}}(M_{\text{th}}) &= 1 - \kappa \exp(\frac{-\lambda}{M_{\text{th}}}) \\ &= 1 - \left(1 + \frac{\lambda}{\sigma_u^2} + \mathcal{O}(\lambda^2)\right)\left(1 - \frac{\lambda}{M_{\text{th}}} + o(\lambda)\right) \\ &= 1 - \left(1 + \frac{\lambda}{\sigma_u^2} + o(\lambda)\right)\left(1 - \frac{\lambda}{M_{\text{th}}} + o(\lambda)\right) \\ &= \left(\frac{1}{M_{\text{th}}} - \frac{1}{\sigma_u^2}\right)\lambda + o(\lambda) \\ &= \left(\frac{1}{M_{\text{th}}} - \frac{1}{\sigma_u^2}\right)\frac{\mathrm{P}_x}{\text{SNR}} + o\left(\frac{1}{\text{SNR}}\right). \end{aligned} \tag{2.42}$$

□

For small λ, $\mathcal{O}(\lambda^2)$ and $o(\lambda)$ vanish faster than λ and as a result it can be claimed that $P_{\text{out}}(M_{\text{th}})$ is approximately a linear function of λ. Due to that, the outage probability

decreases linearly with the inverse of the SNR in the high SNR regime, i.e. resulting to a diversity order of one. This behavior is depicted in Fig. 2.5 for SNR $\in [0\,\text{dB}, 30\,\text{dB}]$.

This behavior is similar to that of bit-error and outage probabilities for digital communication over Rayleigh fading channels. While increasing the SNR helps with outage probability, it does not help significantly and it is necessary to find other techniques for performance improvement, such as diversity techniques.

In addition, (2.42) also implies that the coding for EOP in the high SNR regime is equal to

$$\begin{aligned} G &= \left(\frac{\mathrm{P}_x}{M_{\text{th}}} - \frac{\mathrm{P}_x}{\sigma_u^2} \right)^{-1} \\ &= \left(\frac{\mathrm{P}_x}{M_{\text{th}}} - \frac{1}{1-\rho^2} \right)^{-1} \end{aligned} \tag{2.43}$$

based on (1.10). The notions of diversity order and coding gain are discussed in more details in chapters 3 and 4.

2.4 Summary and Discussion

In this chapter, a recursive integral equation approach was presented for finding the pdf of the instantaneous estimation error variance for linear MMSE estimation, i.e. Kalman filtering, of scalar Gauss-Markov signals sent over scalar fading channels. The pdf is an important tool for calculating other relevant measures of estimation quality, such as the average MSE and the estimation error (distortion) outage probability. It was shown that the steady-state pdf can be written as a two-part function over the domain of instantaneous estimation error variance values. This result holds for any channel distribution, as long as the channels are i.i.d. While the interest was mostly characterizing the steady-state pdf, we have also presented an equation for evolution of the pdf over time. Although, as it appears, such evolution needs to be tracked numerically, as the equations for time-dependent pdf's become quickly intractable. The analysis in this part may be improved and completed in two ways. First, the non-i.i.d. distributions can be explored. In addition, it is interesting to see how quickly the pdf converges to the steady-state distribution. Even if a complete convergence analysis might be cumbersome, a numerical evaluation is possible and still very insightful.

After establishing the recursion for the pdf, we focused on the important case of Rayleigh fading channels. In that case, the first part of the pdf, i.e. the range up to the Gauss-Markov process variance, corresponding to higher outage probabilities, follows a closed-form non-recursive equation. The reason for this simplification lies within a simple property of the exponential function, namely $\exp(a+b) = \exp(a) \cdot \exp(b)$, for any arbitrary a and b. While Rayleigh fading is a widespread fading model, it is still interesting to see how the pdf would develop for other fading models. This will broaden the application

areas of the current work. Also, we have not managed to characterize the pdf in a non-recursive manner beyond the process noise variance position, which is somewhat limiting. However, obtaining the outage probabilities was the main target of this work. As the outage probability decreases with the increasing threshold, given that all other parameters are fixed, it is possible to use the outage values at this point, i.e. the process noise variance, as an upper bound or a rough estimate for the outage probabilities beyond this point.

For the first part of the pdf, upper and lower bounds on the estimation error outage probability were obtained, as obtaining the actual outage probability values became tedious. The bounds were shown to be visibly accurate, especially for higher SNR values and smaller thresholds. The lower bound also shows to be more accurate than the upper bound. Furthermore, the bounds were shown to be tight when the SNR goes to infinity.

Furthermore, it was shown that the outage probability decreases with the inverse of the SNR in the high SNR regime. This is similar to the asymptotic behavior seen in detection over Rayleigh fading channels. This suggests that similar methods, e.g. diversity techniques, might be applicable in order to improve the performance. Note that while this result is only valid for thresholds up to the process noise variance, the outage values at that point can be used as an upper bound for all the thresholds beyond, given that the rest of the system parameters are kept fixed. This implies that for thresholds beyond the process noise variance, the diversity order is *at least* equal to one. Another result was that the diversity order does not depend on the source correlation, and is only dependent on the channel, which was not unexpected.

Finally and as a result of the asymptotic high SNR analysis, the coding gain was obtained as a function of the system parameters. It is interesting to observe how the source correlation affects the coding gain. From (2.43), it is clear that for fixed source power and threshold values, higher correlation, i.e. bigger values for $|\rho|$, results in a higher coding gain value. This means that while correlation does not affect the diversity order, it does in fact affect the coding gain, and thus the performance, in a positive way. The effect is however much more visible for higher correlation coefficients, i.e. $|\rho| \to 1$, than the lower ones.

Chapter 3

Scalar Signals and Multiple Receivers: Diversity

In this chapter, we consider transmission of a scalar first-order Gauss-Markov signal over d parallel i.i.d. Rayleigh fading channels ($K = 1$ and $N = d$) using analog uncoded transmission. As observed in the final part of Chapter 2, the outage probability decreases inversely with only the first power of SNR in the high SNR regime, when a single scalar channel is used. For low latency reliability, when channel state is not available at the transmitter, one needs to avoid fading, and thus design a system where samples experience independent fades. We then analyze the performance and seek additional reliability in parallel channels. We keep the assumption that the whole channel is completely known at the receiver. Thus, the optimal MMSE estimator is still the Kalman filter with random IEV. The main objective of this chapter compared to the previous one is thus studying the effect of multiple receivers on the estimation quality and especially the diversity gain for the outage probability. We use the fundamental results of Chapter 2 on the steady-state distribution, and then extend those results to the case of multiple receivers. We then present diversity results analogous to those existing for digital communication over fading channels. This diversity result confirms the considerable improvement that multiple receivers can provide for analog real-time communication. In addition and for the finite SNR regime, upper and lower bounds are presented for the outage probability. Furthermore, the effect of parameter mismatch on the diversity order is considered. It is shown that for certain parameters, the parameter mismatch will only affect the coding gain, but not the diversity order of the EOP in the high SNR regime.

3.1 Updated System Model and Problem Definition

In this section, the following updated system model is assumed

$$\begin{aligned} x(n) &= \rho x(n-1) + u(n) \ \left(n \geq 1,\, x(0) \sim \mathcal{CN}(0, M_0)\right) \\ \mathbf{y}(n) &= \mathbf{h}(n)x(n) + \mathbf{v}(n). \end{aligned} \tag{3.1}$$

The parameters n, ρ, $T(\cdot)$, $x(n)$ and $u(n)$ are the same as the ones in Chapter 2. Due to the introduction of multiple receivers, the channel is now a random vector, rather than a scalar. The channel vector $\mathbf{h}(n) = [h_1(n), h_2(n), \cdots, h_d(n)]^T$ (The superscript T denotes the transpose operation) is assumed to be a circularly symmetric complex Gaussian vector of dimension d, with independent entries (i.i.d. Rayleigh fading) and with the extra condition that the two degenerate cases $\mathbf{h}(n) = 0$, $\forall n$ and $\mathbf{h}(0) = 0$ are excluded. This assumption is made to fulfill the convergence requirement for the steady-state distribution of the IEV in Sec. 3.2.1, and is similar to the one used in Chapter 2. These conditions are easy to satisfy, because these two events happen with zero probability for Rayleigh fading channels. For the receiver noise, it is assumed that $\mathbf{v}(n)$ is a white circularly symmetric complex Gaussian random vector with the covariance matrix $C_v = \text{diag}(\sigma_{v,1}^2, \sigma_{v,2}^2, \cdots, \sigma_{v,d}^2)$ with $\sigma_{v,i}^2 > 0 \, \forall i$.

Given the previous assumptions and with $\mathbf{h}(n)$ being independent of $u(n)$ and $\mathbf{v}(n)$, the optimal MMSE estimator of $x(n)$ based on the channel outputs $\mathbf{y}(n)$ is still the Kalman filter [52]. It is easy to show that the instantaneous estimation error variance $M(n)$ can be written in a recursive formula in terms of its previous value and $\mathbf{h}(n)$, similar to Chapter 2. Studying the statistical properties of $M(n)$ under such condition is the target of this chapter.

It can be shown that the recursion involving $M(n)$ for a scalar process, but a vector channel, is again obtained through

$$M(n) = \frac{\rho^2 M(n-1) + \sigma_u^2}{1 + \gamma(n)\left(\rho^2 M(n-1) + \sigma_u^2\right)}, \tag{3.2}$$

where in this case, we have that $\gamma(n) = \sum_{i=1}^{d} |h_i(n)|^2 / \sigma_{v,i}^2$.

In order to characterize the random estimation outage event, we follow the same definitions for outage, outage probability and diversity order as in Chapter 1.

3.2 Statistical Properties of the IEV

In this section, the steady-state pdf of the IEV is studied first, and an equation for the outage probability is obtained afterwards.

3.2.1 Derivation of the Steady-State Pdf

In order to find the steady-state outage probability for $M(n)$, one needs to first find the steady-state pdf of $M(n)$. The scalar case ($d = 1$) was already studied in Chapter 2. There, it was shown that $M(n)$ converges in distribution and has the following pdf (2.23)

$$f_M(M) = \begin{cases} \frac{1}{M^2} \int_0^{M_{\max}} f_\gamma \left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m, & M \leqslant \sigma_u^2 \\ \frac{1}{M^2} \int_{\frac{M-\sigma_u^2}{\rho^2}}^{M_{\max}} f_\gamma \left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m, & M > \sigma_u^2, \end{cases} \tag{3.3}$$

where $M_{\max} = P_x = \sigma_u^2/(1-\rho^2)$. The convergence argument presented in Chapter 2, which is based on [10] and [112], does not depend on the dimension of the channel, and holds for general conditionally Gaussian systems, including the system model in chapters 2 and 3, and thus left out for brevity. As a result, (3.3) may be used to obtain the steady-state pdf of $M(n)$ for $d > 1$.

In general, solving (3.3) would yield $f_M(M)$, given in terms of the pdf of $\gamma(n)$ and the parameters of the source, i.e. σ_u^2 and ρ. Note that the d parallel independent channels may be obtained in frequency, time or space, but it is possible to assume without loss of generality that the receiver is equipped with multiple receive antennas. It is assumed that all the channels have the same average SNR, i.e. if we take $\gamma_i(n) = \frac{|h_i(n)|^2}{\sigma_{v,i}^2}$, we then have $f_{\gamma_i}(\gamma_i) = \lambda \mathrm{e}^{-\lambda\gamma_i} \mathcal{U}(\gamma_i)$, where $E(\gamma_i(n)) = \frac{1}{\lambda}$ and therefore $E(\gamma(n)) = \frac{d}{\lambda}$. In that case, the pdf of $\gamma(n)$ is χ^2-distributed and follows $f_\gamma(\gamma) = \frac{\lambda^d}{(d-1)!} \gamma^{d-1} \mathrm{e}^{-\lambda\gamma} \mathcal{U}(\gamma)$. Note that with this definition, channels which are stronger on the average yield smaller values for λ and vice versa. Also, the (total) average SNR at the receiver is equal to $\mathrm{SNR} = \dfrac{d\, P_x}{\sigma_v^2 \lambda}$. Similar to the case of $d = 1$ and given the above assumptions, (3.3) can be rewritten as

$$f_M(M) = \frac{1}{M^2} \frac{\lambda^d}{(d-1)!} \exp(\frac{-\lambda}{M}) \quad \begin{cases} \int_0^{M_{\max}} (\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2})^{d-1} \times \\ \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m, & M \leqslant \sigma_u^2 \\ \int_{\frac{M-\sigma_u^2}{\rho^2}}^{M_{\max}} (\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2})^{d-1} \times \\ \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2} \right) f_M(m)\, \mathrm{d}m, & M > \sigma_u^2. \end{cases} \tag{3.4}$$

The exact solution to (3.4) does not yield easily due to the nature of the equation. However, as it will be shown later, it is possible to get a parametrized equation for the first part of the pdf ($M \leqslant \sigma_u^2$). The pdf in that region then depends on a set of parameters κ_i, through which a formula for $P_{\text{out}}(M_{\text{th}})$ is obtained afterwards.

Using the binomial expansion, i.e. $(a+b)^n = \binom{n}{i} a^{n-i} b^i$, it is possible to simplify (3.4) by first using the equality

$$\left(\frac{1}{M} - \frac{1}{\rho^2 m + \sigma_u^2} \right)^{d-1} = \sum_{i=0}^{d-1} \binom{d-1}{i} \frac{\left(\frac{-1}{\rho^2 m + \sigma_u^2} \right)^i}{M^{d-1-i}}, \tag{3.5}$$

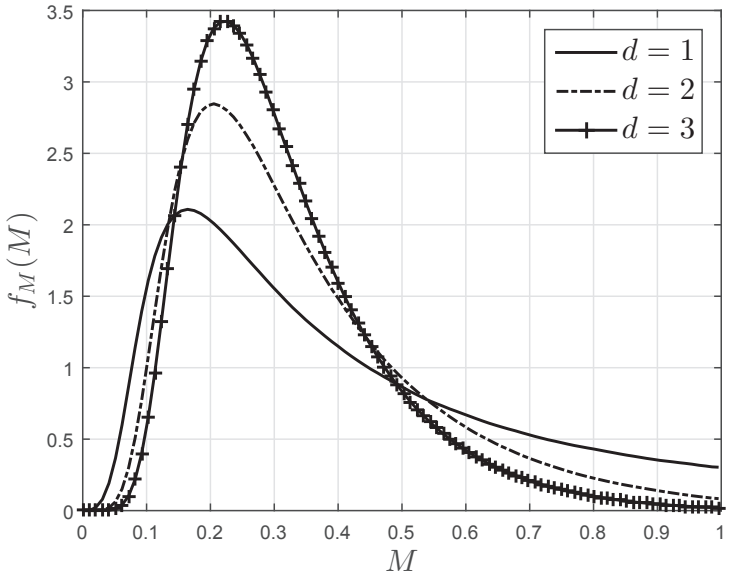

Figure 3.1: Effect of multiple receivers on the pdf of the IEV

where $\binom{d-1}{i} = \frac{(d-1)!}{i!(d-1-i)!}$. Then, the following equality may be obtained

$$f_M(M) = \frac{\lambda^d}{(d-1)!}\frac{1}{M^2}\exp(\frac{-\lambda}{M})\times$$
$$\int_0^{M_{\max}} \sum_{i=1}^{d-1} \binom{d-1}{i} \frac{1}{M^{d-1-i}} \left(\frac{-1}{\rho^2 m + \sigma_u^2}\right)^i \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right) f_M(m)\,\mathrm{d}m, (M \leqslant \sigma_u^2), \tag{3.6}$$

which by integrating over m can be rewritten as

$$f_M(M) = \frac{\lambda^d}{(d-1)!}\frac{1}{M^2}\exp(\frac{-\lambda}{M})\sum_{i=1}^{d-1}(-1)^i\kappa_i\binom{d-1}{i}\frac{1}{M^{d-1-i}} \tag{3.7}$$

with

$$\kappa_i = \int_0^{M_{\max}} \left(\frac{1}{\rho^2 m + \sigma_u^2}\right)^i \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right) f_M(m)\,\mathrm{d}m. \tag{3.8}$$

Note that in general the κ_i constants depend on the pdf itself, and that although they are constants, they depend on the value of d as well, e.g. κ_1 is different for $d = 2$ and $d = 3$. As one increases the SNR, the tail of the pdf shrinks and thus the values of the second part of (3.4) ($M > \sigma_u^2$), especially for outage calculation purpose in the high SNR

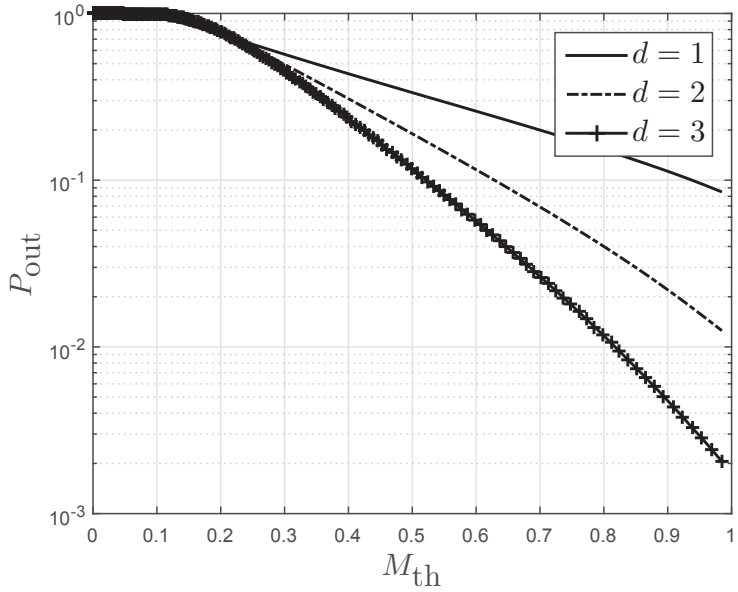

Figure 3.2: Effect of multiple receivers on the outage probability

regime, become quickly insignificant. Figure 3.1 shows some typical pdf's resulting from simulation on 10^7 samples with $\sigma_u^2 = 1, \sigma_{v,i}^2 = 1, \rho = 0.95$, SNR $= 15$ dB ($\lambda/d \approx 0.324$), and for $d = 1, 2, 3$. It is quite visible and also analytically verifiable that values above the break point $M = \sigma_u^2$ are much smaller than those around $M = 0$. Due to this fact, one could in practice neglect the second part of the pdf for high SNR analysis and focus on the first part. Also, λ is normalized over the number of channels to keep the total average SNR the same in all cases and thus only study the effect of diversity and not the extra antenna power gain. Clearly, by increasing d, the tail of the pdf loses its significance, i.e. larger error values can then be decreased by increasing the number of paths. This is the same phenomenon which happens in digital communication over fading channels and one would speculate that similar results on diversity order should hold in this case as well. In the following, it will be proven that this is in fact the case. To begin with the analysis, the reader is provided with Lemma 7, which quantifies the EOP.

Lemma 7. *The expected estimation error outage for any $M_{th} \leqslant \sigma_u^2$ is obtained from*

$$P_{out}(\lambda, M_{th}) = 1 - exp(\frac{-\lambda}{M_{th}}) \sum_{i=0}^{d-1} \kappa_i \frac{(-1)^i \lambda^i}{i!} \sum_{k=0}^{d-1-i} \frac{(\frac{\lambda}{M_{th}})^k}{k!}. \tag{3.9}$$

Proof. It is possible to show that

$$\int_0^M \exp(\frac{-\lambda}{v}) v^{-n} dv = \lambda^{1-n} \Gamma(n-1, \frac{\lambda}{M}), \quad (n > 0) \tag{3.10}$$

where $\Gamma(n, M)$ is the incomplete Γ function with the definition

$$\Gamma(n, M) = (n-1)!\exp(-M)\sum_{k=0}^{n-1}\frac{M^k}{k!}. \tag{3.11}$$

Therefore, we have that

$$\int_0^{M_{\text{th}}} \exp(\frac{-\lambda}{M})M^{-n}\,\mathrm{d}M = \lambda^{1-n}(n-2)!\exp(-\frac{\lambda}{M_{\text{th}}})\sum_{k=0}^{n-1}\frac{(\frac{\lambda}{M_{\text{th}}})^k}{k!}. \tag{3.12}$$

Using the above conclusions, we get the outage probability as follows

$$\begin{aligned}
P_{\text{out}}(M_{\text{th}}) &= 1 - \int_0^{M_{\text{th}}} \frac{\lambda^d}{(d-1)!}\sum_{i=0}^{d-1}\binom{d-1}{i}(-1)^i\kappa_i\frac{\mathrm{e}^{(-\frac{\lambda}{M})}}{M^{d+1-i}}\,\mathrm{d}M \\
&= 1 - \sum_{i=0}^{d-1}\frac{\lambda^d}{(d-1)!}\binom{d-1}{i}(-1)^i\kappa_i\int_0^{M_{\text{th}}}\frac{\mathrm{e}^{(-\frac{\lambda}{M})}}{M^{d+1-i}}\,\mathrm{d}M \\
&= 1 - \sum_{i=0}^{d-1}\frac{\lambda^d}{(d-1)!}\binom{d-1}{i}(-1)^i\kappa_i\lambda^{i-d}\Gamma(d-i, \frac{\lambda}{M_{\text{th}}}) \\
&= 1 - \sum_{i=0}^{d-1}\frac{\lambda^i}{(d-1)!}\binom{d-1}{i}(-1)^i\kappa_i(d-i-1)!\exp(\frac{-\lambda}{M_{\text{th}}})\sum_{k=0}^{d-i-1}\frac{(\frac{\lambda}{M_{\text{th}}})^k}{k!} \\
&= 1 - \exp(\frac{-\lambda}{M_{\text{th}}})\sum_{i=0}^{d-1}\frac{\kappa_i(-1)^i\lambda^i}{i!}\sum_{k=0}^{d-i-1}\frac{(\frac{\lambda}{M_{\text{th}}})^k}{k!}
\end{aligned} \tag{3.13}$$

□

Now the estimation error outage probability for $M \leqslant \sigma_u^2$ is described in terms of $\lambda = \dfrac{d\,\mathrm{P}_x}{\sigma_v^2\mathrm{SNR}}$ and a set of constants κ_i. The use of λ instead of SNR is for convenience. If the EOP is plotted for the same set of parameters as in Fig. 3.1, the results in Fig 3.2 are obtained. Now with EOP at hand (which is a well-defined error function), it is be easier to confront problems such as diversity order. One could see that for a fixed threshold value, increasing the receive diversity substantially impacts and reduces the EOP. Thus, the values of M_{th} closest to σ_u^2 have the minimum outage values.

The analysis is continued in the next section by first using the obtained formula for EOP to provide upper and lower bounds for the outage probability, which are useful for the finite SNR regime. Next, we turn to high SNR analysis and the calculation of the diversity order.

3.2.2 Bounds on the Outage Probability

The analysis which is later provided for asymptotic behavior of the distortion probability does not rely on exact values of κ_i, but rather their asymptotic behavior as SNR goes to infinity. However for finite values of SNR, one might be interested in approximate values for outage probability as a figure of merit for design purposes, given a particular SNR. In this section and for that purpose, upper and lower bounds for the outage probability are provided.

Given the definitions for κ_i, Lemma 8 is presented for calculating upper and lower bounds for κ_i. Through this, upper and lower bounds for $P_{\text{out}}(M_{\text{th}})$ are obtained for $M \leqslant \sigma_u^2$. The bounds are presented in Lemma 9.

Lemma 8. *The coefficients κ_i can be lower and upper bounded as $\kappa_i^l \leqslant \kappa_i \leqslant \kappa_i^u$, where κ_i^l and κ_i^u are obtained from the following linear system of inequalities (order $d+1$). The solution to the inequality*

$$W^u \boldsymbol{\kappa} \leqslant \mathbb{1}_{d+1} \quad (\boldsymbol{\kappa} \geqslant 0), \tag{3.14}$$

namely $\boldsymbol{\kappa}^u = [\kappa_0^u, \kappa_1^u, \cdots, \kappa_{d-1}^u, Z^u]$ results in a set of upper bound values for κ_i and the solution to

$$W^l \boldsymbol{\kappa} \geqslant \mathbb{1}_{d+1} \quad (\boldsymbol{\kappa} \geqslant 0), \tag{3.15}$$

namely $\boldsymbol{\kappa}^l = [\kappa_0^l, \kappa_1^l, \cdots, \kappa_{d-1}^l, Z^l]$ results in a set of lower bound values for κ_i.

The elements of the matrices W^u and W^l are functions of system parameters $\lambda, \rho, \sigma_u^2, d, M_{max}$, and are only explicitly defined in Appendix B in order to avoid repetition. The vector $\mathbb{1}_{d+1}$ is a column vector with size $d+1$, consisting of all zeros except at position $d+1$. Also, $\boldsymbol{\kappa} = [\kappa_0, \kappa_1, \cdots, \kappa_{d-1}, Z]$ with Z as an auxiliary variable is defined in Appendix B (see (B.13)*).*

Proof. See Appendix B.

Lemma 9. *The outage probability $P_{out}(M_{th})$ can be lower and upper bounded for the range of thresholds ($M \leqslant \sigma_u^2$) as follows*

$$P_{out}(M_{th}) > P_{out}^l = 1 - exp(\frac{-\lambda}{M_{th}}) \sum_{i=0}^{d-1} \max\{(-1)^i \kappa_i^u, (-1)^i \kappa_i^l\} \frac{\lambda^i}{i!} \sum_{k=0}^{d-1-i} \frac{(\frac{\lambda}{M_{th}})^k}{k!}, \tag{3.16}$$

$$P_{out}(M_{th}) < P_{out}^u = 1 - exp(\frac{-\lambda}{M_{th}}) \sum_{i=0}^{d-1} \min\{(-1)^i \kappa_i^u, (-1)^i \kappa_i^l\} \frac{\lambda^i}{i!} \sum_{k=0}^{d-1-i} \frac{(\frac{\lambda}{M_{th}})^k}{k!}, \tag{3.17}$$

Proof. It is easy to see that to get a lower bound for $P_{\text{out}}(M_{\text{th}})$, it is sufficient to get an upper bound for the sum $\sum_{i=0}^{d-1} \kappa_i \frac{(-1)^i \lambda^i}{i!} \sum_{k=0}^{d-1-i} \frac{(\frac{\lambda}{M_{\text{th}}})^k}{k!}$. This may be achieved by using an upper bound for κ_i when i is even and a lower bound for κ_i when i is odd, due to the effect of the term $(-1)^i$ in the sum. This is exploited by using the $\max\{\}$ operator in (3.16). The upper bound can be established with similar arguments using the $\min\{\}$ operator instead. □

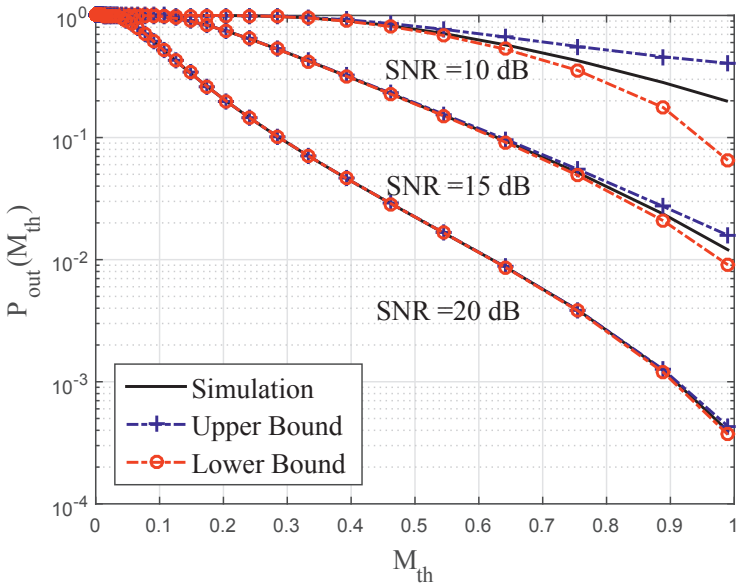

Figure 3.3: Lower and upper bounds on the outage probability for $d = 2$

The bounds from (3.16) and (3.17) obtained by numerical evaluations (including κ_i^l and κ_i^u) and the outage functions obtained via simulating the Kalman filter are depicted in figures 3.3 and 3.4 for $M_{\text{th}} \leqslant \sigma_u^2$. The simulation parameters are $\sigma_u^2 = 1, \sigma_{v,i}^2 = 1, \rho = 0.95, d = 2, 3$ for SNR = 10, 15, 20 dB, equivalent to $\lambda/d \approx 1.026, 0.324, 0.103$, respectively. As it can be seen from the figures, the bounds have visibly a good accuracy further away from $M_{\text{th}} = \sigma_u^2$. Increasing the SNR also improves the bounds, even for M_{th} close to σ_u^2. While we see in figures 3.3 and 3.4 that the bounds start diverging from the actual values if M_{th} approaches σ_u^2, the outage probabilities are relatively lower in that regime. As a result, the bounds may still be satisfactory, depending on the required accuracy. Also note that the discontinuity in lower bound for the case SNR = 10 dB in Figure 3.4 is because the bound delivers essentially a zero value, which cannot be shown in the figure. This suggests that increasing d might reduce the accuracy of the bounds, which is a shortcoming. However, the gaps seem to close for higher SNR values, even for M_{th} close to σ_u^2. In the next section, the high SNR behavior of the outage probability is studied. There, it is shown that when the received signal from d parallel channels is available at the receiver, a diversity order of d can be obtained for the outage probability at high SNR.

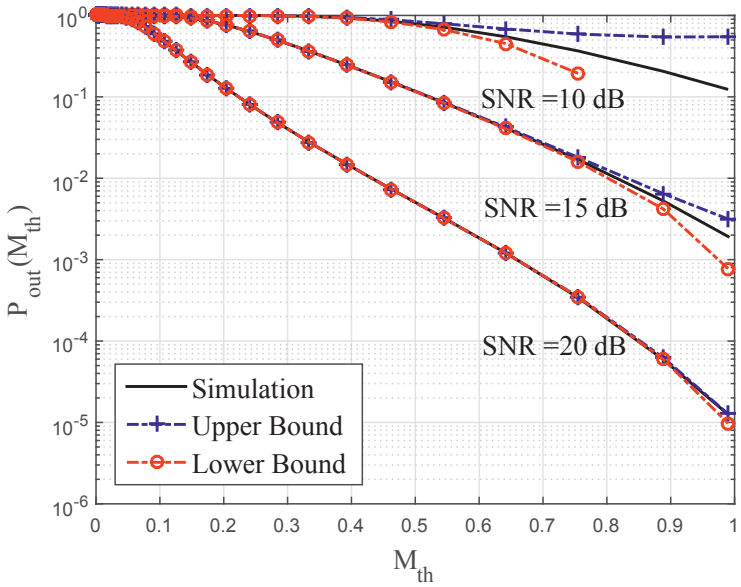

Figure 3.4: Lower and upper bounds on the outage probability for $d = 3$

3.3 Diversity Order Analysis

In this section, it will be shown that in the high SNR regime, there is a diversity order of $d_{\text{ord}} = d$ for the outage probability. That means the EOP will decrease inversely proportionally with the d-th power of SNR, when SNR grows unbounded ($\lambda \to 0$). The definition for the diversity order was presented in Chapter 1. First, another (equivalent) formulation is provided for the diversity order, and then the value of d_{ord} is obtained afterwards.

Consider the error probability function $P_{\text{out}}(\text{SNR}, M_{\text{th}})$ which is a function of SNR and the threshold parameter M_{th}. The following diversity order definition is then used [123]

$$d_{\text{ord}} = -\lim_{\text{SNR}\to\infty} \frac{\log P_{\text{out}}(\text{SNR}, M_{\text{th}})}{\log(\text{SNR})}. \tag{3.18}$$

That means in logarithmic scale for SNR, the probability of error will decrease linearly with the slope of d, i.e. it decreases with the d-th power of SNR in the non-logarithmic scale. Using the above definition, one can write

$$d_{\text{ord}} = -\lim_{\lambda\to 0} \frac{\log P_{\text{out}}(d\,\text{P}_x/(\sigma_v^2\lambda), M_{\text{th}})}{\log(d\,\text{P}_x/(\sigma_v^2\lambda))} \tag{3.19}$$

$$= \lim_{\lambda\to 0} \frac{\log P_{\text{out}}(\lambda, M_{\text{th}})}{\log(\lambda)}, \tag{3.20}$$

where we make a notational reuse of $P_{\text{out}}(\lambda, M_{\text{th}})$ instead of $P_{\text{out}}(d\,\mathrm{P}_x/(\sigma_v^2\lambda), M_{\text{th}})$ for readability and ease of notational use.

Any error function that approaches zero when the SNR goes to infinity makes the limit indefinite. (The error functions that do not behave in that manner, e.g. have an error floor for high SNR, result in a zero order diversity.) Therefore, the limit should be solved to resolve the ambiguity.

If for any function $P_{\text{out}}(\lambda, M_{\text{th}})$, we have that $P_{\text{out}}(\lambda, M_{\text{th}}) = o(\lambda^{d-1}) = \lambda^d(c + o(1))$, we can use the l'Hospital's rule and obtain

$$d_{\text{ord}} = \lim_{\lambda\to 0} \frac{\log \lambda^d(c + o(1))}{\log(\lambda)} \tag{3.21}$$

$$= d + \lim_{\lambda\to 0} \frac{\log\left(c + o(1)\right)}{\log(\lambda)} \tag{3.22}$$

$$= d, \tag{3.23}$$

given that $c \neq 0$. If $c = 0$, then the second limit is still indeterminate and one could expect a higher diversity order than d.

In the rest of this section, it is shown that the outage probability function in this section has this property and therefore satisfies the original condition for $d_{\text{ord}} = d$. To make the way for the proof, Theorem 2 and Theorem 3 are stated as follows.

Theorem 2. *Any error function $P_{out}(\lambda, M_{th})$ which follows the following mathematical form results in a diversity order $d_{ord} = d$*

$$P_{out}(\lambda, M_{th}) = 1 - exp(-\frac{\lambda}{M_{th}})\big(\sum_{i=0}^{d-1} a_i\lambda^i + o(\lambda^{d-1})\big) \tag{3.24}$$

given that

$$a_i = \frac{1}{i!M_{th}^i} \tag{3.25}$$

and that $o(\lambda^{d-1})$ is not exactly equal to zero.

Proof. Take $a_i = 1/(i!M_{\text{th}}^i)$, while having that

$$P_{\text{out}}(\lambda, M_{\text{th}}) = 1 - \exp(-\frac{\lambda}{M_{\text{th}}})\big(\sum_{i=0}^{d-1} a_i\lambda^i + o(\lambda^{d-1})\big). \tag{3.26}$$

We use the Taylor series expansion of $P_{\text{out}}(\lambda, M_{\text{th}})$ around $\lambda = 0$ and get

$$
\begin{aligned}
P_{\text{out}}(\lambda, M_{\text{th}}) &= 1 - (1 + \sum_{j=1}^{d-1}(-1)^j(\frac{\lambda}{M_{\text{th}}})^j/j! + o(\lambda^{d-1}))\big(\sum_{i=0}^{d-1} a_i\lambda^i + o(\lambda^{d-1})\big) \\
&= 1 - \big[a_0(1 + \sum_{j=1}^{d-1}(-1)^j(\frac{\lambda}{M_{\text{th}}})^j/j! + o(\lambda^{d-1})) \\
&+ a_1\lambda(1 + \sum_{j=1}^{d-1}(-1)^j(\frac{\lambda}{M_{\text{th}}})^j/j! + o(\lambda^{d-1})) \\
&+ \cdots \\
&+ a_{d-1}\lambda^{d-1}(1 + \sum_{j=1}^{d-1}(-1)^j(\frac{\lambda}{M_{\text{th}}})^j/j! + o(\lambda^{d-1})) \\
&+ o(\lambda^{d-1})(1 + \sum_{j=1}^{d-1}(-1)^j(\frac{\lambda}{M_{\text{th}}})^j/j! + o(\lambda^{d-1}))\big].
\end{aligned} \tag{3.27}
$$

Then, we have that

$$
\begin{aligned}
P_{\text{out}}(\lambda, M_{\text{th}}) = 1 - \big[&a_0 + \lambda(a_1 - a_0/M_{\text{th}}) + \lambda^2(a_2 - a_1/M_{\text{th}} + a_0/(2!M_{\text{th}}^2)) \\
&+ \lambda^3(a_3 - a_2/M_{\text{th}} + a_1/(2!M_{\text{th}}^2) - a_0/(3!M_{\text{th}}^3)) + \cdots \\
&+ \lambda^{d-1}(a_{d-1} - a_{d-2}/M_{\text{th}} + a_{d-3}/(2!M_{\text{th}}^2) - \cdots \\
&+ (-1)^{d-1}a_0/((d-1)!M_{\text{th}}^{d-1})) + o(\lambda^{d-1})].
\end{aligned} \tag{3.28}
$$

Now we may rewrite (3.28) as

$$
\begin{aligned}
P_{\text{out}}(\lambda, M_{\text{th}}) = (1 - a_0) &+ \lambda(\frac{a_0}{M_{\text{th}}} - a_1) - \lambda^2(\frac{a_0}{2!M_{\text{th}}^2} - \frac{a_1}{M_{\text{th}}} + a_2) \\
&+ \lambda^3(\frac{a_0}{3!M_{\text{th}}^3} - \frac{a_1}{2!M_{\text{th}}^2} + \frac{a_2}{M_{\text{th}}} - a_3) + \cdots \\
&- (-1)^{d-1}\lambda^{d-1}\big(\frac{a_0}{(d-1)!M_{\text{th}}^{d-1}} - \frac{a_{d-2}}{(d-2)!M_{\text{th}}^{d-2}} + \cdots \\
&+ (-1)^{d-1}a_{d-1}\big) + o(\lambda^{d-1}).
\end{aligned} \tag{3.29}
$$

Given that $a_0 = 1$, it is then easy to verify that

$$
P_{\text{out}}(\lambda, M_{\text{th}}) = \sum_{i=1}^{d-1}(-1)^{i-1}D_i + o(\lambda^{d-1}) \tag{3.30}
$$

with

$$
D_i = \sum_{k=0}^{i}\frac{(-1)^k a_k}{(i-k)!M_{\text{th}}^{i-k}}. \tag{3.31}
$$

Now we show that $D_i = 0$ for any $i \geqslant 1$ and therefore prove Theorem 2.

If we replace the assumed values for a_k into (3.28), we get

$$\begin{aligned} D_i &= \sum_{k=0}^{i} \frac{(-1)^k}{(i-k)!k!M_{\text{th}}^{i-k}M_{\text{th}}^{k}} \\ &= \frac{1}{i!M_{\text{th}}^{i}} \sum_{k=0}^{i} \frac{(-1)^k(1)^{i-k}i!}{(i-k)!k!} \\ &= \frac{1}{i!M_{\text{th}}^{i}}(1-1)^i = 0, \end{aligned} \tag{3.32}$$

where the binomial expansion theorem is used in the last step.

□

Note that Theorem 2 is a general statement and holds regardless of the presented system model in this chapter. One should only replace λ with c/SNR for any system dependent constant c and get the diversity order for any other setting where such form may appear. To show that the EOP introduced in this work in fact fits in this category of functions, we need Theorem 3 as stated in the following.

Theorem 3. *For the high SNR regime ($\lambda \to 0$), the Taylor series expansion of κ_i in* (3.8) *is given by*

$$\kappa_i = \sum_{l=0}^{\infty} \frac{\lambda^l}{(\sigma_u^2)^{i+l}l!} \tag{3.33}$$

Proof. See Appendix C.

Equipped with Theorem 2 and Theorem 3, when then show that our outage function, considered as a function of λ satisfies the required condition in Theorem 2 in the limit of $\lambda \to 0$, and therefore achieves $d_{\text{ord}} = d$. This is presented in Theorem 4.

Theorem 4. *For the high SNR regime ($\lambda \to 0$), the expected error outage probability achieves diversity order $d_{ord} = d$, i.e.*

$$\lim_{\lambda \to 0} \frac{\log P_{out}(\lambda, M_{th})}{\log(\lambda)} = d \quad (0 < M_{th} \leqslant \sigma_u^2) \tag{3.34}$$

Proof. See Appendix D.

Theorem 4 shows the connection between the definition of diversity in the conventional wireless fading channel analysis and in the sense of estimation of Gauss-Markov sources sent over fading channels. In other words, Theorem 4 implies that while the probability of detection for independent signal samples achieves a diversity order of d if d independent fading channels are used, the outage probability for optimal MMSE estimator (Kalman filter) for correlated Gauss-Markov exhibits the same behavior, at least for the range of

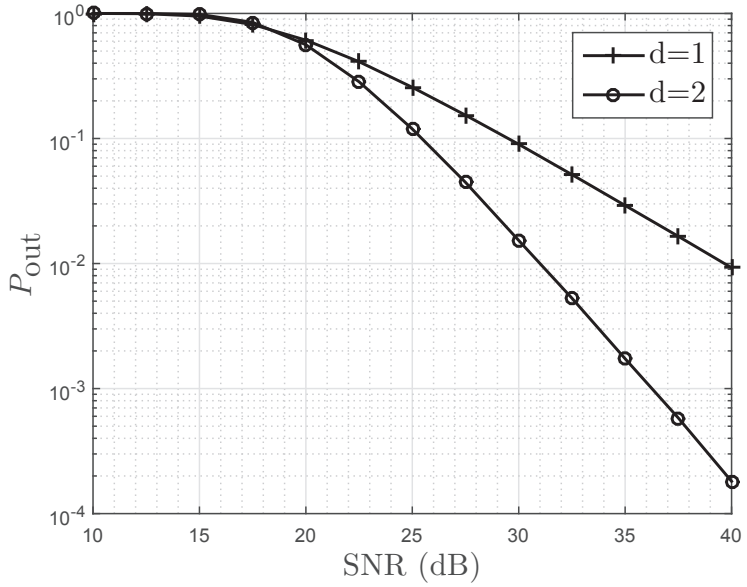

Figure 3.5: Depiction of diversity order for the estimation error outage probability

$0 < M_{\text{th}} \leqslant \sigma_u^2$. Figure 3.5 demonstrates this asymptotic behavior, where outage probability is depicted vs. the SNR in the moderate and high SNR regime. For Fig. 3.5, we keep all the parameters from before, and set $M_{\text{th}} = 0.1$ and $d = 1, 2$. From Fig. 3.5, it is easy to verify a diversity gain of 2, when 2 receive antennas are used.

3.4 Process and Channel Noise Covariance Mismatch

The analysis which has been provided so far in this chapter is based on the assumption that the system parameters are perfectly known at the receiver. If the system parameters are known imperfectly, then there is a mismatch between the assumed values and the actual values of the system parameters. Consequently, the Kalman filter with the current formulation is not optimal. The value of $M(n)$ will not represent the actual instantaneous estimation error variance either. Strictly speaking, one should then apply robust designs, e.g. as in [85] in order to have a better performance. The analysis of such designs in terms of outage probabilities and diversity order is however beyond the scope of this work. Nevertheless, as the system parameters are assumed fixed, it is possible to obtain reasonably good approximates for system parameters during a training phase. With good approximations, it is reasonable to use the proposed design due to its lower design and computational complexity, and one would expect the estimation error and its outage to be close to the case with perfect system knowledge. It is then interesting to see the effect of

imperfect knowledge of system parameters or parameter mismatch on the performance, specifically on the diversity order.

In the following, Kalman filter with mismatched parameters is considered, i.e. when wrong values are used for the system parameters. It is assumed that $\hat{\sigma}_u^2 \neq \sigma_u^2$ and $\hat{\sigma}_{v,i}^2 \neq \sigma_{v,i}^2$ are used in the filter. The choice of σ_u^2 and $\sigma_{v,i}^2$ is for the simplicity of analysis and that the final results are of quite simple forms, which provide relatively good insight to the problem. It should be noted that mismatch analysis for other system parameters, especially the channel values $h(n)$, is also important. Mismatched system parameters can in certain situations degrade the performance significantly or even result in instability.

3.4.1 Mismatched Process Noise Variance

In this subsection, we consider the case when $\hat{\sigma}_u^2 = \sigma_u^2 + \beta$ for some β drawn from a prior pdf $f_\beta(\beta)$, but fixed for the period of interest. For better legibility, accented symbols are used for all the variables in the algorithm which change due to mismatch, e.g. $\hat{M}(n)$ represents the IEV with mismatched parameters, while the true value for IEV is represented by $M(n)$. For simplicity, but without loss of generality (channel branches are already assumed to have equal SNR), it is assumed that $\sigma_{v,i}^2 = \sigma_v^2$.

Borrowing the Kalman filter equations from Chapter 2, we have for the mismatched σ_u^2 case that

$$\hat{x}(n|n-1) = \rho\hat{x}(n-1) \tag{3.35}$$

$$\begin{aligned}\hat{P}(n) &= \rho^2\hat{M}(n-1) + \hat{\sigma}_u^2 \\ &= \rho^2\hat{M}(n-1) + \sigma_u^2 + \beta\end{aligned} \tag{3.36}$$

$$\begin{aligned}\hat{\mathbf{g}}_\mathbf{k}(n) &= \hat{P}(n)\mathbf{h}^*(n)\left(\sigma_v^2 + \mathbf{h}(n)\hat{P}(n)\mathbf{h}^*(n)\right)^{-1} \\ &= \hat{P}(n)\mathbf{h}^*(n)\left(\sigma_v^2 + \|\mathbf{h}(n)\|^2\hat{P}(n)\right)^{-1}\end{aligned} \tag{3.37}$$

$$\hat{x}(n) = \hat{x}(n|n-1) + \hat{\mathbf{g}}_\mathbf{k}(n)(\mathbf{y}(n) - \mathbf{h}(n)\hat{x}(n|n-1)) \tag{3.38}$$

$$\hat{M}(n) = \left(1 - \hat{\mathbf{g}}_\mathbf{k}(n)\mathbf{h}(n)\right)\hat{P}(n). \tag{3.39}$$

Using (3.36), (3.37) and (3.39) will result in the following recursion for $\hat{M}(n)$

$$\hat{M}(n) = \frac{\sigma_v^2\hat{P}(n)}{\left(\sigma_v^2 + \|\mathbf{h}(n)\|^2\hat{P}(n)\right)} \tag{3.40}$$

$$= \frac{\rho^2\hat{M}(n-1) + \hat{\sigma}_u^2}{1 + \frac{\|\mathbf{h}(n)\|^2}{\sigma_v^2}\left(\rho^2\hat{M}(n-1) + \hat{\sigma}_u^2\right)}. \tag{3.41}$$

The two forms for $\hat{\mathbf{g}}_\mathbf{k}(n)$ can be shown to be equal using the matrix inversion lemma. In

order to get $M(n) = E\left(|x(n) - \hat{x}(n)|^2\right)$, it can be seen from (3.38) that

$$\begin{aligned}\hat{x}(n) &= \hat{x}(n|n-1) + \hat{\mathbf{g}}_{\mathbf{k}}(n)\left(\mathbf{y}(n) - \mathbf{h}(n)\hat{x}(n|n-1)\right) \\ &= \hat{x}(n|n-1) \\ &\quad + \hat{\mathbf{g}}_{\mathbf{k}}(n)\left(\mathbf{h}(n)x(n) + \mathbf{v}(n) - \mathbf{h}(n)\hat{x}(n|n-1)\right),\end{aligned} \tag{3.42}$$

which results in

$$\begin{aligned}M(n) &= E\left(|x(n) - \hat{x}(n)|^2\right) \\ &= |1 - \hat{\mathbf{g}}_{\mathbf{k}}(n)\mathbf{h}(n)|^2 E\left(|x(n) - \hat{x}(n|n-1)|^2\right) + \sigma_v^2 \|\hat{\mathbf{g}}_{\mathbf{k}}(n)\|^2 \\ &= \frac{\sigma_v^4 P(n)}{\left(\sigma_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)^2} + \frac{\sigma_v^2 \hat{P}^2(n)\|\mathbf{h}(n)\|^2}{\left(\sigma_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)^2}.\end{aligned} \tag{3.43}$$

Combining (3.40) and (3.43), it is easy to show that

$$\begin{aligned}M(n) - \hat{M}(n) &= \frac{\sigma_v^4\left(P(n) - \hat{P}(n)\right)}{\left(\sigma_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)^2} \\ &= \frac{\sigma_v^4\left(\rho^2\left(M(n-1) - \hat{M}(n-1)\right) + \sigma_u^2 - \hat{\sigma}_u^2\right)}{\left(\sigma_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)^2} \\ &= \frac{\sigma_v^4\left(\rho^2\left(M(n-1) - \hat{M}(n-1)\right) - \beta\right)}{\left(\sigma_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)^2}\end{aligned} \tag{3.44}$$

For positive β, it is easy to show by induction over n that $M(n) \leqslant \hat{M}(n)$, $\forall n$, i.e. $\hat{M}(n)$ is a pessimistic performance measure (see also [47, Ch. 7.8]). One only needs to select a big enough value for $\hat{M}(0)$ to make sure that $M(n) - \hat{M}(n) \leqslant 0$ holds, e.g. one can set $\hat{M}(0) = \frac{\hat{\sigma}_u^2}{1-\rho^2}$. The opposite holds for negative β, i.e. $\hat{M}(n)$ will be an optimistic measure of actual estimation error variance. In this case, one can e.g. set $\hat{M}(0) = 0$ to guarantee that $M(n) \geqslant \hat{M}(n)$, $\forall n$. However, as it will be seen afterwards, $M(n)$ can still be upper bounded by $\hat{M}(n)$ times a constant.

It is known from the analysis in this chapter that $\hat{M}(n)$ converges in distribution for $n \to \infty$ and also achieves a diversity order of d. This is simply due to the fact that the equation for $\hat{M}(n)$ here has the same recursion form as the one with perfect system knowledge. Converting the diversity result from the log-log scale back to linear scale, we should for high SNR have that

$$\Pr(\hat{M} \geqslant M_{\text{th}}) = c\lambda^d, \tag{3.45}$$

where c is a constant not depending on λ, but a function of other system parameters (c is a function of the coding gain). From (3.45), we should also have for $n \to \infty$ that

$$\text{Pr}(\hat{M}(n) \geqslant M_{\text{th}}) = c\lambda^d, \tag{3.46}$$

simply because if (3.46) does not hold, then (3.45) cannot hold either as $\hat{M}(n)$ converges in distribution. From the form of the outage function in Theorem 2, it can be deduced that

$$\text{Pr}(\hat{M}(n) \geqslant M_{\text{th}}) = c^{'} \frac{1}{M_{\text{th}}^d} \lambda^d, \tag{3.47}$$

because all relevant powers of λ only appear in the form λ/M_{th}.

Before we proceed, consider the following lemma which will be used afterwards.

Lemma 10. *For two arbitrary random variables Z_1 and Z_2 and an arbitrary threshold T, if we assume that $Z_1 \leqslant Z_2$ each time an outcome is produced, then we will have that $Pr(Z_1 \geqslant T) \leqslant Pr(Z_2 \geqslant T)$.*

Proof.

$$\begin{aligned}
\text{Pr}(Z_2 \geqslant T) &= \text{Pr}(Z_2 \geqslant T | Z_1 \geqslant T)\text{Pr}(Z_1 \geqslant T) \\
&\quad + \text{Pr}(Z_2 \geqslant T | Z_1 < T)\text{Pr}(Z_1 < T) \\
&\overset{(a)}{=} 1 \cdot \text{Pr}(Z_1 \geqslant T) + \text{Pr}(Z_2 \geqslant T | Z_1 < T)\text{Pr}(Z_1 < T) \\
&\overset{(b)}{\geqslant} \text{Pr}(Z_1 \geqslant T)
\end{aligned} \tag{3.48}$$

where (a) is due to the fact that if $Z_1 \leqslant Z_2$, and $Z_1 \geqslant T$, then certainly we have that $Z_2 \geqslant T$, and (b) holds because any probability are greater than or equal to zero. □

Now we consider the case $\beta \geqslant 0$ and $\beta < 0$ separately for their effects on diversity.

For $\beta \geqslant 0$, it has already been argued that $M(n) \leqslant \hat{M}(n)$. According to Lemma 10, it should hold that $\text{Pr}(M(n) \geqslant M_{\text{th}}) \leqslant \text{Pr}(\hat{M}(n) \geqslant M_{\text{th}})$. As the outage probability for $\hat{M}(n)$ has a diversity order of d, then the outage probability for $M(n)$ must at least have a diversity order of d as well. Due to the fact that mismatched $M(n)$ is supposed to have a worse performance than the original $M(n)$ and it has been shown that the original $M(n)$ has a diversity order of d for its outage probability, then $M(n)$ with mismatched parameter must have a diversity order equal to d. Equal diversity order for $M(n)$ and $\hat{M}(n)$ outage probabilities indicates that in the logarithmic scale for SNR, any gap between the two outage probabilities which is due to mismatch, is equal to a constant at asymptotic SNR.

For $\beta < 0$, we see from (3.40) and (3.44) that

$$\begin{aligned}
\frac{M(n) - \hat{M}(n)}{\hat{M}(n)} &= \frac{\rho^2 \left(M(n-1) - \hat{M}(n-1) \right) + |\beta|}{\hat{P}(n)(1 + \gamma(n)\hat{P}(n))} \\
&\leqslant \frac{\rho^2 C + |\beta|}{\hat{\sigma}_u^2},
\end{aligned} \tag{3.49}$$

where C is a constant such that $M(n) - \hat{M}(n) \leqslant C$. To show that the inequality in (3.49) holds, we first see that $0 \leqslant \hat{\sigma}_u^2 \leqslant \hat{P}(n)$ and $0 \leqslant \gamma(n) < \infty$, which results in $0 \leqslant \left(\hat{P}(n)(1+\gamma(n)\hat{P}(n))\right)^{-1} \leqslant 1/\hat{\sigma}_u^2$. In order to verify that C exists, one may begin by using [47, Theorem 7.7 and corollary] to show that $M(n)$ is bounded. According to [47, Theorem 7.7 and corollary], certain conditions must be satisfied for boundedness of $M(n)$, namely (i) observability and controllability of the mismatched system, (ii) boundedness of the eventual difference between the biases of the actual and the mismatched system, and (iii) boundedness of $M(0)$. For (i), note that with perfectly known Rayleigh fading channel, the systems (actual and mismatched) are uniformly completely observable and controllable. For (ii), note that no bias is assumed in this chapter, so this condition holds as well. Furthermore, (iii) can also be guaranteed by setting $\hat{x}(0)$ to some finite value, such that $M(0) = E(\|x(0) - \hat{x}(0)\|)$ is always bounded. As a consequence and based on [47, Theorem 7.7 and corollary], $M(n)$ is bounded. In addition, we have for the mismatched filter that $\hat{M}(n) \leqslant \frac{\hat{\sigma}_u^2}{1-\rho^2}$, with the equality happening given the worst case (all zero) channel. As a result, $M(n) - \hat{M}(n)$ is also bounded. We can then say that there exists a finite positive constant C, such that $M(n) - \hat{M}(n) \leqslant C$, $\forall n$.

This yields

$$\begin{aligned}
\Pr\left(M(n) \geqslant M_{\text{th}}\right) &\leqslant \Pr(\hat{M}(n)(1+\frac{\rho^2 C + |\beta|}{\hat{\sigma}_u^2}) \geqslant M_{\text{th}}) \\
&= \Pr(\hat{M}(n) \geqslant M_{\text{th}}/(1+\frac{\rho^2 C + |\beta|}{\hat{\sigma}_u^2})) \\
&= c^{'} \frac{1}{(M_{\text{th}}/(1+\frac{\rho^2 C + |\beta|}{\hat{\sigma}_u^2}))^d} \lambda^d, \qquad (3.50)
\end{aligned}$$

and proves a diversity order equal to d for the outage probability of $M(n)$ as well.

At this point, we try to illustrate the diversity result more clearly by means of simulation. As no exact expression is available for $M(n)$ for the mismatched case, $M(n)$ can be obtained by simulating the Kalman filter with mismatched parameters $i_{\max}$ number of times. If we then measure the squared error at time n of the i-th simulation by $e_i(n) = |x_i(n) - \hat{x}_i(n)|^2$, then by the law of large numbers, we have that $M(n) = \lim_{i_{\max}\to\infty} \frac{1}{i_{\max}} \sum_i e_i(n)$. For consistency, the random channel sequence $\mathbf{h}(n)$ is kept constant for all i, but new random values are used for $u(n)$ and $\mathbf{v}(n)$. The following simulation parameters are selected $\sigma_u^2 = \sigma_v^2 = 1, \rho = 0.95, d = 1, M_{\text{th}} = 0.1$ and $\beta = 0, 5, -0.75$. In addition, it was chosen that $i_{\max} = 100$ in order to balance the computational complexity and accuracy of the simulation. The result is depicted in Fig. 3.6. As the figure suggests, the diversity order for the outage probability with mismatched σ_u^2 is still held and there is a constant loss in SNR in the high SNR regime. The figure also suggests that underestimating σ_u^2, i.e. $\beta < 0$ has a more deteriorating effect than overestimating it.

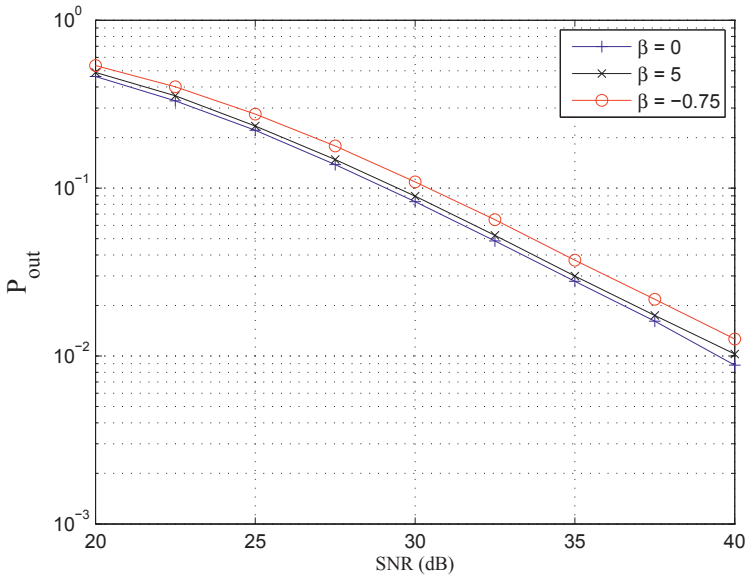

Figure 3.6: Effect of mismatched σ_u^2 on diversity order

Remark 2. A sufficient condition for the derivations of this section to be correct for a random β is that the support of $f_\beta(\beta)$ is bounded for any SNR. This way, one makes sure that $|\beta|$ does not increase with SNR such that (3.50) always holds.

It should be mentioned that similar analysis for the development of $M(n) - \hat{M}(n)$ already exists in [47, Ch. 7.8]. Yet, it was decided to include the necessary derivations for better readability, as [47] uses different notations.

3.4.2 Mismatched Channel Noise Variance

In this subsection, we consider the case when $\hat{\sigma}_v^2 = \sigma_v^2 + \alpha$, with α being a random variable drawn from a prior distribution $f_\alpha(\alpha)$, but constant for the period of interest.

Similarly for mismatched σ_v^2, one should again modify the Kalman equations for $M(n)$ and $\hat{M}(n)$, considering the mismatch. Following the same logic as for mismatched σ_u^2, we obtain that

$$M(n) = \frac{\hat{\sigma}_v^4 P(n) + \sigma_v^2 \hat{P}^2(n) \|\mathbf{h}(n)\|^2}{\left(\hat{\sigma}_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)^2} \tag{3.51}$$

and

$$\hat{M}(n) = \frac{\hat{\sigma}_v^2 \hat{P}(n)}{\left(\hat{\sigma}_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)}$$

and therefore have that

$$\begin{aligned} M(n) - \hat{M}(n) &= \frac{\hat{\sigma}_v^4 \left(P(n) - \hat{P}(n)\right) + (\sigma_v^2 - \hat{\sigma}_v^2)\hat{P}^2(n)\|\mathbf{h}(n)\|^2}{\left(\hat{\sigma}_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)^2} \\ &= \frac{\hat{\sigma}_v^4 \rho^2 \left(M(n-1) - \hat{M}(n-1)\right) - \alpha \hat{P}^2(n)\|\mathbf{h}(n)\|^2}{\left(\hat{\sigma}_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)^2} \end{aligned} \tag{3.52}$$

with $\alpha = \hat{\sigma}_v^2 - \sigma_v^2$.

For $\alpha \geqslant 0$, it is easy to show by induction that $M(n) \leqslant \hat{M}(n)$, $\forall n$, i.e. the actual estimation error variance is upper bounded by the estimation error variance obtained by the Kalman filter. Similar to the case of $\beta \geqslant 0$, we see that the diversity order for the outage probability of $M(n)$ is exactly equal to d.

For $\alpha < 0$, it may also be shown by induction that $M(n) \geqslant \hat{M}(n)$, $\forall n$, i.e. the estimation error variance from the Kalman filter ($\hat{M}(n)$) is an optimistic value for $M(n)$. However, using (3.51) and (3.52), we obtain that

$$\frac{M(n) - \hat{M}(n)}{\hat{M}(n)} = \frac{\hat{\sigma}_v^4 \rho^2 \left(M(n-1) - \hat{M}(n-1)\right) - \alpha \hat{P}^2(n)\|\mathbf{h}(n)\|^2}{\hat{\sigma}_v^2 \hat{P}(n) \left(\hat{\sigma}_v^2 + \|\mathbf{h}(n)\|^2 \hat{P}(n)\right)}. \tag{3.53}$$

Now, we have that

$$0 \leqslant \frac{\hat{P}(n)\|\mathbf{h}(n)\|^2}{\hat{\sigma}_v^2 + \hat{P}(n)\|\mathbf{h}(n)\|^2} \leqslant 1. \tag{3.54}$$

In addition, we can argue that there exists a constant C', such that $M(n) - \hat{M}(n) \leqslant C'$, $\forall n$, similar to the arguments for mismatched σ_u^2. That, combined with (3.53) yields $\frac{M(n)-\hat{M}(n)}{\hat{M}(n)} \leqslant \frac{\rho^2 C'}{\sigma_u^2} + \frac{|\alpha|}{\hat{\sigma}_v^2}$, which results in $M(n) \leqslant (1 + \frac{\rho^2 C'}{\sigma_u^2} + \frac{|\alpha|}{\hat{\sigma}_v^2})\hat{M}(n)$, and consequently

$$\begin{aligned} \text{Pr}(M(n) \geqslant M_{\text{th}}) &\leqslant \text{Pr}\left((1 + \frac{\rho^2 C'}{\sigma_u^2} + \frac{|\alpha|}{\hat{\sigma}_v^2})\hat{M}(n) \geqslant M_{\text{th}}\right) \\ &= \text{Pr}(\hat{M}(n) \geqslant M_{\text{th}}/(1 + \frac{\rho^2 C'}{\sigma_u^2} + \frac{|\alpha|}{\hat{\sigma}_v^2})) \\ &= c' \frac{1}{(M_{\text{th}}/(1 + \frac{\rho^2 C'}{\sigma_u^2} + \frac{|\alpha|}{\hat{\sigma}_v^2}))^d} \lambda^d, \end{aligned} \tag{3.55}$$

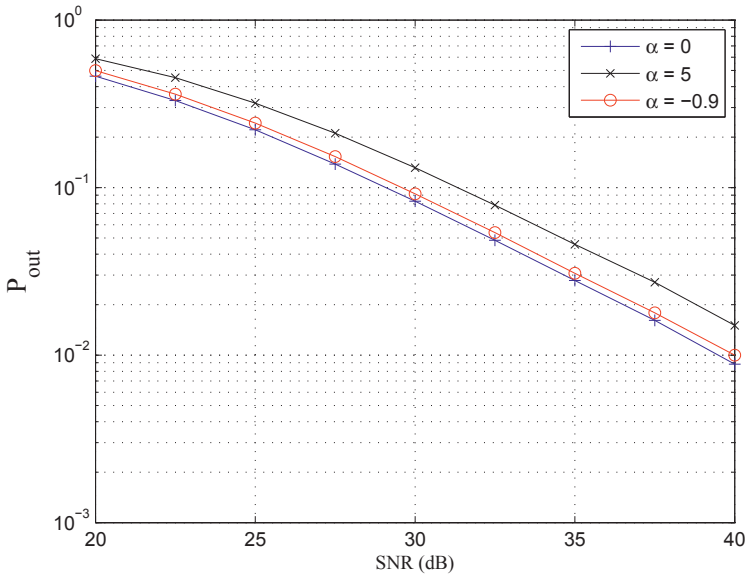

Figure 3.7: Effect of mismatched σ_v^2 on diversity order

which proves a diversity order of d for $M(n)$.

In order to illustrate the mismatched diversity result more clearly for σ_v^2 as well, a numerical simulation was performed using the following parameters $\sigma_u^2 = \sigma_v^2 = 1, \rho = 0.95, d = 1, M_{\text{th}} = 0.1$ and $\alpha = 0, 5, -0.9$. The result is depicted in Fig. 3.7. As the figure suggests, the diversity order for the outage probability with mismatched σ_v^2 is still held and there is a constant loss in SNR in the high SNR regime. The figure also suggests that overestimating σ_v^2, i.e. $\alpha > 0$ has a more deteriorating effect than underestimating it.

Remark 3. It is easy to see from (3.55) that similar to the arguments in Remark 2, a sufficient condition for the derivations of this section to be correct for a random α is that the support of $f_\alpha(\alpha)$ is bounded for any SNR.

3.5 Summary and Discussion

In this chapter, we studied the pdf of the instantaneous estimation error variance resulting from sending a scalar Gauss-Markov process over d parallel independent Rayleigh fading channels. In particular, the focus was on the performance of the Kalman filter with respect to the outage probability of the instantaneous estimation error variance and provided the outage probability using a closed-form formula for the range of thresholds up to the process noise variance. Upper and lower bounds were then proposed for the out-

age probability. Furthermore, it was shown that in the limit of the high SNR, the outage probability achieves a diversity order the same as the number of available independent channels. Finally, the effect of imperfect system knowledge on the diversity order was studied and it was shown that a bounded uncertainty in the process and noise variance will not affect the diversity order.

Regarding the bounds, it should be noted that although not reported here, selecting $d = 1$ for the bounds in this chapter will result in bounds which have very close values to the ones in Chapter 2. However, the bounds of Chapter 2 were needed for diversity analysis. Therefore, they were presented separately.

Regarding the diversity analysis, it should be noted that the diversity order for the outage probability as analyzed in Chapter 2 may be covered as the special case of $d = 1$ in this chapter. However, although the diversity analysis in this chapter is more general and results in new theorems (which might have application in other contexts as well), they fall short of being able to describe the coding gain for $d > 1$. Due to that, it was decided to treat the case of $d = 1$ (with regard to diversity analysis) not as a special case of the material in this chapter. Therefore, it was presented separately.

In this chapter, it was assumed that all the channel branches have the same SNR. It is however not necessary for all the channels to have equal SNR's in order to get a diversity order of d. If the channel branches are unequal, one can always consider another channel with equal SNR's on all branches, but for which the SNR is equal to the minimum of the SNR's of the original channel. This new channel which obviously incurs an inferior outage performance, still has a diversity order of d according to our analysis. Therefore, the original (superior) channel with unequal SNR's should have a minimum diversity order of d. The same line of reasoning can be made using another channel with equal SNR's on all branches, whose values are equal to the maximum SNR of the original channel, which leads to a maximum diversity order of d for the original channel. Therefore, we have a diversity order of d for the channel with unequal SNR's as well. Regarding the bounds for the case when the channels have unequal SNR's, it should be noted that obtaining the bounds is not very straightforward in that case, because $\gamma(n)$ will not have a χ^2 distribution anymore. However, it is possible to obtain the new distribution individually for each case and then extend and adapt the results on the bounds, albeit involving more tedious derivations.

In this chapter and the previous one, transmission of a scalar source over one or several i.i.d. branches is considered. This corresponds to a SISO or SIMO channel model. If the source is a vector, then we deal with MIMO channels and thus matrix-valued $P(n)$ and $M(n)$ rather than scalar ones. Due to the nature of the Kalman equations, specifically the matrix inversions which are involved for the vector case, new tools and techniques may need to be applied for a possible extension of the results of Chapters 2 and 3. From the development of the diversity order analysis, we know that the availability of independent channel branches is a necessity. One may then use channel parallelization techniques in order to create independent (parallel) branches from the MIMO channel. The current bounds however may not directly be extended, because they make a significant use of the

fact that the process and process noise are scalars. The analysis of the estimation error outage probability and its behavior in different SNR regime for vector signals is the topic of the next chapter.

Chapter 4

Vector Signals over MIMO Channels: Full Diversity

In this chapter, the system model is updated for the transmission of first-order vector (arbitrary N and K) ARGM signals over a non-frequency selective MIMO Rayleigh fading channel. The signal is reconstructed at the receiver side with the help of a Kalman filter in order to minimize the instantaneous mean squared error. Complex orthogonal space time codes are utilized in order to increase the quality of estimation and mitigate the destructing effects of the fading channel via achieving the maximum diversity order for the distortion outage probability. In other words, space-time codes and Kalman filtering are jointly incorporated in a common framework, which allows for extra reliability for estimation of ARGM sources over fading channels. The diversity order analysis is performed in the high SNR regime, in line with the previous chapters. Furthermore, upper and lower bounds are obtained for the outage probability as a function of system parameters in the finite SNR regime. That is followed by the high SNR analysis of the bounds, through which achievability of the maximum diversity order for a $N \times K$ MIMO fading channel, equal to KN, is proven. In addition, upper and lower bounds are obtained for the coding gain of the distortion outage probability in the high SNR regime, and outline the correspondence between system parameters and the coding gain.

4.1 Updated System Model and Problem Definition

Consider the following updated system model

$$\begin{aligned} \mathbf{x}(n) &= A\mathbf{x}(n-1) + \mathbf{u}(n) \\ Y(n) &= \sqrt{P_t/(KN)}H(n)T\left(\mathbf{x}(n)\right) + V(n), \end{aligned} \tag{4.1}$$

where $\mathbf{x}(n)$ and $\mathbf{u}(n)$ are column vectors of dimension K and represent the to-be-transmitted signal and the process noise, respectively. In this model, $\mathbf{x}(n)$ is a first order Gauss-Markov process. With respect to that, A is the state-transition matrix and assumed to be

non-singular. This is a sufficient condition for existence of the steady-state outage probability function, and is similar to the $\rho \neq 0$ assumption in Chapters 2 and 3. More details can be found in [10, Theorem 2.4].

The space-time block encoding[1] operation (the pre-filter block of Chapter 1) is represented by the operator $T(\cdot)$. The output of the space-time encoding operation, i.e. $T\left(\mathbf{x}(n)\right)$, is a matrix of dimension $K \times N_c$, which corresponds to N_c channel uses by each of the transmit antennas for each new source symbol $\mathbf{x}(n)$. (The details of the space-time coding operation and the structure of $T\left(\mathbf{x}(n)\right)$ are presented in Sec. 4.2.1). In this work, it is assumed that the number of transmit antennas is equal to K, i.e. the source dimension.

The MIMO channel matrix of dimension $N \times K$ is denoted by $H(n)$, which consists of i.i.d. complex Gaussian elements with zero mean and unit variance (real and imaginary parts have a variance equal to one half), i.e. non-frequency selective Rayleigh fading. For some of the derivations in the upcoming sections, it is required that $H(n)$ are also i.i.d. in time.

At the receiver, the received signals from the channel, and the channel noises for the N_c channel uses are denoted by $Y(n)$ and $V(n)$, which are matrices of dimension $N \times N_c$. The value of P_t is also selected such that required SNR at the receiver is achieved. The elements of $V(n)$ are also considered to be i.i.d. complex Gaussian random variables. The covariance matrices for $\mathbf{u}(n)$ is denoted by C_u and the elements of $V(n)$ have a variance equal to σ_v^2. It is also assumed that the $H(n)$ are perfectly known to the receiver.

The vector source $\mathbf{x}(n)$ is space-time encoded at the transmitter side and sent over the channel. There are then two major operations which should be performed at the receiver. The first operation is the space-time decoding, i.e. the inverse operation for $T\left(\mathbf{x}(n)\right)$, which in turn leads to an equivalent channel and received signal model. The next operation is the estimation of $\mathbf{x}(n)$ from the received signal. The optimal causal minimum mean square error (MMSE) estimator for this setting is the Kalman filter. The Kalman filter provides us with an optimal estimate of the source at the receiver, namely $\hat{\mathbf{x}}(n)$ which minimizes the (normalized) random instantaneous distortion at time n as defined in Chapter 1. The distortion outage probability is also the same as defined in Chapter 1. In this case, the value of SNR may be obtained from

$$\text{SNR} = \frac{P_t}{KN} \frac{E\left(\|H(n)\|^2\right) E\left(\|T(\mathbf{x}(n)\|^2\right)}{E\left(\|V(n)\|^2\right)},$$

In line with the material in Chapters 2 and 3, we are interested in the characterization of the distortion outage probability, albeit via upper and lower bounds, and the high SNR behavior in terms of the diversity order and the coding gain, with the same definitions as in Chapter 1. As it will be seen later, the maximum diversity order is only dependent on the number of available independent individual channel branches, and its achievability

[1] In this chapter, the terms encoding and decoding are used for space-time codes. However, the signals are still uncompressed and the space-time decoding for estimation is different from decoding for detection.

only depends on the space-time code. The coding gain however is a function of the source structure and the selected threshold.

4.2 Joint Space-time Coding and Kalman Filtering

In this section and first, the space-time coding scheme used in this chapter and the Kalman filter which is used in order to estimate the transmitted signal, are introduced. It is then described how these two parts should interact with one another.

4.2.1 Analog Space-time Coding

A space-time block code based on orthogonal designs as defined in [98], is used for transmission of $\mathbf{x} = [x_1, x_2, \ldots, x_K]^T$ over the channel $H_{N\times K}$ (the time index n is dropped in this section). The encoding is adopted (and slightly modified) from [98] as follows. We form a matrix $X = T(\mathbf{x})$ of dimension $K \times N_c$, as instructed in [98] and which consists of elements $\pm x_1, \pm x_2, \ldots, \pm x_K$, their conjugates $\pm x_1^*, \pm x_2^*, \ldots, \pm x_K^*$ or multiples of these elements by $\pm \boldsymbol{i}$ (with $\boldsymbol{i} = \sqrt{-1}$) or, if necessary, other scaling factors. The first column of X can without loss of generality be assumed to be $[x_1, x_2, \ldots, x_K]^T$. The space-time code rate can be defined as

$$r = K/N_c \tag{4.2}$$

source dimensions per channel use. The code rate needs to be maximized in order to minimize the extra incurred channel uses. This is however not the focus of this work and the reader is mainly referred to the current literature on space-time codes for that matter. The code design is such that XX^H (X^H is the conjugate transpose of X) is a diagonal matrix. It is also shown in [98] that if at least one orthogonal design exists, one can always find another design such that

$$XX^H = c\|\mathbf{x}\|^2 I_K,$$

where c is some constant depending on the code [2]. It is also possible to normalize the codewords such that $c = 1$, as it is assumed in the rest of this chapter.

Referring to the system model in (4.1), each row of Y and V corresponds to a particular receive antenna, comprising a total number of N receive antennas, and each column of Y and V corresponds to one channel use, comprising a total number of N_c channel uses for each source symbol transmission, indexed by n. Note that H, although random, is fixed for the transmission of each source symbol.

[2] In this chapter, the transpose of the codes presented in [98] are used, because it is better suited to the generic system model of this thesis.

At the receiver side, the space-time coded signal Y should be decoded first before it is directed to the Kalman filter in order to estimate $\mathbf{x}$. The objective of the space-time decoder, appearing before the Kalman filter, is to provide an equivalent orthogonal channel and allow for a spatial diversity gain. The number of such orthogonal branches for the equivalent channel is at most NK. It is worthwhile mentioning that N has no effect on the code selection as long as the code is orthogonal.

The decoding suggested here is different from what is proposed in [98] due to the different nature of estimation and detection. The approach used in this chapter is similar to the one used in [13], where orthogonal space-time block codes are used for analog channel state information feedback. The basic idea is to convert the channel into an equivalent orthogonal channel and then perform decoding by simply multiplying the received vector by the transpose of the equivalent channel. The method in [13] is applicable for the transmission of real signals only, and uses the 1/2 rate codes based on real orthogonal designs proposed in [98]. While the same design can be used for the purpose of this work as well (by alternating the transmission of real and imaginary parts of the signal $\mathbf{x}(n)$ and therefore having a code of rate 1/4), we propose a different approach which allows for the incorporation of the available complex orthogonal space-time codes, and thus operating at a potentially better rate (e.g. full rate for $K = 2$). This is made possible by converting all the complex vectors into equivalent real vectors with twice the size and finding the equivalent real channel. The proposed space-time decoding can be performed as follows.

Consider the l-th row of H corresponding to the l-th receiver ($l = 1, 2, \ldots, N$), and call that row $\mathbf{h}_l$. Take then the corresponding rows in Y and V to be $\mathbf{y}_l$ and $\mathbf{v}_l$. The received signal for that receiver is $\mathbf{y}_l = \sqrt{P_t/(KN)}\mathbf{h}_l X + \mathbf{v}_l$. Note that it is only enough to analyze the space-time code for one receive antenna. Through that, it is possible to show that each receiver is able to provide K orthogonal channels, involving the corresponding row in H. With N independent rows, a number of NK orthogonal channels can be created by simply summing all the results of the space-time decoding for each receiver.

The next step is to convert all the complex operations to real ones. First extend the source vector $\mathbf{x}$ into a real vector of dimension $2K$ by replacing each complex element by a 2×1 real vector of the real and imaginary part of that element and call this new vector $\mathbf{x}_r$, i.e.

$$\mathbf{x}_r = [x_1^r, x_1^i, x_2^r, x_2^i, \ldots, x_K^r, x_K^i]^T,$$

where the superscripts r and i for each element x_k, $k = 1, 2, \ldots, K$ indicate the real and imaginary part of that element. Then perform the same procedure for X to provide the matrix X_r of dimension $2K \times N_c$. Vectorize (column-wise reshape) the matrix X_r into a real vector of size $2KN_c \times 1$, in the same manner as done for $\mathbf{x}_r$, and call it $\tilde{\mathbf{x}}$. It is now possible to create a linear mapping, i.e. the matrix T with size $2KN_c \times 2K$, which maps $\mathbf{x}^r$ to $\tilde{\mathbf{x}}$ and only consists of real numbers. In other words, we must find a T such that it satisfies $\tilde{\mathbf{x}} = T\mathbf{x}^r$. This can be done by considering that each element in $\tilde{\mathbf{x}}$ can be found in $\mathbf{x}^r$ possibly with a different sign and scaling factor, or as a linear combination of the elements of $\mathbf{x}^r$.

We convert $\mathbf{v}_k$ to the equivalent real vector $\tilde{\mathbf{v}}$ in the same manner.

Next, consider converting the channel into the real and imaginary parts. For this reason, each channel tap in $\mathbf{h}_l$, i.e. $h_{l,k},\ k = 1, 2, \ldots, K$ is converted to the following 2×2 matrix

$$\tilde{H}_{l,k} = \begin{bmatrix} h^r_{l,k} & -h^i_{l,k} \\ h^i_{l,k} & h^r_{l,k} \end{bmatrix}. \tag{4.3}$$

The complex-valued $1 \times K$ vector $\mathbf{h}_l$ is then expanded to the real-valued $2 \times 2K$ channel $\tilde{H} = [\tilde{H}_{l,1}|\tilde{H}_{l,2}, \ldots, |\tilde{H}_{l,K}]$. With these definitions, it can be easily shown that the operation $\mathbf{y}_l = \sqrt{P_t/(KN)}\mathbf{h}_l X + \mathbf{v}_l$ in the domain of complex numbers can be represented by the following operation over the domain of real numbers

$$\begin{aligned} \tilde{\mathbf{y}} &= \sqrt{P_t/(KN)}(I_{N_c} \otimes \tilde{H})\tilde{\mathbf{x}} + \tilde{\mathbf{v}} \\ &= \sqrt{P_t/(KN)}(I_{N_c} \otimes \tilde{H})T\mathbf{x}^r + \tilde{\mathbf{v}} \\ &= \sqrt{P_t/(KN)}H_{\text{eq}}\mathbf{x}^r + \tilde{\mathbf{v}}, \end{aligned} \tag{4.4}$$

where $H_{\text{eq}} = (I_{N_c} \otimes \tilde{H})T$ is the equivalent real channel which acts on the equivalent real source vector $\mathbf{x}_r$ (note the conversion from complex row vectors to equivalent real column vectors). It will be shown in Appendix E that

$$H^T_{\text{eq}}H_{\text{eq}} = \|\mathbf{h}_l\|^2 I_{2K},$$

i.e. the equivalent real channel can be orthogonalized by a simple operation constructed as follows

$$\begin{aligned} H^T_{\text{eq}}\tilde{\mathbf{y}} &= H^T_{\text{eq}}\left(\sqrt{P_t/(KN)}H_{\text{eq}}\mathbf{x}^r + \tilde{\mathbf{v}}\right) \\ &= \sqrt{P_t/(KN)}H^T_{\text{eq}}H_{\text{eq}}\mathbf{x}^r + H^T_{\text{eq}}\tilde{\mathbf{v}} \\ &= \sqrt{P_t/(KN)}\|\mathbf{h}_l\|^2\mathbf{x}^r + H^T_{\text{eq}}\tilde{\mathbf{v}}. \end{aligned} \tag{4.5}$$

Due to the independence of channel noises for each dimension, it can be shown that the variance of each element of $H^T_{\text{eq}}\tilde{\mathbf{v}}$ is equal to $\|\mathbf{h}_l\|^2\sigma_v^2$.

For decoding the whole received signals, one should perform the same procedure for all the receiver antennas, i.e. all the rows of H, and sum the results. If we call the resulting sum $\mathbf{y}_{\text{eq}}$ and convert the real vectors back to the complex domain again, we may finally write

$$\begin{aligned} \mathbf{y}_{\text{eq}} &= \sqrt{P_t/(KN)}\sum_{l=1}^{N}\|\mathbf{h}_l\|^2\mathbf{x} + \mathbf{v}_{\text{eq}} \\ &= \sqrt{P_t/(KN)}\|H\|_F^2\mathbf{x} + \mathbf{v}_{\text{eq}}, \end{aligned} \tag{4.6}$$

where each element in $\mathbf{v}_{\text{eq}}$ has the variance $\|H\|_F^2\sigma_v^2$. The SNR is then equal to $\text{SNR} = P_tP_x/\sigma_v^2$, where $P_x = E(\|\mathbf{x}(n)\|^2)$. After space-time decoding for each time step n, $\mathbf{y}_{\text{eq}}(n)$ is delivered to the Kalman filter in order to estimate $\mathbf{x}(n)$. This is reviewed in the next section.

4.2.2 Kalman Filtering of Space-time Coded Analog Sources

Given that the correct initialization and equivalent channel and received signal model are used, the general equations for the vector Kalman filter (the estimator) adapted from [52] are

$$\begin{aligned}
&\hat{\mathbf{x}}(n|n-1) = A\hat{\mathbf{x}}(n-1|n-1)\\
&P(n) = AM(n-1)A^T + C_u\\
&G_k(n) = \sqrt{P_t/(KN)}P(n)\|\mathbf{H}(n)\|^2\Big(\|\mathbf{H}(n)\|^2\sigma_v^2 + P_t/(KN)P(n)\|\mathbf{H}(n)\|^4\Big)^{-1}\\
&\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1) + G_k(n)\Big(\mathbf{y}_{\text{eq}}(n) - \sqrt{P_t/(KN)}\|\mathbf{H}(n)\|^2\hat{\mathbf{x}}(n|n-1)\Big)\\
&M(n) = \Big(I - K(n)\sqrt{P_t/(KN)}\|\mathbf{H}(n)\|^2\Big)P(n).
\end{aligned} \tag{4.7}$$

The second step in (4.7), i.e. $P(n) = AM(n-1)A^T + C_u$ is the prediction, and it is known that the prediction error covariance matrix $P(n)$ propagates through the random Riccati equation given in (4.8).

It is also possible to simplify (4.8) and obtain

$$\begin{aligned}
P(n+1) &= A\left(P^{-1}(n) + P_t/(KN)\|\mathbf{H}(n)\|^2/\sigma_v^2 I\right)^{-1}A^T + C_u\\
&= AM(n)A^T + C_u
\end{aligned} \tag{4.9}$$

by invoking the Woodbury matrix identity on (4.9). Comparing (4.9) with the second line in (4.7) necessitates that

$$M(n) = \left(P^{-1}(n) + P_t/(KN)\|\mathbf{H}(n)\|^2/\sigma_v^2 I\right)^{-1}. \tag{4.10}$$

We may then rewrite (4.10) as

$$\begin{aligned}
M(n) &= \left(P^{-1}(n) + P_t/(KN)\|\mathbf{H}(n)\|^2/\sigma_v^2 I\right)^{-1}\\
&= \frac{KN}{P_t\|\mathbf{H}(n)\|^2/\sigma_v^2}\left(\left(P_t/(KN)\|\mathbf{H}(n)\|^2/\sigma_v^2 P(n)\right)^{-1} + I\right)^{-1}\\
&= \frac{1}{\gamma_n}\left(\frac{1}{\gamma_n}P^{-1}(n) + I\right)^{-1},
\end{aligned} \tag{4.11}$$

$$\begin{aligned}
&P(n+1) = AP(n)A^T\\
&- P_t/(KN)AP(n)\|\mathbf{H}(n)\|^2\left(\|\mathbf{H}(n)\|^2\sigma_v^2 + P_t/(KN)\|\mathbf{H}(n)\|^4 P(n)\right)^{-1}\|\mathbf{H}(n)\|^2 P(n)A^T\\
&+ C_u.
\end{aligned} \tag{4.8}$$

with

$$\gamma_n = \frac{P_t \|\mathbf{H}(n)\|^2}{\sigma_v^2 K N}. \tag{4.12}$$

One can also rewrite (4.9) (while setting $n-1$ instead of n) as

$$\begin{aligned} P(n) &= AM(n-1)A^T + C_u \\ &= \frac{A}{\gamma_n}\left(\frac{1}{\gamma_n}P^{-1}(n-1) + I\right)^{-1} A^T + C_u. \end{aligned} \tag{4.13}$$

If we denote the k-th diagonal element of $M(n)$ by $M_{kk}(n)$ and define the distortion as $d(n) = \frac{1}{K}\text{tr}\left(M(n)\right)$, the distortion outage probability at time n is equal to

$$\begin{aligned} \text{P}_{\text{out}}(\text{d}_{\text{th}}) &= \text{Pr}\left(d(n) \geqslant \text{d}_{\text{th}}\right) \\ &= \text{Pr}\left(\frac{1}{K}\sum_{k=1}^{K} M_{kk}(n) \geqslant \text{d}_{\text{th}}\right), \end{aligned} \tag{4.14}$$

where d_{th} is an arbitrary threshold value. The analysis of this outage probability as a function of SNR and other system parameters is the topic of the next section.

4.3 Outage Probability Analysis

In this section, we study the achievable diversity order and coding gain for distortion outage probability when the complex orthogonal space-time codes are used in conjunction with the Kalman filter. Upper and lower bounds on the distortion outage probability are also developed. The bounds are used for both obtaining numerical values with application in practical systems and are also a prerequisite tool in obtaining the asymptotic results.

4.3.1 Bounds for the Outage Probability

As the distortion at each time step n is obtained from $M(n)$, we should first try to develop an equation for $M(n)$ which makes distortion calculation possible. One easy way to find the diversity order is finding a closed-form equation for $M(n)$ which is independent of $P^{-1}(n)$ such that it allows for diversity order calculation. This however, proves to be rather complicated, as it was already seen for SISO and SIMO cases in Chapters 2 and 3, respectively. While some initial results were obtained there, even for those simpler cases, we had to resort to finding bounds and approximates for characterization of the outage probability and then its high SNR behavior. The same approach is followed in this chapter as well. In that regard, we first establish upper and lower bounds for the outage

probability and then in the following section obtain the diversity order and coding gain via the high SNR analysis of the bounds.

In order to get upper and lower bounds for $P_{\text{out}}(d_{\text{th}})$, the following fact is used (Lemma 10). If for two random variables X and Y, we have that $X \leqslant Y$, then as a result we obtain that $\text{Pr}(X \geqslant T) \leqslant \text{Pr}(Y \geqslant T)$. As $d(n) = \frac{1}{K}\text{tr}(M(n))$, if we can find $d^l(n)$ and $d^u(n)$ such that we would have $d^l(n) < d(n) < d^u(n)$, then we may bound the outage probability as $\text{Pr}(d^l(n) \geqslant d_{\text{th}}) < \text{Pr}(d(n) \geqslant d_{\text{th}}) < \text{Pr}(d^u(n) \geqslant d_{\text{th}})$, i.e. upper and lower bounds on the outage probability may be established. We would prefer random variables whose cdf have a Taylor series with the first non-zero term equal to that of the original distortion variable. This is for the diversity analysis to be successful and will be explained later in the section.

In the following lemmas (Lemma 11 and 12), we present $d^l(n)$ and $d^u(n)$, used to establish upper and lower bounds on the outage probability.

Lemma 11. *The instantaneous distortion $d(n)$ may be lower bounded by $d^l(n) \leqslant d(n)$ as*

$$d^l(n) = \frac{1}{\gamma_n + \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}},$$

where the $\zeta_l(.)$ function denotes the l-th eigenvalue of its matrix argument.

Proof. Please see Appendix F.

Lemma 12. *The instantaneous distortion $d(n)$ may be upper bounded by $d^u(n) \geqslant d(n)$ as follows*

$$d^u(n) = \frac{1}{\gamma_n + \frac{1}{\frac{1}{K}\sum_{l=1}^{K}\theta_l}}$$

with

$$\theta_l = \zeta_l(C_u) + \frac{|\zeta_{max}(A)|^2}{\frac{1}{\alpha_{max}} + \gamma_{n-1}}, \tag{4.15}$$

where $\zeta_{max}(.)$ denotes the maximum value of the eigenvalues of its matrix argument, and $\alpha_{max} = |\zeta_{max}(A)|^2\zeta_{max}(C_x) + \zeta_{max}(C_u)$, with C_x being the stationary covariance matrix of the source.

Proof. Please see Appendix F.

From Lemmas 11 and 12, and the previous discussion, it is now possible to obtain the bounds for the outage probability as follows. For the lower bound, we have that

$$
\begin{aligned}
P^{l}_{\text{out}}(d_{\text{th}}) &= \Pr(d^{l}(n) \geqslant d_{\text{th}}) \\
&= \Pr\left(\frac{1}{\gamma_n + \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}} \geqslant d_{\text{th}}\right) \\
&= \Pr\left(\gamma_n \leqslant \left(\frac{1}{d_{\text{th}}} - \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}\right)\right) \\
&= \Pr\left(\frac{P_t\|\mathbf{H}(n)\|^2}{\sigma_v^2 KN} \leqslant \left(\frac{1}{d_{\text{th}}} - \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}\right)\right) \\
&= \Pr\left(\|\mathbf{H}(n)\|^2 \leqslant \frac{\sigma_v^2 KN}{P_t}\left(\frac{1}{d_{\text{th}}} - \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}\right)\right) \\
&= F_{\|\mathbf{H}(n)\|^2}\left(\frac{\sigma_v^2 KN}{P_t}\left(\frac{1}{d_{\text{th}}} - \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}\right)\right), \qquad (4.16)
\end{aligned}
$$

where $F_{\|\mathbf{H}(n)\|^2}(.)$ is the cdf of the random variable $\|\mathbf{H}(n)\|^2$ and can be obtained from

$$
F_{\|\mathbf{H}(n)\|^2}(z) = \frac{1}{(NK-1)!}\int_0^z \mathrm{e}^{-t}t^{NK-1}\,\mathrm{d}t. \qquad (4.17)
$$

Note that the only difference between (4.17) and the standard cdf of a χ^2 random variable is a scaling factor of 2 in the integral upper limit z, which is added due to the normalized variance assumption on the individual complex channel paths.

For the upper bound, we similarly have that

$$
\begin{aligned}
P^{u}_{\text{out}}(d_{\text{th}}) &= \Pr(d^{u}(n) \geqslant d_{\text{th}}) \\
&= \Pr\left(\frac{1}{\gamma_n + \frac{1}{1/K\cdot\sum_{l=1}^{K}\theta_l}} \geqslant d_{\text{th}}\right) \\
&= \Pr\left(\gamma_n \leqslant \frac{1}{d_{\text{th}}} - \frac{1}{1/K\cdot\sum_{l=1}^{K}\theta_l}\right). \qquad (4.18)
\end{aligned}
$$

Since θ_l are functions of γ_{n-1}, one can fix the value of γ_{n-1} in order to perform the same procedure as for $P^{l}_{\text{out}}(d_{\text{th}})$ and then integrate over the pdf of γ_{n-1}, in order to obtain the

total outage probability. Note that in order for this procedure to be correct, it is required that γ_n and γ_{n-1} are statistically independent. Assuming that results in

$$\begin{aligned} P_{\text{out}}^{u}(d_{\text{th}}) &= \int_0^\infty \Pr\left(\gamma_n \leqslant \left(\frac{1}{d_{\text{th}}} - \frac{1}{\frac{1}{K}\sum_{l=1}^{K}\theta_l}\right) \middle| \gamma_{n-1} = z\right) f_{\gamma_{n-1}}(z)\,\mathrm{d}z \\ &= \int_0^\infty \Pr\left(\frac{P_t\|\mathbf{H}(n)\|^2}{\sigma_v^2 KN} \leqslant \left(\frac{1}{d_{\text{th}}} - \frac{1}{\frac{1}{K}\sum_{l=1}^{K}\theta_l}\right) \middle| \gamma_{n-1} = z\right) f_{\gamma_{n-1}}(z)\,\mathrm{d}z \\ &= \int_0^\infty F_{\|\mathbf{H}(n)\|^2}\left(\frac{\sigma_v^2 KN}{P_t}\left(\frac{1}{d_{\text{th}}} - \frac{1}{\frac{1}{K}\sum_{l=1}^{K}\theta_l}\right) \middle| \gamma_{n-1} = z\right) f_{\gamma_{n-1}}(z)\,\mathrm{d}z. \end{aligned} \tag{4.19}$$

Note that it is only θ_l which is a function of γ_{n-1}.

For the numerical evaluation of the bounds, please see Sec. 4.4.

4.3.2 Analysis of Diversity Order and Coding Gain

In order to obtain the diversity order with the help of the bounds as previously mentioned, we present the following lemma.

Lemma 13. *Assume that the outage probability $P_{out}(d_{th})$ can be lower and upper bounded by $P_{out}^{l}(d_{th})$ and $P_{out}^{u}(d_{th})$, respectively, i.e. $P_{out}^{l}(d_{th}) < P_{out}(d_{th}) < P_{out}^{u}(d_{th})$ for all system parameters and all n. If $P_{out}^{l}(d_{th})$ and $P_{out}^{u}(d_{th})$ have a diversity order of d_{ord}^{0}, then the outage probability $P_{out}(d_{th})$ has a diversity order of d_{ord}^{0} as well.*

Proof. Since $P_{\text{out}}^{l}(d_{\text{th}}) < P_{\text{out}}(d_{\text{th}}) < P_{\text{out}}^{u}(d_{\text{th}})$ and $\log(.)$ is a monotonic increasing function for all valid (positive) arguments and $\log(\text{SNR})$ is a positive number, then we have that

$$\frac{\log(P_{\text{out}}^{l}(d_{\text{th}}))}{\log(\text{SNR})} < \frac{\log(P_{\text{out}}(d_{\text{th}}))}{\log(\text{SNR})} < \frac{\log(P_{\text{out}}^{u}(d_{\text{th}}))}{\log(\text{SNR})}. \tag{4.20}$$

As we have that $\lim_{\text{SNR}\to\infty} \dfrac{\log(P_{\text{out}}^{l}(d_{\text{th}}))}{\log(\text{SNR})} = -d_{\text{ord}}^{0}$ and that $\lim_{\text{SNR}\to\infty} \dfrac{\log(P_{\text{out}}^{u}(d_{\text{th}}))}{\log(\text{SNR})} = -d_{\text{ord}}^{0}$, then according to the well-known squeeze theorem (Kathy's theorem) for limits, we obtain that

$$\lim_{\text{SNR}\to\infty} \frac{\log(P_{\text{out}}(d_{\text{th}}))}{\log(\text{SNR})} = -d_{\text{ord}}^{0}, \tag{4.21}$$

and the proof is complete. □

At this stage, it suffices to find the diversity order for $P^l_{\text{out}}(d_{\text{th}})$ and $P^u_{\text{out}}(d_{\text{th}})$. By a Taylor series expansion of $F_{\|\mathbf{H}(n)\|^2}(z)$ from (4.17), it is easy to show that the cumulative distribution function (cdf) of this distribution near zero (small z) is of the form

$$F_{\|\mathbf{H}(n)\|^2}(z) = \frac{1}{(NK)!} z^{NK} + o(z^{NK}). \tag{4.22}$$

Then, we obtain that

$$\begin{aligned}
\mathrm{P}^l_{\text{out}}\left(\mathrm{d}_{\text{th}}\right) &= F_{\|\mathbf{H}(n)\|^2}\left(\frac{\sigma_v^2 KN}{P_t}\left(\frac{1}{d_{\text{th}}} - \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}\right)\right) \\
&= \frac{1}{(NK)!}\left(\frac{\sigma_v^2 KN}{P_t}\left(\frac{1}{d_{\text{th}}} - \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}\right)\right)^{NK} + o(P_t^{-NK}) \\
&= \frac{(\sigma_v^2 KN)^{KN}}{(NK)!}\left(\frac{1}{d_{\text{th}}} - \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}\right)^{NK} P_t^{-NK} + o(P_t^{-NK}) \\
&= \frac{\left(\sigma_v^2 KNP_x\right)^{NK}}{(NK)!}\left(\frac{1}{d_{\text{th}}} - \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}\right)^{NK} \mathrm{SNR}^{-NK} + o(\mathrm{SNR}^{-NK}),
\end{aligned} \tag{4.23}$$

with $\zeta_l(.)$ defined in Lemma 11.

Similarly for $P^u_{\text{out}}(d_{\text{th}})$ and taking $\tilde{P} = \dfrac{P_t}{\sigma_v^2 KN}$ and $\bar{\zeta} = \dfrac{1}{K}\sum_{l=1}^{K}\zeta_l(C_u)$, we have that

$$\begin{aligned}
\mathrm{P}^u_{\text{out}}(\mathrm{d}_{\text{th}}) &= \int_0^\infty F_{\|\mathbf{H}(n)\|^2}\left(\frac{1}{\tilde{P}}\left(\frac{1}{d_{\text{th}}} - \frac{1}{\frac{1}{K}\sum_{l=1}^{K}\theta_l}\right)\middle|\gamma_{n-1} = z\right) f_{\gamma_{n-1}}(z)\,\mathrm{d}z \\
&= \int_0^\infty F_{\|\mathbf{H}(n)\|^2}\left(\frac{1}{\tilde{P}}\left(\frac{1}{d_{\text{th}}} - \frac{1}{\bar{\zeta} + \dfrac{|\zeta_{\max}(A)|^2\alpha_{\max}}{1 + z\alpha_{\max}}}\right)\right) f_{\gamma_{n-1}}(z)\,\mathrm{d}z,
\end{aligned}$$

which by substituting the Taylor series expansion results in

$$\begin{aligned}
\mathrm{P}_{\text{out}}^{u}(\mathrm{d}_{\text{th}}) &= \frac{1}{(NK)!}\tilde{P}^{-NK}\int_{0}^{\infty}\left(\frac{1}{d_{\text{th}}}-\frac{1}{\bar{\zeta}+\dfrac{|\zeta_{\max}(A)|^2\alpha_{\max}}{1+z\alpha_{\max}}}\right)^{NK} f_{\gamma_{n-1}}(z)\,\mathrm{d}z \\
&\quad+\int_{0}^{\infty} o(\tilde{P}^{-NK}) f_{\gamma_{n-1}}(z)\,\mathrm{d}z. \\
&= \frac{(\sigma_v^2 NKP_x)^{NK}}{(NK)!}\mathrm{SNR}^{-NK}\int_{0}^{\infty}\left(\frac{1}{d_{\text{th}}}-\frac{1}{\bar{\zeta}+\dfrac{|\zeta_{\max}(A)|^2\alpha_{\max}}{1+z\alpha_{\max}}}\right)^{NK} f_{\gamma_{n-1}}(z)\,\mathrm{d}z \\
&\quad+\int_{0}^{\infty} o(\mathrm{SNR}^{-NK}) f_{\gamma_{n-1}}(z)\,\mathrm{d}z. \qquad (4.24)
\end{aligned}$$

It is now possible to calculate the diversity order for the bounds in order to show their equality and thus prove the diversity result for the outage probability function. For the lower bound, we have that

$$\begin{aligned}
\mathrm{d}_{\text{ord}}^{l} &= -\lim_{\mathrm{SNR}\to\infty}\frac{\log(P_{\text{out}}^{l}(d_{\text{th}}))}{\log(\mathrm{SNR})} \\
&= KN. \qquad (4.25)
\end{aligned}$$

This is due to the fact that when $\mathrm{SNR}\to\infty$, $o(\mathrm{SNR}^{-NK}) \ll \mathrm{SNR}^{-NK}$ and thus the $o(\mathrm{SNR}^{-NK})$ term in (4.23) vanishes before the first term containing SNR^{-NK}, resulting in a diversity order of NK.

The analysis for the upper bound is also similar. When $\mathrm{SNR}\to\infty$, the term $o(\mathrm{SNR}^{-NK})$ can be upper-bounded by $\kappa\mathrm{SNR}^{-NK}$, with κ being an arbitrary constant, but going to zero as SNR goes to infinity. Therefore, it is possible to deduce that

$$\begin{aligned}
\int_{0}^{\infty} o(\mathrm{SNR}^{-NK}) f_{\gamma_{n-1}}(z)\,\mathrm{d}z &< \int_{0}^{\infty}\kappa\mathrm{SNR}^{-NK} f_{\gamma_{n-1}}(z)\,\mathrm{d}z \\
&= \kappa\mathrm{SNR}^{-NK}\int_{0}^{\infty} f_{\gamma_{n-1}}(z)\,\mathrm{d}z \\
&= \kappa\mathrm{SNR}^{-NK}.
\end{aligned}$$

As a result, we can say that

$$\begin{aligned}
\mathrm{d}_{\text{ord}}^{u} &= -\lim_{\mathrm{SNR}\to\infty}\frac{\log(P_{\text{out}}^{u}(d_{\text{th}}))}{\log(\mathrm{SNR})} \\
&= KN. \qquad (4.26)
\end{aligned}$$

As $\mathrm{d}_{\text{ord}}^{u} = \mathrm{d}_{\text{ord}}^{l} = NK$, it can be deduced from Lemma 13 that the diversity order for the outage probability is also equal to NK and the analysis is complete. It is worthwhile mentioning that the maximum diversity order is dependent on the MIMO channel and its achievability only on the space-time code. The source structure does not play any role on the diversity order. However, as it will be discussed next, the source structure plays an important role on the coding gain.

Coding gain for the outage probability is considered next. We know that both $P_{\text{out}}^{l}(d_{\text{th}})$ and $P_{\text{out}}^{u}(d_{\text{th}})$ have the same diversity order, but possibly different coding gains. We may then write

$$\begin{aligned} P_{\text{out}}^{l}(d_{\text{th}}) < P_{\text{out}}(d_{\text{th}}) < P_{\text{out}}^{u}(d_{\text{th}}) \Rightarrow \\ (G_1 \cdot \text{SNR})^{-KN} + o(\text{SNR}^{-KN}) < (G \cdot \text{SNR})^{-KN} + o(\text{SNR}^{-KN}) \\ < (G_2 \cdot \text{SNR})^{-KN} + o(\text{SNR}^{-KN}) \end{aligned} \tag{4.27}$$

This results in

$$\begin{aligned} G_1^{-NK} + o(1) < G^{-NK} + o(1) \\ < G_2^{-NK} + o(1). \end{aligned} \tag{4.28}$$

In the high SNR regime, the term $o(1)$ vanishes quickly compared to the constants G_1, G, G_2. As a result, the relationship in (4.28) simplifies to

$$G_2 < G < G_1, \quad \text{SNR} \to \infty, \tag{4.29}$$

which provides upper and lower bounds for the coding gain by setting $G^l = G_2$ and $G^u = G_1$. Note that a higher coding gain means a better SNR performance, i.e. a lower outage probability for a given SNR. That is why the upper bound for coding gain is obtained from the lower bound on the outage probability and vice versa. The constants G^u and G^l may in turn be extracted from (4.23) and (4.24) as

$$G^u = \frac{(NK)!^{1/(NK)}}{(\sigma_v^2 KNP_x)} \left(\frac{1}{d_{\text{th}}} - \frac{1}{K} \sum_{l=1}^{K} \frac{1}{\zeta_l(C_u)} \right)^{-1} \tag{4.30}$$

and

$$G_1^l = \frac{(NK)!^{1/(NK)}}{(\sigma_v^2 NKP_x)} \left(\int_0^{\infty} \left(\frac{1}{d_{\text{th}}} - \frac{1}{\bar{\zeta} + \dfrac{|\zeta_{\max}(A)|^2 \alpha_{\max}}{1 + z\alpha_{\max}}} \right)^{NK} f_{\gamma_{n-1}}(z)\, \mathrm{d}z \right)^{-1/(NK)}. \tag{4.31}$$

Remark 4. The value of G^l introduced in (4.31) (subscripted by 1) is computationally more demanding to calculate than the value for G^u, especially because one should also

consider the limit behavior of $f_{\gamma_{n-1}(z)}$ when $P \to \infty$. It is relatively easy to show that G_1^l itself may be lower bounded by the following value

$$G_2^l = \frac{(NK)!^{1/(NK)}}{(\sigma_v^2 NKP_x)} \left(\frac{1}{d_{\text{th}}} - \frac{1}{\bar{\zeta} + |\zeta_{\max}(A)|^2 \alpha_{\max}} \right)^{-1}, \tag{4.32}$$

which is less accurate, but is of a much simpler form than G_1^l.

As we can see from (4.30), (4.31), and (4.32), the coding gain depends on the source structure and the threshold. It is the eigenvalues of A and C_u which play a significant role. We observe e.g. that smaller d_{th} leads to smaller coding gain. This is due to the fact that lower thresholds lead to higher outage probabilities and for fixed diversity order, this leads to smaller coding gains. Also, if the eigenvalues of C_u are large, the coding gain decreases, i.e. the outage probabilities increase in the asymptotic limit. Heuristically, such values imply more randomness in the process, resulting in higher distortion for the Kalman filter and consequently higher outage value.

Remark 5. There are limits for d_{th} for which (4.30), (4.31) and (4.32) are valid. In (4.30), it is required that

$$\frac{1}{d_{\text{th}}} - \frac{1}{K} \sum_{l=1}^{K} \frac{1}{\zeta_l(C_u)} \geqslant 0 \tag{4.33}$$

so that the lower bound is meaningful. This leads to

$$d_{\text{th}} \leqslant \frac{K}{\sum_{l=1}^{K} \frac{1}{\zeta_l(C_u)}}, \tag{4.34}$$

which equals the harmonic mean of the eigenvalues of C_u. Similarly, it is also possible to show that a sufficient condition for (4.31) and (4.32) to be valid is that

$$d_{\text{th}} \leqslant \frac{1}{K} \sum_{l=1}^{K} \zeta_l(C_u). \tag{4.35}$$

Therefore, a sufficient condition on d_{th} in order to get valid bounds can be obtain from

$$d_{\text{th}} \leqslant \min \left\{ \frac{K}{\sum_{l=1}^{K} \frac{1}{\zeta_l(C_u)}}, \frac{1}{K} \sum_{l=1}^{K} \zeta_l(C_u) \right\}, \tag{4.36}$$

which requires that d_{th} is smaller than the minimum of mean and harmonic mean of the eigenvalues of the process noise covariance matrix. The limiting regime in both cases is when the eigenvalues of C_u are small. This happens when the randomness in the process from $\mathbf{u}(n)$ is too slow compared to the process memory from A. In realistic applications,

this can be solved by adjusting the sampling rate of the original continuous-time process, if necessary. One might argue that lowering the sampling rate in very slow varying processes in order to get better bounds would eventually increase the outage probability. However, if the application is critically sensitive in that regard, the interesting regime is already small d_{th}, because it is the regime which results in higher outage probabilities. For small d_{th} however, the bounds would be functioning. This shows that the limiting behavior in (4.36) is not a serious issue for the bounds in most practical cases.

Remark 6. For well-conditioned C_u, the bounds perform better than the case when C_u is ill-conditioned. In fact, for the case when $C_u = \sigma_u^2 I$, the bounds are tight. This is elaborated more in Appendix G.

Remark 7. Although the upper and lower bounds for the outage probability and coding gain have been presented for joint Kalman filtering and space-time coding, they can be used to obtain bounds for the scalar case as well. It is then enough to select the correct parameters for the scalar case in order to use the presented bounds. While no general conclusion can be made about the relative accuracy of these two cases, a numerical comparison is performed in Sec. 4.4.

For the numerical evaluation of the accuracy of the coding gain expressions and related discussions, please see Sec. 4.4.

4.4 Numerical Evaluation of the Bounds and Diversity Results

In this section, simulation results are provided to accompany the presented theory in the previous sections.

We begin the numerical evaluations with the following system parameters. We take $K = 2$ and $N = 1$ to keep the simulated outage values practically calculable. This necessitates a maximum diversity order of $KN = 2$. We select

$$A = \begin{bmatrix} 0.6 & -0.8 \\ 0.7 & 0.6 \end{bmatrix}$$

has the eigenvalues $\{0.6 \pm j\sqrt{0.56}\}$. This corresponds to the case where $x_1(n)$ and $x_2(n)$ are relatively highly cross-correlated in time. We select $\sigma_v^2 = 1$ and $d_{\text{th}} = 0.1$. As for the orthogonal space-time code we use the Alamouti code from [2], while for simplicity we calculate P_x from simulations. We consider two cases for C_u, namely C_u^1 and C_u^2 as follows

$$C_u^1 = \begin{bmatrix} 0.25 & 0 \\ 0 & 1.44 \end{bmatrix}, \quad C_u^2 = \begin{bmatrix} 0.53 & 0.28 \\ 0.28 & 0.53 \end{bmatrix}.$$

The choice is mainly to show how the accuracy of the bounds will change as different values are used for C_u, and also that different values for C_u will result in different coding

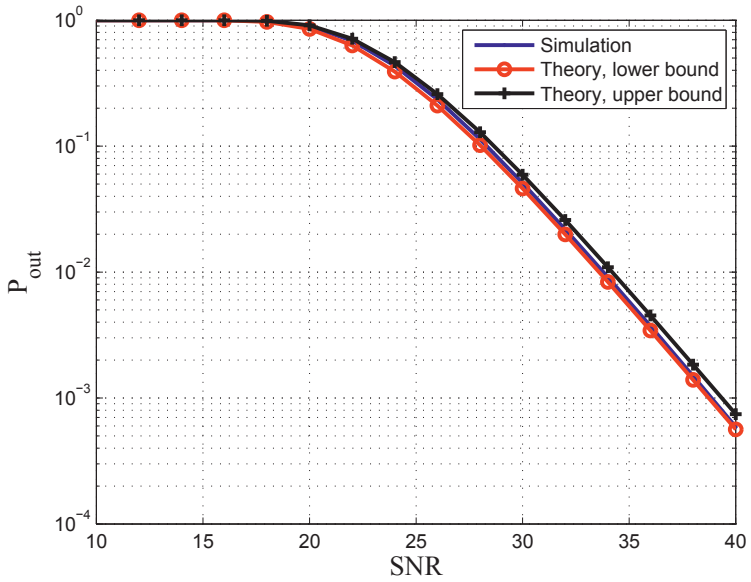

Figure 4.1: Outage probability and the corresponding bounds for $K = 2$, $N = 1$, $d_{\text{th}} = 0.1$, $\sigma_v^2 = 1$, and for $\zeta_{1,2}(C_{u,1}) = 0.25, 1.44$.

gains. Also, the eigenvalues of C_u^1 are equal $\{1.44, 0.25\}$ and the eigenvalues of C_u^2 are equal to $\{0.81, 0.25\}$. The P_{out} vs. SNR graph is depicted in Figures 4.1 and 4.2. We simulate the Kalman filter for $n \leqslant 10^7$ and numerically calculate the outage probabilities after discarding the first 400 samples. The upper and lower bounds are visibly good for both cases and seem to be quite accurate for a large range of SNR values, compared to the simulated result from the Kalman filter. The numerical evaluation for the coding gain bounds is also depicted in Figures 4.3 and 4.4. The simulated outage probabilities are plotted along with a linear function with a slope of 2 and with calculated values for G^l, G_1^u, G_2^u. The upper and lower bounds for the coding gain become visibly accurate from SNR's close to 30 dB. This shows that for the high SNR analysis to be correct, one needs at least an SNR of the same value or higher. A slope of 2 corresponding to the diversity order $d_{\text{ord}} = 2$ is quite visible in both cases as well. One can also notice in Figures 4.3 and 4.4 that the lower bound for coding gain is more accurate than the upper bound and that G_1^l is a much better lower bound that G_2^l, as expected.

In order to observe the performance of the system for higher dimensions, we now take $K = 3$, $N = 1$ and use the space-time code construction from [98, Eq. 39]. This leads to $N_c = 4$ and $r = 3/4$ as well. We also modify other system parameters to $C_u = \text{diag}\{0.5, 0.75, 0.65\}$ and $A = D\,\text{diag}\{0.95, 0.9, 0.98\}\,D^{-1}$, with D representing the normalized discrete cosine transform matrix, and then simulate the system for $n \leqslant 10^8$. The results for outage probability bounds and the coding gain bounds and the

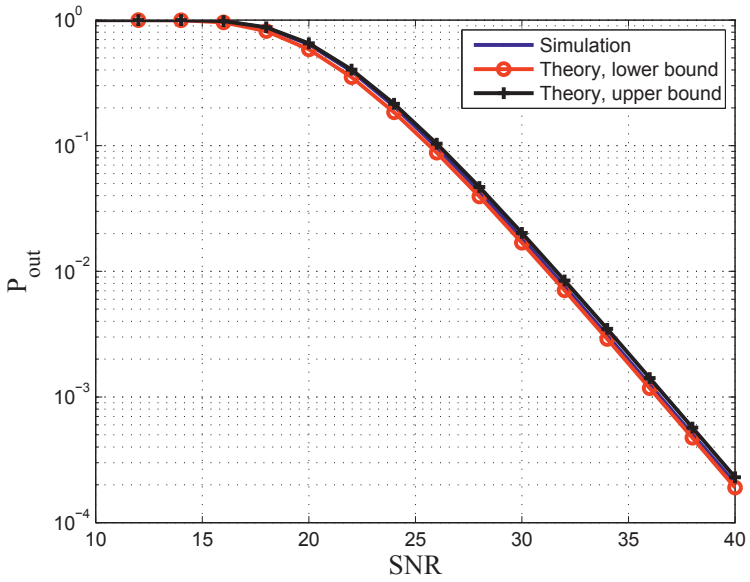

Figure 4.2: Outage probability and the corresponding bounds for $K = 2$, $N = 1$, $d_{\text{th}} = 0.1$, $\sigma_v^2 = 1$, and for $\zeta_{1,2}(C_{u,2}) = 0.25, 0.81$.

diversity order are presented in Figures 4.5 and 4.6. A diversity order of 3 is visible in both figure and the bounds are visibly very accurate.

In order to compare the relative performance of the presented bounds in this chapter with the bounds in Chapter 2, we select the system parameters from Chapter 2 as $\rho = 0.95$, $\sigma_u^2 = \sigma_v^2 = 1$, and $d_{\text{th}} = 0.1, 0.8$. The results are depicted in Figures 4.7 and 4.8. The figures suggest that the performance for the bounds are very close for small d_{th}, while the bounds from Chapter 2, perform better (significantly better for lower SNR values) for the higher threshold values. It is then recommended to continue to use the bounds from Chapter 2 for the scalar case.

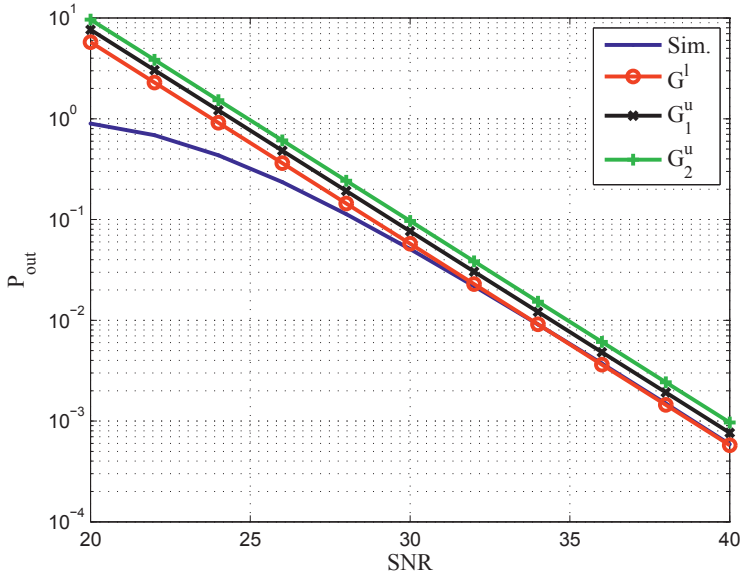

Figure 4.3: Comparison of accuracy of the coding gain bounds for $K = 2$, $N = 1$, $d_{\text{th}} = 0.1$, $\sigma_v^2 = 1$, and for $\zeta_{1,2}(C_{u,1}) = 0.25, 1.44$.

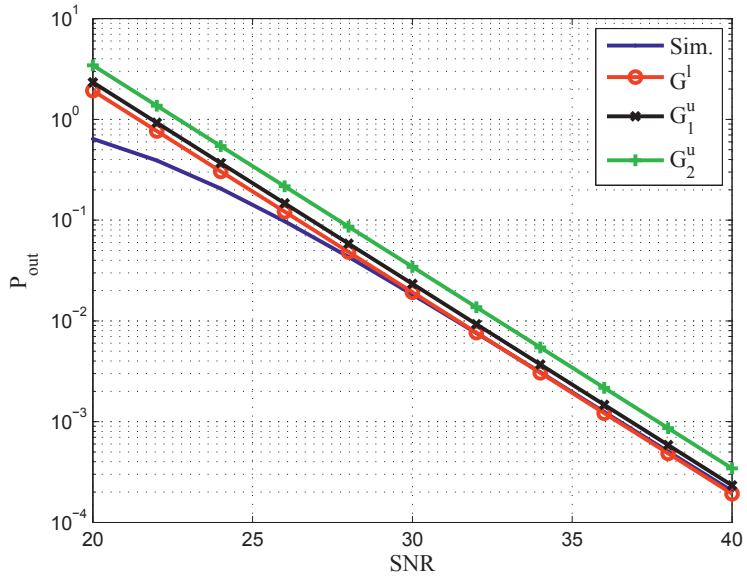

Figure 4.4: Comparison of accuracy of the coding gain bounds for $K = 2$, $N = 1$, $d_{\text{th}} = 0.1$, $\sigma_v^2 = 1$, and for $\zeta_{1,2}(C_{u,2}) = 0.25, 0.81$.

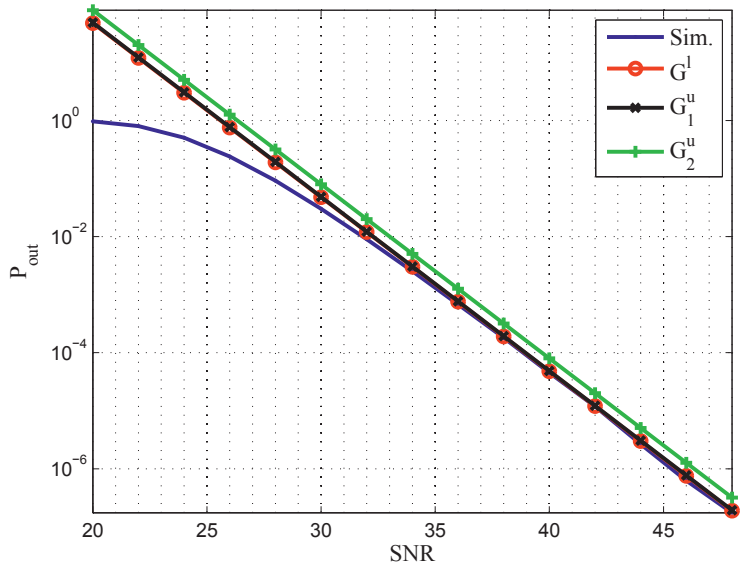

Figure 4.5: Comparison of accuracy of the coding gain bounds for $K = 3$, $N = 1$, $d_{\text{th}} = 0.1$, $\sigma_v^2 = 1$.

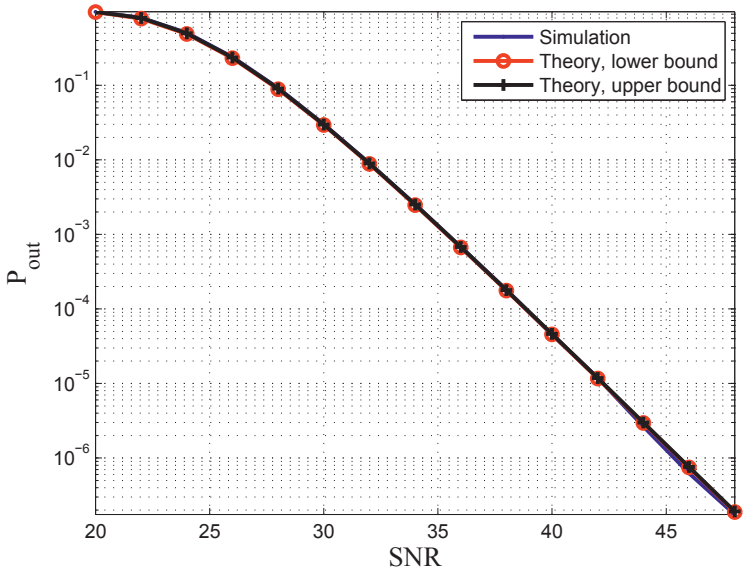

Figure 4.6: Outage probability and the corresponding bounds for $K = 3$, $N = 1$, $d_{\text{th}} = 0.1$, $\sigma_v^2 = 1$.

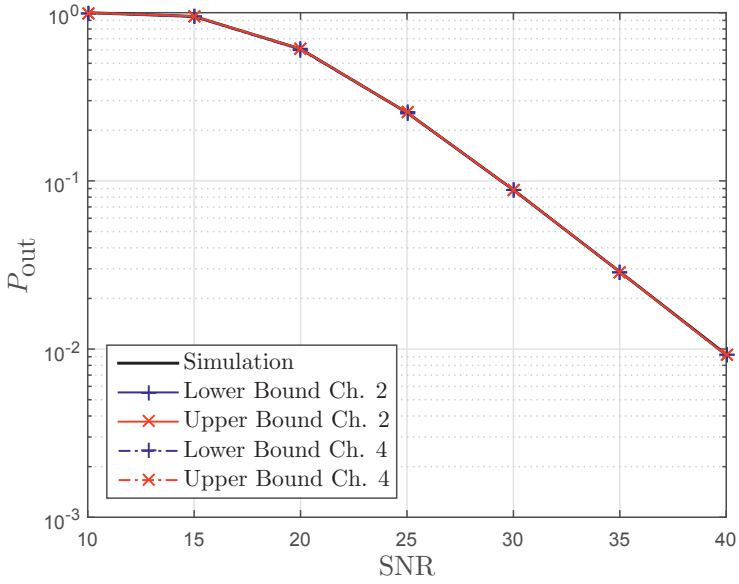

Figure 4.7: Comparison of accuracy of the outage probability bounds for scalar sources for $d_{\text{th}} = 0.1$.

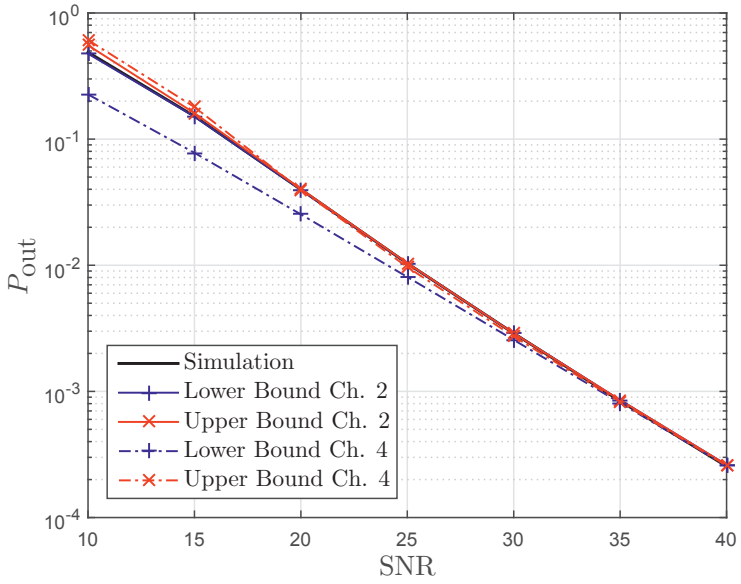

Figure 4.8: Comparison of accuracy of the outage probability bounds for scalar sources for $d_{\text{th}} = 0.8$.

4.5 Summary and Discussion

In this chapter, a new method for analog transmission of Gauss-Markov sources over MIMO fading channels was proposed. The proposed method incorporates complex orthogonal space-time codes. By decoding the real and imaginary parts of the code separately, we allow any complex orthogonal space-time code with arbitrary rate to be used for analog transmission. We then showed that for Rayleigh fading channels, the distortion outage probability can achieve the maximum diversity order allowed by the MIMO channel. By considering process memory only limited to two previous steps, we are able to provide bounds for the distortion outage probability which are applicable for any SNR, and also present bounds for the coding gain in the high SNR regime. We also outline how the coding gain depends on the eigenvalues of the state transition and the process noise covariance matrices, and the outage threshold.

There is no independence requirement for the consecutive channel instances for obtaining the lower bound for outage. Therefore, the lower bound is always valid, even if the i.i.d. assumption (in time) does not hold. In addition, we have shown that the diversity order for the lower bound is also the same for outage probability. This means that KN is an upper bound for the diversity order of any Rayleigh flat fading channel, regardless of the correlation between different channel instances. The i.i.d. assumption must however hold for the derivation of the upper bound. Due to that, it is not possible to draw any conclusion about the actual value of the diversity order for non i.i.d. channels based on the derivations in this chapter.

The coding gain results of this chapter can actually be compared with those of Chapter 2. Reducing the system dimensions to one, results in a lower bound for the coding gain, which is equal to the coding gain obtained in that chapter. Comparing the bounds of this chapter with the ones in the previous chapters, one could claim that, while the bounds in Chapter 2 are asymptotically tight and the bounds in Chapter 3 show improvement with increasing SNR, the bounds in this chapter show a different behavior. The tightness of the bounds in this chapter does not only depend on the SNR, but also on the structure of the process noise covariance matrix. The bounds in this chapter are tight if the process noise covariance matrix is a scaled version of the identity matrix. This means that the bounds in this chapter are tight in the high SNR regime for the scalar case as well. We could not however establish any general theoretical conclusion about their finite SNR performance.

Chapter 5

Large Dimensional Signals

In this chapter, Kalman filtering of a large dimensional Gauss-Markov vector process over a fading channel is considered. The large dimension assumption can both model a source with many co-located elements (centralized) or a collection of physically separated, but still correlated sources with smaller dimensions (distributed). This problem may appear in sensor networks, where a large number of signals need to be transmitted and then estimated in real time. For the centralized model, it is easily possible to use the diversity scheme introduced in the previous chapter. For the distributed source, using the space-time diversity schemes requires separate communication channels among the sensors. This might be hard to achieve in practice and the overhead is also significant. The simplest option is to transmit the signals over the channel without any extra pre-processing. With the signals following an ARGM model, the optimal estimator at the receiver is the Kalman filter with random estimation error covariance matrix. Then, a method is needed to evaluate the quality of estimation for such a large dimensional system.

In order to measure the quality of estimation, we evaluate the instantaneous random distortion, i.e. the normalized trace of the instantaneous estimation error covariance matrix. The instantaneous random covariance matrix from the Kalman filter follows a recursive matrix equation, which is difficult to track. In order to simplify the analysis, we make two extra assumptions. First, we assume that source dimensions is *infinitely* large. This assumption allows us to incorporate tools from large system analysis and random matrix theory. In addition, we make the assumption that the SNR is asymptotically high. Using these assumptions, it is possible to provide an approximation of the performance of large, but finite dimensional systems. By using tools such as the Stieltjes transform, we are able to find an approximation to the pdf of the eigenvalue distribution of the random estimation error covariance matrix of the Kalman filter in the high SNR regime. Although the approximated pdf seems to be accurate for a large range of possible distortion values, it cannot approximate the tail of the pdf well enough for outage calculations similar to what we did in the previous chapters. However, the approximated pdf can still be used to obtain the average MSE of the Kalman filter. The average MSE then can provide us with an overall performance measure as a function of system parameters.

5.1 System Model and Theoretical Background

In this section, the updated system model used in the rest of this section is introduced first. Then, some background is provided from random matrix theory and Stieltjes transform. These two theories are later used in the analysis in Sec. 5.2.

5.1.1 Updated System Model and Problem Definition

Consider the following updated system model

$$\begin{aligned}\mathbf{x}(n) &= A\mathbf{x}(n-1) + \mathbf{u}(n)\\ \mathbf{y}(n) &= \sqrt{\gamma}H(n)\mathbf{x}(n) + \mathbf{v}(n).\end{aligned} \tag{5.1}$$

In this section, it is assumed that $\mathbf{x}(n)$ and $\mathbf{u}(n)$ are column vectors of dimension K, $H(n)_{N\times K}$ is a matrix consisting of i.i.d Gaussian elements with variance equal to $\frac{1}{K}$ (necessary for large system analysis), and $\mathbf{y}(n)$ and $\mathbf{v}(n)$ are column vectors of dimension N, and γ is used for transmit power adjustment or the equivalently the pre-filtering operation denoted by $T(\cdot)$, in line with the system model in Chapter 1. The covariance matrices for $\mathbf{u}(n)$ and $\mathbf{v}(n)$ are respectively C_u, C_v. In this section, it is assumed that all signals are real (for a discussion on complex signals please see Section 5.4).The assumption that $H(n)$'s are independent from one another is kept as well. In addition, it is also assumed that A is non-singular and that $H(0) \neq 0$. Note that the probability of the event $H(0) = 0$ is equal to 0 for the Rayleigh fading channel model.

Assuming perfect channel knowledge at the receiver, the optimal MMSE filter at the receiver is the Kalman filter. Similar to (4.7), the vector Kalman equations adapted from [52] are

$$\begin{aligned}&\hat{\mathbf{x}}(n|n-1) = A\hat{\mathbf{x}}(n-1|n-1)\\ &P(n) = AM(n-1)A^T + C_u\\ &G_k(n) = \sqrt{\gamma}P(n)H^T(n)\left(C_v + \gamma H(n)P(n)H^T(n)\right)^{-1}\\ &\hat{\mathbf{x}}(n|n) = \hat{\mathbf{x}}(n|n-1) + G_k(n)\big(\mathbf{y}(n) - \sqrt{\gamma}H(n)\hat{\mathbf{x}}(n|n-1)\big)\\ &M(n) = \big(I - \sqrt{\gamma}G_k(n)H(n)\big)P(n),\end{aligned} \tag{5.2}$$

the prediction error covariance matrix $P(n)$ then evolves through the following RRE

$$\begin{aligned}P(n+1) = AP(n)A^T - \gamma AP(n)H^T(n)\left(C_v + \gamma H(n)P(n)H^T(n)\right)^{-1}H(n)P(n)A^T\\ + C_u.\end{aligned} \tag{5.3}$$

For large system analysis, it is assumed that K, N grow large while $\beta = N/K \geqslant 1$ is fixed. In order to simplify the problem furthermore, it is assumed that $C_v = \sigma_v^2 I$ and

$C_u = I$ (the extension to $C_u = \sigma_u^2 I$ is straightforward). Then, an approximation for the eigenvalue distribution of the estimation error covariance matrix is developed. We note that in a related work, the case with $A = I$ and $\gamma = 1$ was studied in [105], where they obtain a set of coupled Stieltjes transform pairs for the prediction error covariance matrix (but not the estimation error covariance matrix).

With the aforementioned assumptions, the random Riccati equation in (5.3) can be shown to be equal to

$$P(n+1) = A(P^{-1}(n) + \gamma H^T(n)H(n))^{-1}A^T + I \tag{5.4}$$

by invoking the matrix inversion lemma on (5.4).

By using arguments similar to those in Theorem 1, it can be shown that the eigenvalue distributions of $P(n)$ and $M(n)$ converge to a steady state distribution as n goes to infinity. For the convergence of the eigenvalues, it is required (according to [10, Theorem 2.4]) that the system is conditionally Gaussian, weakly controllable and weakly observable and that $\log\log^+ \|A\|$, $\log\log^+ \|A^{-1}\|$, $\log\log^+ \|H(0)\|$, $\log\log^+ \|C_v\|$ are integrable. The system model in this thesis is conditionally Gaussian. Weak controllability and weak observability are easy to show for the Rayleigh fading channel (see [10, Definition 2.1]), and the aforementioned variables are integrable if A is non-singular and $H(0) \neq 0$, as assumed in the system model.

This convergence is regardless of the system dimensions. However, if the number of dimensions is high, assuming that the matrix dimensions go to infinity provides us with tools from random matrix theory such as the Stieltjes transform that significantly can simplify the analysis.

5.1.2 Stieltjes Transform

The interested reader is referred to [103] for further information on the following definitions and lemmas regarding the Stieltjes transform for large system analysis involving random matrices.

Definition 1. *The Stieltjes transform $S(z)$ of a pdf $f(\zeta)$ is defined as follows*

$$S(z) = \int_{\zeta \in \mathcal{Z}} \frac{f(\zeta)\,\mathrm{d}\zeta}{\zeta - z} \tag{5.5}$$

where $\mathcal{Z}$ is the support of ζ. The pdf can be obtained from the inverse Stieltjes transform as

$$f(\zeta) = \lim_{\omega \to 0^+} \frac{1}{\pi} \mathrm{Im}[S(\zeta + j\omega)] \tag{5.6}$$

Lemma 14. *For the symmetric matrix $Q_{K\times K}$ and random matrix $H_{N\times K}$ as described in Sec. 5.1, and diagonal matrix $D_{N\times N}$ with empirical eigenvalue distribution $g_D(w)$, the*

Stieltjes transform of the eigenvalue distribution of the following random matrix

$$R = Q + H^T D H, \tag{5.7}$$

i.e. $S_R(z)$*, is obtained from*

$$S_R(z) = S_Q\left(z - \beta \int \frac{w g_D(w) dw}{1 + w S_R(z)}\right), \tag{5.8}$$

where $S_Q(z)$ *is the Stieltjes transform of the empirical eigenvalue distribution of* Q*. For the analysis afterwards, the case where* $D = \gamma I$ *is used. Given that,*

$$S_R(z) = S_Q\left(z - \frac{\beta\gamma}{1 + \gamma S_R(z)}\right). \tag{5.9}$$

Lemma 15. *For random matrix* $Q_{K\times K}$*, if* $R = Q^{-1}$*, then*

$$S_R(z) = -\frac{1}{z} - \frac{S_Q(z^{-1})}{z^2}, \tag{5.10}$$

Lemma 16. *For random matrix* $Q_{K\times K}$ *and scalar* a*, if* $R = aQ$*, then*

$$S_R(z) = \frac{1}{a} S_Q(z/a). \tag{5.11}$$

5.2 High SNR Approximation and Analysis

In this section, we apply a high SNR assumption to obtain an approximation of the Stieltjes transform and the eigenvalue distribution of the prediction and estimation error covariance matrices. For our system model, the (total) SNR at the receiver is equal to $\text{SNR} = \dfrac{\gamma P_x}{\sigma_v^2}$. High SNR is also referred to the case where $\gamma \to \infty$. This is done to simplify the analysis procedure, which otherwise is very tedious to develop. We begin with obtaining a simple formula for the estimation error covariance matrix.

We had from the RRE for the prediction error covariance matrix $P(n)$ in (5.4) that

$$P(n+1) = A(P^{-1}(n) + \gamma H^T(n) H(n))^{-1} A^T + I.$$

As $\gamma \to \infty$, it is expected that the term $\gamma H^T(n) H(n)$ dominates the term $P^{-1}(n)$, (given that $\beta \geqslant 1$) and therefore the matrix $P(n+1)$ can be approximated by the matrix I. For the steady-state case when $n \to \infty$, $P(n)$ should also be a matrix close to I, and therefore $P^{-1}(n)$ is close to I as well. Using this assumption, it is possible to get an approximate equation for $P(n+1)$ in the high SNR regime and through that an approximate equation for $M(n)$. With that, one obtains an approximate equation for $S_M(z)$ and as a result for $f_M(\zeta)$ as well. Assuming that $P(n) \to I$, one obtains that

$$P(n+1) = A(I + \gamma H^T(n) H(n))^{-1} A^T + I. \tag{5.12}$$

Also, (5.2) results in

$$P(n+1) = AM(n)A^T + I. \tag{5.13}$$

Combining this with (5.12) leads to the fact that the matrix $M(n)$ should have the same eigenvalue distribution as the matrix $(I + \gamma H^T(n)H(n))^{-1}$, and as a result the same Stieltjes transform. Using (5.12), the following set of Stieltjes transform equations can then be obtained

$$R(n) = I + \gamma H^T(n)H(n) \Leftrightarrow S_R(z) = \frac{1}{1 - \left(z - \frac{\beta\gamma}{1+\gamma S_R(z)}\right)} \tag{5.14}$$

$$M(n) = R^{-1}(n) \Leftrightarrow S_M(z) = -\frac{1}{z} - \frac{1}{z^2} S_R(z^{-1}). \tag{5.15}$$

Solving $S_R(z)$ from (5.14) leads to

$$(1-z)\gamma S_R^2(z) + (\beta\gamma - \gamma + 1 - z)S_R(z) - 1 = 0, \tag{5.16}$$

which results in

$$S_R(z) = \frac{-(\beta\gamma - \gamma + 1 - z) \pm \sqrt{(\beta\gamma - \gamma + 1 - z)^2 + 4\gamma(1-z)}}{2\gamma(1-z)}. \tag{5.17}$$

Therefore, we have that

$$S_M(z) = \\ -z^{-1} - z^{-2}\left(\frac{-(\beta\gamma - \gamma + 1 - z^{-1})}{2\gamma(1-z^{-1})} \pm \frac{\sqrt{(\beta\gamma - \gamma + 1 - z^{-1})^2 + 4\gamma(1-z^{-1})}}{2\gamma(1-z^{-1})}\right), \tag{5.18}$$

where the correct sign of $\pm$ is chosen such that $f_M(\zeta) \geqslant 0$ is satisfied. The next step is to take the inverse Stieltjes transform to obtain $f_M(\zeta)$. This is done as follows. First, $S_M(z)$ is rewritten as

$$S_M(z) = -z^{-1} - z^{-2}\Theta(z), \tag{5.19}$$

where

$$\Theta(z) = \left(\frac{-(\beta\gamma - \gamma + 1 - z^{-1})}{2\gamma(1-z^{-1})} \pm \frac{\sqrt{(\beta\gamma - \gamma + 1 - z^{-1})^2 + 4\gamma(1-z^{-1})}}{2\gamma(1-z^{-1})}\right) \tag{5.20}$$

One can then obtain $f_M(\zeta)$ by inverse Stieltjes transform as

$$\begin{aligned} f_M(\zeta) &= \lim_{\omega\to 0^+} \frac{1}{\pi} \operatorname{Im}\left[S_M(\zeta+j\omega)\right] \\ &= \lim_{\omega\to 0^+} \frac{1}{\pi} \operatorname{Im}\left[-(\zeta+j\omega)^{-1} - (\zeta+j\omega)^{-2}\Theta(\zeta+j\omega)\right] \\ &= \lim_{\omega\to 0^+} \frac{1}{\pi} \operatorname{Im}\left[\frac{-(\zeta-j\omega)}{\zeta^2+\omega^2}\right] \\ &- \lim_{\omega\to 0^+} \frac{1}{\pi} \operatorname{Im}\left[\frac{(\zeta-j\omega)(\zeta-j\omega)}{(\zeta^2+\omega^2)^2}\right] \operatorname{Re}\left[\Theta(\zeta+j\omega)\right] \\ &- \lim_{\omega\to 0^+} \frac{1}{\pi} \operatorname{Re}\left[\frac{(\zeta-j\omega)(\zeta-j\omega)}{(\zeta^2+\omega^2)^2}\right] \operatorname{Im}\left[\Theta(\zeta+j\omega)\right]. \end{aligned} \tag{5.21}$$

It is easy to show that

$$\begin{aligned} \lim_{\omega\to 0^+} \frac{1}{\pi} \operatorname{Im}\left[\frac{(\zeta-j\omega)}{\zeta^2+\omega^2}\right] &= 0, \\ \lim_{\omega\to 0^+} \frac{1}{\pi} \operatorname{Im}\left[\frac{(\zeta-j\omega)(\zeta-j\omega)}{(\zeta^2+\omega^2)^2}\right] &= 0. \end{aligned} \tag{5.22}$$

Therefore, it is possible to say that

$$f_M(\zeta) = \lim_{\omega\to 0^+} \frac{1}{\pi} \operatorname{Re}\left[\frac{(\zeta-j\omega)(\zeta-j\omega)}{(\zeta^2+\omega^2)^2}\right] \operatorname{Im}\left[\Theta(\zeta+j\omega)\right], \tag{5.23}$$

which itself is equal to

$$f_M(\zeta) = \lim_{\omega\to 0^+} \frac{1}{\pi}\frac{1}{\zeta^2} \operatorname{Im}\left[\Theta(\zeta+j\omega)\right]. \tag{5.24}$$

Furthermore, it is also possible to show that

$$\lim_{\omega\to 0^+} \operatorname{Im}\left(\frac{-(\beta\gamma-\gamma+1-z^{-1})}{2\gamma(1-z^{-1})}\right) = 0. \tag{5.25}$$

This leads to

$$f_M(\zeta) = \pm\frac{1}{\pi}\frac{1}{\zeta^2} \lim_{\omega\to 0^+} \operatorname{Im}\left[\frac{\sqrt{(\kappa-(\zeta+j\omega)^{-1})^2+4\gamma(1-(\zeta+j\omega)^{-1})}}{2\gamma(1-(\zeta+j\omega)^{-1})}\right], \tag{5.26}$$

with $\kappa = (\beta-1)\gamma+1$, which by correct selection of the sign results in

$$f_M(\zeta) = \frac{-1}{2\pi\gamma\zeta(\zeta-1)} \operatorname{Im}\sqrt{(\kappa-\frac{1}{\zeta})^2+4\gamma(1-\frac{1}{\zeta})}. \tag{5.27}$$

One can also remove the $\mathrm{Im}()$ operator and rewrite (5.27) as

$$f_M(\zeta) = \frac{c}{2\pi\gamma\zeta^2(1-\zeta)}\sqrt{(\zeta_u-\zeta)(\zeta-\zeta_l)},\ (\zeta_l \leqslant \zeta \leqslant \zeta_u) \tag{5.28}$$

with

$$\begin{aligned} c &= \sqrt{(\beta-1)^2\gamma^2 + 2\gamma(\beta+1)+1} \\ \zeta_l &= \frac{\gamma(1+\beta-2\sqrt{\beta}+1)}{c^2} \\ \zeta_u &= \frac{\gamma(1+\beta+2\sqrt{\beta}+1)}{c^2}. \end{aligned} \tag{5.29}$$

The following results can also be obtained from (5.28).

Theorem 5. *The average MSE, i.e.* $\overline{MSE} = \frac{1}{K}E_H\left(\|\mathbf{x}(n)-\hat{\mathbf{x}}(n)\|^2\big|H\right)$ *is equal to the first moment of the eigenvalue distribution of the random estimation error covariance matrix, i.e.*

$$\overline{MSE} = \int_\Lambda \zeta f_M(\zeta)\,\mathrm{d}\zeta, \tag{5.30}$$

and can be obtained from

$$\begin{aligned} \overline{MSE} &= \int_{\zeta_l}^{\zeta_u} \frac{c}{2\pi\gamma\zeta(1-\zeta)}\sqrt{(\zeta_u-\zeta)(\zeta-\zeta_l)}\mathrm{d}\zeta \\ &= \frac{c}{2\gamma}\left(1-\sqrt{\zeta_l\zeta_u}-\sqrt{1-\zeta_l-\zeta_u+\zeta_l\zeta_u}\right). \end{aligned} \tag{5.31}$$

Note that the averaging operation is over all channel realizations. It is possible to show that in the high SNR regime, the average MSE decreases with the inverse of the square root of γ for $\beta = 1$ and inverse of γ for $\beta > 1$, i.e. when $\gamma \to \infty$

$$\overline{\mathrm{MSE}} = \begin{cases} (\sqrt{\gamma})^{-1}, & \beta = 1 \\ \big((\beta-1)\gamma\big)^{-1}, & \beta > 1 \end{cases} \tag{5.32}$$

Theorem 5 is the main result of this chapter regarding the MSE for Kalman filtering over fading channels in the high SNR regime. Equations (5.31) and (5.32) can be used to obtain the average MSE and analyze the quality of estimation. In addition, it is also easy to show the following corollary

Corollary 1. *The eigenvalue distribution for the prediction error covariance matrix for the special case of* $A = \sqrt{a}I$*, namely* $f_P(\zeta)$ *can be obtained from*

$$f_P(\zeta) = \frac{ac\sqrt{(a\zeta_u+1-\zeta)(\zeta-a\zeta_l-1)}}{2\pi\gamma(\zeta-1)^2(a+1-\zeta)},\ (a\zeta_l+1 \leqslant \zeta \leqslant a\zeta_u+1). \tag{5.33}$$

Proof. Given that in this case $P(n) = aM(n) + I$, one could easily obtain the eigenvalue distribution of the prediction error covariance matrix using algebraic manipulations as given in (5.33). □

With our system model, the special case of $P = \sqrt{a}I$, combined with $C_u = I$ models a collection of independent time-correlated processes, which share a common communication channel and which have similar state-space models.

Corollary 2. *The eigenvalue distribution for the estimation error covariance matrix for the case when $C_u = \sigma_u^2 I$, can be obtained from*

$$f_M(\zeta) = \frac{c^{'}}{2\pi\gamma\zeta^2(1-\zeta/\sigma_u^2)}\sqrt{(\zeta_u^{'}-\zeta/\sigma_u^2)(\zeta/\sigma_u^2-\zeta_l^{'})},\ (\zeta_l^{'} \leqslant \zeta/\sigma_u^2 \leqslant \zeta_u^{'}) \tag{5.34}$$

with

$$\begin{aligned} c^{'} &= \sqrt{(\beta-1)^2\gamma^2\sigma_u^4 + 2\gamma\sigma_u^2(\beta+1)+1} \\ \zeta_l^{'} &= \frac{\gamma\sigma_u^2(1+\beta-2\sqrt{\beta}+1)}{c^{'2}} \\ \zeta_u^{'} &= \frac{\gamma\sigma_u^2(1+\beta+2\sqrt{\beta}+1)}{c^{'2}}. \end{aligned} \tag{5.35}$$

Proof. Given that in this case

$$\begin{aligned} M(n) &= (1/\sigma_u^2 + \gamma H^T(n)H(n))^{-1} \\ &= \sigma_u^2(I + \gamma\sigma_u^2 H^T(n)H(n))^{-1} \end{aligned} \tag{5.36}$$

one could easily obtain the eigenvalue distribution of the prediction error covariance matrix using algebraic manipulations as given in (5.34). □

In the next section and by numerical evaluations, we evaluate the accuracy of the approximations for the pdf of the instantaneous estimation error and instantaneous prediction error covariance matrices, as well as the average MSE.

5.3 Numerical Evaluation of Results

In this section, the theoretical results from Sec. 5.2, namely the results regarding $f_M(\zeta)$, $f_P(\zeta)$, and $\overline{\text{MSE}}$ are evaluated.

We begin with $f_M(\zeta)$ and for the purpose of evaluation select $K = 8$ and $N = 8,\ 12$ corresponding to the values of $\beta = \frac{N}{K} = 1,\ 3/2$. The matrix A is created by taking

$$A = \frac{1}{K}Q_H D_A Q_H^T,$$

where Q_H represents the Hadamard matrix of dimension $K = 8$, and

$$D_A = \text{diag}\{0.95,\ 0.9,\ 0.5,\ -0.6,\ -0.75,\ 0.97,\ 0.33,\ 0.87\}.$$

We note that this choice of A was only for convenience and any other A leading to a stable process could have been used. In addition, we take $\gamma = 10\,\text{dB}$, which given the current parameters, results in SNR $\approx 26.320\,\text{dB}$.

We simulate a Kalman filter with the given parameters for $n \leqslant 10^5$, and plot and compare $f_M(\zeta)$ from (5.28) vs. the $f_M(\zeta)$ from the simulation in Figures 5.1 and 5.2. Clearly, the main body of the pdf from (5.28) fits well into the profile obtained by the simulation. However, it can be noted that the behavior of the theory and simulation plots are slightly different for the limits of the support of the pdf, especially for the tail. Though this is generally a drawback especially for outage calculation which is considered in the previous chapters, it does not impose a significant problem for evaluating the average MSE, as will be shown later in this section.

Next, $f_P(\zeta)$ is evaluated from (5.33). We select the same set of parameters as before, except for a which is taken to be $a = 0.8$. In that case, we would have that SNR $\approx$ $26.021\,\text{dB}$. The results are plotted in figures 5.3 and 5.4. It is quite visible that the same trends mentioned for $f_M(\zeta)$ hold also for $f_P(\zeta)$.

Finally, we evaluate the average MSE (with the original A), i.e. $\overline{\text{MSE}}$, from (5.31), and its simple high SNR approximation from (5.32). $K, N = 8, 16$ are used respectively and varied SNR $\in [10, 60]\,\text{dB}$. The results are plotted in figures 5.5 and 5.6. Clearly increasing β and SNR decreases the average MSE as predicted by (5.32). It is also visible from the figures that the approximate average MSE obtained from (5.31) is quite accurate even for finite dimensions and moderate SNR. The accuracy improves when bigger β, higher SNR, or a larger dimension is applied.

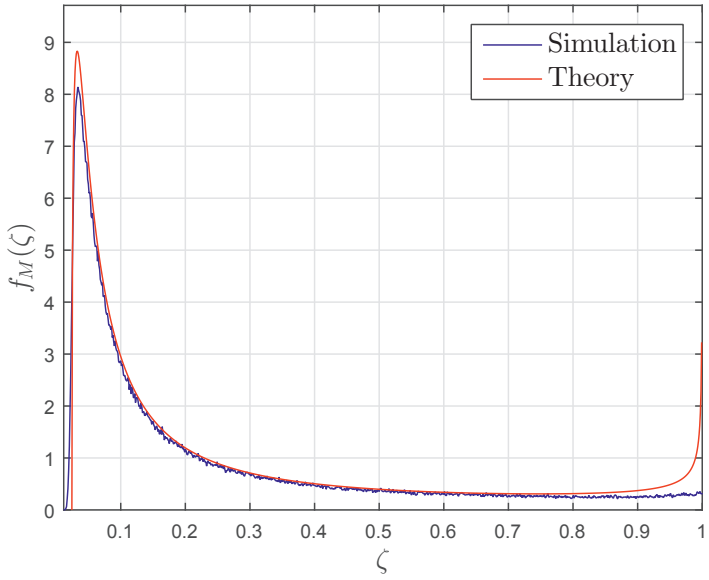

Figure 5.1: Comparison of theory from (5.27) and simulation for $N = 8$, $K = 8$ and $\gamma = 10\,\text{dB}$.

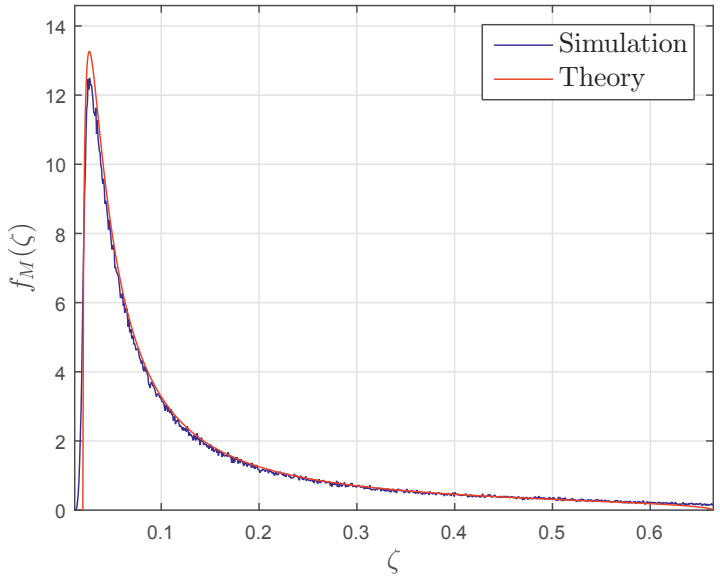

Figure 5.2: Comparison of theory from (5.27) and simulation for $N = 12$, $K = 8$ and $\gamma = 10\,\text{dB}$.

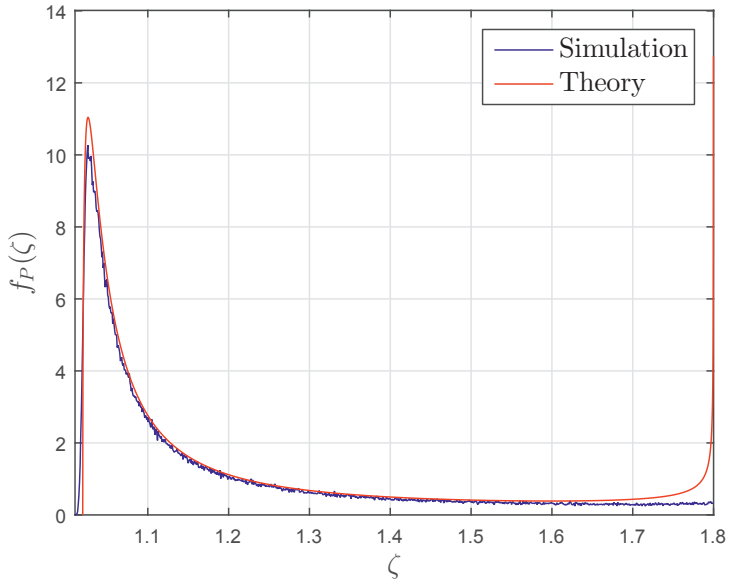

Figure 5.3: Comparison of theory from (5.33) and simulation for $N = 8$, $K = 8$ and $\gamma = 10\,\text{dB}$

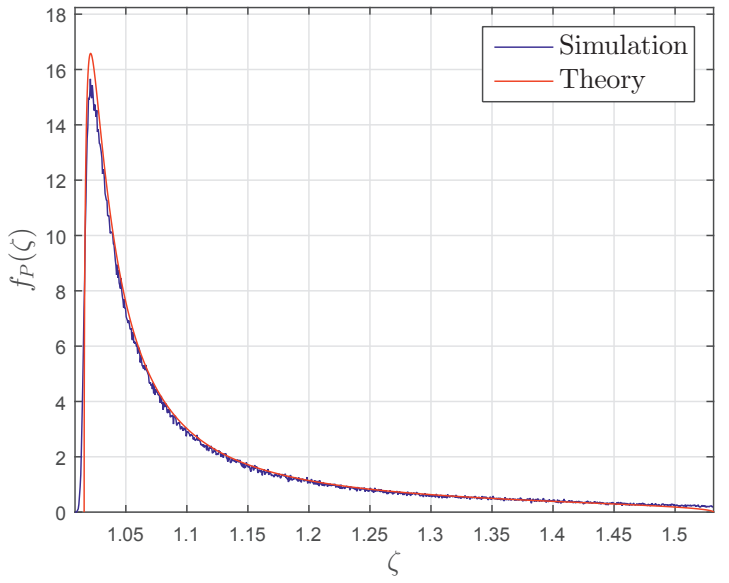

Figure 5.4: Comparison of theory from (5.33) and simulation for $N = 12$, $K = 8$ and $\gamma = 10\,\text{dB}$

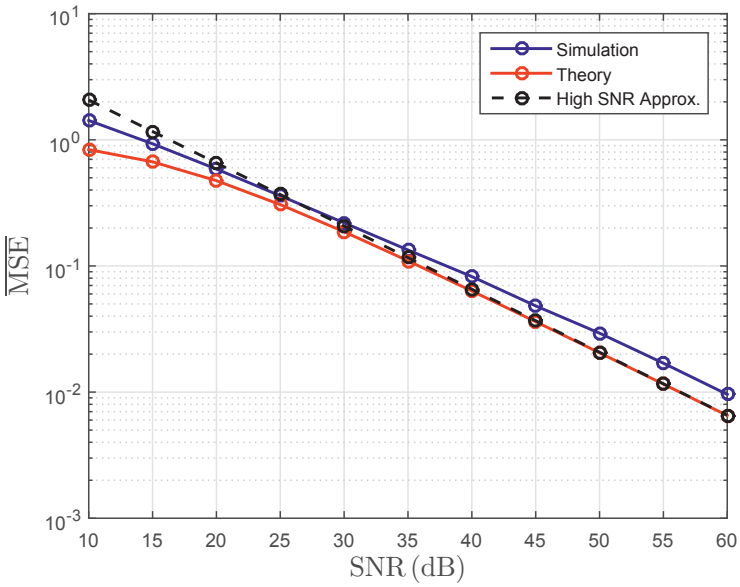

Figure 5.5: Comparison of theory from (5.31) and (5.32) and simulation for $N = 8,\ K = 8$

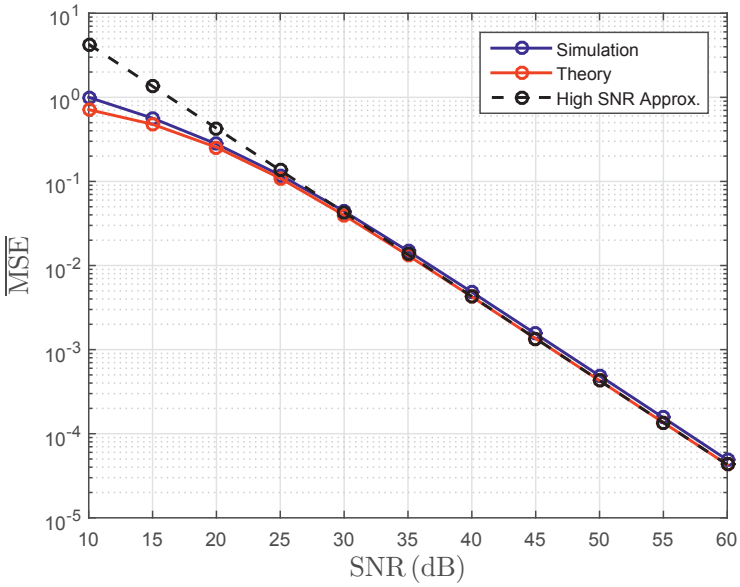

Figure 5.6: Comparison of theory from (5.31) and (5.32) and simulation for $N = 16,\ K = 8$

5.4 Summary and Discussion

In this chapter, we studied the steady state eigenvalue distribution of the prediction and estimation error covariance matrices, for Kalman filtering of certain large dimensional sources over fading channels. With the help of the Stieltjes transform, approximate formulas are found for the distributions in the high SNR regime. We also obtained a simple formula for the average MSE in the high SNR regime. In order to make the analysis tractable, some simplifying assumptions were made, which are discussed in more details in the following.

One assumption was that the signals are real. Note that the formulation for complex Kalman filtering for large dimensional signals is no different from what was presented in this chapter, apart from the fact that all matrix transpose operations must be done as conjugate transpose operations. All the theoretical results regarding $f_M(\zeta)$, $f_P(\zeta)$, and $\overline{\text{MSE}}$ will hold as well. As for compliance with numerical simulations, good accuracy even for dimensions as low as $N = K = 8$ can be observed. However, it was noticed that for the case of complex signals, when $\beta = 1$ and for small dimensions, the accuracy is affected such that the actual average MSE from the simulations is lower than what is predicted from the analysis. Although, the gap between theory and simulation closes as dimensions grow bigger, and it is the regime of large dimensions for which the theory is actually supposed to hold, this is a phenomenon we currently cannot fully explain. As a result, we have decided to limit the system model to real signals.

In this chapter, a Gaussian system model is assumed. In the limit of large dimensions, this assumption is not necessary for any of the asymptotic results. The Gaussian assumption was kept in order to be close to the original system model as much as possible. For a Gaussian system model, the Kalman filter is the optimal filter. For other distributions, the Kalman filter is the optimal linear filter. Given that the asymptotic results are the same for any source distribution, it is safe to say that the Kalman filter is the optimal filter with respect to average MSE for any source distribution, if system dimensions grow infinitely large.

For our analysis, we assume that $\beta \geqslant 1$. Consider only one of the source dimensions in this system model, which could be the signal from one of the assumed sensors. The received signal consists of the transmitted signal plus noise and the signals from other source dimensions, which for our purpose can be considered as interference. If no coding or transform is used, e.g. space-time coding, when $\beta < 1$, asymptotically increasing the SNR results in unlimited increase of the interference from other dimensions. This will ruin the performance, as the effective signal-to-interference-plus-noise-ratio will actually approach zero for $\beta < 1$. For that matter, one should always have $\beta \geqslant 1$.

The infinite dimension assumption was made in order to simplify the analysis, and it was already observed that it has interesting consequences, e.g. the Gaussian assumption becomes irrelevant. However, if the high SNR assumption is kept, one may obtain the eigenvalue distribution of $M(n)$ via manipulating the eigenvalue distribution of $H^T H$,

which is a Wishart matrix. The eigenvalue distribution problem for Wishart matrices is a well-studied one, especially in the context of digital MIMO communication. A good review of the existing work, plus new theory on ordered eigenvalues of different Wishart matrices can be found in [121]. The analysis using eigenvalue distribution of Wishart matrices can however be more involved than the analysis presented in this chapter, and for our purpose, we have chosen to apply the Stieltjes transform tool instead.

Chapter 6

Oversampling Diversity for Band-limited Signals

In this chapter, uncoded transmission of band-limited signals over fading channels is considered. So far, we have considered different cases of first order ARGM signals, with the corresponding estimator, i.e. the Kalman filter. When a general band-limited signal is to be recovered from noise, and when there is no state-space model, the Wiener filter can be used. When the signals are sent over random fading channels, the estimation error covariance matrix of the filter for each channel realization is also random. Considering the normalized trace of that matrix as the instantaneous random distortion, one is faced with a similar problem as in the last two chapters, i.e. estimation quality assessment based on the statistical behavior of the random distortion. In order to evaluate the quality of estimation, we consider the outage probability as a criterion. In addition and in order to improve the performance, we propose using additional antennas at the receiver in order to obtain a diversity gain for the distortion outage probability.

In this chapter, it is assumed that a K-dimensional signal is transmitted over a $N \times K$ fading channel. The mathematical model of the system can be interpreted in two different ways. If the measured process already is a vector process of dimension K, the mathematical model represents transformation and dimension expansion in order to obtain a diversity gain. This is also a zero-delay transmission scheme. On the other hand, it is also possible to look at the scheme in such a way that the N values are consecutive samples of an oversampled band-limited scalar process, with the oversampling ratio of N/K. In that case, a processing delay of $N-1$ samples will be incurred due to buffering. In both cases however, there is an increase in the number of original signal samples by a factor of N/K. We have thus selected the name *oversampling diversity* for the proposed scheme. To achieve the diversity gain, it is assumed that the N fading channels are parallel, and also perfectly known at the receiver. As the final result of this chapter, we will show that the distortion outage probability vanishes inversely polynomially with SNR, with the exponent of $N-K+1$ in the high SNR regime, thus achieving a diversity order of $N-K+1$.

6.1 Updated System Model and Problem Definition

Consider the following signal model[1]

$$\mathbf{x}(n) = FD\mathbf{x_s}(n) \tag{6.1}$$

where $\mathbf{x_s}_{K\times 1}(n)$ is a zero-mean, white (complex Gaussian) vector whose elements have unit power, i.e. $E\left(\mathbf{x_s}^H(n)\mathbf{x_s}(n)\right) = I_{K\times K}$. The matrix $D_{N\times K}$ is all zeros except at K positions, whose values are equal to one. It should be noted that the elements of D as $D[k,l]$, and non-zero elements must be selected in such a way that each column of the matrix must have only one non-zero element (equal to one) and no non-zero elements (ones) should occur in the same row. The indexes of the rows k where $D[k,l]$ are equal to one are denoted by $\bar{k}$. We call D as the oversampling matrix. Two such D's (D_1 and D_2) for the setting where $N = 3, K = 2$ are shown below as

$$D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}, \text{ or } \qquad D_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}. \tag{6.2}$$

It is possible to show that $DD^T = \bar{D}$ is diagonal and is only non-zero (equal to one) at positions $\bar{k}$ in the main diagonal ($\bar{D}[\bar{k},\bar{k}] = 1$). We have for instance that

$$D_1D_1^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ and } \qquad D_2D_2^T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \tag{6.3}$$

It is also easy to show that $D^TD = I_{K\times K}$.

The matrix F is a unitary matrix and is used to model/shape the spectrum of $\mathbf{x_s}$. In particular, a normalized inverse discrete Fourier transform (iDFT) matrix of dimension N can be used, i.e. $F[k,l] = \frac{1}{\sqrt{N}}\exp(\frac{\boldsymbol{i}2\pi kl}{N})$, $k,l = 0,\ldots,N-1$, with $\boldsymbol{i} = \sqrt{-1}$. The matrix F^H is then the normalized DFT matrix.

Selecting different matrices for F results in different interpretations of the signal model. If $\mathbf{x_s}$ is originally a vector source, then the signal model is a dimension expansion in order to enable transmission over a large number of channels. However, by selecting F to be the normalized iDFT matrix, equation (6.1) can actually approximate a band-limited signal $\mathbf{x}(n)$ as shown later. The approximation improves as the dimensions N and K increase. However, this would incur larger delay for scalar sources. In order to show that the band-limitedness in fact holds, we first calculate the covariance matrix for $\mathbf{x}(n)$ (note that $\mathbf{x_s}(n)$

[1]The variable names for the system model are selected such that they best fit the oversampled scalar case, rather than the K-dimensional vector case.

is zero-mean) as

$$\begin{aligned} C_{xx} &= E\left(\mathbf{x}(n)\mathbf{x}(n)^H\right) \qquad (6.4)\\ &= E\left(FD\mathbf{x_s}(n)\mathbf{x_s}^H(n)D^TF^H\right)\\ &= FDE\left(\mathbf{x_s}(n)\mathbf{x_s}^H(n)\right)D^TF^H\\ &= FDID^TF^H\\ &= F\bar{D}F^H. \qquad (6.5)\end{aligned}$$

We can then take the DFT of the covariance (correlation) matrix to obtain the power spectrum matrix, i.e.

$$\begin{aligned} S_{xx} &= E\left(F^H\mathbf{x}(F^H\mathbf{x})^H\right)\\ &= F^HE\left(\mathbf{x}\mathbf{x}^H\right)F\\ &= F^HF\bar{D}F^HF\\ &= \bar{D}. \qquad (6.6)\end{aligned}$$

Now, $\bar{D}$ has only K non-zero elements and is diagonal as mentioned before. By selecting the correct placement for non-zero elements of D, the model can (approximately) represent a band-limited signal due its particular spectrum shape. Note that $\mathbf{x}(n)$ is only truly band-limited if $N, K \to \infty$ at fixed rate K/N. However, smaller sizes result in a signal with most of the energy of the spectrum in certain bands. In this way, this signal model can also be used to analyze the transmission of a signal over fading channels while the signal has been oversampled when going through discretization. This provides us with a model which is useful for two applications which are different in nature, but have the same mathematical model.

For transmission, it is assumed that the signal $\mathbf{x}(n)$ is sent over a frequency-flat fading channel. The received signal will then be equal to

$$\mathbf{y}(n) = \sqrt{\frac{\gamma}{K}}H(n)\mathbf{x}(n) + \mathbf{v}(n). \qquad (6.7)$$

It is assumed that $H(n)$ is a diagonal matrix of dimension $N \times N$, is random, but is fixed for L channel uses, i.e. the block fading model. It is also assumed that $\mathbf{y}(n)$ and $\mathbf{v}(n)$ are column vectors of length N. With this model, $H(n)$ represents N parallel scalar channels. The diagonal elements of H are assumed to be unit variance zero mean complex Gaussian random variables (Rayleigh fading). Furthermore, it is assumed that entries of $\mathbf{v}(n)$ are white. With this definition, the SNR at the receiver is also equal to SNR $= \gamma$.

At the receiver, MMSE estimation is performed on channel outputs $\mathbf{y}(n)$. It is assumed that $H(n)$ is known to the receiver, but not the transmitter. Given known $H(n)$, the MMSE estimator is linear (Wiener filter) and follows [49]

$$\hat{\mathbf{x}}(n) = C_{xy}C_{yy}^{-1}\mathbf{y}(n) \qquad (6.8)$$

and the corresponding estimation error covariance matrix is equal to

$$C_{x|y} = C_{xx} - \frac{\gamma}{K} C_{xy} C_{yy}^{-1} C_{yx}. \tag{6.9}$$

The random distortion $d(n)$ is then defined as

$$d(n) = \frac{1}{N} \mathrm{tr}(C_{x|y}). \tag{6.10}$$

Note that $d(n)$ is random due to the random channel H (the time index in $H(n)$ is dropped hereafter). Furthermore, the distortion outage probability is equal to

$$\mathrm{P}_{\text{out}} = \Pr\left(d(n) \geqslant d_{\text{th}}\right), \tag{6.11}$$

where d_{th} is an arbitrary distortion threshold. Considering the high SNR regime, the objective of this chapter is to show how the distortion outage probability varies with the SNR and N, K as the SNR goes to infinity, and find the eventual diversity order.

6.2 Diversity Order for Distortion Outage Probability

In this section, we analyze the distortion outage probability in the high SNR regime and find the diversity order. First, we obtain a formula for the random distortion of the estimator. Then we use that formula in order to find the diversity order.

6.2.1 Distortion Calculation

We begin the analysis by obtaining simple equations for the MMSE estimator and the corresponding MSE and distortion. It is easy to show that $C_{xy} = \sqrt{\frac{\gamma}{K}} C_{xx} H^H$ and $C_{yy} = \frac{\gamma}{K} H C_{xx} H^H + \sigma_v^2 I$. This results in

$$C_{xy} C_{yy}^{-1} = \sqrt{\frac{\gamma}{K}} C_{xx} H^H \left(\frac{\gamma}{K} H C_{xx} H^H + \sigma_v^2 I \right)^{-1},$$

which can also be written as

$$C_{xy} C_{yy}^{-1} = \sqrt{\frac{\gamma}{K}} F \bar{D} F^H H^H \left(\frac{\gamma}{K} H F \bar{D} F^H H^H + \sigma_v^2 I \right)^{-1}. \tag{6.12}$$

We take the corresponding columns of F where $\bar{D}$ has a one, and arrange them in a $N \times K$ matrix $\Omega = FD$. It is easy to see that $C_{xx} = F \bar{D} F^H = \Omega \Omega^H$. One can therefore rewrite the expression for the estimator as

$$C_{xy} C_{yy}^{-1} = \sqrt{\frac{\gamma}{K}} \Omega \Omega^H H^H \left(\frac{\gamma}{K} H \Omega \Omega^H H^H + \sigma_v^2 I \right)^{-1}. \tag{6.13}$$

Next, $C_{x|y}$ matrix is simplified as

$$C_{x|y} = \Omega\Omega^H - \frac{\gamma}{K}\Omega\Omega^H H^H \left(\frac{\gamma}{K} H\Omega\Omega^H H^H + \sigma_v^2 I\right)^{-1} H\Omega\Omega^H. \tag{6.14}$$

Then, the Woodbury matrix identity [76] is used to obtain

$$\begin{aligned}\left(\frac{\gamma}{K} H\Omega\Omega^H H^H + \sigma_v^2 I\right)^{-1} &= \left(\frac{\gamma}{K}(H\Omega)(H\Omega)^H + \sigma_v^2 I\right)^{-1} \\ &= \frac{1}{\sigma_v^2} I - \frac{\frac{\gamma}{K}}{\sigma_v^4} H\Omega \left(I + \frac{\frac{\gamma}{K}}{\sigma_v^2}(H\Omega)^H(H\Omega)\right)^{-1} (H\Omega)^H.\end{aligned} \tag{6.15}$$

Inserting (6.15) into (6.14) results in

$$\begin{aligned}C_{x|y} &= \Omega\Omega^H - \frac{\gamma}{K}\Omega\Omega^H H^H \left(\frac{1}{\sigma_v^2} I - \frac{\frac{\gamma}{K}}{\sigma_v^4} H\Omega \left(I + \frac{\frac{\gamma}{K}}{\sigma_v^2}(H\Omega)^H(H\Omega)\right)^{-1} (H\Omega)^H\right) H\Omega\Omega^H \\ &= \Omega\Omega^H - \Omega\left(\frac{\frac{\gamma}{K}}{\sigma_v^2}(H\Omega)^H(H\Omega)\right)\Omega^H \\ &+ \Omega\left(\frac{\frac{\gamma}{K}}{\sigma_v^2}(H\Omega)^H(H\Omega)\right)\left(I + \frac{\frac{\gamma}{K}}{\sigma_v^2} H\Omega)^H(H\Omega)\right)^{-1}\left(\frac{\frac{\gamma}{K}}{\sigma_v^2}(H\Omega)^H(H\Omega)\right)\Omega^H.\end{aligned} \tag{6.16}$$

Taking $S = \frac{\frac{\gamma}{K}}{\sigma_v^2}(H\Omega)^H(H\Omega)$ results in

$$C_{x|y} = \Omega\Omega^H - \Omega S\Omega^H + \Omega S(I+S)^{-1}\Omega^H. \tag{6.17}$$

Now we try to find $\text{tr}(C_{x|y})$ as follows

$$\begin{aligned}\text{tr}(C_{x|y}) &= \text{tr}(\Omega\Omega^H) - \text{tr}\left(\Omega S\Omega^H\right) + \text{tr}\left(\Omega S(I+S)^{-1} S\Omega^H\right) \\ &= \text{tr}(\Omega^H\Omega) - \text{tr}\left(\Omega^H\Omega S\right) + \text{tr}\left(\Omega^H\Omega S(I+S)^{-1} S\right) \\ &\overset{(a)}{=} K - \text{tr}(S) + \text{tr}\left(S^2(I+S)^{-1}\right) \\ &\overset{(b)}{=} K - \text{tr}(S) + \text{tr}\left(S^2 S^{-1}(I+S^{-1})^{-1}\right) \\ &= K - \text{tr}(S) + \text{tr}\left(S(I+S^{-1})^{-1}\right) \qquad (6.18) \\ &= \text{tr}(I - S + S(I+S^{-1})^{-1}), \qquad (6.19)\end{aligned}$$

where (a) is due to $\Omega^H\Omega = I_{K\times K}$ and (b) comes from the following matrix equality [86, Page 151]

$$(I + A^{-1})^{-1} = A(A+I)^{-1} \tag{6.20}$$

for any arbitrary invertible matrix A.

In the following lemma, we provide a representation of $d(n)$ which we will use in the diversity analysis.

Lemma 17. *The distortion for a given realization of H can be obtained from*

$$d(n) = \frac{1}{N} tr\left((I+S)^{-1}\right). \tag{6.21}$$

Proof. We can rewrite $\text{tr}(C_{x|y})$ from (6.19) as

$$\begin{aligned}
\text{tr}(C_{x|y}) &= \text{tr}(I - S + S(I+S^{-1})^{-1}) \\
&= \text{tr}(I - S + S(I - S^{-1} + S^{-2} - S^{-3} + \ldots)) \\
&= \text{tr}(I - S + S - I + S^{-1} - S^{-2} + \ldots) \\
&= \text{tr}(S^{-1} - S^{-2} + S^{-3} - S^{-4} + \ldots) \\
&= \text{tr}\sum_{i=1}^{\infty}(-1)^{i-1}S^{-i}.
\end{aligned} \tag{6.22}$$

We now take $Z = \sum_{i=1}^{\infty}(-1)^{i-1}S^{-i}$ and find a form which only depends on S^{-1}.

We begin with

$$\begin{aligned}
Z &= S^{-1} - S^{-2} + S^{-3} - S^{-4} + \ldots \\
&= S^{-1}(I - S^{-1} + S^{-2} - S^{-3} + \ldots \\
&= S^{-1}(I+S^{-1})^{-1}.
\end{aligned} \tag{6.23}$$

We can also write

$$\begin{aligned}
Z &= S^{-1} - S^{-2} + S^{-3} - S^{-4} + \ldots \\
&= I - I + S^{-1}(I - S^{-1} + S^{-2} - S^{-3} + \ldots \\
&= S^{-1}(I+S^{-1})^{-1}.
\end{aligned} \tag{6.24}$$

Combining (6.23) and (6.24) results in

$$I - (I+S^{-1})^{-1} = S^{-1}(I+S^{-1})^{-1}. \tag{6.25}$$

We multiply both sides in (6.25) with S. As a result, it is obtained that

$$S - S(I+S^{-1}) = (I+S^{-1})^{-1}. \tag{6.26}$$

Then, we get that

$$\begin{aligned}
S &= S(I+S^{-1}) + (I+S^{-1})^{-1} \\
&= (I+S)(I+S^{-1}),
\end{aligned} \tag{6.27}$$

which results in

$$(I + S^{-1})^{-1} = S(I + S)^{-1}. \tag{6.28}$$

Combining (6.28) with (6.23) results in

$$\begin{aligned} Z &= S^{-1}(I + S^{-1})^{-1} \\ &= S^{-1}S(I + S)^{-1} \\ &= (I + S)^{-1}, \end{aligned} \tag{6.29}$$

which means that

$$\begin{aligned} d(n) &= \frac{1}{N}\text{tr}(Z) \\ &= \frac{1}{N}\text{tr}\left((I + S)^{-1}\right), \end{aligned} \tag{6.30}$$

and proof of lemma 17 is complete. □

Lemma 17 will be used for finding the diversity order in the next section.

6.2.2 Diversity Order Analysis for Parallel Channels

For parallel channels, $H^H H$ is diagonal. We take $\zeta_i = |h_i|^2$, so that the diagonal elements of $H^H H$ are equal to ζ_i. Let $\zeta_1 \leq \zeta_2 \leq \ldots \leq \zeta_N$ be the ordered diagonal elements (also eigenvalues) of $H^H H$ in increasing order.

In order to find the diversity order for the outage probability, it must be such that if the SNR goes to infinity, the outage probability must go to zero, unless S becomes singular. For small enough threshold d_{th}, outage may occur even though S is non-singular. For that case, the event that S is close to singular is the event which leads to outage. With that reasoning, the outage probability may be considered to be approximately equal to the (almost) singularity probability of S, which will be studied in the following. We note that if the eigenvalues of S are ordered, it is the smallest eigenvalues which have the most significant effect on that event.

To perform the outage analysis, two key components are required. First, the near-zero behavior of the cdf of the ordered eigenvalues of $H^H H$ is needed. In addition, a high SNR approximation for the distortion can also be very useful. These are presented in Lemmas 18 and 19.

The cdf of the unordered eigenvalues, i.e. $\zeta_i, i = 1, 2, \ldots, N$ near zero is of the form $F_{\zeta_i}(\zeta) \approx c \cdot \zeta$ for small $\zeta > 0$ and a SNR-dependent constant c, assuming ζ_i is the magnitude squared of a complex Gaussian variable (see [111] for more details). This fact may be used in order to obtain the cdf of the ordered eigenvalues as presented in Lemma 18.

Lemma 18. *The cdf of the $(N-K+1)$-st ordered i.i.d. exponential random variable, among N i.i.d. exponential random variables for small ζ follows $F_{\zeta_{N-K+1}}(\zeta) \approx c_{N-K+1} \cdot \zeta^{N-K+1}$ for some constant c_{N-K+1}.*

Proof. Consider N ordered i.i.d. exponential random variables, namely $\zeta_1 \leqslant \zeta_2, \ldots, \zeta_N$. We are interested in the near zero behavior of the cdf of ζ_{N-K+1}. As all the variables are i.i.d., then the cdf of ζ_i is equal to

$$\begin{aligned} F_\zeta(\zeta) &= \text{Pr}(\zeta_i < \zeta) \\ &= 1 - \exp(\zeta/\gamma), \end{aligned}$$

for some γ. If, ζ is the magnitude square of a Rayleigh fading channel, i.e. $\zeta_i = |h_i|^2$, then γ is equal to the SNR of the channel. We have in addition that the pdf as $f_\zeta(\zeta) = 1/\gamma \exp(\zeta/\gamma)$. We also denote the pdf (cdf) of the m-smallest random variable as $f_m(\zeta)$ ($F_m(\zeta)$), e.g. the pdf of the minimum of this random variables is denoted by $f_1(\zeta)$.

From [18], know that the pdf of the $(N-K+1)$'st variable in this set can be obtained from

$$f_{N-K+1}(\zeta) = \frac{N!}{(N-K)!(K-1)!} F_\zeta(\zeta)^{N-K} (1 - F_\zeta(\zeta))^{K-1} f_\zeta(\zeta). \tag{6.31}$$

The near zero behavior of $F_{N-K+1}(\zeta)$ can now be obtained from (6.31). Simply, from Taylor series expansion of $F_\zeta(\zeta)$ around zero, one can see that $F_\zeta(\zeta) = \zeta/\gamma + o(\zeta)$. Now, the Taylor series expansion of $F_{N-K+1}(\zeta)$ around zero can be written as

$$F_{N-K+1}(\zeta) = \sum_{i=0}^{\infty} \frac{F_{N-K+1}^{(i)}(0)}{i!} \zeta^i, \tag{6.32}$$

where $F_{N-K+1}^{(i)}(0)$ is the i'th derivative of $F_{N-K+1}(\zeta)$ at $\zeta = 0$. Given that $\zeta_i \geqslant 0$,

$$\begin{aligned} F_{N-K+1}^{(i)}(0) &= \lim_{\zeta \to 0} \frac{\partial^i}{\partial \zeta^i} \int_0^\zeta f_{N-K+1}(u)\, \mathrm{d}u \\ &= f_{N-K+1}^{(i-1)}(0). \end{aligned} \tag{6.33}$$

Now, given that $F_\zeta(\zeta) = \zeta/\gamma + o(\zeta)$ and consequently $f_\zeta(\zeta) = 1/\gamma + o(1)$, it is possible to rewrite (6.31) as

$$\begin{aligned} & f_{N-K+1}(\zeta) \\ &= \frac{N!}{(N-K)!(K-1)!} (\zeta/\gamma + o(\zeta))^{N-K} (1 - (\zeta/\gamma + o(\zeta)))^{K-1} (1/\gamma + o(1)) \\ &= \frac{N!}{(N-K)!(K-1)!} (\zeta/\gamma + o(\zeta))^{N-K} (1 - (K-1)(\zeta/\gamma) + o(\zeta))(1/\gamma + o(1)) \\ &= \frac{N!}{(N-K)!(K-1)!} (\zeta/\gamma + o(\zeta))^{N-K} (1/\gamma + o(1)) \\ &= \frac{N!}{\gamma(N-K)!(K-1)!} ((\zeta/\gamma)^{N-K} + o(\zeta^{N-K})) \end{aligned} \tag{6.34}$$

Now, clearly from (6.34), we have that

$$f^{(l)}_{N-K+1}(0) = 0,\ l = 0, 1, \ldots, N-K-1, \tag{6.35}$$

while

$$f^{(l)}_{N-K+1}(0) \neq 0,\ l \geqslant N-K. \tag{6.36}$$

As a result

$$F^{(l)}_{N-K+1}(0) = 0,\ l = 0, 1, \ldots, N-K, \tag{6.37}$$

while we have that

$$F^{(l)}_{N-K+1}(0) \neq 0,\ l \geqslant N-K+1. \tag{6.38}$$

Consequently, we obtain that

$$F_{N-K+1}(\zeta) = c_{N-K+1}\zeta^{N-K+1} + o(\zeta^{N-K+1}) \tag{6.39}$$

for some constant c_{N-K+1} as claimed. □

Regarding the high SNR approximation for the distortion, consider the following lemma.
Lemma 19. *In the high SNR regime, the instantaneous random distortion may be approximated as*

$$d(n) \approx \frac{1/N}{\left(1 + c^{'}_{N,K}\gamma\zeta_{N-K+1}\right)}, \tag{6.40}$$

for some constant $c^{'}_{N,K}$.

Proof. Assume that $\beta_i,\ i = 1, 2, \ldots, K$ are the eigenvalues of S. Then we have that

$$\text{tr}(I+S)^{-1} = \sum_{i=1}^{K} \frac{1}{1+\beta_i}. \tag{6.41}$$

It is already assumed that $K-1$ lowest ζ_i are equal to zero ($\zeta_1 \leqslant \zeta_2 \leqslant \ldots \leqslant \zeta_{K-1} = 0$), but ζ_{N-K+1} has a small but non-negligible value ϵ. In this case, S is still of rank K. Now if $\epsilon \to 0$, then the rank becomes $K-1$. That means, we will have one eigenvalue of S equal to 0. We denote this eigenvalue by β_j. Other eigenvalues of S should be non-negligible positive values. S is a normal matrix, i.e. $S^H S = SS^H$. It can be shown from [23] that an ϵ change in $H^H H$, will at most incur $\epsilon K\gamma/(N\sigma_v^2)$ change in the eigenvalues. This means that β_j was at most equal to ϵ before the change. We use that value as an approximation

and for when $\zeta_{N-K+1} = \epsilon$, write

$$\begin{aligned}
\text{tr}(I+S)^{-1} &\approx \frac{1}{1+\epsilon K\gamma/(N\sigma_v^2)} + \sum_{\substack{i=1\\ i\neq j}}^{K} \frac{1}{1+\beta_i} \\
&\approx \frac{1}{1+\zeta_{N-K+1} K\gamma/(N\sigma_v^2)} + \sum_{\substack{i=1\\ i\neq j}}^{K} \frac{1}{1+\beta_i} \\
&\overset{(a)}{\approx} \frac{1}{1+\zeta_{N-K+1} K\gamma/(N\sigma_v^2)} \\
&= \frac{1}{\left(1+c_{N,K}^{'}\gamma\zeta_{N-K+1}\right)},
\end{aligned} \tag{6.42}$$

with $c_{N,K}^{'} = \frac{K}{N\sigma_v^2}$, and where (a) comes from the fact that β_i, $i \neq j$ are considerably bigger than β_j. Otherwise, the rank would become less than $K-1$, which would then contradict our assumptions. As a result, $d(n)$ may be obtained as suggested in the lemma. □

Continuing the diversity analysis, we turn to the singularity event (corresponding to the outage event), and consider the scenario in which $\zeta_1, \ldots, \zeta_{N-K}$ are close to zero, ζ_{N-K+1} has small value and $\zeta_{N-K+2}, \ldots, \zeta_N$ are big enough (due to the high SNR), so that singularity of S only comes from ζ_{N-K+1} going to zero. Then, using Lemma 19, results in

$$\begin{aligned}
d(n) &\approx \frac{1/N}{\left(1+c_{N,K}^{'}\gamma\zeta_{N-K+1}\right)} \\
&\approx 1/N(1-c_{N,K}^{'}\gamma\zeta_{N-K+1}).
\end{aligned} \tag{6.43}$$

The behavior of the distortion $d(n)$ in the neighborhood of the normalized sum power of the source, i.e. the maximum possible distortion, is determined by the behavior of ζ_{N-K+1} near zero. This neighborhood near zero is magnified by γ. So, it is a neighborhood of size of order $1/\gamma$. One may then write the outage probability as

$$\begin{aligned}
P_{\text{out}} &= \text{Pr}\left(d(n) \geqslant d_{\text{th}}\right) \\
&= \text{Pr}(d(n) \in [1/N - d_{\text{th}}, 1/N]),
\end{aligned} \tag{6.44}$$

for some arbitrary d_{th}. Using Lemma 18, the outage probability can then be approximated as

$$\begin{aligned}
P_{\text{out}} &= \text{Pr}(d(n) \in [1/N - d_{\text{th}}, 1/N]) \\
&\approx F_{\zeta_{N-K+1}}(\frac{1-Nd_{\text{th}}}{c_{N,K}^{'}\gamma}) \\
&= c_{N-K+1}\gamma^{-(N-K+1)},
\end{aligned} \tag{6.45}$$

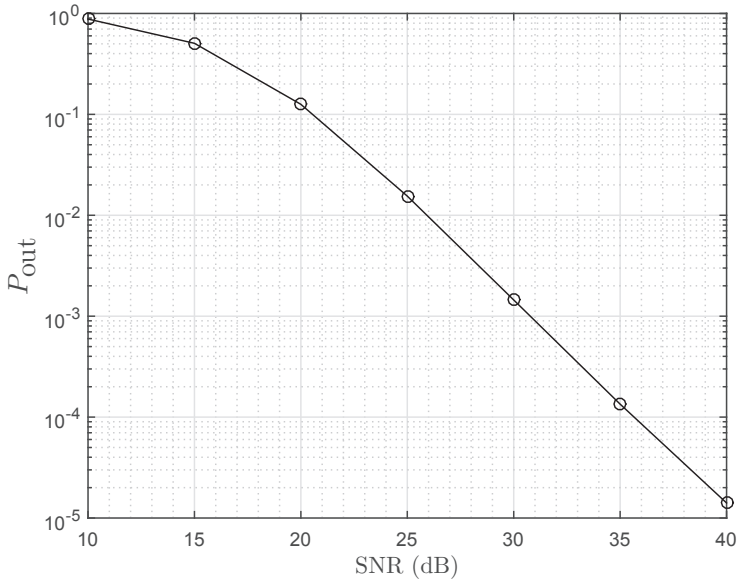

Figure 6.1: Distortion outage probability for $N = 4$, $K = 3$, $d_{\text{th}} = 0.1$, $\sigma_x^2 = \sigma_v^2 = 1$, block size = 1.

which results in a diversity order of $N - K + 1$ for the outage probability. Numerical simulations are provided in the following, in order to evaluate the accuracy of the results.

We select several values for N, K and plot the P_{out} vs. SNR curve in Figures 6.1 - 6.4.

we begin with selecting $N = 4$, $K = 3$ and $d_{\text{th}} = 0.1$, $\sigma_x^2 = \sigma_v^2 = 1$. We also try different block sizes of $L = 1, 10$ over which the channel H is assumed fixed. The simulation is performed for 10^7 blocks. The results are depicted in Figures 6.1 and 6.2. The expected diversity order for distortion outage probability is equal to $d_{\text{ord}} = N-K+1 = 4-3+1 = 2$. It is quite visible in both figures that the slope of both curves in the high SNR regime follows the theory.

Next, we take $K = 2$, while $L = 1, 10$, and keep the other system parameters the same. The simulation was performed for 10^8 blocks and the results are presented in Figures 6.3 and 6.4. A diversity order equal to $N - K + 1 = 4 - 2 + 1 = 3$ is expected. The results in these figures also indicate compliance with the theory.

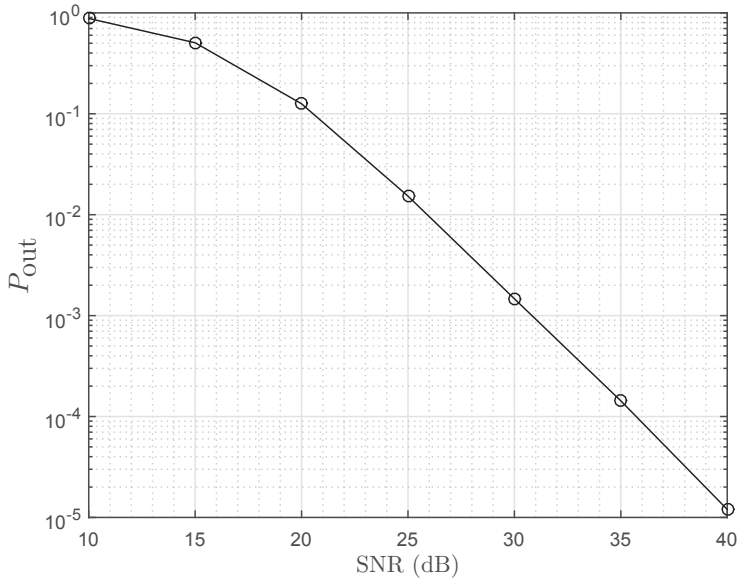

Figure 6.2: Distortion outage probability for $N = 4$, $K = 3$, $d_{\text{th}} = 0.1$, $\sigma_x^2 = \sigma_v^2 = 1$, block size = 10.

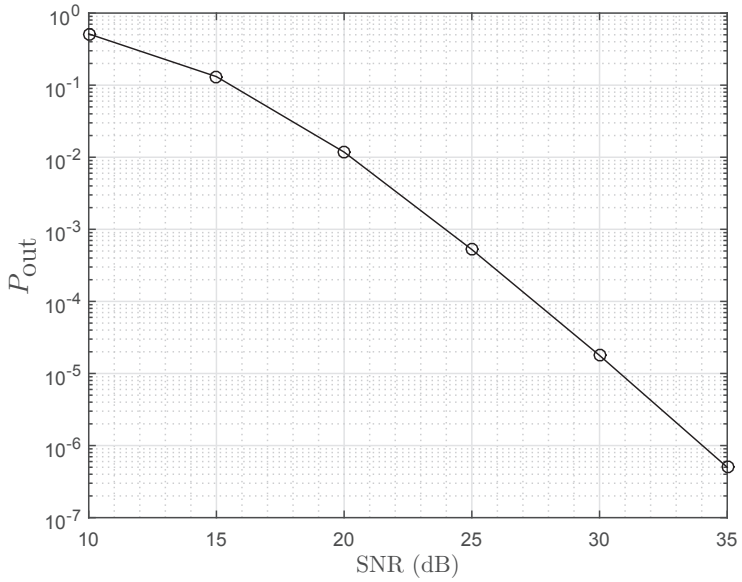

Figure 6.3: Distortion outage probability for $N = 4$, $K = 2$, $d_{\text{th}} = 0.1$, $\sigma_x^2 = \sigma_v^2 = 1$, block size = 1.

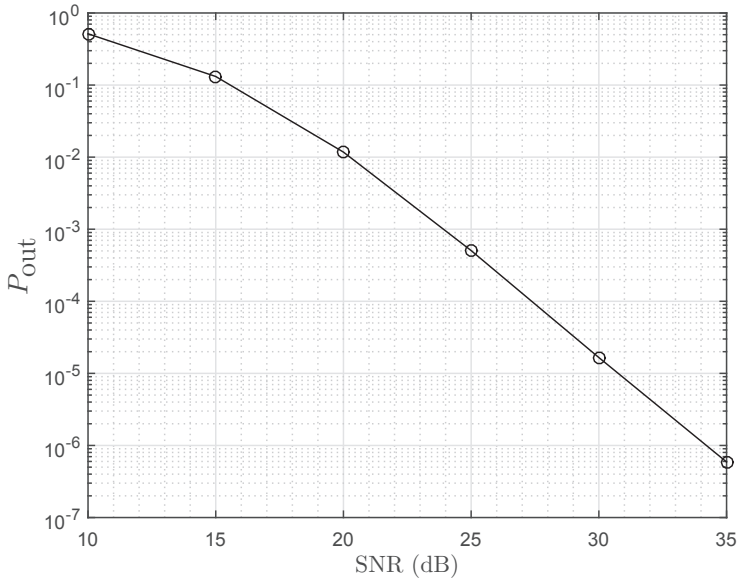

Figure 6.4: Distortion outage probability for $N = 4$, $K = 2$, $d_{\text{th}} = 0.1$, $\sigma_x^2 = \sigma_v^2 = 1$, block size = 10.

6.3 Summary and Discussion

In this chapter, we proposed a simple framework to perform uncoded transmission of a band-limited signal over N parallel fading channels with an N/K oversampling rate, and incorporated the distortion outage probability as a measure of estimation quality. We applied techniques from Taylor series analysis and perturbation of eigenvalues of normal matrices, and showed that in the high SNR regime, the estimation error outage probability vanishes inversely polynomially with SNR, with the exponent of $N-K+1$, and therefore achieves a diversity order of $N - K + 1$.

A basic part of the system model is the (inverse) discrete Fourier transform matrix. The mathematical system model can in fact model two different types of signals. For the scalar ones, the matrix F is used to model the spectrum of an oversampled band-limited signal. The transmission part will then simply consist of sending N samples over N i.i.d. channels. For the vector case, F can be considered to be the pre-filter $T(\cdot)$ in the original system model of Chapter 1, along with the matrix D. In that sense, the role of matrix F is only to spread each of the K source dimensions over N channels. Other transformations may be used as well, and as long as unitary matrices are concerned, the diversity results will remain the same.

Chapter 7

Conclusion

7.1 Summary and Conclusion

In this thesis, the performance of two specific estimation schemes, namely Kalman and Wiener filters, has been studied for analog uncoded transmission over fading channels. With random channels, the instantaneous mean estimation error, or simply distortion, is also random. The instantaneous random distortion has then a steady-state probability distribution function, given that certain conditions on the system model are satisfied. The estimation quality was afterwards analyzed with respect to distortion outage probability and average MSE measures. One major distinction between the current work and related previous works on estimation quality analysis for analog transmission over fading channels, is considering time-correlated sources, especially those following a state-space model. We also introduced modified and new diversity schemes for analog communications in order to improve the outage probability performance. For estimation quality analysis, two main approaches were followed, namely developing bounds on the outage probability, as well as asymptotic high SNR analysis.

As finding a complete analytical solution for the outage probability turned out to be tedious even for the simplest cases, upper and lower bounds were developed for the outage probability for certain SISO, SIMO, and MIMO settings in Chapters 2, 3, and 4. The significance of the bounds is their potential to be used for system design. Given a system model and a threshold, one can incorporate the bounds to find the minimum required SNR and diversity order. In addition to the bounds, asymptotic high SNR analysis was also performed on both outage and average MSE measures throughout this work. The application of high SNR analysis is its ability to provide a simple means for estimation quality assessment. The high SNR analysis of the SISO case shows that the outage probability decreases with the inverse of the SNR in the high SNR regime. This performance was improved by incorporating diversity achieving techniques in Chapters 3 and 4 for time-correlated signals and in Chapter 6 for band-limited signals. The performance improvement is reflected in an increased diversity order of the outage probability function.

The asymptotic high SNR performance also lead to characterizing the asymptotic (relative) average power gain, i.e. the coding gain. The coding gain was derived as a function of source parameters in Chapters 2 and 4, which together with the diversity order, completely describes the high SNR performance of the outage probability function.

The analysis starts in Chapter 2 by considering transmission of scalar sources over scalar i.i.d. channels. This is a rather simple system model, which allows for simpler analysis, which in turn, could (and did) lead to better understanding of the fundamental behavior of the estimation error outage function for estimation over fading channels. In Chapter 3, we studied the effect of having several independent versions of the transmitted signal at the receiver, e.g. via having multiple antennas, in order to achieve a diversity gain for the outage probability. The effect of parameter mismatch on the diversity results was also considered afterwards. The analysis in that chapter shows that for process and channel noise variance, using mismatched parameters does not effect the diversity results. The loss is rather in the coding gain and for small mismatches, i.e. good parameter estimation techniques, the SNR loss due to mismatch is not significant. This important result shows that the standard Kalman filtering may still be useful, even though all the simplifying assumptions made in Chapter 2 may not always hold.

The encouraging diversity results of Chapter 3 lead to designing a joint space-time coding and Kalman filtering scheme in Chapter 4, in order to obtain transmit diversity gains in addition to the receive diversity gain in Chapter 3. We incorporated complex orthogonal space-time codes in series with the Kalman filter. The decoding procedure was modified accordingly to allow for inclusion of codes with any rates. The analysis in this chapter shows that it is possible for the Kalman filter to achieve the maximum allowable diversity order for distortion outage probability. This interesting result is despite the fact that such scheme is not necessarily optimal, as it does not exploit the memory in the source at the transmitter side. The outage probability, as well as the coding gain, were also bounded. The results of this chapter show the potential for the practical application of the proposed scheme in order to provide high quality delay-free reliable analog communication schemes.

Further on, we studied estimation over fading channels for large dimensional systems. We found an expression for the average MSE of the Kalman filter for estimation of samples transmitted over a large MIMO fading channel. The results of this chapter can be used in order to get an overall performance metric for large sensor networks performing simple analog communication. The high SNR behavior indicates also the importance of additional antennas at the receiver. If the number of antennas at the receiver is the same as the number of transmitters, or alternatively transmit antennas, the average MSE decreases with the inverse of the square root of the SNR. The asymptotic behavior changes significantly, even only one additional antenna is applied. In that case, i.e. when the number of receive antennas is bigger than the number of transmit antennas, the average MSE decreases with the inverse of the SNR in the high SNR regime.

Finally, we introduced a new scheme to allow for high quality, delay-free Wiener filtering of bandlimited signals over MIMO block fading channels. We considered transmission

of a K dimensional signal over N parallel fading channels, and proved a diversity order of $N - K + 1$ for the distortion outage probability. If, N is selected large enough, high quality estimates can be obtained in practice. There is also no limit on the block length, which makes the results of this chapter even more interesting for practical applications.

We believe that the estimation and diversity schemes presented and analyzed in this work have the potential to be used for providing simple but high performance delay-free communication schemes. While some aspects and some of the pros and cons of these schemes were explored in this work, there is still plenty of unanswered questions which require further investigation. Some of these open problems are reviewed in the next section.

7.2 Further Thoughts and Future Work

The results of this thesis show that it is possible to perform low-complexity, high quality, and at the same time delay-free estimation over fading channels. The suggested estimation scheme is based on analog communication, rather than digital. There are no claims about the optimality of analog communication in the settings which are considered in this work. However, a comparison with relevant zero- or low-delay digital communication schemes can be a very useful and informative future work. In addition, these results are based on certain simplifications, most significant of which are known system parameters, known channel at the receiver, and i.i.d. channels. While we studied certain deviations from these assumptions, a full analysis for unknown parameters, considering channel estimation error and the effect of channels with memory still remains. Such analysis also allows for better understanding the fundamental behavior of the random estimation error variance/covariance matrix in more realistic settings.

In order to evaluate the consequences of unknown system parameters, one may take the analysis in this thesis in Chapter 3 as a starting point and extend it to other system parameters and other system models. A natural extension to the analysis in Chapter 3 is analyzing the effect of process and channel noise mismatch for higher dimensions. In addition, studying the effect of modeling error/mismatch for the state transition matrix should be considered. The effect of modeling error/ estimation error for the channel is a slightly different problem, as the channel is time-varying and needs to be estimated for each time instant.

To analyze the effect of channel memory, one approach is to start from Chapter 2 and try to develop new theory for both the non-i.i.d. fading channel model and the i.i.d. block fading model. We believe that one key issue is selecting a good model for the channel memory, in order to be able to develop relevant theory by modifying the fundamental equations of the estimation error variance distribution in Chapter 2. A first order Markov model can be a good starting point. Other than that, a block fading model with a large block size assumption may also lead to interesting results, as the analysis becomes simpler in that case. One may then assume that the Kalman filter converges to a steady-state value for that block, which is random but then almost constant over each block. Obtaining

the distribution of the random estimation error variance in that case, provides a tool for evaluating the outage performance of finite block sizes, as their performance is expected to be limited by the two extreme cases of block sizes of one and infinity.

Another possible direction for future research is practical schemes which can employ a trade-off between source-to-channel bandwidth ratio and the estimation quality either in terms of outage probability or the average MSE. A similar concept and a simple example was provided in Chapter 6. There and for the specifically presented scheme, a trade off exists between the oversampling ratio N/K, and the diversity order given by $N - K + 1$. One could see that decreasing source-to-channel bandwidth ratio (increasing N/K) results in increasing diversity order and vice versa. While theoretical limits already exist for analog communication, and practical schemes for digital communication are already available, practical implementations, especially for the Kalman filter, is a an interesting problem to consider.

In this thesis, the channels are assumed known at the receiver, but not at the transmitter. It is interesting to study the potentials that full or partial channel information at the transmitter can create for improving the estimation quality. As one of the application areas of this can be within the wireless control framework, it is expected that a feedback channel exists between the control center and the point of observation, which might be co-located with the control plant. One could then use this channel in order make the feed-forward channel information available at the transmitter. One could then study how this extra information can help improve the performance.

Finally, more general source models can also be considered. The first-order ARGM model allowed for simplification in deriving new theory, helping understand some of the fundamental behavior of the instantaneous random estimation error. It is interesting to see how more general signal models will affect the performance. It was already shown in this thesis that certain source parameters appear in the coding gain of the outage probability in the high SNR regime. It remains to see how a higher AR model would affect the coding gain.

Bibliography

[1] Ian F Akyildiz and Mehmet Can Vuran. *Wireless sensor networks*, volume 4. John Wiley & Sons, 2010.

[2] Siavash Alamouti. A simple transmit diversity technique for wireless communications. *IEEE Journal on Selected Areas in Communications*, 16(8):1451–1458, 1998.

[3] Animashree Anandkumar, Lang Tong, and Ananthram Swami. Detection of Gauss–Markov random fields with nearest-neighbor dependency. *IEEE Transactions on Information Theory*, 55(2):816–827, 2009.

[4] Panos Antsaklis and John Baillieul. Special issue on technology of networked control systems. *Proceedings of the IEEE*, 95(1):5–8, 2007.

[5] Israfil Bahceci and Amir K. Khandani. Energy-efficient estimation of correlated data in wireless sensor networks. In *40th Annual Conference on Information Sciences and Systems*, pages 973–978. IEEE, 2006.

[6] Stephan Bäro, Gerhard Bauch, and Axel Hansmann. Improved codes for space-time trellis coded modulation. 2000.

[7] Richard H Battin. *Astronautical guidance*. McGraw-Hill, 1964.

[8] Dimitri P Bertsekas and Ian B Rhodes. Recursive state estimation for a set-membership description of uncertainty. *IEEE Transactions on Automatic Control*, 16(2):117–128, 1971.

[9] Subir Biswas, Raymond Tatchikou, and Francois Dion. Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety. *IEEE Communications Magazine*, 44(1):74–82, 2006.

[10] Philippe Bougerol. Kalman filtering with random coefficients and contractions. *SIAM Journal on Control and Optimization*, 31(4):942–959, 1993.

[11] Shraga Bross, Amos Lapidoth, Stephan Tinguely, et al. Broadcasting correlated Gaussians. *IEEE Transactions on Information Theory*, 56(7):3057–3068, 2010.

[12] Giuseppe Caire and Krishna Narayanan. On the distortion SNR exponent of hybrid digital–analog space–time coding. *IEEE Transactions on Information Theory*, 53(8):2867–2878, 2007.

[13] Jinhui Chen and Dirk T.M. Slock. Orthogonal space-time block codes for analog channel feedback. In *IEEE International Symposium on Information Theory (ISIT)*, pages 1473–1477. IEEE, 2008.

[14] Giacomo Como, Fabio Fagnani, and Sandro Zampieri. Anytime reliable transmission of real-valued information through digital noisy channels. *SIAM Journal on Control and Optimization*, 48(6):3903–3924, 2010.

[15] Thomas M. Cover and Joy A. Thomas. *Elements of information theory*. John Wiley & Sons, 2012.

[16] Helge Coward and Tor Ramstad. Quantizer optimization in hybrid digital-analog transmission of analog source signals. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 5, pages 2637–2640. IEEE, 2000.

[17] Shuguang Cui, Jin-Jun Xiao, Andrea J. Goldsmith, Zhi-Quan Luo, and H. Vincent Poor. Estimation diversity and energy efficiency in distributed sensing. *IEEE Transactions on Signal Processing*, 55(9):4683–4695, 2007.

[18] Herbert Aron David and Haikady Navada Nagaraja. *Order statistics*. Wiley Online Library, 1970.

[19] Subhrakanti Dey, Alex S. Leong, and Jamie S. Evans. Kalman filtering with faded measurements. *Automatica*, 45(10):2223–2233, 2009.

[20] Dahir H. Dini, Danilo P. Mandic, and Simon J. Julier. A widely linear complex unscented Kalman filter. *IEEE Signal Processing Letters*, 18(11):623–626, 2011.

[21] Stark C. Draper and Anant Sahai. Universal anytime coding. In *5th International Symposium on Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks and Workshops (WiOpt)*, pages 1–5. IEEE, 2007.

[22] Garry A. Einicke and Langford B. White. Robust extended Kalman filtering. *IEEE Transactions on Signal Processing*, 47(9):2596–2599, 1999.

[23] Stanley C. Eisenstat and Ilse C.F. Ipsen. Three absolute perturbation bounds for matrix eigenvalues imply relative bounds. *SIAM Journal on Matrix Analysis and Applications*, 20(1):149–158, 1998.

[24] Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research: Oceans (1978–2012)*, 99(C5):10143–10162, 1994.

[25] Samuel L. Fagin. Recursive linear regression theory, optimal filter theory, and error analysis of optimal systems. In *IEEE International Convention Record*, volume 12, pages 216–245, 1964.

[26] Pål Anders Floor, Anna N. Kim, Niklas Wernersson, Tor Ramstad, Mikael Skoglund, and Ilangko Balasingham. Zero-delay joint source-channel coding for a bivariate Gaussian on a Gaussian MAC. *IEEE Transactions on Communications*, 60(10):3091–3102, 2012.

[27] Yang Gao and Ertem Tuncel. New hybrid digital/analog schemes for transmission of a Gaussian source over a Gaussian channel. *IEEE Transactions on Information Theory*, 56(12):6014–6019, 2010.

[28] William A. Gardner. Cyclic Wiener filtering: theory and method. *IEEE Transactions on Communications*, 41(1):151–163, 1993.

[29] Michael Gastpar. Uncoded transmission is exactly optimal for a simple Gaussian sensor network. *IEEE Transactions on Information Theory*, 54(11):5247–5251, 2008.

[30] Michael Gastpar, Bixio Rimoldi, and Martin Vetterli. To code, or not to code: Lossy source–channel communication revisited. *IEEE Transactions on Information Theory*, 49(5):1147–1158, 2003.

[31] Michael Gastpar and Martin Vetterli. On the capacity of large Gaussian relay networks. *IEEE Transactions on Information Theory*, 51(3):765–779, 2005.

[32] Sandeep P. Ghael, Akbar M. Sayeed, and Richard G. Baraniuk. Improved wavelet denoising via empirical Wiener filtering. In *Optical Science, Engineering and Instrumentation'97*, pages 389–399. International Society for Optics and Photonics, 1997.

[33] Andrea Goldsmith. *Wireless communications*. Cambridge University Press, 2005.

[34] Neil J. Gordon, David J. Salmond, and Adrian F.M. Smith. Novel approach to nonlinear/non-gaussian bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, volume 140, pages 107–113. IET, 1993.

[35] Robert E. Griffin and Andrew P. Sage. Sensitivity analysis of discrete filtering and smoothing algorithms. *AIAA Journal*, 7(10):1890–1897, 1969.

[36] Deniz Gunduz and Elza Erkip. Source and channel coding for cooperative relaying. *IEEE Transactions on Information Theory*, 53(10):3454–3475, 2007.

[37] Deniz Gunduz and Elza Erkip. Joint source–channel codes for MIMO block-fading channels. *IEEE Transactions on Information Theory*, 54(1):116–134, 2008.

[38] Fredrik Gustafsson and Gustaf Hendeby. Some relations between extended and unscented Kalman filters. *IEEE Transactions on Signal Processing*, 60(2):545–555, 2012.

[39] H. Heffes. The effect of erroneous models on the Kalman filter response. *IEEE Transactions on Automatic Control*, 11(3):541–543, 1966.

[40] Onésimo Hernández-Lerma and Jean B. Lasserre. *Markov chains and invariant probabilities*. Springer, 2003.

[41] Peter Hoeher, Stefan Kaiser, and Patrick Robertson. Two-dimensional pilot-symbol-aided channel estimation by Wiener filtering. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 3, pages 1845–1848. IEEE, 1997.

[42] Tim Holliday, Andrea J. Goldsmith, and H. Vincent Poor. Joint source and channel coding for MIMO systems: Is it better to be robust or quick? *IEEE Transactions on Information Theory*, 54(4):1393–1405, 2008.

[43] Roger A. Horn and Charles R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[44] Peter L. Houtekamer and Herschel L. Mitchell. Data assimilation using an ensemble Kalman filter technique. *Monthly Weather Review*, 126(3):796–811, 1998.

[45] Hamid Jafarkhani. A quasi-orthogonal space-time block code. *IEEE Transactions on Communications*, 49(1):1–4, 2001.

[46] Hamid Jafarkhani and Nambi Seshadri. Super-orthogonal space-time trellis codes. *IEEE Transactions on Information Theory*, 49(4):937–950, 2003.

[47] Andrew H. Jazwinski. *Stochastic processes and filtering theory*. Academic Press, 1970.

[48] Simon J. Julier and Jeffrey K. Uhlmann. New extension of the Kalman filter to nonlinear systems. In *AeroSense'97*, pages 182–193. International Society for Optics and Photonics, 1997.

[49] Thomas Kailath, Ali H. Sayed, and Babak Hassibi. *Linear estimation*, volume 1. Prentice Hall New Jersey, 2000.

[50] Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of Fluids Engineering*, 82(1):35–45, 1960.

[51] Soummya Kar, Bruno Sinopoli, and José M.F. Moura. Kalman filtering with intermittent observations: Weak convergence to a stationary distribution. *IEEE Transactions on Automatic Control*, 57(2):405–420, 2012.

[52] Steven M. Kay. *Fundamentals of statistical signal processing*. PTR Prentice-Hall, Englewood Cliffs, NJ, 1993.

[53] Gökhan Korkmaz, Eylem Ekici, Füsun Özgüner, and Ümit Özgüner. Urban multi-hop broadcast protocol for inter-vehicle communication systems. In *Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*, pages 76–85. ACM, 2004.

[54] Xing-Hong Kuang, Shao Hui-He, and Feng Rui. A new distributed localization scheme for wireless sensor networks. *Acta Automatica Sinica*, 34(3):344–348, 2008.

[55] Harold J. Kushner. Approximations to optimal nonlinear filters. *IEEE Transactions on Automatic Control*, 12(5):546–556, 1967.

[56] Harold J. Kushner. Dynamical equations for optimal nonlinear filtering. *Journal of Differential Equations*, 3(2):179–190, 1967.

[57] J. Nicholas Laneman, Emin Martinian, Gregory W. Wornell, and John G. Apostolopoulos. Source-channel diversity for parallel channels. *IEEE Transactions on Information Theory*, 51(10):3518–3539, 2005.

[58] Amos Lapidoth and Stephan Tinguely. Sending a bivariate Gaussian over a Gaussian MAC. *IEEE Transactions on Information Theory*, 56(6):2714–2752, 2010.

[59] Amos Lapidoth and Stephan Tinguely. Sending a bivariate Gaussian source over a Gaussian MAC with feedback. *IEEE Transactions on Information Theory*, 56(4):1852–1864, 2010.

[60] Franck L. Lewis. Wireless sensor networks. *Smart environments: technologies, protocols, and applications*, pages 11–46, 2004.

[61] Xie Lihua. Control over communication networks: Trend and challenges in integrating control theory and information theory. In *30th Chinese Control Conference (CCC)*, pages 35–39. IEEE, 2011.

[62] D. Magill. Optimal adaptive estimation of sampled stochastic processes. *IEEE Transactions on Automatic Control*, 10(4):434–439, 1965.

[63] Peter S. Maybeck. *Stochastic models, estimation, and control*, volume 3. Academic Press, 1982.

[64] Joerg Meyer and Klaus Uwe Simmer. Multi-channel speech enhancement in a car environment using Wiener filtering and spectral subtraction. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 2, pages 1167–1170. IEEE, 1997.

[65] Udar Mittal and Nam Phamdo. Hybrid digital-analog (HDA) joint source-channel codes for broadcasting and robust communications. *IEEE Transactions on Information Theory*, 48(5):1082–1102, 2002.

[66] Maryam Moayedi, Yung Kuan Foo, and Yeng Chai Soh. Adaptive Kalman filtering in networked systems with random sensor delays, multiple packet dropouts and missing measurements. *IEEE Transactions on Signal Processing*, January 2010.

[67] T. Nishimura. On the a priori information in sequential estimation problems. *IEEE Transactions on Automatic Control*, 11(2):197–204, 1966.

[68] Athanasios Papoulis and S. Unnikrishna Pillai. *Probability, random variables, and stochastic processes*. McGraw-Hill Europe, fourth edition, January 2002.

[69] Reza Parseh, Santiago Sanchez Acevedo, Kimmo Kansanen, Marta Molinas, and Tor A Ramstad. Real-time compression of measurements in distribution grids. In *IEEE Third International Conference on Smart Grid Communications (SmartGridComm)*, pages 223–228, 2012.

[70] Reza Parseh and Kimmo Kansanen. On estimation error outage for scalar Gauss–Markov signals sent over fading channels. *IEEE Transactions on Signal Processing*, 62(23):6225–6234, Dec. 2014.

[71] Reza Parseh and Kimmo Kansanen. Diversity effects in Kalman filtering over Rayleigh fading channels. *IEEE Transactions on Signal Processing*, 63(23):6329–6342, Dec. 2015.

[72] Reza Parseh, Kimmo Kansanen, and Dirk Slock. Distortion outage analysis for joint space-time coding and Kalman filtering. *Submitted to IEEE Transactions on Signal Processing*, 2016.

[73] Reza Parseh, Dirk Slock, and Kimmo Kansanen. Mean estimation MSE for Kalman filtering of large dimensional sources sent over fading channels. In *IEEE 15th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 499 – 503, June 2014.

[74] Reza Parseh, Dirk Slock, and Kimmo Kansanen. Oversampling diversity for uncoded transmission of bandlimited sources over parallel fading channels. In *IEEE International Conference on Communications (ICC)*, pages 4653 – 4658, 2015.

[75] Li Peng and A. Guillen i Fabregas. Distortion outage probability in MIMO block-fading channels. In *IEEE International Symposium on Information Theory (ISIT)*, pages 2223–2227, 2010.

[76] Kaare Brandt Petersen and Michael Syskind Pedersen. The matrix cookbook. *Technical University of Denmark*, pages 7–15, 2008.

[77] Gregory J. Pottie. Wireless sensor networks. In *Information Theory Workshop (ITW)*, pages 139–140. IEEE, 1998.

[78] Charles F. Price. An analysis of the divergence problem in the Kalman filter. *IEEE Transactions on Automatic Control*, 13(6):699–702, 1968.

[79] Daniel E. Quevedo, Anders Ahlén, Alex S. Leong, and Subhrakanti Dey. On Kalman filtering over fading wireless channels with controlled transmission powers. *Automatica*, 48(7):1306–1316, 2012.

[80] Cauligi S Raghavendra, Krishna M Sivalingam, and Taieb Znati. *Wireless sensor networks*. Springer Science & Business Media, 2004.

[81] Herbert E. Rauch, C.T. Striebel, and F. Tung. Maximum likelihood estimates of linear dynamic systems. *AIAA journal*, 3(8):1445–1450, 1965.

[82] Eduardo Rohr, Damian Marelli, and Minyue Fu. Kalman filtering for a class of degenerate systems with intermittent observations. In *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 2422–2427. IEEE, 2011.

[83] Anant Sahai and Sanjoy Mitter. The necessity and sufficiency of anytime capacity for control over a noisy communication link–Part II: Vector systems. *arXiv.org*, 2006.

[84] Anant Sahai and Sanjoy Mitter. The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link–Part I: Scalar systems. *IEEE Transactions on Information Theory*, 52(8):3369–3395, 2006.

[85] Ali H. Sayed. A framework for state-space estimation with uncertain models. *IEEE Transactions on Automatic Control*, 46(7):998–1013, 2001.

[86] Shayle R Searle. Matrix algebra useful for statistics. *New York*, 1982, 1982.

[87] Uri Shaked and Yahali Theodor. H_∞-optimal estimation: A tutorial. In *Proceedings of the 31st IEEE Conference on Decision and Control (CDC)*, pages 2278–2286. IEEE, 1992.

[88] Shlomo Shamai, Sergio Verdú, and Ram Zamir. Systematic lossy source/channel coding. *IEEE Transactions on Information Theory*, 44(2):564–579, 1998.

[89] Ling Shi and Lihua Xie. Optimal sensor power scheduling for state estimation of Gauss–Markov systems over a packet-dropping network. *IEEE Transactions on Signal Processing*, 60(5):2701–2705, 2012.

[90] Ling Shi, Lihua Xie, and Richard M. Murray. Kalman filtering over a packet-delaying network: A probabilistic approach. *Automatica*, 45(9):2134–2140, 2009.

[91] Amirpasha Shirazinia, Subhrakanti Dey, Domenico Ciuonzo, and Pierluigi Salvo Rossi. Massive mimo for decentralized estimation over coherent multiple access channels. In *IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 241–245, June 2015.

[92] Peng-Lang Shui. Image denoising algorithm via doubly local Wiener filtering with directional windows in wavelet domain. *IEEE Signal Processing Letters*, 12(10):681–684, 2005.

[93] Harold W. Sorenson. Least-squares estimation: from Gauss to Kalman. *IEEE Spectrum*, 7(7):63–68, 1970.

[94] Harold W. Sorenson and Allen R. Stubberud. Non-linear filtering by approximation of the a posteriori density. *International Journal of Control*, 8(1):33–51, 1968.

[95] John A. Stankovic. Wireless Sensor Networks. *Computer*, 41(10):92–95, October 2008.

[96] A. Subramanian and A. H. Sayed. Joint rate and power control algorithms for wireless networks. *IEEE Transactions on Signal Processing*, 53(11):4204–4214, November 2005.

[97] Vikas Taliwal, Daniel Jiang, Heiko Mangold, Chi Chen, and Raja Sengupta. Empirical determination of channel characteristics for DSRC vehicle-to-vehicle communication. In *Proceedings of the 1st ACM international workshop on Vehicular ad hoc networks*, pages 88–88. ACM, 2004.

[98] Vahid Tarokh, Hamid Jafarkhani, and A. Robert Calderbank. Space-time block codes from orthogonal designs. *IEEE Transactions on Information Theory*, 45(5):1456–1467, 1999.

[99] Vahid Tarokh, Nambi Seshadri, and A. Robert Calderbank. Space-time codes for high data rate wireless communication: Performance criterion and code construction. *IEEE Transactions on Information Theory*, 44(2):744–765, 1998.

[100] Olav Tirkkonen and Ari Hottinen. Improved MIMO performance with non-orthogonal space-time block codes. In *Global Telecommunications Conference (GLOBECOM)*, volume 2, pages 1122–1126, 2001.

[101] Marc Torrent-Moreno, Jens Mittag, Paolo Santi, and Hannes Hartenstein. Vehicle-to-vehicle communication: fair transmit power control for safety-critical information. *IEEE Transactions on Vehicular Technology*, 58(7):3684–3703, 2009.

[102] David Tse and Pramod Viswanath. *Fundamentals of wireless communication*. Cambridge University Press, 2005.

[103] Antonia M. Tulino and Sergio Verdú. *Random matrix theory and wireless communications*, volume 1. Now Publishers Inc., 2004.

[104] Ali Vakili and Babak Hassibi. A Stieltjes transform approach for analyzing the RLS adaptive filter. In *46th Annual Allerton Conference on Communication, Control, and Computing*, pages 432–437. IEEE, 2008.

[105] Ali Vakili and Babak Hassibi. A Stieltjes transform approach for studying the steady-state behavior of random Lyapunov and Riccati recursions. In *IEEE Conference on Decision and Control (CDC)*, pages 453–458. IEEE, 2008.

[106] Rudolph Van Der Merwe and Eric A. Wan. The square-root unscented Kalman filter for state and parameter-estimation. In *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, volume 6, pages 3461–3464. IEEE, 2001.

[107] Luxmiram Vijayandran, Maïté Brandt-Pearce, Kimmo Kansanen, and Torbjörn Ekman. Efficient state estimation with energy harvesting and fairness control us-

ing stochastic optimization. In *Global Telecommunications Conference (GLOBECOM)*, pages 1–6. IEEE, 2011.

[108] Eric A. Wan and Rudolph Van Der Merwe. The unscented Kalman filter for nonlinear estimation. In *The IEEE Adaptive Systems for Signal Processing, Communications, and Control Symposium (AS-SPCC)*, pages 153–158. IEEE, 2000.

[109] Chih-Hong Wang, Alex S. Leong, and Subhrakanti Dey. Distortion outage minimization and diversity order analysis for coherent multiaccess. *IEEE Transactions on Signal Processing*, 59(12):6144–6159, 2011.

[110] Yuan Wang and Lei Guo. On stability of random Riccati equations. *Science in China Series E: Technological Sciences*, 42(2):136–148, 1999.

[111] Zhengdao Wang and Georgios B. Giannakis. A simple and general parameterization quantifying performance in fading channels. *IEEE Transactions on Communications*, 51(8):1389–1398, 2003.

[112] David Williams. *Probability with martingales*. Cambridge University Press, 1991.

[113] Alan S. Willsky. A survey of design methods for failure detection in dynamic systems. *Automatica*, 12(6):601–611, 1976.

[114] William M. Wonham. On a matrix Riccati equation of stochastic control. *SIAM Journal on Control*, 6(4):681–697, 1968.

[115] Li Xie and Lihua Xie. Stability of a random Riccati equation with Markovian binary switching. *IEEE Transactions on Automatic Control*, 53(7):1759–1764, 2008.

[116] Lihua Xie, Yeng Chai Soh, and Carlos E. de Souza. Robust Kalman filtering for uncertain discrete-time systems. *IEEE Transactions on Automatic Control*, 39(6):1310–1314, 1994.

[117] Xue Yang, Jie Liu, N.F. Vaidya, and Feng Zhao. A vehicle-to-vehicle communication protocol for cooperative collision warning. In *The First Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services (MOBIQUITOUS)*, pages 114–123. IEEE, 2004.

[118] Keyou You, Minyue Fu, and Lihua Xie. Mean square stability for Kalman filtering with Markovian packet losses. *Automatica*, 47(12):2647–2657, 2011.

[119] M. Karbalaye Zadeh, Reza Parseh, Marta Molinas, and Kimmo Kansanen. Bifurcation in PWM converter-based systems with wireless communication-based current controller. In *4th IEEE/PES Innovative Smart Grid Technologies Europe (ISGT EUROPE)*, pages 1–5, 2013.

[120] M. Karbalaye Zadeh, Reza Parseh, Marta Molinas, and Kimmo Kansanen. Modeling and simulation of wireless communication based robust controller for multiconverter systems. In *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 738–743, 2013.

[121] Alberto Zanella, Marco Chiani, and Moe Z. Win. On the marginal distribution of the eigenvalues of wishart matrices. *IEEE Transactions on Communications*, 57(4):1050–1060, 2009.

[122] Wei Zhang, Michael S. Branicky, and Stephen M. Phillips. Stability of networked control systems. *IEEE Control Systems*, 21(1):84–99, 2001.

[123] Lizhong Zheng and David N. C. Tse. Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels. *IEEE Transactions on Information Theory*, 49(5):1073–1096, 2003.

[124] Zhi Zhong, Luo Da-Yong, Liu Shao-Qiang, Fan Xiao-Ping, and Qu Zhi-Hua. An adaptive localization approach for wireless sensor networks based on Gauss–Markov mobility model. *Acta Automatica Sinica*, 36(11):1557–1568, 2010.

[125] Hao Zhu, I. D. Schizas, and G. B. Giannakis. Power-efficient dimensionality reduction for distributed channel-aware Kalman tracking using WSNs. *IEEE Transactions on Signal Processing*, 57(8):3193–3207, August 2009.

Appendix A

Upper and lower bounds for κ

We begin to rewrite $f_M(M)$ in the following manner for simplicity

$$f_M(M) = \begin{cases} \frac{\kappa\lambda}{M^2}\text{exp}(\frac{-\lambda}{M}) & M \leqslant \sigma_u^2 \\ g(M) & M > \sigma_u^2. \end{cases} \tag{A.1}$$

We have that $f_M(M)$ is a pdf, therefore $\int_0^{M_{\text{max}}} f_M(M)\,\mathrm{d}M = 1$. As a result, we have that

$$\begin{aligned} 1 &= \int_0^{M_{\text{max}}} f_M(M)\,\mathrm{d}M \\ &= \int_0^{\sigma_u^2} \frac{\kappa\lambda}{M^2}\text{exp}(\frac{-\lambda}{M}) + \int_{\sigma_u^2}^{M_{\text{max}}} g(M)\,\mathrm{d}M \qquad \text{(A.2)} \\ &= \kappa\text{exp}(\frac{-\lambda}{M})\Big|_0^{\sigma_u^2} + \int_{\sigma_u^2}^{M_{\text{max}}} g(M)\,\mathrm{d}M \\ &= \kappa\text{exp}(\frac{-\lambda}{\sigma_u^2}) + \int_{\sigma_u^2}^{M_{\text{max}}} g(M)\,\mathrm{d}M, \qquad \text{(A.3)} \end{aligned}$$

which gives

$$\int_{\sigma_u^2}^{M_{\text{max}}} g(M)\,\mathrm{d}M = 1 - \kappa\text{exp}(\frac{-\lambda}{\sigma_u^2}). \tag{A.4}$$

Now we take

$$\sigma_u^2 < m < M_{\text{max}}.$$

Then we have

$$(\rho^2+1)\sigma_u^2 < \rho^2 m + \sigma_u^2 < \rho^2 M_{\text{max}} + \sigma_u^2$$

and

$$\exp\left(\frac{\lambda}{\rho^2 M_{\max}+\sigma_u^2}\right) < \exp\left(\frac{\lambda}{\rho^2 m+\sigma_u^2}\right) < \exp\left(\frac{\lambda}{(\rho^2+1)\sigma_u^2}\right). \tag{A.5}$$

Now we have that

$$\int_{\sigma_u^2}^{M_{\max}} \exp\left(\frac{\lambda}{\rho^2 m+\sigma_u^2}\right) g(m)\,\mathrm{d}m > \int_{\sigma_u^2}^{M_{\max}} \exp\left(\frac{\lambda}{\rho^2 M_{\max}+\sigma_u^2}\right) g(m)\,\mathrm{d}m \tag{A.6}$$

$$\int_{\sigma_u^2}^{M_{\max}} \exp\left(\frac{\lambda}{\rho^2 m+\sigma_u^2}\right) g(m)\,\mathrm{d}m < \int_{\sigma_u^2}^{M_{\max}} \exp\left(\frac{\lambda}{(\rho^2+1)\sigma_u^2}\right) g(m)\,\mathrm{d}m. \tag{A.7}$$

Next, if we use the definition of $f_M(M)$ for $M \leqslant \sigma_u^2$, we obtain the following

$$\begin{aligned} f(M) &= \frac{\lambda}{M^2}\exp(\frac{-\lambda}{M}) \int_0^{M_{\max}} \exp(\frac{\lambda}{\rho^2 m+\sigma_u^2}) f_M(m)\,\mathrm{d}m \\ &= \frac{\kappa\lambda}{M^2}\exp(\frac{-\lambda}{M}) \qquad (A.8) \\ &= \frac{\lambda}{M^2}\exp(\frac{-\lambda}{M}) \left(\int_0^{\sigma_u^2} \exp(\frac{\lambda}{\rho^2 m+\sigma_u^2}) \left(\frac{\kappa\lambda}{m^2}\right) \exp(\frac{-\lambda}{m})\,\mathrm{d}m \right. \\ &\quad \left. + \int_{\sigma_u^2}^{M_{\max}} \exp(\frac{\lambda}{\rho^2 m+\sigma_u^2}) g(m)\,\mathrm{d}m \right), \qquad (A.9) \end{aligned}$$

from which, we may deduce that

$$\kappa = \kappa \int_0^{\sigma_u^2} \exp(\frac{\lambda}{\rho^2 m+\sigma_u^2}) \left(\frac{\lambda}{m^2}\right) \exp(\frac{-\lambda}{m})\,\mathrm{d}m + \int_{\sigma_u^2}^{M_{\max}} \exp(\frac{\lambda}{\rho^2 m+\sigma_u^2}) g(m)\,\mathrm{d}m.$$

And then we obtain

$$\kappa = \frac{\int_{\sigma_u^2}^{M_{\max}} \exp(\frac{\lambda}{\rho^2 m+\sigma_u^2}) g(m)\,\mathrm{d}m}{1 - \int_0^{\sigma_u^2} \exp(\frac{\lambda}{\rho^2 m+\sigma_u^2}) \left(\frac{\lambda}{m^2}\right) \exp(\frac{-\lambda}{m})\,\mathrm{d}m}. \tag{A.10}$$

Now by letting

$$a_\kappa = 1 - \int_0^{\sigma_u^2} \exp(\frac{\lambda}{\rho^2 m+\sigma_u^2}) \left(\frac{\lambda}{m^2}\right) \exp(\frac{-\lambda}{m})\,\mathrm{d}m \tag{A.11}$$

and combining (A.6) into (A.10) while using (A.4), we get

$$\begin{aligned} \kappa a_\kappa &> \int_{\sigma_u^2}^{M_{\max}} \exp(\frac{\lambda}{\rho^2 M_{\max}+\sigma_u^2}) g(m)\,\mathrm{d}m \\ &> \exp(\frac{\lambda}{\rho^2 M_{\max}+\sigma_u^2}) \int_{\sigma_u^2}^{M_{\max}} g(m)\,\mathrm{d}m \qquad (A.12) \\ &> \exp(\frac{\lambda}{\rho^2 M_{\max}+\sigma_u^2})(1 - \kappa\exp(\frac{-\lambda}{\sigma_u^2})), \qquad (A.13) \end{aligned}$$

which leads to

$$\kappa > \frac{1}{(a_\kappa \exp(\frac{-\lambda}{\rho^2 M_{\max}+\sigma_u^2}) + \exp(\frac{-\lambda}{\sigma_u^2}))}. \tag{A.14}$$

So, we finally get

$$\kappa_l = \frac{1}{\left(a_\kappa \exp\left(\frac{-\lambda}{\rho^2 M_{\max}+\sigma_u^2}\right) + \exp(\frac{-\lambda}{\sigma_u^2})\right)}. \tag{A.15}$$

The same procedure also holds for κ_u by integrating (A.7) into (2.26) while using (A.4). We then get

$$\kappa_u = \frac{1}{\left(a_\kappa \exp\left(\frac{-\lambda}{\sigma_u^2(1+\rho^2)}\right) + \exp(-\frac{\lambda}{\sigma_u^2})\right)}. \tag{A.16}$$

Appendix B

Proof of Lemma 8

The proof is based on converting the definition of each κ_i into a linear inequality on other κ_j's. Then we will have d similar inequalities. We also obtain another equality by taking into account that the pdf should integrate to one over its domain. This last equality also involves κ_i's and we then get $d+1$ linear simple inequalities which result in upper and lower bounds for κ_i.

We begin the proof by first noting that from (3.7), $f_M(M)$ can be written as

$$f_M(M) = \begin{cases} f(M), & M \leqslant \sigma_u^2 \\ g(M), & M > \sigma_u^2, \end{cases} \tag{B.1}$$

where

$$f(M) = \sum_{i=0}^{d-1} \kappa_i (-1)^i \binom{d-1}{i} \frac{\lambda^d}{(d-1)!} \frac{\exp(\frac{-\lambda}{M})}{M^{d+1-i}}. \tag{B.2}$$

Given the definition for κ_i from (3.8), we have that

$$\begin{aligned} \kappa_i &= \int_0^{M_{\max}} \frac{1}{(\rho^2 m + \sigma_u^2)^i} \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right) f_M(m)\, \mathrm{d}m \\ &\overset{(a)}{=} \int_0^{\sigma_u^2} \frac{1}{(\rho^2 m + \sigma_u^2)^i} \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right) f(m)\, \mathrm{d}m \\ &\quad + \int_{\sigma_u^2}^{M_{\max}} \frac{1}{(\rho^2 m + \sigma_u^2)^i} \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right) g(m)\, \mathrm{d}m, \end{aligned} \tag{B.3}$$

where in (a) we partition the domain of integration and use (B.1). Inserting (B.2) into (B.3) results in

$$\kappa_i = \sum_{j=0}^{d-1} \kappa_j w_{ij} + \int_{\sigma_u^2}^{M_{\max}} \frac{1}{(\rho^2 m + \sigma_u^2)^i} \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right) g(m)\, \mathrm{d}m, \tag{B.4}$$

where

$$w_{ij} = (-1)^j \binom{d-1}{j} \frac{\lambda^d}{(d-1)!} \int_0^{\sigma_u^2} \frac{1}{(\rho^2 m + \sigma_u^2)^i} \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right) \frac{\exp(\frac{-\lambda}{m})}{m^{d+1-i}} \, \mathrm{d}m, \tag{B.5}$$

when $0 \leqslant i, j \leqslant d-1$. Next, we note that we should have

$$\begin{aligned} 1 &= \int_0^{M_{\max}} f_M(m) \, \mathrm{d}m = \int_0^{\sigma_u^2} f(m) \, \mathrm{d}m + \int_{\sigma_u^2}^{M_{\max}} g(m) \, \mathrm{d}m \\ &\overset{(a)}{=} \int_0^{\sigma_u^2} \sum_{j=0}^{d-1} \kappa_j (-1)^j \binom{d-1}{j} \frac{\lambda^d}{(d-1)!} \frac{\exp(\frac{-\lambda}{m})}{m^{d+1-i}} \, \mathrm{d}m + \int_{\sigma_u^2}^{M_{\max}} g(m) \, \mathrm{d}m, \end{aligned} \tag{B.6}$$

where (a) is obtained by using the definition of $f(M)$ from (B.2) for $0 \leqslant M \leqslant \sigma_u^2$. Now we can rewrite (B.6) as

$$1 = \sum_{j=0}^{d-1} \kappa_j c_j + \int_{\sigma_u^2}^{M_{\max}} g(m) \, \mathrm{d}m, \tag{B.7}$$

where

$$c_j = (-1)^j \binom{d-1}{j} \frac{\lambda^d}{(d-1)!} \int_0^{\sigma_u^2} \frac{\exp(\frac{-\lambda}{m})}{m^{d+1-i}} \, \mathrm{d}m. \tag{B.8}$$

At this stage, it is possible to get d equalities involving κ_i's from (B.4) in addition to another equality involving κ_i's from (B.7). At the next stage of proof we convert the aforementioned equalities (also involving $g(M)$ which is at this stage unknown) to a set of linear inequalities.

Now note that due to the fact that $\sigma_u^2 \leqslant m \leqslant M_{\max}$, we have $(1+\rho^2)\sigma_u^2 \leqslant \rho^2 m + \sigma_u^2 \leqslant \rho^2 M_{\max} + \sigma_u^2$. As a result, it is easy to show that

$$\frac{\exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right)}{(\rho^2 m + \sigma_u^2)^i} \leqslant \frac{\exp\left(\frac{\lambda}{(1+\rho^2)\sigma_u^2}\right)}{((1+\rho^2)\sigma_u^2)^i} \tag{B.9}$$

and

$$\frac{\exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right)}{(\rho^2 m + \sigma_u^2)^i} \geqslant \frac{\exp\left(\frac{\lambda}{\rho^2 M_{\max} + \sigma_u^2}\right)}{(\rho^2 M_{\max} + \sigma_u^2)^i}, \tag{B.10}$$

which in turn results in

$$\int_{\sigma_u^2}^{M_{\max}} \frac{\exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right)}{(\rho^2 m + \sigma_u^2)^i} g(m) \, \mathrm{d}m \leqslant \frac{\exp\left(\frac{\lambda}{(1+\rho^2)\sigma_u^2}\right)}{((1+\rho^2)\sigma_u^2)^i} \int_{\sigma_u^2}^{M_{\max}} g(m) \, \mathrm{d}m \tag{B.11}$$

and

$$\int_{\sigma_u^2}^{M_{\max}} \frac{\exp\left(\frac{\lambda}{\rho^2 m+\sigma_u^2}\right)}{(\rho^2 m+\sigma_u^2)^i} g(m)\,\mathrm{d}m \geqslant \frac{\exp\left(\frac{\lambda}{\rho^2 M_{\max}+\sigma_u^2}\right)}{(\rho^2 M_{\max}+\sigma_u^2)^i} \int_{\sigma_u^2}^{M_{\max}} g(m)\,\mathrm{d}m. \tag{B.12}$$

If we now define $Z \geqslant 0$ as

$$Z = \int_{\sigma_u^2}^{M_{\max}} g(m)\,\mathrm{d}m \tag{B.13}$$

(Z will also play the role of the auxiliary random variable in determining bounds for κ_i's), we may rewrite (B.4) as a set of inequalities over κ_i given by

$$\kappa_i \geqslant \sum_{j=0}^{d-1} \kappa_j w_{ij} + \frac{\exp\left(\frac{\lambda}{\rho^2 M_{\max}+\sigma_u^2}\right)}{(\rho^2 M_{\max}+\sigma_u^2)^i} \int_{\sigma_u^2}^{M_{\max}} g(m)\,\mathrm{d}m \tag{B.14}$$

and

$$\kappa_i \leqslant \sum_{j=0}^{d-1} \kappa_j w_{ij} + \frac{\exp\left(\frac{\lambda}{(1+\rho^2)\sigma_u^2}\right)}{((1+\rho^2)\sigma_u^2)^i} \int_{\sigma_u^2}^{M_{\max}} g(m)\,\mathrm{d}m. \tag{B.15}$$

Now we show that the obtained sets of inequalities may be converted into a set of linear matrix inequalities as stated in Lemma 8. We begin by defining the constants b_i^u and b_i^l as

$$b_i^u = \frac{\exp\left(\frac{\lambda}{\rho^2 M_{\max}+\sigma_u^2}\right)}{(\rho^2 M_{\max}+\sigma_u^2)^i} \tag{B.16}$$

$$b_i^l = \frac{\exp\left(\frac{\lambda}{(1+\rho^2)\sigma_u^2}\right)}{((1+\rho^2)\sigma_u^2)^i}. \tag{B.17}$$

We then rewrite the inequalities in (B.14) and (B.15) as

$$\kappa_i \geqslant \sum_{j=0}^{d-1} \kappa_j w_{ij} + b_i^u Z \tag{B.18}$$

$$\kappa_i \leqslant \sum_{j=0}^{d-1} \kappa_j w_{ij} + b_i^l Z \tag{B.19}$$

Also, (B.7) may be rewritten as

$$1 = \sum_{j=0}^{d-1} \kappa_j c_j + Z. \tag{B.20}$$

Now we define the vectors, $\mathbf{b}^l = [b_0^l, b_1^l, \cdots, b_{d-1}^l]^T$ and $\mathbf{b}^u = [b_0^u, b_1^u, \cdots, b_{d-1}^u]^T$ and the augmented vectors $\boldsymbol{\kappa} = [\kappa_0, \kappa_1, \cdots, \kappa_{d-1}, Z]^T$ and $\mathbf{c} = [c_0, c_1, \cdots, c_{d-1}, 1]^T$. Using

the aforementioned vectors, it is easy to write the i-th inequality in (B.18) and (B.19) respectively as

$$(\mathbf{w}_i^u)^T \boldsymbol{\kappa} \leqslant \mathbb{1}_i^T \boldsymbol{\kappa}, \quad i = 0, 1, \cdots, d-1 \tag{B.21}$$
$$(\mathbf{w}_i^l)^T \boldsymbol{\kappa} \geqslant \mathbb{1}_i^T \boldsymbol{\kappa}, \quad i = 0, 1, \cdots, d-1, \tag{B.22}$$

where

$$\mathbf{w}_i^u = [w_{i,1}, w_{i,2}, \cdots, w_{i,d-1}, b_i^u]^T \tag{B.23}$$
$$\mathbf{w}_i^l = [w_{i,1}, w_{i,2}, \cdots, w_{i,d-1}, b_i^l]^T \tag{B.24}$$

and $\mathbb{1}_i$ is a column vector of all zeros except at position i, which there, it is equal to one. Also the equality in (B.20) may be rewritten as $\mathbf{c}^T \boldsymbol{\kappa} = 1$. Furthermore, we use this equality as two inequalities to incorporate it into the previous inequalities when necessary, i.e.

$$\mathbf{c}^T \boldsymbol{\kappa} \leqslant 1 \tag{B.25}$$
$$\mathbf{c}^T \boldsymbol{\kappa} \geqslant 1 \tag{B.26}$$

Then we may combine (B.21) and (B.25) to get

$$\begin{aligned} (\mathbf{w}_i^u)^T \boldsymbol{\kappa} - \mathbb{1}_i^T \boldsymbol{\kappa} \leqslant 0, \quad i = 0, 1, \cdots, d-1 \\ \mathbf{c}^T \boldsymbol{\kappa} \leqslant 1, \end{aligned} \tag{B.27}$$

and combine (B.22) and (B.26) to get

$$\begin{aligned} (\mathbf{w}_i^l)^T \boldsymbol{\kappa} - \mathbb{1}_i^T \boldsymbol{\kappa} \geqslant 0, \quad i = 0, 1, \cdots, d-1 \\ \mathbf{c}^T \boldsymbol{\kappa} \geqslant 1. \end{aligned} \tag{B.28}$$

Finally, we rewrite the set of inequalities in (B.27) and (B.28) as matrix inequalities given by

$$[[W - I_{d\times d}|\mathbf{b}^u]^T|\mathbf{c}]^T \boldsymbol{\kappa} \leqslant \mathbb{1}_{d+1} \tag{B.29}$$
$$[[W - I_{d\times d}|\mathbf{b}^l]^T|\mathbf{c}]^T \boldsymbol{\kappa} \geqslant \mathbb{1}_{d+1}, \tag{B.30}$$

with $W = [w_{ij}]_{d\times d}$, $0 \leqslant i, j \leqslant d-1$, which is the same form as presented in Lemma 8 if we we take

$$W^u = [[W - I_{d\times d}|\mathbf{b}^u]^T|\mathbf{c}]^T \tag{B.31}$$
$$W^l = [[W - I_{d\times d}|\mathbf{b}^l]^T|\mathbf{c}]^T. \tag{B.32}$$

Note that in (B.29) and (B.30) for vectors $\mathbf{z}$ and $\mathbf{w}$ of arbitrary size N, $\mathbf{z} \leqslant \mathbf{w}$ holds if and only if $z_i \leqslant w_i$ for $i = 0, 1, \cdots, N-1$.

Solving (B.29) and (B.30) results in $\boldsymbol{\kappa} \leqslant \boldsymbol{\kappa}^u$ and $\boldsymbol{\kappa} \geqslant \boldsymbol{\kappa}^l$ respectively. The elements of $\boldsymbol{\kappa}^u$ and $\boldsymbol{\kappa}^l$ then constitute the upper and lower bound for κ_i. One simple and immediate way to solve the inequalities is by by calculating

$$\boldsymbol{\kappa}^u = (W^u)^{-1}\mathbb{1}_{d+1} \tag{B.33}$$

$$\boldsymbol{\kappa}^l = (W^l)^{-1}\mathbb{1}_{d+1}, \tag{B.34}$$

which is the formulation used in Lemma 8. If any negative elements are encountered, one could then simply replace them with zero and still have valid upper or lower bound. Note that we have made no claims on the performance of the bounds obtained using this method, so any valid bound for κ_i's will suffice to prove our claim. However, simulation in Sec. 3.2.2 show that the bounds obtained by using the bounds on κ_i in fact perform well under the proposed system model and selected parameters.

Appendix C

Proof of Theorem 3

We begin the proof by first showing that the cumulative distribution function of IEV, i.e. $F_M(M)$ approaches the step function when $\lambda \to 0$. We have that

$$\begin{aligned} F_M(M_{\text{th}}) &= 1 - P_{\text{out}}(M_{\text{th}}) \\ &= \exp(\frac{-\lambda}{M_{\text{th}}}) \sum_{i=0}^{d-1} \frac{\kappa_i (-1)^i \lambda^i}{i!} \sum_{k=0}^{d-i-1} \frac{(\frac{\lambda}{M_{\text{th}}})^k}{k!} \end{aligned} \tag{C.1}$$

Now for any $M_{\text{th}} > 0$, we have

$$\begin{aligned} \lim_{\lambda \to 0} F_M(M_{\text{th}}) &= \lim_{\lambda \to 0} \exp(\frac{-\lambda}{M_{\text{th}}}) \lim_{\lambda \to 0} (\sum_{i=0}^{d-1} \frac{\kappa_i (-1)^i \lambda^i}{i!} \sum_{k=0}^{d-i-1} \frac{(\frac{\lambda}{M_{\text{th}}})^k}{k!}) \\ &= \lim_{\lambda \to 0} \kappa_0, \end{aligned} \tag{C.2}$$

where in the last step we used the fact that $\kappa_i < \exp(\frac{\lambda}{\sigma_u^2})$ over different values of λ and all the terms in the second limit with $i \geqslant 1$ vanish when $\lambda \to 0$.

Now we have that,

$$\begin{aligned}
\lim_{\lambda\to 0}\kappa_0 &= \lim_{\lambda\to 0}\int_0^{M_{\max}} \exp(\frac{\lambda}{\rho^2 m+\sigma_u^2}) f_M(m)\,\mathrm{d}m \\
&= \lim_{\lambda\to 0}\int_0^{M_{\max}} \sum_{l=0}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\rho^2 m+\sigma_u^2)^l} f_M(m)\,\mathrm{d}m \\
&= \lim_{\lambda\to 0}\int_0^{M_{\max}} f_M(m)\,\mathrm{d}m + \lim_{\lambda\to 0}\int_0^{M_{\max}} \sum_{l=1}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\rho^2 m+\sigma_u^2)^l} f_M(m)\,\mathrm{d}m \\
&= \lim_{\lambda\to 0} 1 + \lim_{\lambda\to 0}\int_0^{M_{\max}} \sum_{l=1}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\rho^2 m+\sigma_u^2)^l} f_M(m)\,\mathrm{d}m \\
&= 1 + \lim_{\lambda\to 0}\int_0^{M_{\max}} \sum_{l=1}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\rho^2 m+\sigma_u^2)^l} f_M(m)\,\mathrm{d}m,
\end{aligned} \tag{C.3}$$

but for $m \geqslant 0$, we have that $\frac{1}{\rho^2 m+\sigma_u^2} \leqslant \frac{1}{\sigma_u^2}$. As a result

$$\lim_{\lambda\to 0}\int_0^{M_{\max}} \sum_{l=1}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\rho^2 m+\sigma_u^2)^l} f_M(m)\,\mathrm{d}m \leqslant \lim_{\lambda\to 0}\sum_{l=1}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\sigma_u^2)^l}\int_0^{M_{\max}} f_M(m)\,\mathrm{d}m \tag{C.4}$$

and thus

$$\lim_{\lambda\to 0}\int_0^{M_{\max}} \sum_{l=1}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\rho^2 m+\sigma_u^2)^l} f_M(m)\,\mathrm{d}m \leqslant \lim_{\lambda\to 0}\sum_{l=1}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\sigma_u^2)^l}, \tag{C.5}$$

but

$$\lim_{\lambda\to 0}\sum_{l=1}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\sigma_u^2)^l} = \lim_{\lambda\to 0}(\exp(\frac{\lambda}{\sigma_u^2}) - 1) = 0. \tag{C.6}$$

Therefore we obtain that,

$$\lim_{\lambda\to 0}\int_0^{M_{\max}} \sum_{l=1}^{\infty}\frac{\lambda^l}{l!}\frac{1}{(\rho^2 m+\sigma_u^2)^l} f_M(m)\,\mathrm{d}m = 0, \tag{C.7}$$

and finally $\lim_{\lambda\to 0}\kappa_0 = 1$. That shows that $F_M(M)$ approaches the unit step function when $\lambda \to 0$ and as a result $f_M(m)$ approaches the Dirac's delta function when $\lambda \to 0$.

With this assumption we have

$$\begin{aligned}
\kappa_i &= \int_0^{M_{\max}} \left(\frac{1}{\rho^2 m + \sigma_u^2}\right)^i \exp\left(\frac{\lambda}{\rho^2 m + \sigma_u^2}\right) f_M(m)\,\mathrm{d}m \\
&= \int_0^{M_{\max}} \frac{1}{(\rho^2 m + \sigma_u^2)^i} \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} \frac{1}{(\rho^2 m + \sigma_u^2)^l} f_M(m)\,\mathrm{d}m \\
&= \int_0^{M_{\max}} \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} \frac{1}{(\rho^2 m + \sigma_u^2)^{l+i}} f_M(m)\,\mathrm{d}m \\
&= \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} \int_0^{M_{\max}} \frac{1}{(\rho^2 m + \sigma_u^2)^{l+i}} f_M(m)\,\mathrm{d}m \\
&= \sum_{l=0}^{\infty} \frac{\lambda^l}{l!} \int_0^{M_{\max}} \frac{1}{(\sigma_u^2)^{l+i}}.
\end{aligned} \tag{C.8}$$

for $\lambda \to 0$ and Theorem 3 is proved.

Appendix D

Proof of Theorem 4

We first rewrite the IEV outage function $P_{\text{out}}(M_{\text{th}})$ as

$$P_{\text{out}}(M_{\text{th}}) = 1 - \exp(\frac{\lambda}{M_{\text{th}}}) \sum_{i=0}^{d-1} \frac{\kappa_i (-1)^i Q_i(M_{\text{th}})}{i!}, \tag{D.1}$$

where

$$\begin{aligned} Q_i(M_{\text{th}}) &= \lambda^i \sum_{k=0}^{d-1-i} \frac{(\lambda/M_{\text{th}})^k}{k!} \\ &= \lambda^i + \frac{\lambda^{i+1}}{M_{\text{th}}} + \frac{\lambda^{i+2}}{2! M_{\text{th}}^2} + \cdots + \frac{\lambda^{d-1}}{(d-1-i)! M_{\text{th}}^{d-1-i}}. \end{aligned} \tag{D.2}$$

Now, we may rewrite $P_{\text{out}}(M_{\text{th}})$ as follows

$$P_{\text{out}}(M_{\text{th}}) = 1 - \exp(\frac{\lambda}{M_{\text{th}}}) \sum_{i=0}^{d-1} b_i \lambda^i, \tag{D.3}$$

where

$$\begin{aligned} &b_0 = \kappa_0, \quad b_1 = \frac{\kappa_0}{M_{\text{th}}} - \kappa_1, \quad b_2 = \frac{\kappa_0}{2! M_{\text{th}}^2} - \frac{\kappa_1}{M_{\text{th}}} + \frac{\kappa_2}{2!}, \\ &\cdots, \quad b_{d-1} = \frac{(-1)^{d-1} \kappa_{d-1}}{(d-1)! M_{\text{th}}^{d-1}} + \cdots + \frac{(-1)^{d-1} \kappa_0}{(d-1)!} \end{aligned} \tag{D.4}$$

and in general

$$b_i = \sum_{j=0}^{i} \frac{(-1)^j \kappa_j}{j!(i-j)! M_{\text{th}}^{i-j}}, \quad (i = 0, 1, \cdots, d-1). \tag{D.5}$$

Now using Theorem 3, we may claim

$$\kappa_i = \sum_{l=0}^{d-1} \frac{\lambda^l}{(\sigma_u^2)^{i+l} l!} + \mathcal{O}(\lambda^d), \tag{D.6}$$

which may alternatively be written as

$$\kappa_i = \sum_{l=0}^{d-1} \frac{\lambda^l}{(\sigma_u^2)^{i+l} l!} + o(\lambda^{d-1}), \tag{D.7}$$

in order to be compatible with the basic definition of diversity order.

With that at hand, we rewrite $P_{\text{out}}(M_{\text{th}})$ as follows

$$\begin{aligned}
P_{\text{out}}(M_{\text{th}}) &= 1 - \exp(\frac{-\lambda}{M_{\text{th}}})\big[(1 + \lambda/\sigma_u^2 + \cdots + o(\lambda^{d-1})) \\
&+ (\frac{1}{M_{\text{th}}} - \frac{1}{\sigma_u^2})\lambda(1 + \lambda/\sigma_u^2 + \cdots + o(\lambda^{d-1})) \\
&+ (\frac{1}{2!M_{\text{th}}^2} - \frac{1}{\sigma_u^2 M_{\text{th}}} + \frac{1}{2!(\sigma_u^2)^2})\lambda^2(1 + \lambda/\sigma_u^2 + \cdots + o(\lambda^{d-1})) \\
&+ \cdots \\
&+ \big(\sum_{m=0}^{i} \frac{(-1)^m}{(i-m)!M_{\text{th}}^{i-m} m!(\sigma_u^2)^m}\big)\lambda^i(1 + \lambda/\sigma_u^2 + \cdots + o(\lambda^{d-1})) \\
&+ \cdots \\
&+ \big(\sum_{m=0}^{d-1} \frac{(-1)^m}{(d-1-m)!M_{\text{th}}^{d-1-m} m!(\sigma_u^2)^m}\big)\lambda^{d-1}(1 + \lambda/\sigma_u^2 + \cdots + o(\lambda^{d-1})) \\
&+ o(\lambda^{d-1}),
\end{aligned}$$

which may be summarized for legibility as

$$P_{\text{out}}(M_{\text{th}}) = 1 - \exp(\frac{-\lambda}{M_{\text{th}}})\big(\sum_{i=0}^{d-1} c_i\lambda^i(1 + \lambda/\sigma_u^2 + \cdots + o(\lambda^{d-1}))\big) \tag{D.8}$$

with

$$c_i = \sum_{m=0}^{i} \frac{(-1)^m}{(i-m)!m!M_{\text{th}}^{i-m}(\sigma_u^2)^m}. \tag{D.9}$$

Now, we can simplify $P_{\text{out}}(M_{\text{th}})$ and write it as

$$P_{\text{out}}(M_{\text{th}}) = 1 - \exp(\frac{-\lambda}{M_{\text{th}}})\big(\sum_{i=0}^{d-1} d_i\lambda^i + o(\lambda^{d-1})\big). \tag{D.10}$$

If we could show that $d_i = 1/(i!M_{\text{th}}^i)$, then we can apply Theorem 3 and deduce that $P_{\text{out}}(M_{\text{th}})$ has diversity order d. This is done in the next step.

By simplifying (D.8) and isolating powers of λ, it is straightforward to show that $d_i = \sum_{j=0}^{i} \frac{c_j}{(i-j)!(\sigma_u^2)^{i-j}}$. Then, we may simplify c_i and get

$$\begin{aligned} c_i &= \sum_{m=0}^{i} \frac{(-1)^m}{(i-m)!m!M_{\text{th}}^{i-m}(\sigma_u^2)^m} \\ &= \frac{1}{i!M_{\text{th}}^i} \sum_{m=0}^{i} \frac{(-1)^m i!}{(i-m)!m!} (M_{\text{th}}/\sigma_u^2)^m \\ &= \frac{1}{i!M_{\text{th}}^i} \sum_{m=0}^{i} \frac{(-M_{\text{th}}/\sigma_u^2)^m (1)^{i-m} i!}{(i-m)!m!} \\ &= \frac{1}{i!M_{\text{th}}^i} (1 - M_{\text{th}}/\sigma_u^2)^i \\ &= \frac{1}{i!M_{\text{th}}^i (\sigma_u^2)^i} (\sigma_u^2 - M_{\text{th}})^i. \end{aligned} \tag{D.11}$$

Finally, by inserting the values of c_i into the equation for d_i, we obtain

$$\begin{aligned} d_i &= \sum_{j=0}^{i} \frac{c_j}{(i-j)!(\sigma_u^2)^{i-j}} \\ &= \sum_{j=0}^{i} \frac{(\sigma_u^2 - M_{\text{th}})^{i-j}}{(i-j)!j!(\sigma_u^2)^{i-j} M_{\text{th}}^j (\sigma_u^2)^j} \\ &= \frac{1}{i!(\sigma_u^2)^i} \sum_{j=0}^{i} \frac{i!(\sigma_u^2 - M_{\text{th}})^{i-j}}{(i-j)!j!M_{\text{th}}^j} \\ &= \frac{1}{i!(\sigma_u^2)^i} \sum_{j=0}^{i} (\sigma_u^2/M_{\text{th}} - 1)^j 1^{i-j} \mathcal{C}(i,j) \\ &= \frac{1}{i!(\sigma_u^2)^i} ((\sigma_u^2/M_{\text{th}} - 1 + 1)^i \\ &= \frac{(\sigma_u^2)^i}{i!(\sigma_u^2)^i M_{\text{th}}^i} \\ &= \frac{1}{i!M_{\text{th}}^i} \end{aligned} \tag{D.12}$$

and the proof is complete.

Appendix E

Orthogonality for the Equivalent Real Channel Matrix

In this section, we denote the elements of the arbitrary matrix W as $W[i,j]$, where i denotes the row position and j the column position. The i-th row of W from columns j_1 to j_2 is denoted by $W[i, j_1 : j_2]$ and the whole row is denoted by $W[i,:]$. Similar rules hold for the j-th column.

We first define $T_l = T[2(l-1)K + 1 : 2lK, 1 : 2K]$, $l = 1, 2, \ldots, N_c$, i.e.

$$T = [T_1^T | T_2^T | \cdots | T_{N_c}^T]^T. \tag{E.1}$$

Note that each block T_l then maps the variable $\mathbf{x}_r$ to the l–th column of the matrix X_r, i.e. $X_r[:, l] = T_l \mathbf{x}_r$. In order to show that H_{eq} can in fact be orthogonalized, we first see that

$$\begin{aligned}
\mathbf{x}_r^T H_{\text{eq}}^T H_{\text{eq}} \mathbf{x}_r &= \mathbf{x}_r^T T^T (I_{N_c} \otimes \tilde{H})^T (I_{N_c} \otimes \tilde{H}) T \mathbf{x}_r \\
&= \mathbf{x}_r^T T^T \left(I_{N_c} \otimes \left(\tilde{H}^T \tilde{H} \right) \right) T \mathbf{x}_r \\
&= \sum_{l=1}^{N_c} \mathbf{x}_r^T T_l^T \left(\tilde{H}^T \tilde{H} \right) T_l \mathbf{x}_r
\end{aligned} \tag{E.2}$$

The matrix $\tilde{H}^T \tilde{H}$ is of dimension $2K \times 2K$, whereas the matrix $\tilde{H}$ is of dimension $2 \times 2K$. For that reason, the rank of $\tilde{H}^T \tilde{H}$ is equal to the rank of $\tilde{H}$, which for i.i.d. Rayleigh fading is equal to 2 with probability one. In addition, the eigenvalues of $\tilde{H}^T \tilde{H}$ are the same as those of $\tilde{H} \tilde{H}^T$, with additional $2K - 2$ zeros. It is easy to see that

$$\tilde{H} \tilde{H}^T = \|\mathbf{h}_k\|^2 I_2. \tag{E.3}$$

Therefore, we can write the eigenvalue decomposition of $\tilde{H}^T \tilde{H}$ as

$$\begin{aligned}
\tilde{H}^T \tilde{H} &= Q \, \text{diag}\{0, 0, \ldots, 0, \|\mathbf{h}_k\|^2, \|\mathbf{h}_k\|^2\} Q^T \\
&= \|\mathbf{h}_k\|^2 Q Z Q^T,
\end{aligned} \tag{E.4}$$

with $QQ^T = I_{2K}$ and $Z = \text{diag}\{0, 0, \ldots, 0, 1, 1\}$. Note that the position of ones in Z is only for the simplification of the proof, as the equivalent matrix QZQ^T will be the same for any positioning. Inserting (E.4) into (E.2), we obtain that

$$\begin{aligned}
\mathbf{x}_r^T H_{\text{eq}}^T H_{\text{eq}} \mathbf{x}_r &= \|\mathbf{h}_k\|^2 \sum_{l=1}^{N_c} \mathbf{x}_r^T T_l^T Q Z Q^T T \mathbf{x}_r \\
&= \|\mathbf{h}_k\|^2 \sum_{l=1}^{N_c} \mathbf{x}_r^T (Q^T T_l)^T Z (Q^T T_l) \mathbf{x}_r \\
&= \|\mathbf{h}_k\|^2 \sum_{l=1}^{N_c} \mathbf{x}_r^T T_l^{'T} Z T_l^{'} \mathbf{x}_r \\
&\overset{(a)}{=} \|\mathbf{h}_k\|^2 \sum_{l=1}^{N_c} \mathbf{x}_r^T (T_l^{'T} Z)(Z T_l^{'}) \mathbf{x}_r
\end{aligned} \tag{E.5}$$

where $T_l^{'}, l = 1, 2, \ldots, N_c$ may be assumed to be the building blocks of a matrix $T^{'}$(same as for T), which maps $\mathbf{x}_r$ to the matrix $Q^T X_r$ in the same way as T maps $\mathbf{x}_r$ to X_r. In addition, (a) holds because $Z^2 = Z$. The positioning of the ones in Z is such that it is only the last two columns of $T_l^{'T}$ which remain non-zero after multiplication by Z. The last two columns of $T_l^{'T}$ correspond to the last two rows of $T_l^{'}$ (also visible in the structure of $ZT_l^{'}$ in (E.5)). In order to provide better intuition into (E.5), we see that (E.5) can be rewritten as

$$\begin{aligned}
\mathbf{x}_r^T H_{\text{eq}}^T H_{\text{eq}} \mathbf{x}_r &= \|\mathbf{h}_k\|^2 \sum_{l=1}^{N_c} \mathbf{x}_r^T T_l^{'T} (Z_{2K-1} + Z_{2K}) T_l^{'} \mathbf{x}_r \\
&= \|\mathbf{h}_k\|^2 \sum_{l=1}^{N_c} \mathbf{x}_r^T T_l^{'T} Z_{2K-1} T_l^{'} \mathbf{x}_r \\
&+ \|\mathbf{h}_k\|^2 \sum_{l=1}^{N_c} \mathbf{x}_r^T T_l^{'T} Z_{2K} T_l^{'} \mathbf{x}_r,
\end{aligned} \tag{E.6}$$

where Z_{2K-1} and Z_{2K} are all-zero matrices, except for $Z_{2K-1}[2K-1, 2K-1] = Z_{2K}[2K, 2K] = 1$. We can then rewrite (E.6) as

$$\begin{aligned}
\mathbf{x}_r^T H_{\text{eq}}^T H_{\text{eq}} \mathbf{x}_r &= \|\mathbf{h}_k\|^2 \sum_{l=1}^{N_c} \mathbf{x}_r^T (T_l^{'T} Z_{2K-1})(Z_{2K-1} T_l^{'}) \mathbf{x}_r \\
&+ \|\mathbf{h}_k\|^2 \sum_{l=1}^{N_c} \mathbf{x}_r^T (T_l^{'T} Z_{2K})(Z_{2K} T_l^{'}) \mathbf{x}_r.
\end{aligned} \tag{E.7}$$

Given that, the result of (E.7) and equivalently (E.5), is the sum of the squares of the last

two rows (row $2K-1$ and $2K$) of $X_r^{'} = Q^T X_r$. This can be written as

$$\mathbf{x}_r^T H_{\text{eq}}^T H_{\text{eq}} \mathbf{x}_r = \|\mathbf{h}_k\|^2 \left(\sum_{l=1}^{N_c} X_r^{'2}[2K-1, l] + X_r^{'2}[2K, l] \right). \tag{E.8}$$

It is easy to see that $X_r^{'} = Q_r X_r$, where Q_r is a matrix of dimension $2 \times 2K$, which consists of the last two rows of the matrix Q^T (transpose of the last two columns of Q). From the definition of the eigenvalue decomposition for $\tilde{H}^T\tilde{H}$ and given that $\tilde{H}^T\tilde{H}$ only has two non-zero eigenvalues (with the corresponding eigenvectors as the last two columns of Q, equal to Q_r^T), it is readily established that

$$\tilde{H}^T\tilde{H} = \|\mathbf{h}_k\|^2 Q_r^T Q_r, \tag{E.9}$$

which means that $\tilde{H} = \pm\|\mathbf{h}_k\| Q_r$, i.e. Q_r is a normalized version of $\tilde{H}$. This is a property we will use later on in the proof.

In order to be able to calculate the value of the sum in (E.8), we may convert the matrices X_r and Q_r to complex equivalent matrices. In that case, we are able to finally use the orthogonality of the complex space-time code in order to prove the orthogonality of the equivalent channel. This step is necessary, as the code's orthogonality is best described in the domain of the complex numbers. We perform this procedure as follows. For X_r, take all the consecutive odd and even real rows, add the even row multiplied by $\boldsymbol{i}$ to the previous odd row and then remove the even rows. The resulting matrix, of dimension $K \times N_c$ is equal to X, the original space-time code. For the matrix Q_r, perform the same procedure with odd and even columns, and call the resulting matrix, of dimension $2 \times K$ as Q_c. We further take the first row of the matrix Q_c to be equal to $\mathbf{q}_1^T$ and the second row as $\mathbf{q}_2^T$. Note that $\mathbf{q}_1$ and $\mathbf{q}_2$ are still orthogonal to one another and we still have that $\mathbf{q}_1^H\mathbf{q}_1 = \mathbf{q}_2^H\mathbf{q}_2 = 1$. With these definitions, the sum in (E.8) can be rewritten as

$$\begin{aligned}
&\mathbf{x}_r^T H_{\text{eq}}^T H_{\text{eq}} \mathbf{x}_r \\
&= \|\mathbf{h}_k\|^2 \left(\operatorname{Re}(\mathbf{q}_1^H X) \operatorname{Re}(\mathbf{q}_1^H X)^T + \operatorname{Re}(\mathbf{q}_2^H X) \operatorname{Re}(\mathbf{q}_2^H X)^T \right)
\end{aligned}$$

or equivalently

$$\begin{aligned}
&\mathbf{x}_r^T H_{\text{eq}}^T H_{\text{eq}} \mathbf{x}_r \\
&= \|\mathbf{h}_k\|^2 \left(\operatorname{Re}(\mathbf{q}_1^H X) \operatorname{Re}(X^H \mathbf{q}_1) + \operatorname{Re}(\mathbf{q}_2^H X) \operatorname{Re}(X^H \mathbf{q}_2) \right).
\end{aligned} \tag{E.10}$$

The following equalities are also easy to verify

$$\begin{aligned}
\|\mathbf{x}\|^2 &= \mathbf{q}_2^H X X^H \mathbf{q}_2 \\
&= \mathbf{q}_1^H X X^H \mathbf{q}_1 \\
&= \operatorname{Re}(\mathbf{q}_1^H X) \operatorname{Re}(X^H \mathbf{q}_1) - \operatorname{Im}(\mathbf{q}_1^H X) \operatorname{Im}(X^H \mathbf{q}_1) \\
&= \operatorname{Re}(\mathbf{q}_1^H X) \operatorname{Re}(X^H \mathbf{q}_1) + \operatorname{Im}(\mathbf{q}_1^H X) \operatorname{Im}(\mathbf{q}_1^H X)^T.
\end{aligned} \tag{E.11}$$

If we convert the matrix $\tilde{H}$ to a complex equivalent matrix in the same manner as for Q_r and obtain $\tilde{H}_c$, we see that due to the structure of $\tilde{H}$ (see (4.3)), we should have

$$\mathrm{Re}(\tilde{H}_c[1,:]) = \mathrm{Im}(\tilde{H}_c[2,:])$$

and

$$\mathrm{Re}(\tilde{H}_c[2,:]) = -\mathrm{Im}(\tilde{H}_c[1,:])$$

As Q_r is only a scaled version of $\tilde{H}$, then Q_c is only a scaled (by a real number) version of $\tilde{H}_c$. Due to that, we should also have for Q_c that

$$\mathrm{Re}(Q_c[1,:]) = \mathrm{Im}(Q_c[2,:]) \tag{E.12}$$

and

$$\mathrm{Re}(Q_c[2,:]) = -\mathrm{Im}(Q_c[1,:]). \tag{E.13}$$

It is then straightforward to show using (E.12) and (E.13) to that

$$\mathrm{Im}(\mathbf{q}_1^H X) = \mathrm{Re}(\mathbf{q}_2^H X). \tag{E.14}$$

That can be understood better by considering that if $Q_r[1,:] = [q_{1,1}, q_{1,2}, \ldots, q_{1,2K-1}, q_{1,2K}]^T$ and $Q_r[2,:] = [q_{2,1}, q_{2,2}, \ldots, q_{2,2K-1}, q_{2,2K}]^T$, then we have from (E.12) and (E.13) that e.g. $q_{1,1} = q_{2,1}$ and $q_{1,2} = -q_{2,1}$ and so on. This can be considered within the multiplication operations of $\mathbf{q}_1^H X$ and $\mathbf{q}_2^H X$, in order to produce the result in (E.14).

Inserting (E.14) into (E.11) results in

$$\|\mathbf{x}\|^2 = \mathrm{Re}(\mathbf{q}_1^H X)\,\mathrm{Re}(X^H \mathbf{q}_1) + \mathrm{Re}(\mathbf{q}_2^H X)\,\mathrm{Re}(X^H \mathbf{q}_2), \tag{E.15}$$

which by comparing to (E.10) confirms that

$$\begin{aligned} \mathbf{x}_r^T H_{\text{eq}}^T H_{\text{eq}} \mathbf{x}_r &= \|\mathbf{h}_k\|^2 \|\mathbf{x}\|^2 \\ &= \|\mathbf{h}_k\|^2 \mathbf{x}_r^T \mathbf{x}_r \\ &= \mathbf{x}_r^T \|\mathbf{h}_k\|^2 \mathbf{x}_r \end{aligned} \tag{E.16}$$

for all $\mathbf{x}_r$, which effectively results in

$$H_{\text{eq}}^T H_{\text{eq}} = \|\mathbf{h}_k\|^2 I_{2K} \tag{E.17}$$

Appendix F

Bounds for the Instantaneous Distortion

From (4.11), we have that $M(n) = \frac{1}{\gamma_n}\left(\frac{1}{\gamma_n}P^{-1}(n) + I\right)^{-1}$. If we denote the eigenvalues $M(n)$ by $\zeta_l(M(n))$, $l = 1, 2, \ldots, K$, and the eigenvalues of $P(n)$ by $\zeta_l(P(n))$, $l = 1, 2, \ldots, K$, then we have that

$$\begin{aligned} d(n) &= \frac{1}{K}\text{tr}(M(n)) \\ &= \frac{1}{K}\sum_{l=1}^{K}\zeta_l(M(n)) \\ &= \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\gamma_n}(\frac{1}{\gamma_n}\frac{1}{\zeta_l(P(n))} + 1)^{-1} \\ &= \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\frac{1}{\zeta_l(P(n))} + \gamma_n} \end{aligned} \tag{F.1}$$

As $P(n)$ is a covariance matrix, it is (semi)-positive definite. Therefore, we have that $\zeta_l(P(n)) \geqslant 0$, $\forall l$. If we denote the ordered eigenvalues of C_u by $\zeta_l(C_u)$, $\forall l$ ($\zeta_1(C_u) \geqslant \zeta_2(C_u) \geqslant \ldots \geqslant \zeta_K(C_u)$) and also order $\zeta_l(P(n))$ such that $\zeta_1(P(n)) \geqslant \zeta_2(P(n)) \geqslant \ldots \geqslant \zeta_K(P(n))$, we know from Weyl's inequalities [43, Ch. 3] that

$$\zeta_l(P(n)) \geqslant \zeta_l(C_u). \tag{F.2}$$

This is due to the fact that $P(n) = AM(n-1)A^T + C_u$ and $AM(n-1)A^T$ is a positive definite matrix (because $M(n-1)$ is positive-definite). Now combining (F.1) and (F.2),

we obtain that

$$d(n) = \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\dfrac{1}{\zeta_l(P(n))}+\gamma_n} \geqslant \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\dfrac{1}{\zeta_l(C_u)}+\gamma_n}.$$

It is easy to show that the function $f(z) = \dfrac{1}{z+c}$ is convex in z for any positive c. Now, invoking Jensen's inequality from [15, Ch. 2.6] leads to

$$\sum_l p_l f(z_l) \geqslant f\left(\sum_l p_l z_l\right). \tag{F.3}$$

Assuming then $p_l = 1/K$ and $z_l = 1/\zeta_l(C_u)$, we would have that

$$\begin{aligned}\frac{1}{K}\sum_{l=1}^{K}\frac{1}{\dfrac{1}{\zeta_l(C_u)}+\gamma_n} &= \sum_{l=1}^{K}\frac{1}{K}\frac{1}{\dfrac{1}{\zeta_l(C_u)}+\gamma_n} \\ &\geqslant \frac{1}{\dfrac{1}{K}\sum_{l=1}^{K}\dfrac{1}{\zeta_l(C_u)}+\gamma_n},\end{aligned} \tag{F.4}$$

which establishes the lower bound $d^l(n)$ as stated in Lemma 11 as

$$d^l(n) = \frac{1}{\dfrac{1}{K}\sum_{l=1}^{K}\dfrac{1}{\zeta_l(C_u)}+\gamma_n}. \tag{F.5}$$

For the upper bound, if we manage to find a series of random variables which are greater than or equal to $\zeta_l(P(n))$, we can then obtain the upper bound $d^u(n)$ in the same manner as we found $d^l(n)$. $\zeta_l(P(n))$ are functions of all filter memory, and it is a cumbersome task to track the filter memory. Instead, we decide to consider only the two previous time steps, i.e. $n-1$ and $n-2$, and show that we are able to find reasonably good bounds.

First, we consider $n-1$. Given that $P(n) = AM(n-1)A^T + C_u$, it is possible to obtain an upper bound on $\zeta_l(P(n))$ based on $\zeta_l(M(n-1))$. Based on Weyl's theorem on eigenvalues of sum of positive definite Hermitian symmetric matrices, we can state that

$$\zeta_l(P(n)) \leqslant \zeta_{\max}(AM(n-1)A^T) + \zeta_l(C_u). \tag{F.6}$$

It is also easy to show from Weyl's inequalities [43, Ch. 3] that for two symmetric matrices A and B, we have that

$$\zeta_{\max}(AB) \leqslant \zeta_{\max}(A)\zeta_{\max}(B). \tag{F.7}$$

Based on (F.7), we may extend (F.6) to

$$\begin{aligned}\zeta_l(P(n)) &\leqslant \zeta_{\max}(A^T)\zeta_{\max}(AM(n-1)) + \zeta_l(C_u) \\ &\leqslant \zeta_{\max}(A^T)\zeta_{\max}(A)\zeta_{\max}(M(n-1)) + \zeta_l(C_u) \\ &\leqslant |\zeta_{\max}(A)|^2\zeta_{\max}(M(n-1)) + \zeta_l(C_u). \end{aligned} \tag{F.8}$$

The next step is to find an upper bound for $\zeta_{\max}(M(n-1))$. We know that

$$\zeta_l(M(n-1)) = \frac{1}{\dfrac{1}{\zeta_l(P(n-1))} + \gamma_{n-1}}. \tag{F.9}$$

From that we conclude that

$$\zeta_{\max}(M(n-1)) = \frac{1}{\dfrac{1}{\zeta_{\max}(P(n-1))} + \gamma_{n-1}}. \tag{F.10}$$

The next step is to find an upper bound for $\zeta_{\max}(P(n-1))$. Now, if we consider one more time step backwards, i.e. $n-2$, we know that $P(n-1) = AM(n-2)A^T + C_u$. Therefore, we have as before that

$$\zeta_{\max}(P(n-1)) \leqslant |\zeta_{\max}(A)|^2\zeta_{\max}(M(n-2)) + \zeta_{\max}(C_u). \tag{F.11}$$

In order to get an upper bound for $\zeta_{\max}(M(n-2))$, we assume the worst case scenario for $M(n-2)$. Obviously $M(n-2)$ cannot be worse that C_x. That happens when $h(n^{'}) = 0$, $\forall\, n^{'} < n-2$. As a result, an upper bound for $\zeta_{\max}(M(n-2))$ is $\zeta_{\max}(C_x)$.

$$\zeta_{\max}(P(n-1)) \leqslant |\zeta_{\max}(A)|^2\zeta_{\max}(C_x) + \zeta_{\max}(C_u). \tag{F.12}$$

So far, we have proven that

$$d(n) \leqslant \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\dfrac{1}{\theta_l} + \gamma_n} \tag{F.13}$$

with

$$\theta_l = \zeta_l(C_u) + \frac{|\zeta_{\max}(A)|^2}{\dfrac{1}{\alpha_{\max}} + \gamma_{n-1}}, \tag{F.14}$$

and

$$\alpha_{\max} = |\zeta_{\max}(A)|^2\zeta_{\max}(C_x) + \zeta_{\max}(C_u). \tag{F.15}$$

Now, similar to the approach used to establish the lower bound, we may consider the function $f(z) = \dfrac{1}{1/z + c}$ for arbitrary positive c. It is easy to show that $f(z)$ is a concave function. Again invoking the Jensen's inequality, we may say that

$$\sum_l p_l f(z_l) \leqslant f\left(\sum_l p_l z_l\right). \tag{F.16}$$

Assuming then $p_l = 1/K$ and $z_l = \theta_l$, we would have that

$$\begin{aligned} \frac{1}{K}\sum_{l=1}^{K} \frac{1}{\dfrac{1}{\theta_l} + \gamma_n} &= \sum_{l=1}^{K} \frac{1}{K} \frac{1}{\dfrac{1}{\theta_l} + \gamma_n} \\ &\leqslant \frac{1}{\dfrac{1}{\dfrac{1}{K}\sum_{l=1}^{K}\theta_l} + \gamma_n}, \end{aligned} \tag{F.17}$$

which finally establishes the upper bound $d^u(n)$ as stated in Lemma 12 as

$$d^u(n) = \frac{1}{\dfrac{1}{\dfrac{1}{K}\sum_{l=1}^{K}\theta_l} + \gamma_n}, \tag{F.18}$$

with θ_l defined Lemma 12.

Appendix G

Tightness of the Coding Gain Bounds

In this section, we outline how the gap between G^u and G^l behaves as a function of system parameters and if it is eventually tight. We try to evaluate the terms $\log(G_u/G_2^l)$ and $\log(G^u/G_1^l)$, which correspond to the gap (in dB) in the $\log$(SNR)-scale. For the simple lower bound G_2^l, we have that

$$\log(G^u/G_2^l) = \log\left(\frac{\dfrac{1}{d_{\text{th}}} - \dfrac{1}{\bar{\zeta} + |\zeta_{\max}(A)|^2\alpha_{\max}}}{\dfrac{1}{d_{\text{th}}} - \dfrac{1}{K}\sum_{l=1}^{K}\dfrac{1}{\zeta_l(C_u)}}\right). \tag{G.1}$$

This is a constant gap independent of the average SNR and only a function of system parameters. We take $\left(\overline{\frac{1}{\zeta}}\right) = \frac{1}{K}\sum_{l=1}^{K}\frac{1}{\zeta_l(C_u)}$. It is possible to show that for $\zeta_l(C_u) > 0$, we have that

$$\frac{1}{d_{\text{th}}} - \left(\overline{\frac{1}{\zeta}}\right) \leqslant \frac{1}{d_{\text{th}}} - \frac{1}{\bar{\zeta}} < \frac{1}{d_{\text{th}}} - \frac{1}{\bar{\zeta} + |\zeta_{\max}(A)|^2\alpha_{\max}}. \tag{G.2}$$

Now for fixed $\bar{\zeta} = 1/K\text{tr}(C_u)$, it is easy to verify that $\log(G^u/G_2^l)$ can be minimized if $\zeta_l(C_u) = \text{const.}$, i.e. the gap is minimized when $C_u = \sigma_u^2 I$, for some $\sigma_u^2 > 0$.

Performing similar analysis for the other lower bound G_1^l, we obtain that

$$NK\log(G^u/G_1^l) = \lim_{\text{SNR}\to\infty} \log \int_0^\infty \left(\frac{\frac{1}{d_{\text{th}}} - \frac{1}{\bar{\zeta} + \frac{|\zeta_{\max}(A)|^2\alpha_{\max}}{1+z\alpha_{\max}}}}{\frac{1}{d_{\text{th}}} - \left(\frac{\overline{1}}{\zeta}\right)} \right)^{NK} f_{\gamma_{n-1}}(z)\,\mathrm{d}z$$

$$\leqslant \lim_{\text{SNR}\to\infty} \log \int_0^\infty \left(\frac{\frac{1}{d_{\text{th}}} - \frac{1}{\bar{\zeta} + |\zeta_{\max}(A)|^2 w}}{\frac{1}{d_{\text{th}}} - \left(\frac{\overline{1}}{\zeta}\right)} \right)^{NK} f_{\gamma_{n-1}^{-1}}(w)\,\mathrm{d}w. \qquad \text{(G.3)}$$

It is possible to show (similar to Appendix C) that when SNR $\to \infty$, then $f_{\gamma_{n-1}^{-1}}(w) \to \delta(w)$, i.e. Dirac's delta function. Consequently, we will have that

$$\log(G_1^u/G^l) \leqslant \frac{1}{NK} \lim_{\text{SNR}\to\infty} \log \int_0^\infty \left(\frac{\frac{1}{d_{\text{th}}} - \frac{1}{\bar{\zeta} + |\zeta_{\max}(A)|^2 w}}{\frac{1}{d_{\text{th}}} - \left(\frac{\overline{1}}{\zeta}\right)} \right)^{NK} \delta(w)\,\mathrm{d}w$$

$$= \log \left(\frac{\frac{1}{d_{\text{th}}} - \frac{1}{\bar{\zeta}}}{\frac{1}{d_{\text{th}}} - \left(\frac{\overline{1}}{\zeta}\right)} \right), \qquad \text{(G.4)}$$

which is equal to zero iff $\zeta_l(C_u) = \text{const.}$ or equivalently $C_u = \sigma_u^2 I$. The best bounds are thus achieved for well-conditioned C_u.