Parameter Estimation with Additional Information

Von der Fakultät Informatik, Elektrotechnik und Informationstechnik
der Universität Stuttgart
zur Erlangung der Würde eines Doktor-Ingenieurs (Dr.-Ing.)
genehmigte Abhandlung

vorgelegt von

Stefan Uhlich

aus Chemnitz

Hauptberichter:	Prof. Dr.-Ing. Bin Yang
Mitberichter:	Prof. Dr.-Ing. Abdelhak M. Zoubir
Tag der Einreichung:	02.04.2012
Tag der mündlichen Prüfung:	03.09.2012

Institut für Signalverarbeitung und Systemtheorie

Universität Stuttgart

2012

Acknowledgements

This thesis evolved from my research activities at the Institute for Signal Processing and System Theory (ISS), University of Stuttgart. I would like to express my gratitude to my supervisor, Prof. Dr.-Ing. Bin Yang, for giving me the opportunity to work under his guidance at the ISS. As a mentor, his door was always open and he contributed in many ways to my work. Thank you for this – I could learn from you many tricks to solve signal processing problems which I am sure will also influence my future work.

I also would like to thank Prof. Dr.-Ing. Abdelhak Zoubir for taking over the assessment of this thesis and his valuable comments.

Thanks to all colleagues from the ISS for many fruitful discussions – I had a great time working with you. Furthermore, I thank my family and parents for their manifold support which eventually enabled me to do my studies. Especially, I have fond memories of the time when I started programming Turbo Pascal with my father.

Last but not least, however, I would like to thank the one which helped me most during all that time – my Father in Heaven.

Renningen, September 2012 Stefan Uhlich

Contents

Zusammenfassung

Diese Arbeit untersucht die Schätzung von Parametern aus verrauschten Beobachtungen mit zusätzlicher Information. Mit zusätzlicher Information ist dabei jegliches Wissen gemeint, welches neben den eigentlichen Beobachtungen zur Verfügung steht. In praktischen Anwendungen taucht solches Wissen meist auf ganz natürliche Art und Weise auf, nämlich dann, wenn der Kontext der Anwendung mit berücksichtigt wird. Die Motivation zur Betrachtung solcher Parameterschätzprobleme ist zum einen, dass dadurch sichergestellt werden kann, dass physikalische Randbedingungen auch eingehalten werden. Zum anderen erlauben sie es, den Schätzfehler zu reduzieren, insbesondere in Fällen mit hoher Ungenauigkeit in den Beobachtungen. Im Folgenden sollen kurz Beispiele für solches zusätzliche Wissen gegeben werden, bevor dann die zwei Lösungsansätze, die dieser Dissertation zu Grunde liegen, näher erläutert werden.

Parameterschätzprobleme mit zusätzlicher Information haben in den letzten Jahren vermehrt an Aufmerksamkeit gewonnen. Besonders *Compressed Sensing* ist ein bekanntes Beispiel hierfür, bei dem das Wissen, dass der zu schätzende Parametervektor dünn besetzt ist, es auch im Fall von Unterabtastung noch erlaubt, das ursprüngliche Signal zu rekonstruieren [Candes and Wakin, 2008]. Unterabtastung kann entweder bewusst verwendet werden, um Zeit bzw. Kosten zu sparen, siehe z. B. [Duarte et al., 2008; Lustig et al., 2008], oder tritt aufgrund von physikalischen Randbedingungen auf [Lee et al., 2011]. Ein weiteres Beispiel für zusätzliches Wissen ergibt sich für die Aufgabe der Bildentrauschung. Gewöhnlich wird der mittlere quadratische Fehler als Optimierungskriterium zum Entwurf von Verfahren verwendet. Dies resultiert jedoch in Verfahren, die alle das gemeinsame Problem haben, dass der mittlere quadratische Fehler kein gutes Gütekriterium (engl. *loss function*) zum Vergleich von Bildern ist. Ein großer mittlerer quadratischer Fehler zwischen zwei Bildern bedeutet nicht immer, dass sich beide für einen menschlichen Betrachter stark unterscheiden [Wang and Bovik, 2009]. Es ist deshalb für den Entwurf von Bildentrauschungsmethoden interessant, andere Gütekriterien, die besser auf die menschliche Wahrnehmung zugeschnitten sind, zu verwenden, da das Ergebnis letztendlich auch von einem Menschen beurteilt wird.

Beide Beispiele motivieren die Betrachtung von Parameterschätzproblemen mit zusätzlicher Information, wie sie in dieser Arbeit erfolgt. Aufgrund der unterschiedlichen Art des verfügbaren zusätzlichen Wissens, werden zwei verschiedene Wege aufgezeigt, wie diese Information verwendet werden kann, um die Schätzung zu verbessern. Die vorgestellten Methoden gehen

dabei über den traditionell verwendeten Bayes-Ansatz mit verschiedenen A-priori-Verteilungen hinaus und ermöglichen es so, möglichst flexibel das zusätzliche Wissen in die Schätzung zu integrieren.

Eine erste Möglichkeit wird in Kapitel 3 vorgestellt, indem die zusätzliche Information in Form von Bedingungen formuliert wird, die der Schätzwert erfüllen muss. Das vorgestellte Beispiel des Compressed Sensings ist z. B. ein Problem, welches so beschrieben werden kann – es wird der dünn besetzte Parametervektor gesucht, der die Messgleichungen erfüllt. Insbesondere die konvexe 1-Norm hat sich als geeignet für Compressed Sensing gezeigt [Tibshirani, 1996]. In Kapitel 3 wird aufgezeigt, wie ein zeitvarianter Parametervektor in einem normalverteilten, linearen Signalmodell mit allgemeinen Bedingungen geschätzt werden kann und es werden drei verschiedene Schätzer für dieses Problem vorgestellt. Diese sind

- der *Recursive Constrained Maximum Likelihood* (RCML) Schätzer, der die Log-Likelihood Funktion unter der gegebenen Bedingung maximiert,
- der *Recursive Affine Minimax* (RAMX) Schätzer, der den maximalen Fehler innerhalb der erlaubten Lösungsmenge, die durch die Bedingungen definiert wird, minimiert, sowie
- der *Recursive Minimum Mean Squared Error* (RMMSE) Schätzer, welcher das ursprüngliche Problem in ein Bayes-Schätzproblem umformuliert. Dabei wird die A-priori-Verteilung des Parametervektors als gleichverteilt auf der erlaubten Lösungsmenge angenommen.

Alle drei Schätzer können in einem ersten Schritt durch den *Recursive Weighted Least Squares* (RWLS) Algorithmus vereinfacht werden, da RWLS die suffiziente Statistik für das betrachtete Schätzproblem berechnet. Die vorgestellten Verfahren werden anschließend an drei Beispielen näher erläutert und miteinander verglichen.

Eine zweite Möglichkeit, um zusätzliches Wissen zu berücksichtigen, wird in Kapitel 4 vorgestellt und besteht in der Verwendung von anderen Gütekriterien als dem bekannten quadratischen Fehler. Der verbreitete Gebrauch des quadratischen Gütekriteriums ist meist nicht darin begründet, dass sie gut zu dem betrachteten Problem passt, sondern dass es oft – zumindest unter Verwendung von *Markov-Chain-Monte-Carlo*-Methoden – eine numerisch zugängliche Lösung erlaubt. Andere, für die Praxis besser geeignete, Gütekriterien sind z. B. der *Structural-Similiarity*-Index (SSIM) zum Vergleich von Bildern [Wang and Bovik, 2009] und das *Perceptual-Evaluation-of-Speech-Quality*-Kriterium (PESQ) für Sprachdaten [Hu and Loizou, 2008]. Um eine numerisch effiziente Lösung für Parameterschätzprobleme mit solchen Gütekriterien zu erhalten, wird in Kapitel 4 eine neue parametrische Familie von Schätzern vorgestellt. Durch die Verwendung dieser Familie wird das Schätzproblem in zwei Teile aufgeteilt. In einem ersten Schritt muss für ein gegebenes Signalmodell und ein gegebenes Gütekriterium der beste Schätzer bezüglich des Bayes-Risikos innerhalb der Familie gefunden werden. Dazu

wird in dieser Arbeit ein Gradientenverfahren vorgeschlagen und die Elemente des Gradientenvektors hergeleitet. Um in einem zweiten Schritt den Schätzwert für eine neue Beobachtung zu finden, wird *Importance Sampling* verwendet. Durch diesen Ansatz kann das allgemeine Problem, welches meist mit einer Bayes-Parameterschätzung mit nicht gebräuchlichen Gütekriterien einhergeht, nämlich dass für jede Beobachtung ein Optimierungsproblem gelöst werden muss, umgangen werden. Abschließend werden in Kapitel 4 ebenfalls drei Beispiele betrachtet, um die parametrische Familie von Schätzern und ihre Eigenschaften näher zu untersuchen.

List of Symbols

Notation

x	Scalar
$\mathbf{x}$	Column vector
$\mathbf{e}_i$	ith unit vector
$\mathbf{X}$	Matrix
$\mathbf{I}$	Identity matrix
$\mathbf{J}^{i,j}$	Single-entry matrix which is zero everywhere except a one at (i, j)
$\mathbb{X}$	Set
$\mathbb{N}, \mathbb{Z}, \mathbb{R}, \mathbb{C}$	Set of all natural, integer, real and complex numbers
$\mathbb{R}[\mathbf{z}]$	Set of all polynomials in $\mathbf{z}$ with real-valued coefficients

Mathematical Operations

$\|\mathbf{x}\|$	Euclidean norm of $\mathbf{x}$, i.e. $\|\mathbf{x}\| = \|\mathbf{x}\|_2$
$\|\mathbf{x}\|_p$	p-norm of $\mathbf{x}$
$\|\mathbf{x}\|_\mathbf{W}$	Generalized Euclidean norm of $\mathbf{x}$ with respect to $\mathbf{W} \succ \mathbf{0}$, i.e. $\|\mathbf{x}\|_\mathbf{W} = \sqrt{\mathbf{x}^T\mathbf{W}\mathbf{x}}$
$\mathbf{X}^T, \mathbf{X}^H$	Transpose and Hermitian transpose of a vector or matrix
$\mathrm{tr}\{\mathbf{X}\}, \det\{\mathbf{X}\}$	Trace and determinant of a matrix
$\mathbf{X} \succ \mathbf{0}$, $\mathbf{X} \succeq \mathbf{0}$	$\mathbf{X}$ is symmetric and positive definite/nonnegative definite
$\mathbf{X} \otimes \mathbf{Y}, \mathbf{X} \circ \mathbf{Y}$	Kronecker and elementwise (Hadamard) product of $\mathbf{X}$ and $\mathbf{Y}$
$\mathbf{x} * \mathbf{y}$, $\mathbf{x} \circledast \mathbf{y}$	Linear and circular convolution of $\mathbf{x}$ and $\mathbf{y}$

$\mathrm{diag}\{\mathbf{x}\}$	Diagonal matrix with $\mathbf{x}$ on its diagonal
$\mathrm{vec}\{\mathbf{X}\}$	Vector which is obtained after column stacking of $\mathbf{X}$
$\mathrm{DFT}\{.\}$	Discrete Fourier transform operator
$\mathrm{E}[.]$	Expectation operator

Special Functions

$\mathrm{erf}(x)$	Error function, $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$
$\Gamma(x)$	Gamma function, $\Gamma(x) = \int_0^\infty e^{-t} t^{x-1} dt$
$I_n(x)$	Modified Bessel function of the first kind and nth order
$M(a, b, z)$	Kummer's function, see Appendix C.3.4
$\mathcal{V}(.)$	Valley-filling function, see Section 2.2.2

Acronyms

AMX	Affine minimax
BIM	Bayesian information matrix
BLinEx	Bounded LinEx
BLS	Bayes least squares
BLUE	Best linear unbiased estimator
BR	Bayes risk
CDF	Cumulative distribution function
CML	Constrained maximum likelihood
CRB	Cramér-Rao bound
DFT	Discrete Fourier transform
DOA	Direction of arrival
EM	Expectation-maximization
FFT	Fast Fourier transform
FIM	Fisher information matrix
FS	Feasible set
GMRF	Gaussian Markov random field
GSM	Gaussian scale mixture
KKT	Karush-Kuhn-Tucker
LASSO	Least absolute shrinkage and selection operator
LinEx	Linear-exponential
LMI	Linear matrix inequality
LP	Linear program
LS	Least squares
MAP	Maximum a posteriori
MC	Monte Carlo
MCMC	Markov chain Monte Carlo
MELE	Mean likelihood estimator
MIE	Method of interval errors
ML	Maximum likelihood
MMSE	Minimum mean squared error
MPM	Marginal posterior median

MSE	Mean squared error
MVUE	Minimum variance unbiased estimator
OBE	Optimal Bayesian estimator
ODE	Ordinary differential equation
PAMX	Projected affine minimax
PESQ	Perceptual evaluation of speech quality
PLS	Projected least squares
QCQP	Quadratically constrained quadratic program
QP	Quadratic program
RAMX	Recursive affine minimax
RCML	Recursive constrained ML
RIR	Room impulse response
RMMSE	Recursive MMSE
RWLS	Recursive weighted least squares
PDF	Probability density function
ScLS	Scaled least squares
SDP	Semidefinite program
SLS	Spherical least squares
SNR	Signal-to-noise ratio
SOS	Sum of squares
SSIM	Structural similarity
STSA	Short-time spectral attenuation
VAD	Voice activity detector
WE-STSA	Weighted Euclidean STSA
WLS	Weighted least squares

Chapter 1

Introduction

This thesis deals with the problem of estimating unknown parameters from noisy data using additional information. By additional information we denote any domain knowledge that is available to us except the data itself. Such prior domain knowledge very often arises naturally from the estimation problem at hand if the context is taken into account. The motivation to incorporate the prior domain knowledge is twofold: First, it allows us to ensure that the estimate will fulfill physical constraints which are given and second, we can also expect a better estimation performance. In the following, we will give examples of this domain knowledge and then outline the general way how we could deal with such estimation problems.

Considering estimation problems with additional a priori information has seen an increased interest over the last years. Especially *Compressed Sensing* is a prominent example where the knowledge that the parameter vector is sparse, i.e. it only contains a small number of nonzero elements, allows a reconstruction of the original signal even in the case of undersampling, see e.g. [Candes and Wakin, 2008]. Such schemes are interesting from a practical point of view in a real system: We could either arbitrarily undersample the original signal to save time/costs as was shown in [Duarte et al., 2008; Lustig et al., 2008] or due to physical boundary conditions, we are facing the problem of undersampling [Lee et al., 2011]. Another example where we have prior information is image denoising where the goal is to estimate a (clean) undistorted image from the noisy image itself. Traditionally, the *mean squared error* (MSE) is used to derive image denoising algorithms. However, these approaches suffer from the fact that the MSE is not a good quality measure of image similarity if compared to human perception, see e.g. [Wang and Bovik, 2009]. Therefore, using other quality measures, called loss functions in the context of estimation theory, we expect better results as we exploit the knowledge of the human perception, i.e. the prior domain knowledge is that the image denoising result will be assessed by a human.

Both examples motivate the study of general schemes that allow to incorporate additional prior information into the estimation procedure. Traditionally, a Bayesian approach with different a priori distributions is used to exploit this information. However, this approach is not well suited for many practical problems, e.g. the image denoising example given above, as the additional

information can not be directly represented by the a priori distribution. We will therefore discuss in this thesis two other approaches to deal with prior domain knowledge.

The first possibility is to express the prior information in terms of constraints that have to be satisfied by the estimator. The first example of compressed sensing is such a case where we want to find the sparsest parameter vector which fulfills the measurement equations. Especially the convex 1-norm has turned out to be a suitable constraint for compressed sensing [Tibshirani, 1996]. We will discuss in this thesis the problem of estimating a time-varying parameter vector in a Gaussian linear signal model with general constraints. Three different estimators are proposed for this problem:

- the *recursive constrained maximum likelihood* (RCML) estimator, which maximizes the log-likelihood function such that the constraints are fulfilled,
- the *recursive affine minimax* (RAMX) estimator, which minimizes the worst case error on the feasible set defined by the constraints, and
- the *recursive minimum mean squared error* (RMMSE) estimator, which recasts the frequentists problem into a Bayesian one and uses the prior with the maximum entropy on the feasible set.

All three estimators can be substantially simplified by using the *recursive weighted least squares* (RWLS) algorithm in a first step as RWLS computes a sufficient statistic for this estimation problem. The recursive constrained ML needs to solve an optimization problem in the second step for the case that the RWLS solution does not fulfill the constraint. In case of affine minimax, we have to solve an optimization problem and to perform an affine transform. The RMMSE estimator needs to calculate the mean of a truncated Gaussian density in the second step which is done by Monte Carlo integration. A simple rejection scheme is used to take general constraints for the parameter vector into account.

A second possibility to incorporate additional information is to use other quality measures (i.e. loss functions) than the MSE for the design of estimators. The prominent use of the MSE is often not its suitability to the application at hand but that it allows tractable solutions, at least in terms of *Markov chain Monte Carlo* (MCMC) methods [Liu, 2008]. Other loss functions that are more suited for particular applications are e.g. the *structural similarity* (SSIM) index for image comparison [Wang and Bovik, 2009] or the *perceptual evaluation of speech quality* (PESQ) measure for speech data [Hu and Loizou, 2008]. To obtain estimators for such nonstandard loss functions – opposed to the mean squared error or the hit-or-miss error – we introduce a parametric family of estimators. By restricting the estimator to lie in this family, we split the estimation problem into two parts: In a first step, we find the best estimator with respect to the Bayes risk for a given loss function, which has to be done only once. The second step then calculates the estimate for a particular observation. The advantage of the proposed parametric family

is that *importance sampling* allows the efficient computation of estimates for new observations. It thus overcomes the general problem of Bayesian estimation with nonstandard loss functions, namely that the Bayesian estimator for such loss functions can often only be stated in terms of an optimization problem, which has to be solved for each new observation.

The outline of this thesis is as follows: Chapter 2 will introduce the basic notions and methods from estimation theory which lay the foundations for later chapters. Based on these concepts, Chapter 3 will consider the problem of recursive estimation using a priori information in terms of a set where we know that the time-varying parameter vector has to lie in. It summarizes the main ideas which we presented in [Uhlich and Yang, 2009, 2010b, 2011]. Chapter 4 will reveal another possibility of introducing application domain knowledge by using suitable loss functions. The material for this chapter is mainly drawn from [Uhlich and Yang, 2010a, 2012]. Motivated by the parametric family of estimators from Chapter 4, we will turn in Chapter 5 to the problem of finding the loss function for a given estimator and signal model. There, we look at the inverse problem opposed to what is usually encountered and we will give solutions in terms of convex optimization problems which have to be solved. Finally, Chapter 6 concludes this thesis by restating the main points of the presented work and summarizing further work that was published in [Uhlich and Yang, 2008; Blind et al., 2009; Uhlich et al., 2009]. The last part of this final chapter also shows issues which are interesting for future work.

Please note that Appendix A and B summarize the used concepts from matrix calculus and optimization theory, which are used throughout this thesis.

Chapter 2

Parameter Estimation Methods

This chapter briefly summarizes the estimation methods and tools which we will use in later chapters. The interested reader is referred to [Lehmann and Casella, 1998] and [Robert, 2001] for a more detailed treatment.

2.1 Overview

The goal of estimation theory is to infer the value of an unknown parameter from noisy data, i.e. data with a random component. More formally, let $\mathbf{x} \in \mathbb{R}^K$ denote the data, which we use to estimate the value of the unknown parameter $\boldsymbol{\theta} \in \mathbb{R}^M$. Then, the goal is to find an estimator $\hat{\boldsymbol{\theta}}(\mathbf{x}) : \mathbb{R}^K \to \mathbb{R}^M$ which is optimal with respect to a specific design criterion.

In general, two different approaches can be distinguished:

(i) The *frequentist* approach assumes that the parameter vector is deterministic but unknown. Using the data $\mathbf{x}$ which is known to have the parametrized *probability density function* (PDF) $p(\mathbf{x}; \boldsymbol{\theta})$, the value of the unknown parameter vector $\boldsymbol{\theta}$ has to be estimated.

(ii) The *Bayesian* approach treats $\boldsymbol{\theta}$ as a random parameter vector which exhibits an *a priori density* $p(\boldsymbol{\theta})$. The parameter which we want to estimate is then one realization of the random vector $\boldsymbol{\theta}$. All information, which we can use to estimate the value of $\boldsymbol{\theta}$, is described by $p(\mathbf{x}, \boldsymbol{\theta}) = p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})$. The PDF $p(\mathbf{x}|\boldsymbol{\theta})$ is called the *likelihood density* and it represents our knowledge about the data conditioned on a fixed $\boldsymbol{\theta}$. This density is often derived from the signal model underlying our application. The a priori PDF $p(\boldsymbol{\theta})$ summarizes our knowledge about $\boldsymbol{\theta}$ before any data is observed.

Both approaches are used in practice and have their advantages as well as disadvantages. In [Samaniego, 2010], a profound comparison of both philosophies is performed. Considering different examples, Samaniego shows that depending on the application, the Bayesian or frequentist approach gives better results and the statistician should ask himself the question whether the problem is better treated as a Bayesian or frequentist one.

Before introducing the estimators that will be used in this thesis, we first review performance bounds on the covariance matrix and the *mean squared error* (MSE) matrix of an estimator.

2.2 Performance Bounds

Having lower bounds on the estimation error $\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}$ is very useful in practice [Kay, 1993; Van Trees and Bell, 2007]. They allow the evaluation and design of estimators, e.g. to find the *minimum variance unbiased estimator* (MVUE), but are also useful for the design of experiments if some degrees of freedom can be used to obtain estimates with a higher precision, see e.g. [Pukelsheim, 2006].

In the following, we will derive lower bounds on the estimation error for the deterministic and the Bayesian case. In the literature, it is common to speak of a *deterministic/Bayesian* bound if the underlying estimation problem is a frequentist/Bayesian one. If the parameter vector contains a deterministic part and a random part then the corresponding bounds are called *hybrid*.

The estimator bias $\mathbf{b}_{\hat{\boldsymbol{\theta}}}$, covariance matrix $\mathbf{C}_{\hat{\boldsymbol{\theta}}}$ and MSE matrix $\mathbf{S}_{\hat{\boldsymbol{\theta}}}$ are defined as

Deterministic parameter:

$$
\begin{aligned}
\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) &= \mathrm{E}\left[\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right] = \int \hat{\boldsymbol{\theta}}(\mathbf{x}) p(\mathbf{x};\boldsymbol{\theta}) d\mathbf{x} - \boldsymbol{\theta}, \\
\mathbf{C}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) &= \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)^T\right] - \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})^T \\
&= \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \mathrm{E}\left[\hat{\boldsymbol{\theta}}(\mathbf{x})\right]\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \mathrm{E}\left[\hat{\boldsymbol{\theta}}(\mathbf{x})\right]\right)^T\right], \\
\mathbf{S}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) &= \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)^T\right] = \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})^T + \mathbf{C}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}),
\end{aligned}
$$

Random parameter:

$$
\begin{aligned}
\mathbf{b}_{\hat{\boldsymbol{\theta}}} &= \mathrm{E}\left[\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right] = \iint \left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right) p(\mathbf{x},\boldsymbol{\theta}) d\mathbf{x} d\boldsymbol{\theta}, \\
\mathbf{S}_{\hat{\boldsymbol{\theta}}} &= \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)^T\right] = \iint \left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)^T p(\mathbf{x},\boldsymbol{\theta}) d\mathbf{x} d\boldsymbol{\theta}.
\end{aligned}
$$

In general, at least three families of bounds can be distinguished: the *covariance-inequality* bounds, the *Ziv-Zakai* bounds and the *interval estimation* methods which will now be successively introduced. Note that the interval estimation method only provides an approximation of the MSE and not a strict lower bound as the two other families.

Furthermore, also other families of lower bounds exist in the literature. For example in [Todros and Tabrikian, 2010a,b] a new class of lower bounds was introduced recently. However, in this thesis we will only introduce the three families, which we mentioned before as they are the most often used ones in practice.

2.2.1 Covariance-Inequality Family

The covariance-inequality family is the most prominent family of lower bounds, especially as the Cramér-Rao bound, which is often applied due to its simple form yielding analytical results, is included. The name *covariance-inequality family* stems from the fact that all following bounds rely on the same covariance inequality [Weinstein and Weiss, 1988; Abel, 1993]

$$\mathrm{E}\left[\left(\mathbf{f}(\mathbf{x},\boldsymbol{\theta})-\mathbf{A}\mathbf{g}(\mathbf{x},\boldsymbol{\theta})\right)\left(\mathbf{f}(\mathbf{x},\boldsymbol{\theta})-\mathbf{A}\mathbf{g}(\mathbf{x},\boldsymbol{\theta})\right)^T\right]\succeq\mathbf{0} \tag{2.1}$$

for arbitrary functions $\mathbf{f}(\mathbf{x},\boldsymbol{\theta}):\mathbb{R}^K\times\mathbb{R}^M\to\mathbb{R}^P$, $\mathbf{g}(\mathbf{x},\boldsymbol{\theta}):\mathbb{R}^K\times\mathbb{R}^M\to\mathbb{R}^Q$ and $\mathbf{A}\in\mathbb{R}^{P\times Q}$. $\mathbf{X}\succeq\mathbf{Y}$ means that $\mathbf{X}-\mathbf{Y}$ is nonnegative definite. Note that $\mathrm{E}[.]$ either assumes $\boldsymbol{\theta}$ to be deterministic if we derive a deterministic bound or assumes $\boldsymbol{\theta}$ to be a random vector for the Bayesian case.[1] Expanding (2.1), we obtain

$$\mathrm{E}\left[\mathbf{f}(\mathbf{x},\boldsymbol{\theta})\mathbf{f}(\mathbf{x},\boldsymbol{\theta})^T\right]\succeq\mathbf{T}\mathbf{A}^T+\mathbf{A}\mathbf{T}^T-\mathbf{A}\mathbf{G}\mathbf{A}^T\quad\forall\mathbf{A} \tag{2.2}$$

with $\mathbf{T}=\mathrm{E}\left[\mathbf{f}(\mathbf{x},\boldsymbol{\theta})\mathbf{g}(\mathbf{x},\boldsymbol{\theta})^T\right]$ and $\mathbf{G}=\mathrm{E}\left[\mathbf{g}(\mathbf{x},\boldsymbol{\theta})\mathbf{g}(\mathbf{x},\boldsymbol{\theta})^T\right]$. As the right-hand side has to hold for all $\mathbf{A}$, we can obtain a good lower bound by using the special choice $\mathbf{A}_{\text{opt}}=\mathbf{T}\mathbf{G}^{-1}$ as shown in Appendix C.1.1. Using $\mathbf{f}_{\text{D}}(\mathbf{x},\boldsymbol{\theta})=\hat{\boldsymbol{\theta}}-\mathrm{E}[\hat{\boldsymbol{\theta}}]$ and $\mathrm{E}[(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})^T]=\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})^T+\mathrm{E}[(\hat{\boldsymbol{\theta}}-\mathrm{E}[\hat{\boldsymbol{\theta}}])(\hat{\boldsymbol{\theta}}-\mathrm{E}[\hat{\boldsymbol{\theta}}])^T]$ in the deterministic case and $\mathbf{f}_{\text{B}}(\mathbf{x},\boldsymbol{\theta})=\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}$ in the Bayesian case, we finally can establish the following bounds on the MSE matrix of any estimator

$$\textit{Deterministic case:}\quad \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{x})-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x})-\boldsymbol{\theta}\right)^T\right]\succeq\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})^T+\mathbf{T}\mathbf{G}^{-1}\mathbf{T}^T, \tag{2.3a}$$

$$\textit{Bayesian case:}\quad \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{x})-\boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x})-\boldsymbol{\theta}\right)^T\right]\succeq\mathbf{T}\mathbf{G}^{-1}\mathbf{T}^T. \tag{2.3b}$$

Choosing different functions $\mathbf{g}(\mathbf{x},\boldsymbol{\theta})$ yields different bounds which will now be discussed.

Cramér-Rao Bound

Using the special choice

$$\mathbf{g}_{\text{D}}(\mathbf{x},\boldsymbol{\theta})=\frac{\partial\ln p(\mathbf{x};\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}^T,\qquad\qquad \mathbf{g}_{\text{B}}(\mathbf{x},\boldsymbol{\theta})=\frac{\partial\ln p(\mathbf{x},\boldsymbol{\theta})}{\partial\boldsymbol{\theta}}^T, \tag{2.4}$$

[1] Deterministic case: $\mathrm{E}[.]=\int . \; p(\mathbf{x};\boldsymbol{\theta})d\mathbf{x}$, Bayesian case: $\mathrm{E}[.]=\iint . \; p(\mathbf{x},\boldsymbol{\theta})d\mathbf{x}d\boldsymbol{\theta}$.

the bounds in (2.3) turn into the *Cramér-Rao bound* (CRB) [Rao, 1945; Cramer, 1946; Van Trees, 1968]. Please note that the transpose operator $(.)^T$ was used in (2.4) to obtain column vectors as $\partial \ln p(\mathbf{x};\boldsymbol{\theta})/\partial\boldsymbol{\theta}$ results by definition in a row vector, see Appendix A.

The matrices $\mathbf{T}$ and $\mathbf{G}$ take the form $\mathbf{T}_{\mathrm{D}} = \mathbf{I} + \frac{\partial \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$, $\mathbf{T}_{\mathrm{B}} = \mathbf{I}$ and

$$\mathbf{G}_{\mathrm{D}} = \mathrm{E}\left[\frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^T \frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right], \qquad \mathbf{G}_{\mathrm{B}} = \mathrm{E}\left[\frac{\partial \ln p(\mathbf{x},\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^T \frac{\partial \ln p(\mathbf{x},\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right], \tag{2.5}$$

where $\mathbf{G}_{\mathrm{D}}$ is the *Fisher information matrix* (FIM) and $\mathbf{G}_{\mathrm{B}}$ the *Bayesian information matrix* (BIM) [Van Trees and Bell, 2007].[2]

In the deterministic case, for example, we obtain for the MSE matrix

$$\mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \boldsymbol{\theta}\right)^T\right] \succeq \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})^T + \left(\mathbf{I} + \frac{\partial \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)\mathbf{G}_{\mathrm{D}}^{-1}\left(\mathbf{I} + \frac{\partial \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right)^T \tag{2.6}$$

which is sometimes called the *biased CRB* [Eldar, 2008c]. This bound will be useful in the derivation of the affine minimax estimator in Section 2.3.2. If we restrict our consideration to unbiased estimators with $\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \mathbf{0}$ only, (2.6) simplifies to

$$\mathrm{E}\left[\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \mathrm{E}[\hat{\boldsymbol{\theta}}(\mathbf{x})]\right)\left(\hat{\boldsymbol{\theta}}(\mathbf{x}) - \mathrm{E}[\hat{\boldsymbol{\theta}}(\mathbf{x})]\right)^T\right] = \mathbf{C}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) \succeq \mathbf{G}_{\mathrm{D}}^{-1} \tag{2.7}$$

which is the most often used form of the CRB [Kay, 1993] as the bias term $\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})$ is typically not known. The popularity of the CRB is due to its simplicity which allows an analytical derivation in many situations.

It is important to note that the CRB is a *small error bound*, i.e. it bounds the performance of an estimator by considering the behavior of the log-likelihood function[3] $\ln p(\mathbf{x};\boldsymbol{\theta})$ in the vicinity of the true parameter value $\boldsymbol{\theta}$. Such errors are called *small errors*. However, for a low SNR, it is also important to consider the case of *large errors*, which occurs if the log-likelihood function has its maximum at a different position and thus the MSE of an estimator is not only locally determined by the log-likelihood function at $\boldsymbol{\theta}$. For many estimation problems, one can therefore observe that the CRB is not a tight bound at lower SNR values and hence can not be used to predict the threshold region, i.e. the SNR region where the MSE of an estimator increases more rapidly. This threshold phenomena can for example be found in frequency estimation [Van Trees and

[2] For the derivation of $\mathbf{T}_{\mathrm{D}}$ and $\mathbf{T}_{\mathrm{B}}$, some regularity conditions have to be fulfilled, see e.g. [Kay, 1993; Van Trees and Bell, 2007]. E.g. in the deterministic case, it is assumed that the differentiation with respect to $\boldsymbol{\theta}$ and the expectation operator $E[.]$ can be interchanged.

[3] The same is also true in the Bayesian case as the CRB only considers the joint PDF $p(\mathbf{x},\boldsymbol{\theta})$ in the vicinity of the true parameter value $\boldsymbol{\theta}$.

Bell, 2007; Knockaert, 1997]. Due to this problem of the CRB, tighter bounds were studied in the literature, which we will now introduce.

Bhattacharyya Bound

The *Bhattacharyya bound* [Bhattacharyya, 1946; Van Trees, 1968] is obtained from the covariance inequality (2.3) for the special choice

$$\mathbf{g}_{\mathrm{D}}(\mathbf{x},\boldsymbol{\theta}) = \frac{1}{p(\mathbf{x};\boldsymbol{\theta})} \begin{bmatrix} \frac{\partial p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^T \\ \frac{\partial^2 p(\mathbf{x};\boldsymbol{\theta})}{\mathrm{Rem}\{\partial\boldsymbol{\theta}\otimes\partial\boldsymbol{\theta}\}}^T \\ \vdots \end{bmatrix}, \qquad \mathbf{g}_{\mathrm{B}}(\mathbf{x},\boldsymbol{\theta}) = \frac{1}{p(\mathbf{x},\boldsymbol{\theta})} \begin{bmatrix} \frac{\partial p(\mathbf{x},\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^T \\ \frac{\partial^2 p(\mathbf{x},\boldsymbol{\theta})}{\mathrm{Rem}\{\partial\boldsymbol{\theta}\otimes\partial\boldsymbol{\theta}\}}^T \\ \vdots \end{bmatrix}, \tag{2.8}$$

where "$\otimes$" denotes the Kronecker product. In comparison to the CRB, $\mathbf{g}(\mathbf{x},\boldsymbol{\theta})$ is augmented with higher-order derivatives. The operator $\mathrm{Rem}\{.\}$ is used to remove redundant elements in $\mathbf{g}(\mathbf{x},\boldsymbol{\theta})$ as e.g. $\frac{\partial p(\mathbf{x},\boldsymbol{\theta})}{\partial\theta_i\partial\theta_j} = \frac{\partial p(\mathbf{x},\boldsymbol{\theta})}{\partial\theta_j\partial\theta_i}$. If the redundant elements are not removed then $\mathbf{G}$ is singular and can therefore not be inverted.

It can be shown that [Abel, 1993; Van Trees and Bell, 2007] under some suitable regularity conditions

$$\mathbf{T}_{\mathrm{D}} = \begin{bmatrix} \mathbf{I} & 0 & \cdots \end{bmatrix} + \begin{bmatrix} \frac{\partial \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})}{\partial\boldsymbol{\theta}} & \frac{\partial^2 \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})}{\mathrm{Rem}\{\partial\boldsymbol{\theta}\otimes\partial\boldsymbol{\theta}\}} & \cdots \end{bmatrix}, \qquad \mathbf{T}_{\mathrm{B}} = \begin{bmatrix} \mathbf{I} & 0 & \cdots \end{bmatrix}. \tag{2.9}$$

Although it can be proved that the Bhattacharyya bound yields a tighter bound than the CRB, its derivation is quite involved and also the difference in comparison to the CRB is often only small which does not justify the increased computational complexity. In [Forster and Larzabal, 2002] for example, the Bhattacharyya bound of third order was derived and the results show that the improvement does not help to predict the threshold phenomena that often occurs in nonlinear estimation problems. This stems from the fact that the Bhattacharyya bound is again a small error bound that does not explicitly consider the case of large errors.

Bobrovsky-Zakai Bound

Another bound that can be derived from the covariance inequality (2.3) is the *Bobrovsky-Zakai bound* [Bobrovsky and Zakai, 1975; Reuven and Messer, 1997]. It follows from the choice

$$\mathbf{g}_{\mathrm{D}}(\mathbf{x},\boldsymbol{\theta}) = \begin{bmatrix} \frac{p(\mathbf{x};\boldsymbol{\theta}+\mathbf{h}_1)}{p(\mathbf{x};\boldsymbol{\theta})} - 1 \\ \vdots \\ \frac{p(\mathbf{x};\boldsymbol{\theta}+\mathbf{h}_Q)}{p(\mathbf{x};\boldsymbol{\theta})} - 1 \end{bmatrix}, \qquad \mathbf{g}_{\mathrm{B}}(\mathbf{x},\boldsymbol{\theta}) = \begin{bmatrix} \frac{p(\mathbf{x},\boldsymbol{\theta}+\mathbf{h}_1)}{p(\mathbf{x},\boldsymbol{\theta})} - 1 \\ \vdots \\ \frac{p(\mathbf{x},\boldsymbol{\theta}+\mathbf{h}_Q)}{p(\mathbf{x},\boldsymbol{\theta})} - 1 \end{bmatrix}, \tag{2.10}$$

which can be seen as a finite difference approximation of (2.4) that does not require the differentiability of $p(\mathbf{x};\boldsymbol{\theta})$ and $p(\mathbf{x},\boldsymbol{\theta})$ with respect to $\boldsymbol{\theta}$. It results in

$$\mathbf{T}_{\mathrm{D}} = \left[\mathbf{h}_1 + \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}+\mathbf{h}_1) - \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) \quad \cdots \quad \mathbf{h}_Q + \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}+\mathbf{h}_Q) - \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\right],$$
$$\mathbf{T}_{\mathrm{B}} = \left[\mathbf{h}_1 \quad \cdots \quad \mathbf{h}_Q\right].$$

The Q test points $\{\mathbf{h}_1,\ldots,\mathbf{h}_Q\}$ can be freely chosen. The bound which is obtained from using the optimal set of test points is also called the *Hammersley-Chapman-Robbins bound*.

The Bobrovsky-Zakai bound is the most widely used Barankin-style[4] bound [Abel, 1993] and belongs to the class of *large error bounds*. Note that the Cramér-Rao bound is included in the Bobrovsky-Zakai bound for the special choice $Q = M$ and $\mathbf{h}_i = \epsilon_i \mathbf{e}_i$ with $\epsilon_i \to 0$ for all $i = 1,\ldots,M$ and assuming that $p(\mathbf{x};\boldsymbol{\theta})$ and $p(\mathbf{x},\boldsymbol{\theta})$ are differentiable at $\boldsymbol{\theta}$.

Weiss-Weinstein Bound and Composite Bounds

For completeness, we would like to mention the *Weiss-Weinstein bound* which was introduced in [Weiss and Weinstein, 1985]. It can also be obtained from (2.3) and is an extension of the Bobrovsky-Zakai bound resulting in tighter bounds which have proven to be useful for predicting the threshold region in direction-of-arrival (DOA) estimation problems [Bell et al., 1993].

Moreover, also compositions of bounds derived from (2.3) have been studied in the literature, e.g. in [Abel, 1993]. Abel proposed a combination of the Bhattacharyya bound and the Hammersley-Chapman-Robbins bound to benefit from the advantages of both of them, i.e. to develop a bound which considers both small and large errors.

2.2.2 Other Approaches

Ziv-Zakai Family

Another approach to obtain a lower bound was developed in [Ziv and Zakai, 1969] for the scalar case and later extended in [Bell et al., 1997] to the case of a parameter vector.

Starting from the observation

$$\mathrm{E}\left[X^2\right] = \frac{1}{2}\int_0^\infty h \Pr\left\{|X| \geq \frac{h}{2}\right\} dh, \tag{2.11}$$

[4]The term *Barankin bound* refers to the more general choice (for the scalar case $Q = 1$)

$$\mathbf{g}_{\mathrm{D}}(\mathbf{x},\boldsymbol{\theta}) = \frac{1}{p(\mathbf{x};\boldsymbol{\theta})}\left(\int p(\mathbf{x};\boldsymbol{\theta}+\mathbf{h})p(\mathbf{h})d\mathbf{h} - \int p(\mathbf{x};\boldsymbol{\theta}+\tilde{\mathbf{h}})p(\tilde{\mathbf{h}})d\tilde{\mathbf{h}}\right),$$
$$\mathbf{g}_{\mathrm{B}}(\mathbf{x},\boldsymbol{\theta}) = \frac{1}{p(\mathbf{x},\boldsymbol{\theta})}\left(\int p(\mathbf{x},\boldsymbol{\theta}+\mathbf{h})p(\mathbf{h})d\mathbf{h} - \int p(\mathbf{x},\boldsymbol{\theta}+\tilde{\mathbf{h}})p(\tilde{\mathbf{h}})d\tilde{\mathbf{h}}\right).$$

which is proved in Appendix C.1.2 and holds for any random variable X, the Ziv-Zakai bound is derived by using $X = \mathbf{a}^T(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})$ where $\mathbf{a} \in \mathbb{R}^M$ is an arbitrary vector. It can be shown that in the Bayesian case[5]

$$\mathbf{a}^T \mathbf{S}_{\hat{\boldsymbol{\theta}}} \mathbf{a} \geq \frac{1}{2} \int_0^\infty \mathcal{V}\left(\max_{\mathbf{a}^T\boldsymbol{\delta}=h} \int \left(p(\boldsymbol{\theta}) + p(\boldsymbol{\theta}+\boldsymbol{\delta})\right) P_{\min}(\boldsymbol{\theta}, \boldsymbol{\theta}+\boldsymbol{\delta}) d\boldsymbol{\theta} \right) h \, dh \tag{2.12}$$

can be established where $\mathcal{V}(.)$ is the *valley-filling function*[6] [Bellini, 1978; Bell et al., 1997] and $P_{\min}(\boldsymbol{\theta}, \boldsymbol{\theta}+\boldsymbol{\delta})$ is the minimum error probability of a corresponding decision problem. In particular, $\mathbf{a} = \mathbf{e}_i$ is interesting as (2.12) returns a bound on the MSE of the ith parameter $\boldsymbol{\theta}_i$. Appendix C.1.3 shows the basic steps to derive (2.12).

To find $P_{\min}(\boldsymbol{\theta}, \boldsymbol{\theta}+\boldsymbol{\delta})$, results from detection theory can be used. Nevertheless is the overall computation of the Ziv-Zakai bound often challenging but also worthwile as, for example, [Nguyen and Van Trees, 1994] showed. There, a comparison of the Weiss-Weinstein bound and the Ziv-Zakai bound was performed and the simulation results show that the Ziv-Zakai bound is a tighter bound for this particular estimation problem.

Method of Interval Errors

The last approach we would like to mention is the *method of interval errors* (MIE) [Van Trees, 1968]. This method directly models the two kinds of error, namely small errors and large errors, which contribute together to the estimator's MSE. Therefore, the MSE matrix can be approximated by

$$\mathbf{S}_{\hat{\boldsymbol{\theta}}} \approx \Pr(\mathcal{I}) \, \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^T \middle| \mathcal{I} \right] + (1 - \Pr(\mathcal{I})) \, \mathrm{E}\left[\left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^T \middle| \bar{\mathcal{I}} \right] \tag{2.13}$$

where $\mathcal{I}$ denotes the case of a large error[7]. A good approximation of the MSE matrix in the case of no large error is given by the Cramér-Rao bound. Therefore, it remains to compute $\Pr(\mathcal{I})$ and $\mathrm{E}[(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^T | \mathcal{I}]$ which are application dependent, see e.g. the work in [Athley, 2005] where the DOA problem is studied. Note that the method of interval errors does not provide a strict lower bound which cannot be beaten by any estimator but gives in many cases a good approximation of the MSE matrix.

[5] A similar bound can be derived for the deterministic case, see [Gu and Wong, 1991] and Appendix C.1.3 for details.

[6] The valley-filling function $\mathcal{V}(f(h))$ is a nonincreasing function of h which is obtained by filling all valleys in $f(h)$. Mathematically, it is defined as [Bellini, 1978]

$$\mathcal{V}(f(h)) = \max_{\epsilon \geq 0} f(h+\epsilon).$$

[7] In the context of MIE, one often uses the notion of *interval errors* if one refers to large errors. We therefore use the symbol $\mathcal{I}$ to denote such an event.

After introducing the most important families of lower estimation bounds, we now turn to the different possibilities to design estimators, which will be used in this thesis.

2.3 Frequentist Approach to Estimation

The frequentist approach assumes the unknown parameter $\boldsymbol{\theta}$ to be deterministic but unknown and is therefore purely data driven. In the following, we will introduce the *maximum likelihood* and the *affine minimax estimator*. Beside them, also other frequentist estimators are known, e.g. the *minimum variance unbiased estimator* (MVUE), the *best linear unbiased estimator* (BLUE) and the *method of moments* [Kay, 1993].

2.3.1 Maximum Likelihood Estimation

The *maximum likelihood* (ML) method is probably the most prominent method used in estimation theory. The ML estimator is given by

$$\hat{\boldsymbol{\theta}}_{\mathrm{ML}}(\mathbf{x}) = \arg\max_{\boldsymbol{\theta}} \; p(\mathbf{x};\boldsymbol{\theta}) = \arg\max_{\boldsymbol{\theta}} \; \ln\big(p(\mathbf{x};\boldsymbol{\theta})\big)\,, \tag{2.14}$$

i.e. it estimates the value of $\boldsymbol{\theta}$ that is most likely given the data $\mathbf{x}$. Assuming that $\boldsymbol{\theta}$ is continuously valued, a necessary condition to find the ML estimator is

$$\left.\frac{\partial p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right|_{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\mathrm{ML}}} = \mathbf{0}, \tag{2.15}$$

which often cannot be solved analytically. Then, numerical techniques like the Newton-Raphson method, sometimes used with scoring where the Hessian is replaced by the inverse FIM (2.5), and the *Expectation-Maximization* (EM) algorithm have to be used [Dempster et al., 1977; Kay, 1993].

The popularity of the ML estimator stems not only from its straightforward derivation using the necessary condition (2.15) but also from its asymptotic properties [Kay, 1993] if an infinite number of data is available. The ML estimator is consistent, i.e. it converges in probability to the true parameter $\boldsymbol{\theta}$. Furthermore, it can be established that the ML estimator converges in distribution to a Gaussian distribution with mean $\boldsymbol{\theta}$ and covariance matrix equal to the inverse FIM (2.5). Note that these asymptotic properties only hold under some regularity conditions which have to be fulfilled [Lehmann and Casella, 1998].

For the special case of a linear Gaussian model

$$\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{z}, \tag{2.16}$$

where $\mathbf{z}$ is zero-mean Gaussian noise with covariance $\mathbf{C}$, i.e. $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$, the ML estimator is well known to be $\hat{\boldsymbol{\theta}}_{\text{ML}}(\mathbf{x}) = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{x}$ assuming a full column rank of $\mathbf{H}$ [Kay, 1993]. This estimator is also the MVUE and it achieves the CRB.

2.3.2 Affine Minimax Estimation

The second frequentist approach we consider is the *affine minimax* (AMX) estimator, which relies on the approach

$$\hat{\boldsymbol{\theta}}_{\text{AMX}}(\mathbf{x}) = \mathbf{U}\mathbf{x} + \mathbf{u}, \tag{2.17}$$

i.e. it performs an affine transform of the data $\mathbf{x}$ where $\mathbf{U}$ and $\mathbf{u}$ are found by solving the minimax problem

$$\begin{aligned}\min_{\mathbf{U},\mathbf{u}} \max_{\boldsymbol{\theta}\in\mathbb{T}} \operatorname{E}\left[\left\|\hat{\boldsymbol{\theta}}_{\text{AMX}}(\mathbf{x}) - \boldsymbol{\theta}\right\|^2\right] &= \min_{\mathbf{U},\mathbf{u}} \max_{\boldsymbol{\theta}\in\mathbb{T}} \operatorname{tr}\left\{\mathbf{C}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\right\} + \left\|\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\right\|^2 \\ &= \min_{\mathbf{U},\mathbf{u}} \max_{\boldsymbol{\theta}\in\mathbb{T}} \operatorname{tr}\left\{\mathbf{U}\operatorname{E}\left[(\mathbf{x}-\operatorname{E}[\mathbf{x}])(\mathbf{x}-\operatorname{E}[\mathbf{x}])^T\right]\mathbf{U}^T\right\} + \left\|\mathbf{U}\operatorname{E}[\mathbf{x}] + \mathbf{u} - \boldsymbol{\theta}\right\|^2. \end{aligned} \tag{2.18}$$

$\mathbb{T} \subset \mathbb{R}^M$ is the set of possible parameter values. Hence, the affine minimax estimator seeks those $\mathbf{U}$ and $\mathbf{u}$ that minimize the worst-case MSE[8].

For the special case of the linear model (2.16), we can rewrite (2.18) as

$$\min_{\mathbf{U},\mathbf{u}} \max_{\boldsymbol{\theta}\in\mathbb{T}} \operatorname{tr}\left\{\mathbf{U}\mathbf{C}\mathbf{U}^T\right\} + \left\|\left(\mathbf{U}\mathbf{H} - \mathbf{I}\right)\boldsymbol{\theta} + \mathbf{u}\right\|^2. \tag{2.19}$$

This problem was intensively studied in the literature for the case of affine and/or quadratic constraints defining $\mathbb{T}$, see e.g. [Pinsker, 1980; Pilz, 1986; Eldar et al., 2005; Eldar, 2008b]. Especially, [Eldar, 2008b] is interesting as it shows how to calculate $\mathbf{U}$ and $\mathbf{u}$ for the special quadratic constraint $\mathbb{T} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^T\mathbf{A}\boldsymbol{\theta} + 2\mathbf{b}^T\boldsymbol{\theta} + c \leq 0\}$ using a *semidefinite program*[9] (SDP) formulation. For other constraints, standard techniques from convex optimization like Finsler's lemma or the S-procedure can be used [Boyd and Vandenberghe, 2007; Ben-Tal et al., 2009]. However, the calculation of $\mathbf{U}$ and $\mathbf{u}$ is in general, i.e. for an arbitrary and possibly nonconvex $\mathbb{T}$, difficult.

Eldar's Motivation for the Affine Minimax Estimator

In [Eldar, 2006a, 2008a,c], a motivation for the affine minimax estimator in terms of achieving the biased CRB was given. In contrast to the traditional concept of CRB for the variance of

[8] Note that the MSE cannot be minimized directly as it is a function of the unknown parameter $\boldsymbol{\theta}$. Therefore, either the worst-case MSE has to be minimized, which is done by the affine minimax estimator, or the MSE is minimized with the constraint that the estimator is unbiased. This leads to the best linear unbiased (BLUE) estimator if $\mathbf{u} = \mathbf{0}$ is chosen. See [Kay, 1993] for further details.

[9] A brief introduction to convex optimization and semidefinite programs can be found in Appendix B.

unbiased estimators and efficient estimators achieving this bound, Eldar proposed a framework to derive estimators achieving the biased CRB (2.6) which is actually an MSE bound. This bound can be made, by allowing for a suitable bias $\mathbf{b}$, often smaller than the CRB for all $\boldsymbol{\theta} \in \mathbb{T}$. In particular, Eldar presented a linear bias $\mathbf{M}\boldsymbol{\theta}$ in [Eldar, 2006a] and extended this idea to an affine bias $\mathbf{M}\boldsymbol{\theta} + \mathbf{u}$ in [Eldar, 2008b]. Moreover, it was shown that, given an unbiased and efficient estimator $\hat{\boldsymbol{\theta}}_{\text{eff}}$, the following affine transform of $\hat{\boldsymbol{\theta}}_{\text{eff}}$

$$\hat{\boldsymbol{\theta}}_{\text{BE}} = (\mathbf{I} + \mathbf{M})\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{x}) + \mathbf{u}, \tag{2.20}$$

with bias $\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \mathbf{M}\boldsymbol{\theta} + \mathbf{u}$ and covariance $\mathbf{C}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = (\mathbf{I} + \mathbf{M})\,\mathbf{G}(\boldsymbol{\theta})^{-1}\,(\mathbf{I} + \mathbf{M})^T$ achieves the MSE bound (2.6) as $\partial \mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})/\partial \boldsymbol{\theta} = \mathbf{M}$ and has thus, if $\mathbf{M} \neq \mathbf{0}$ or/and $\mathbf{u} \neq \mathbf{0}$, a smaller MSE than $\hat{\boldsymbol{\theta}}_{\text{eff}}$ for all $\boldsymbol{\theta} \in \mathbb{T}$. Thus, there is no other estimator with fixed bias $\mathbf{b}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta}) = \mathbf{M}\boldsymbol{\theta} + \mathbf{u}$ that *dominates*[10] it.

Introducing the short-hand notation

$$\text{MSE}(\mathbf{M}, \mathbf{u}, \boldsymbol{\theta}) = \text{tr}\{\mathbf{S}_{\hat{\boldsymbol{\theta}}}\} = \|\mathbf{M}\boldsymbol{\theta} + \mathbf{u}\|^2 + \text{tr}\{(\mathbf{I} + \mathbf{M})\,\mathbf{G}(\boldsymbol{\theta})^{-1}\,(\mathbf{I} + \mathbf{M})^T\},$$

Eldar proposed the minimax problem

$$\min_{\mathbf{M},\mathbf{u}} \max_{\boldsymbol{\theta}\in\mathbb{T}} \text{MSE}(\mathbf{M}, \mathbf{u}, \boldsymbol{\theta}) - \text{MSE}(\mathbf{0}, \mathbf{0}, \boldsymbol{\theta}) \tag{2.21}$$

to find a biased estimator which dominates the efficient (unbiased) estimator $\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{x})$ for all $\boldsymbol{\theta} \in \mathbb{T}$. She showed that (2.21) has a unique solution which is *admissible*[11] on the set $\mathbb{T}$ and dominates $\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{x})$ if $\mathbf{M} \neq \mathbf{0}$ and/or $\mathbf{u} \neq \mathbf{0}$.

For the special case of a linear model (2.16), the efficient estimator is given by $\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{x}) = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{x}$ with the covariance matrix $\mathbf{G}(\boldsymbol{\theta})^{-1} = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}$. Since $\hat{\boldsymbol{\theta}}_{\text{eff}}(\mathbf{x})$ is linear in $\mathbf{x}$, $\hat{\boldsymbol{\theta}}_{\text{BE}}(\mathbf{x})$ in (2.20) becomes affine. The corresponding minimax problem (2.21) is

$$\min_{\mathbf{M},\mathbf{u}} \max_{\boldsymbol{\theta}\in\mathbb{T}} \|\mathbf{M}\boldsymbol{\theta} + \mathbf{u}\|^2 + \text{tr}\left\{(\mathbf{I} + \mathbf{M})\left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}(\mathbf{I} + \mathbf{M})^T\right\} \tag{2.22}$$

which is equivalent to (2.19) using the identity $\mathbf{U} = (\mathbf{I} + \mathbf{M})(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1}$. Hence, the use of (2.17) and (2.19) for a Gaussian linear model can be motivated by the fact that (2.17) is an admissible estimator, i.e. there is no other estimator with affine bias that dominates it.

[10] An estimator $\hat{\boldsymbol{\theta}}_1(\mathbf{x})$ is said to *dominate* another estimator $\hat{\boldsymbol{\theta}}_2(\mathbf{x})$ if the MSE of $\hat{\boldsymbol{\theta}}_1(\mathbf{x})$ is never larger than the MSE of $\hat{\boldsymbol{\theta}}_2(\mathbf{x})$ for all $\boldsymbol{\theta} \in \mathbb{T}$ and is strictly smaller for some $\boldsymbol{\theta} \in \mathbb{T}$ [Eldar, 2006b].

[11] An estimator $\hat{\boldsymbol{\theta}}(\mathbf{x})$ is said to be *admissible* if there is no other estimator that dominates it.

Pilz's Motivation for the Affine Minimax Estimator

Pilz gave in [Pilz, 1986] another interpretation of the affine minimax estimator for a Gaussian linear model which we would like to briefly mention.

Pilz showed that the affine minimax estimator is equivalent to a linear Bayesian estimator which is found by considering the least favourable prior on the set $\mathbb{T}$, i.e. that prior which results in the largest Bayes risk among all densities with support $\mathbb{T}$. He used this result to explicitely calculate the affine minimax estimator for ellipsoidal and linear inequality constraints as this analogy allows the application of results known from Bayesian design of experiments.[12]

2.4 Bayesian Approach to Estimation

The Bayesian approach, in contrast to the frequentist approach, assumes that $\boldsymbol{\theta}$ is a random vector with a priori density $p(\boldsymbol{\theta})$. In the following, we will derive the *optimal Bayesian estimator* (OBE) for a given loss function $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, in particular for the case of a quadratic loss $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathbf{W} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})$ with $\mathbf{W} \succ \mathbf{0}$ which yields the *minimum mean squared error* (MMSE) estimator.

2.4.1 Loss Functions

Suppose we have an estimator $\hat{\boldsymbol{\theta}}(\mathbf{x})$ that estimates the unknown, random parameter $\boldsymbol{\theta}$ from the data $\mathbf{x}$. To evaluate the quality of the estimator, we assign a loss $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x})) \geq 0$ to the error of estimating $\hat{\boldsymbol{\theta}}(\mathbf{x})$ from the data $\mathbf{x}$ when the true value is $\boldsymbol{\theta}$.

Following different types of loss functions can be distinguished:

Definition. *A loss function $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))$ is called*

(i) symmetric, if $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = L(-\boldsymbol{\theta}, -\hat{\boldsymbol{\theta}})$;

(ii) spherical, if $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \tilde{L}(\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|)$.

Note that a spherical loss function is also symmetric but the converse is in general not true. An example is the loss which is symmetric but not spherical for $\mathbf{W} \neq \alpha \mathbf{I}$ in the quadratic loss case $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathbf{W} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})$, $\mathbf{W} \succ \mathbf{0}$. Besides these two properties, *scale invariance* [Stein, 1964; Brown, 1968] and *boundedness* [Kaminska and Porosinski, 2009; Wen and Levy, 2001] are other characteristics of the loss function that may be desired in practical applications.

[12] Note that Eldar discusses the results of Pilz in [Eldar, 2006b] and it was shown that the proposed estimator for an ellipsoidal constraint $\mathbb{T}$ in [Pilz, 1986] can be further improved as it is not always the minimax estimator. The interested reader is referred to Appendix C in [Eldar, 2006b].

2.4.2 Optimal Bayesian Estimator

Averaging the loss $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))$ with respect to the joint PDF $p(\boldsymbol{\theta}, \mathbf{x})$ yields an important characteristic value for an estimator. It is called the *Bayes risk* (BR) and given by [Scharf, 1990]

$$\text{BR} = \iint L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))p(\boldsymbol{\theta}, \mathbf{x})d\boldsymbol{\theta}d\mathbf{x}. \tag{2.23}$$

The *optimal Bayesian estimator* (OBE) is now that estimator that minimizes the Bayes risk, i.e.

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\text{OBE}}(\mathbf{x}) &= \arg\min_{\hat{\boldsymbol{\theta}}(\mathbf{x})} \text{BR} = \arg\min_{\hat{\boldsymbol{\theta}}(\mathbf{x})} \iint L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))p(\boldsymbol{\theta}, \mathbf{x})d\boldsymbol{\theta}d\mathbf{x} \\ &= \arg\min_{\hat{\boldsymbol{\theta}}(\mathbf{x})} \int L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}, \end{aligned} \tag{2.24}$$

where we used in the last line of (2.24) the fact that $p(\mathbf{x}) \geq 0$ and thus it is sufficient to minimize the inner integral for each $\mathbf{x}$. Hence, $\arg\min_{\hat{\boldsymbol{\theta}}} \text{BR}$ is equivalent to minimizing the loss averaged over the a posteriori distribution $p(\boldsymbol{\theta}|\mathbf{x})$ and we immediately see that all information to find the OBE is included in the a posteriori density $p(\boldsymbol{\theta}|\mathbf{x})$.

Assuming that the loss function $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))$ is differentiable, we can calculate the first-order derivative with respect to the estimate and equate it to zero to obtain a necessary condition[13] to find the OBE, i.e.

$$\frac{\partial}{\partial \hat{\boldsymbol{\theta}}} \int L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta} = \int \frac{\partial L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))}{\partial \hat{\boldsymbol{\theta}}} p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta} \stackrel{!}{=} \mathbf{0}. \tag{2.25}$$

Solving (2.25) can often not be done analytically and therefore Bayesian estimation with most loss functions is difficult. We will therefore propose a parametric family of estimators in Chapter 4 to circumvent this problem.

In the following, we will now give the OBE for various loss functions. We start with three well known, symmetric loss functions and then discuss the OBE for the asymmetric *linear-exponential* (LinEx) loss and the *monomial splined* loss.

Absolute Error Loss

The absolute error loss is defined as

$$L_{\text{AE}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x})) = \left\| \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}(\mathbf{x}) \right\|_1 = \sum_{m=1}^{M} \left| \theta_m - \hat{\theta}_m(\mathbf{x}) \right| \tag{2.26a}$$

[13] We assume here that the parameter space is open. Otherwise, the OBE could also lie on the boundary of the parameter space and (2.25) is not necessary anymore.

and the corresponding OBE is $\hat{\boldsymbol{\theta}}_{\text{MPM}}(\mathbf{x}) \in \mathbb{R}^M$ with [Kay, 1993]

$$\int_{-\infty}^{\hat{\theta}_{\text{MPM},m}(\mathbf{x})} p(\theta_m|\mathbf{x})d\theta = \int_{\hat{\theta}_{\text{MPM},m}(\mathbf{x})}^{\infty} p(\theta_m|\mathbf{x})d\theta \quad \forall\, m = 1,\ldots,M \tag{2.26b}$$

where $p(\theta_m|\mathbf{x})$ is the marginal a posteriori distribution of the mth element of $\boldsymbol{\theta}$. From (2.26b) we see that the OBE is the median of the marginal a posteriori densities $p(\theta_m|\mathbf{x})$ and therefore this estimator is called *marginal posterior median* (MPM) estimator.

Squared Error Loss

The (weighted) squared error loss is defined as

$$L_{\text{SE}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x})) = \left\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}(\mathbf{x})\right\|_{\mathbf{W}}^2 = \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}(\mathbf{x})\right)^T \mathbf{W} \left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}(\mathbf{x})\right), \quad \mathbf{W} \succ \mathbf{0} \tag{2.27a}$$

and the corresponding estimator is the well known *minimum mean squared error* (MMSE) estimator [Kay, 1993]

$$\hat{\boldsymbol{\theta}}_{\text{MMSE}}(\mathbf{x}) = \int \boldsymbol{\theta}\, p(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta} = \frac{\int \boldsymbol{\theta}\, p(\boldsymbol{\theta}, \mathbf{x})d\boldsymbol{\theta}}{\int p(\boldsymbol{\theta}, \mathbf{x})d\boldsymbol{\theta}} \tag{2.27b}$$

which is the mean of the a posteriori density $p(\boldsymbol{\theta}|\mathbf{x})$. We will use the MMSE estimator throughout this thesis and therefore Appendix C.1.4 proves (2.27b) for the loss function (2.27a) and also derives some important properties of the MMSE estimator.

Hit-or-miss Loss

The hit-or-miss loss is defined as

$$L_{\text{HM}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x})) = \begin{cases} 1 & \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}(\mathbf{x})\| > \epsilon \\ 0 & \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}(\mathbf{x})\| < \epsilon \end{cases}, \tag{2.28a}$$

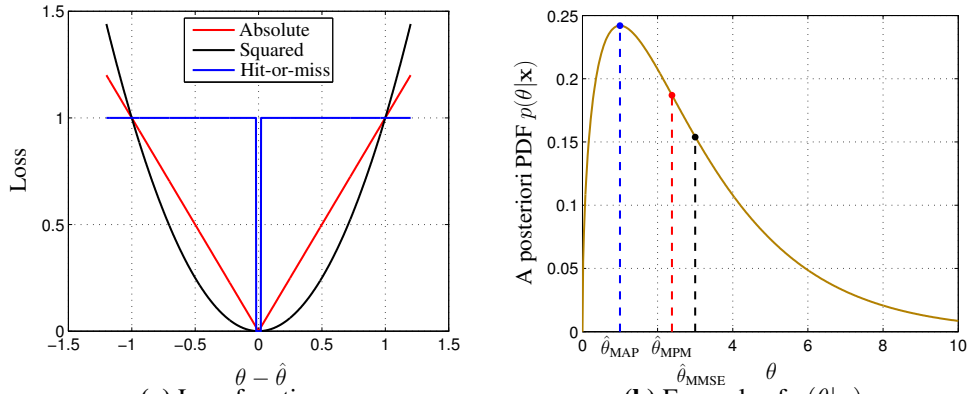
where ϵ is chosen such that $\epsilon \to 0^+$.[14] The corresponding OBE is the *maximum a posteriori* (MAP) estimator

$$\hat{\boldsymbol{\theta}}_{\text{MAP}}(\mathbf{x}) = \arg\max_{\boldsymbol{\theta}}\, p(\boldsymbol{\theta}|\mathbf{x}) \tag{2.28b}$$

which is the position of the maximum of the a posteriori density $p(\boldsymbol{\theta}|\mathbf{x})$.

Note that the MAP estimator is the solution of an optimization problem which is often easier to solve then computing the mean of the a posteriori density. Hence, the MAP is often used instead of the MMSE estimator as it is computational tractable. Furthermore, we would like to point out

[14] Note that the case $\epsilon \nrightarrow 0^+$ is also interesting as minimizing this loss corresponds to minimizing the ϵ-outage error [Routtenberg and Tabrikian, 2009].



(a) Loss functions **(b)** Example of $p(\theta|\mathbf{x})$

Figure 2.1: Comparison of the MPM, MMSE and MAP estimator

that similar but not identical asymptotic properties can be established for the MAP estimator as for the ML estimator [Van Trees and Bell, 2007]. One main difference is that the MAP estimator might not be efficient, i.e. it does in general not achieve the Bayesian CRB asymptotically.

Figure 2.1 shows a comparison of the MPM, MMSE and MAP estimator for the univariate case $M = 1$ which summarizes the estimators that we derived so far. Note that there are posterior distributions which exhibit the same mean, median and mode and therefore the corresponding OBEs coincide. This is e.g. the case if the posterior distribution is Gaussian.

LinEx Loss

The *linear-exponential* (LinEx) loss was introduced by Varian in [Varian, 1974] as an asymmetric loss function for the study of real estate assessment and is frequently used in Bayesian estimation, see e.g. [Zellner, 1986]. It rises approximately linear on one side and exponential on the other. This was motivated by Varian by the fact that underassessment of real estate results in a linear loss of revenue whereas overassessment results in appeals with litigation and other costs.

The univariate LinEx loss function is given by

$$L_{\text{LinEx}}(\theta, \hat{\theta}) = b\left(e^{a\Delta} - a\Delta - 1\right), \tag{2.29}$$

where $\Delta = \hat{\theta} - \theta$, $a \neq 0$ and $b > 0$. An example for the LinEx loss is shown in Figure 4.2 in Chapter 4. The multivariate LinEx loss is defined as a straightforward extension and given by [Zellner, 1986]

$$L_{\text{LinEx}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x})) = \sum_{m=1}^{M} b_m \left(e^{a_m \Delta_m} - a_m \Delta_m - 1\right), \tag{2.30a}$$

where $\Delta_m = \hat{\theta}_m - \theta_m$, $a_m \neq 0$ and $b_m > 0$. To calculate the OBE, we can use (2.25) with $\partial L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))/\partial\hat{\theta}_m = b_m a_m \left(e^{a_m \Delta_m} - 1\right)$ and finally obtain

$$\hat{\theta}_{\text{LinEx},m} = -\frac{1}{a_m} \ln \int e^{-a_m \theta_m} p(\theta_m|\mathbf{x}) d\theta_m = -\frac{1}{a_m} \ln M_{\theta_m}(-a_m|\mathbf{x}), \quad m = 1, \ldots, M \tag{2.30b}$$

where $M_{\theta_m}(t|\mathbf{x}) = \mathrm{E}[\exp(t\theta_m)|\mathbf{x}]$ is the (marginal) moment generating function of θ_m with respect to the marginal a posteriori PDF $p(\theta_m|\mathbf{x})$.

Monomial-Splined Loss

The last general class of loss functions, which we would like to briefly mention, is the *monomial-splined* loss which was proposed by [Thompson and Basu, 1996]. For a univariate setup, it is defined as

$$L_{\text{MonSpl}}(\theta, \hat{\theta}) = \begin{cases} a_1 |\Delta|^{p_1} & \Delta > 0 \\ a_2 |\Delta|^{p_2} & \Delta \leq 0 \end{cases}, \tag{2.31}$$

where $\Delta = \hat{\theta} - \theta$; $p_1, p_2 \in \mathbb{N}$ and $a_1, a_2 > 0$. Similar to the LinEx loss, this definition can be extended to the multivariate case

$$L_{\text{MonSpl}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \sum_{m=1}^{M} L_{\text{MonSpl}}(\theta_m, \hat{\theta}_m). \tag{2.32}$$

Two important cases are $p_1 = p_2 = 1$ which is called the *LinLin* loss and $p_1 = p_2 = 2$ which is the *QuadQuad* loss. The LinLin loss was already considered by Laplace in 1773 [Robert, 2001] and it can be shown that the OBE is the $\frac{a_2}{a_1+a_2}$-fractile of the marginal a posteriori density $p(\theta_m|\mathbf{x})$. This can be shown by considering the Bayes risk

$$\text{BR} = \sum_{m=1}^{M} \left[\int_{-\infty}^{\hat{\theta}_m} a_1(\hat{\theta}_m - \theta_m) p(\theta_m|\mathbf{x}) d\theta_m + \int_{\hat{\theta}_m}^{\infty} a_2(\theta_m - \hat{\theta}_m) p(\theta_m|\mathbf{x}) d\theta_m \right].$$

Using the Leibniz integral identity [Kay, 1993]

$$\frac{\partial}{\partial u} \int_{\phi_1(u)}^{\phi_2(u)} h(u,v) dv = \int_{\phi_1(u)}^{\phi_2(u)} \frac{\partial}{\partial u} h(u,v) dv + \frac{\partial \phi_2(u)}{\partial u} h(u, \phi_2(u)) - \frac{\partial \phi_1(u)}{\partial u} h(u, \phi_1(u)),$$

we obtain

$$\frac{\partial \text{BR}}{\partial \hat{\theta}_m} = \int_{-\infty}^{\hat{\theta}_m} a_1 p(\theta_m|\mathbf{x}) d\theta_m - \int_{\hat{\theta}_m}^{\infty} a_2 p(\theta_m|\mathbf{x}) d\theta_m = 0,$$

which finally implies $\Pr\left\{\theta_m \leq \hat{\theta}_m|\mathbf{x}\right\} = \frac{a_2}{a_1+a_2}$. Note that the absolute loss (2.26a) is a special case of the LinLin loss with $a_1 = a_2$ and the corresponding OBE is the 50%-fractile, i.e. the median, as stated in (2.26b).

2.4.3 Practical Considerations

After this introduction of the Bayesian estimation concept, we will now turn to two important issues which have to be considered in order to use these methods. The first issue is how the a priori density $p(\boldsymbol{\theta})$ should be chosen. Very often, the likelihood density $p(\mathbf{x}|\boldsymbol{\theta})$ is known from the signal model but the choice of the prior is a design parameter which has to be specified.

The second issue is how the OBE can be calculated. In particular, the calculation of the MMSE estimator is difficult as it involves two M-dimensional integrations as can be seen from (2.27b). We will therefore briefly introduce Monte Carlo integration methods, especially importance sampling, which will later be useful.

Choosing the A Priori Distribution

The choice of the a priori density $p(\boldsymbol{\theta})$ is a critical point for Bayesian estimation as $p(\boldsymbol{\theta})$ is often not known. In the following, we will discuss two possible ways to choose the a priori density, namely *conjugate priors* and *maximum entropy priors*. Beside these methods, many other approaches have been discussed in the literature and a good summary of them is given in [Robert, 2001]. Especially *noninformative prior*, e.g. the *Jeffreys prior* which is invariant with respect to reparameterization, is another interesting approach to choose the a priori density which is briefly reviewed in Appendix C.1.5.

(a) Conjugate Priors: A prior distribution $p(\boldsymbol{\theta})$ is called conjugate to the likelihood $p(\mathbf{x}|\boldsymbol{\theta})$ if the a posteriori density $p(\boldsymbol{\theta}|\mathbf{x})$ belongs to the same parametric family as $p(\boldsymbol{\theta})$. One example is the Gaussian distribution which is conjugate to a Gaussian likelihood as $\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\mu}_k, \mathbf{C}_k)$ and $\mathbf{x}|\boldsymbol{\theta} \sim \mathcal{N}(\mathbf{H}\boldsymbol{\theta}, \mathbf{C})$ implies $\boldsymbol{\theta}|\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}_{k+1}, \mathbf{C}_{k+1})$ with

$$\begin{aligned}\boldsymbol{\mu}_{k+1} &= \boldsymbol{\mu}_k + \mathbf{C}_k\mathbf{H}^T(\mathbf{H}\mathbf{C}_k\mathbf{H}^T + \mathbf{C})^{-1}(\mathbf{x} - \mathbf{H}\boldsymbol{\mu}_k),\\ \mathbf{C}_{k+1} &= \mathbf{C}_k - \mathbf{C}_k\mathbf{H}^T(\mathbf{H}\mathbf{C}_k\mathbf{H}^T + \mathbf{C})^{-1}\mathbf{H}\mathbf{C}_k.\end{aligned}$$

Using conjugate priors can be motivated in the following two ways [Robert, 2001]: First, the amount of information of an observation $\mathbf{x}$ is limited. Therefore, it is reasonable of the statistician to not change the structure of the density but only to update its parameters. This justification is called the *invariance reasoning*. Another motivation, and probably the more important one, is that using conjugate priors yields tractable a posteriori densities as the calculation of the a posteriori density becomes merely an update of its parameters. This allows to use recursive Bayesian

approaches. If e.g. Gaussian distributions for the prior and likelihood are used, then the Kalman filter equations are obtained [Mendel, 1995].

(b) Maximum Entropy Priors: The motivation for the maximum entropy prior is that for a given set of distributions, we should choose the distribution with maximum entropy as this will minimize the amount of prior information [Levine and Tribus, 1978]. Let $\boldsymbol{\theta}$ be a random vector, then the entropy of $\boldsymbol{\theta}$ is defined as (see e.g. [Cover and Thomas, 2006])

$$H(\boldsymbol{\theta}) = -\int p(\boldsymbol{\theta}) \ln p(\boldsymbol{\theta}) d\boldsymbol{\theta}. \tag{2.33}$$

It quantifies the average amount of information contained in one "message" (i.e. sample) $\boldsymbol{\theta}_i \sim p(\boldsymbol{\theta})$. Therefore, choosing the maximum entropy prior shows that each "message" $\boldsymbol{\theta}_i$ contains the most information and hence the PDF itself contains the least prior information. If, for example, the mean and the variance of $\boldsymbol{\theta}$ are known, then the distribution with the maximum entropy is a Gaussian distribution with these first and second moments.[15] If it is only known that $\boldsymbol{\theta} \in \mathbb{T}$, then the uniform distribution on $\mathbb{T}$ is the distribution with the maximum entropy. Such a prior will be used in Section 3.3.3 for the recursive constrained MMSE estimator.

Numerical Calculation of the Optimal Bayesian Estimator

In Bayesian estimation, often integrals occur which have to be solved to obtain the OBE. For example the MMSE estimator requires to evaluate two M-dimensional integrals as can be seen from (2.27b). Thus, many different methods were developed in the literature and a summary of approaches which are especially suited for Bayesian integrals can be found in [Evans and Swartz, 1995; Shaw et al., 1996; Robert, 2001]. In particular, three different families of methods have proven to be useful: *Laplace approximation*, *numerical integration* and *Monte Carlo integration*. We will briefly discuss these three approaches in the following paragraphs.

(a) Laplace Approximation: Consider the integral

$$\int_{\mathbb{T}} f_1(\boldsymbol{\theta}) e^{f_2(\boldsymbol{\theta})} d\boldsymbol{\theta} \tag{2.34}$$

which we want to evaluate. Then, the idea of the Laplace approximation is to approximate the integrand by a Gaussian curve, i.e. we do a Taylor series expansion of $f_2(\boldsymbol{\theta})$ up to terms with second order, see e.g. [de Bruijn, 1961]. Let $\boldsymbol{\theta}^*$ denote the location of the maximum of $f_2(\boldsymbol{\theta})$, i.e. $\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} f_2(\boldsymbol{\theta})$. If we assume that $f_1(\boldsymbol{\theta}) \approx const$ in the region where $f_2(\boldsymbol{\theta})$ has its maximum, we obtain the approximation

$$\int_{\mathbb{T}} f_1(\boldsymbol{\theta}) e^{f_2(\boldsymbol{\theta})} d\boldsymbol{\theta} \approx f_1(\boldsymbol{\theta}^*) \int_{\mathbb{T}} e^{f_2(\boldsymbol{\theta}^*) + \frac{1}{2}(\boldsymbol{\theta}-\boldsymbol{\theta}^*)^T \boldsymbol{\nabla}\boldsymbol{\nabla}^T f_2(\boldsymbol{\theta}^*)(\boldsymbol{\theta}-\boldsymbol{\theta}^*)} d\boldsymbol{\theta} \tag{2.35}$$

[15] This can be shown using *calculus of variations* [Cover and Thomas, 2006].

$$= f_1(\boldsymbol{\theta}^*)e^{f_2(\boldsymbol{\theta}^*)}\sqrt{\frac{(2\pi)^M}{|-\boldsymbol{\nabla}\boldsymbol{\nabla}^T f_2(\boldsymbol{\theta}^*)|}}, \tag{2.36}$$

where $|-\boldsymbol{\nabla}\boldsymbol{\nabla}^T f_2(\boldsymbol{\theta}^*)|$ denotes the determinant of the negative Hessian of $f_2(\boldsymbol{\theta})$ at $\boldsymbol{\theta}$. The underlying assumptions for this approximation are that $f_2(\boldsymbol{\theta})$ has only one dominant mode in the set $\mathbb{T}$ and that the maximum $\boldsymbol{\theta}^*$ is not located on the boundary of $\mathbb{T}$. These assumptions are often fulfilled for $K \to \infty$, i.e. if the number of samples is large enough.

Applying the Laplace approximation to the MMSE estimator in (2.27b), we obtain

$$\hat{\boldsymbol{\theta}}(\mathbf{x}) = \frac{\int \boldsymbol{\theta}\, p(\boldsymbol{\theta})p(\mathbf{x}|\boldsymbol{\theta})d\boldsymbol{\theta}}{\int p(\boldsymbol{\theta})p(\mathbf{x}|\boldsymbol{\theta})d\boldsymbol{\theta}} \approx \boldsymbol{\theta}^* \frac{p(\boldsymbol{\theta}^*)p(\mathbf{x}|\boldsymbol{\theta}^*)\sqrt{\frac{(2\pi)^M}{|-\boldsymbol{\nabla}\boldsymbol{\nabla}^T \ln\{p(\boldsymbol{\theta}^*)p(\mathbf{x}|\boldsymbol{\theta}^*)\}|}}}{p(\boldsymbol{\theta}^*)p(\mathbf{x}|\boldsymbol{\theta}^*)\sqrt{\frac{(2\pi)^M}{|-\boldsymbol{\nabla}\boldsymbol{\nabla}^T \ln\{p(\boldsymbol{\theta}^*)p(\mathbf{x}|\boldsymbol{\theta}^*)\}|}}} = \boldsymbol{\theta}^* \tag{2.37}$$

where $\boldsymbol{\theta}^* = \arg\max_{\boldsymbol{\theta}} \ln\{p(\boldsymbol{\theta})p(\mathbf{x}|\boldsymbol{\theta})\}$, i.e. the MMSE estimator with the Laplace approximation is given by the MAP estimator (2.28b). To obtain a better approximation, [Tierney and Kadane, 1986] proposed the *fully-exponential Laplace approximation* which does a separate Taylor series expansion of the numerator and denominator integrand.

Summarizing the Laplace method, we see that it provides good approximations of the integrals in the asymptotic case[16].

(b) Numerical Integration: The idea of numerical integration is to use

$$\int_{\mathbb{T}} f(\boldsymbol{\theta})d\boldsymbol{\theta} \approx \sum_{i=1}^{I} w_i f(\boldsymbol{\theta}_i), \tag{2.38}$$

where $\boldsymbol{\theta}_i$ are *node points* and w_i are the corresponding *weights*, see e.g. [Dahlquist and Björck, 2008]. To obtain the node positions and the weights, i.e. the integration rule, two different approaches can be used. One possibility is to use repeated one-dimensional integrations, especially the *product rule* is interesting if the integration boundary does not depend on the integration variable $\boldsymbol{\theta}$. Another possibility is to use a similar approach as done for the one-dimensional case where the node positions and the weights are chosen such that monomials up to a certain degree are exactly integrated. We review this approach for the one-dimensional case briefly in the Appendix C.1.6.

A numerical integration is often only feasible for a small value of M, typically $M \lessapprox 10$. This is due to the *curse of dimensionality* which states that the error scales with $\mathcal{O}(I^{-c/M})$ where c is given by the particular integration rule, e.g. $c = 2$ if the trapezoidal rule is used, see Appendix C.1.6. Therefore, we can only expect a slow convergence of the numerical integration with increasing I if we have a high-dimensional problem.

[16] By asymptotic case we refer here to the case that either K and/or the SNR is very large and therefore $p(\boldsymbol{\theta}|\mathbf{x})$ has one sharp peak at the true parameter value.

Furthermore, to apply (2.38) to the approximation of the MMSE estimator, we also need to ensure that we have a small number of data K and a low *signal-to-noise ratio* (SNR) because otherwise a sharp peak of $p(\mathbf{x}|\boldsymbol{\theta})$ would make the numerical integration difficult and would yield wrong estimates. Hence, we can conclude that numerical integration is interesting for small values of M and the non-asymptotic case.

(c) Monte Carlo (MC) Integration: As we have seen in the paragraph before, numerical integration quickly becomes intractable if the dimension M is to large. MC integration is then the method of choice and very often the only possibility to solve such high-dimensional integrals. In Bayesian estimation for biomedical applications for example, $M \lessapprox 10^6$ is not uncommon, see [Gelman et al., 2003].

The idea is to use the well known fact that the expectation operator can be replaced by the average, i.e.

$$\int_{\mathbb{T}} f(\boldsymbol{\theta}) d\boldsymbol{\theta} = V(\mathbb{T}) \int_{\mathbb{T}} f(\boldsymbol{\theta}) \frac{1}{V(\mathbb{T})} d\boldsymbol{\theta} = V(\mathbb{T})\, \mathrm{E}\left[f(\bar{\boldsymbol{\theta}})\right] \approx \frac{V(\mathbb{T})}{I} \sum_{i=1}^{I} f(\bar{\boldsymbol{\theta}}_i), \tag{2.39}$$

where $V(\mathbb{T})$ denotes the volume of $\mathbb{T}$, $\bar{\boldsymbol{\theta}}$ is a random vector which is uniformly distributed on $\mathbb{T}$ and $\bar{\boldsymbol{\theta}}_i$ with $i = 1, \ldots, I$ are independent samples generated from this uniform distribution. The approximation (2.39) can be motivated by the *law of large numbers*, which states that $\frac{1}{I} \sum_{i=1}^{I} f(\bar{\boldsymbol{\theta}}_i)$ converges almost surely to $\mathrm{E}[f(\bar{\boldsymbol{\theta}})]$. Let $e = \frac{V(\mathbb{T})}{I} \sum_{i=1}^{I} f(\bar{\boldsymbol{\theta}}_i) - \int_{\mathbb{T}} f(\boldsymbol{\theta}) d\boldsymbol{\theta}$ denote the integration error, then it is obvious that $\mathrm{E}[e] = 0$ and the variance is given by

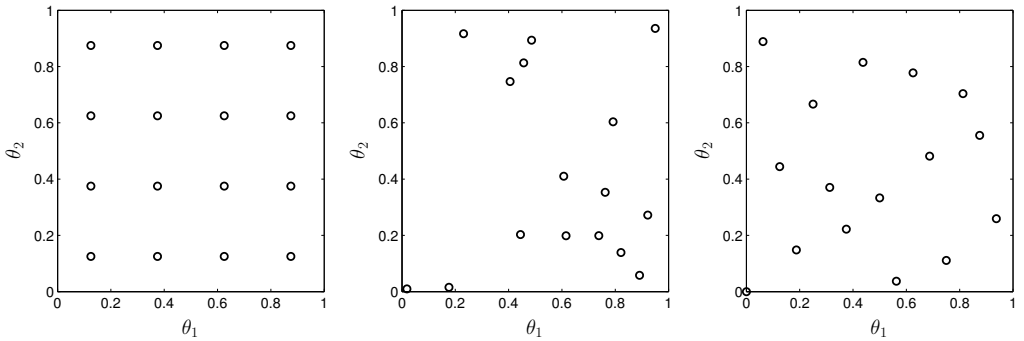
$$\mathrm{E}[e^2] = \frac{V(\mathbb{T})^2}{I^2} \sum_{i=1}^{I} \mathrm{E}\left[\left(f(\bar{\boldsymbol{\theta}}_i) - \frac{1}{V(\mathbb{T})} \int_{\mathbb{T}} f(\boldsymbol{\theta}) d\boldsymbol{\theta}\right)^2\right] = \frac{V(\mathbb{T})}{I} \sigma^2(f) \sim \frac{1}{I} \tag{2.40}$$

where we introduced

$$\sigma^2(f) = \int_{\mathbb{T}} \left[f(\boldsymbol{\theta}) - \frac{1}{V(\mathbb{T})} \int_{\mathbb{T}} f(\boldsymbol{\theta}) d\boldsymbol{\theta}\right]^2 d\boldsymbol{\theta}, \tag{2.41}$$

i.e. $\sigma^2(f)$ measures the fluctuation of $f(\boldsymbol{\theta})$ around the mean value $\frac{1}{V(\mathbb{T})} \int_{\mathbb{T}} f(\boldsymbol{\theta}) d\boldsymbol{\theta}$. Eq. (2.40) shows that the error decreases by $\mathcal{O}(I^{-1/2})$ independent of the dimension M. This result is in contrast to what we obtained for the numerical integration above and is also the reason why MC methods are so popular. Note however that $\sigma^2(f)$ in (2.40) might be quite large and therefore, for smaller dimensions, numerical integration might be a better choice.

We will now briefly explain the $\mathcal{O}(I^{-1/2})$ behaviour of MC integration. Going back to (2.39), we see that we can think of MC integration as a randomized rectangular quadrature (*midpoint rule*, see Appendix C.1.6). Figure 2.2 shows the node positions for the rectangular rule and



(a) Regular node positions **(b)** Random node positions **(c)** Quasi-random node positions

Figure 2.2: Comparison of node positions

the MC integration. Assume that $f(\boldsymbol{\theta})$ depends more strongly on θ_2 than on θ_1, then we have only four nodes in the direction of θ_2 for the regular node grid, i.e. the regular grid is not a good choice for this function and the randomized grid is better. The requirement is therefore to have a distribution of the node positions which is as uniform as possible for any projection onto a subspace [Sobol, 1979]. However, one drawback of using random node positions is that the nodes may not be homogeneously distributed on the region of integration. This is the case in Figure 2.2(b) where there is a cluster in the lower right corner. Therefore, quasi-random methods have been studied in the literature which generate samples according to a mathematical formula and spread them homogeneously over $\mathbb{T}$. An example is shown in Figure 2.2(c) where we used the *Halton sequence* [Halton, 1960]. Using such quasi-random node positions allows us to obtain convergence rates for the error which scale almost with $\mathcal{O}(I^{-1})$ in contrast to the $\mathcal{O}(I^{-1/2})$ which we observed before using random node positions, see e.g. [Cranley and Patterson, 1976; Sobol, 1979].

As we have already pointed out, the variance $\sigma^2(f)$ can be quite large. Therefore, different *variance reduction methods* have been proposed in the literature, see e.g. [Liu, 2008]. Among them, *importance sampling* is the most prominent one. It introduces a new trial function $g(\boldsymbol{\theta}) \geq 0$ from which we generate samples $\boldsymbol{\theta}_i$ instead of the uniform distribution that was used before. The trial function $g(\boldsymbol{\theta})$ should be chosen such that it is easy to generate random samples with this distribution but at the same time, it should be as close as possible to $f(\boldsymbol{\theta})$. It holds

$$\int_{\mathbb{T}} f(\boldsymbol{\theta}) d\boldsymbol{\theta} = \int_{\mathbb{T}} \frac{f(\boldsymbol{\theta})}{g(\boldsymbol{\theta})} g(\boldsymbol{\theta}) d\boldsymbol{\theta} \approx V(\mathbb{T}) \frac{\sum_{i=1}^{I} f(\bar{\boldsymbol{\theta}}_i) / g(\bar{\boldsymbol{\theta}}_i)}{\sum_{i=1}^{I} 1 / g(\bar{\boldsymbol{\theta}}_i)} \tag{2.42}$$

where $\bar{\boldsymbol{\theta}}_i$ are now $i.i.d.$ random vectors with $\bar{\boldsymbol{\theta}}_i \sim g(\boldsymbol{\theta})$. Using importance sampling now yields

the new variance $\sigma^2(f/g)$ which is minimum if $g(\boldsymbol{\theta})$ has the same shape as $f(\boldsymbol{\theta})$. Note that $g(\boldsymbol{\theta})$ does not have to be normalized to one as can be seen by (2.42).

These MC integration techniques can be applied to compute the MMSE estimate for given data $\mathbf{x}$. For completeness, we would like to mention another approach which uses the fact that the MMSE estimator is the mean of the a posteriori density. Therefore, the MMSE estimator can also be computed if we can generate samples from the a posteriori density $p(\boldsymbol{\theta}|\mathbf{x})$. This is in general a difficult task. However, if we do not restrict ourselves to generate independent samples from $p(\boldsymbol{\theta}|\mathbf{x})$ but allow them to have a weak correlation, then there are efficient methods which are called *Markov chain Monte Carlo* (MCMC) methods, see e.g. [Spall, 2003; Doucet and Wang, 2005]. The two most prominent methods are the *Metropolis-Hastings algorithm* and *Gibbs sampling*. Note that the approximation of the mean with the average is still valid although more samples may be needed to obtain an accurate result due to the correlation.

2.5 Sufficient Statistics

Up to now, we used the data $\mathbf{x}$ directly to estimate the parameter $\boldsymbol{\theta}$. However, instead of using $\mathbf{x}$, it is often interesting to use a statistic $\mathbf{t}(\mathbf{x})$ of smaller dimension which gives the same result as using the data $\mathbf{x}$. Such a statistic $\mathbf{t}(\mathbf{x})$ is called *sufficient* and was first considered by Fisher [Fisher, 1922]. In the following, we will discuss the definition of sufficient statistics for the deterministic and the Bayesian case and state the Fisher-Neyman factorization theorem. This theorem is very useful as it will allow us to show that a statistic is sufficient.

2.5.1 Deterministic Case

Let $\mathbf{x}$ be the random data about the unknown, deterministic parameter $\boldsymbol{\theta}$. Furthermore, let $\mathbf{t}(\mathbf{x})$ be a statistic of $\mathbf{x}$. Then $\mathbf{t}(\mathbf{x})$ is a sufficient statistic if $p(\mathbf{x}|\mathbf{t}(\mathbf{x}) = \mathbf{t}_0; \boldsymbol{\theta}) = p(\mathbf{x}|\mathbf{t}(\mathbf{x}) = \mathbf{t}_0)$. Therefore, given $\mathbf{t}(\mathbf{x}) = \mathbf{t}_0$, we can not infer additional information about $\boldsymbol{\theta}$ from the data $\mathbf{x}$. A convenient way to find the sufficient statistic is to use the Fisher-Neyman factorization theorem which was first introduced by Fisher in [Fisher, 1922] and later extended by Neyman in [Neyman, 1935].

Theorem 1 (Fisher-Neyman factorization)**.** *A sufficient and necessary condition for* $\mathbf{t}(\mathbf{x})$ *being a sufficient statistic for* $\boldsymbol{\theta}$ *is the factorization*

$$p(\mathbf{x}; \boldsymbol{\theta}) = g(\mathbf{t}(\mathbf{x}), \boldsymbol{\theta}) h(\mathbf{x}). \tag{2.43}$$

A proof of this theorem can e.g. be found in [Kay, 1993]. Furthermore, two additional concepts are often useful if working with sufficient statistics. They are the *minimality* and *completeness*

of a sufficient statistic. The interested reader is referred to [Mendel, 1995; Lehmann and Casella, 1998] for further details.

For the linear signal model in additive Gaussian noise in (2.16), a well known sufficient statistic is $\mathbf{t}(\mathbf{x}) = \mathbf{H}^T\mathbf{C}^{-1}\mathbf{x}$, see e.g. [Scharf, 1990]. As any one-to-one function of $\mathbf{t}(\mathbf{x})$ is also a sufficient statistic, we use the following sufficient statistic

$$\mathbf{t}(\mathbf{x}) = \mathbf{Q}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{x} \tag{2.44}$$

where $\mathbf{Q}$ is any square invertible matrix. In other words, the transform $\mathbf{T}^T\mathbf{x}$ of the signal model in (2.16) with $\mathbf{T} = \mathbf{C}^{-1}\mathbf{H}\mathbf{Q}^T$ does not result in an information loss for the estimation of $\boldsymbol{\theta}$. The new signal model is now

$$\tilde{\mathbf{x}} = \mathbf{T}^T\mathbf{x} = \mathbf{T}^T\mathbf{H}\boldsymbol{\theta} + \mathbf{T}^T\mathbf{z} = \tilde{\mathbf{H}}\boldsymbol{\theta} + \tilde{\mathbf{z}}. \tag{2.45}$$

with $\tilde{\mathbf{z}} \sim \mathcal{N}(\mathbf{0}, \tilde{\mathbf{C}})$ and $\tilde{\mathbf{C}} = \mathbf{T}^T\mathbf{C}\mathbf{T} = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}$.

2.5.2 Bayesian Case

Bayesian sufficiency is the analog concept of classical sufficiency in the deterministic case [Scharf, 1990; Ghosh et al., 2006]. Let $\mathbf{x}$ be the data about the unknown, random parameter $\boldsymbol{\theta}$. Furthermore, let $\mathbf{t}(\mathbf{x})$ be a statistic of $\mathbf{x}$. Then $\mathbf{t}(\mathbf{x})$ is a Bayesian sufficient statistic if for all priors $p(\boldsymbol{\theta})$ the a posteriori density $p(\boldsymbol{\theta}|\mathbf{x})$ can be written as $p(\boldsymbol{\theta}|\mathbf{x}) = p(\boldsymbol{\theta}|\mathbf{t}(\mathbf{x}))$. This concept was introduced by Kolmogorov in [Kolmogorov, 1942] and is called *Bayesian sufficiency*. Therefore, given $\mathbf{t}(\mathbf{x})$, we can not infer additional information about $\boldsymbol{\theta}$ from the data $\mathbf{x}$. The concept of Bayesian sufficiency is directly related to classical sufficiency in the deterministic case. Using the Fisher-Neyman factorization theorem, it is easy to proof that the classical sufficiency implies Bayesian sufficiency, as we will now show: Using the Fisher-Neyman factorization theorem in (2.43), we obtain

$$p(\boldsymbol{\theta}|\mathbf{x}) = \frac{p(\boldsymbol{\theta})p(\mathbf{x}|\boldsymbol{\theta})}{\int p(\boldsymbol{\theta})p(\mathbf{x}|\boldsymbol{\theta})d\theta} = \frac{p(\boldsymbol{\theta})g(\mathbf{t}(\mathbf{x}),\boldsymbol{\theta})}{\int p(\boldsymbol{\theta})g(\mathbf{t}(\mathbf{x}),\boldsymbol{\theta})d\theta} = p(\boldsymbol{\theta}|\mathbf{t}(\mathbf{x})) \tag{2.46}$$

for all priors $p(\boldsymbol{\theta})$. This means, $\mathbf{t}(\mathbf{x})$ is also sufficient in the Bayesian sense. The converse is also true under some regularity conditions [Heyer, 1982].

For example the MMSE estimator $\hat{\boldsymbol{\theta}}_{\text{MMSE}}(\mathbf{x})$ can then be written as

$$\hat{\boldsymbol{\theta}}_{\text{MMSE}}(\mathbf{x}) = \frac{\int \boldsymbol{\theta}\, g(\mathbf{t}(\mathbf{x}),\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}}{\int g(\mathbf{t}(\mathbf{x}),\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}}. \tag{2.47}$$

This result will be useful in Section 3.3.3 when we derive the recursive constrained MMSE estimator.

Chapter 3

Incorporation of Additional Information using Constraints

After the introduction of the basic principles from estimation theory that will be needed in this thesis, we now turn to the first possibility to improve the performance of an estimator which is to use additional information in terms of constraints. The constraints may either naturally arise from physical problems or are artificially introduced, e.g. in regularized regression problems. An example of natural constraints is a power limitation on $\boldsymbol{\theta}_0(n)$, i.e. $\|\boldsymbol{\theta}_0(n)\|^2 \leq E_{\text{max}}$ or an energy passivity constraint for the estimation of a channel impulse response (c.f. Section (3.5.3). Another important class of constraints are 1-norm constraints that are used to obtain sparse results. In [Tibshirani, 1996], for example, Tibshirani introduced the LASSO (least absolute shrinkage and selection operator) which is a least squares approach with the 1-norm constraint $\mathbb{T} = \{\boldsymbol{\theta}_0 : \|\boldsymbol{\theta}_0\|_1 \leq t\}$.

3.1 Overview

In this chapter, we will study the estimation of an unknown, time-variant parameter vector

$$\boldsymbol{\theta}_0(n) \in \mathbb{T} \subset \mathbb{R}^M \tag{3.1}$$

from observations $\mathbf{x}(n) \in \mathbb{R}^K$ of a linear, time-variant Gaussian model

$$\mathbf{x}(n) = \mathbf{H}(n)\boldsymbol{\theta}_0(n) + \mathbf{z}(n), \quad n \geq 0. \tag{3.2}$$

$\mathbf{H}(n) \in \mathbb{R}^{K\times M}$ is a known model matrix, $\mathbf{z}(n)$ is Gaussian noise with $\mathbf{z}(n) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}(n))$ which is temporally uncorrelated, i.e. $\mathrm{E}[\mathbf{z}(n_1)\mathbf{z}(n_2)^T] = \mathbf{0}$ for all $n_1 \neq n_2$, and n denotes the discrete time. In addition, we assume to know a priori $\boldsymbol{\theta}_0(n) \in \mathbb{T}$, where $\mathbb{T}$ is an arbitrary subset of $\mathbb{R}^M$. Thus, the task is to estimate $\boldsymbol{\theta}_0(n)$ for all $n \geq 0$ subject to $\boldsymbol{\theta}_0(n) \in \mathbb{T}$ given all observations $\mathbf{x}(0), \ldots, \mathbf{x}(n)$.

The aim of this chapter is to derive and compare three recursive estimators that reduce the mean squared error for the estimation of $\boldsymbol{\theta}_0(n)$ from the observations $\mathbf{x}(0), \ldots, \mathbf{x}(n)$ by exploiting the a priori knowledge $\boldsymbol{\theta}_0(n) \in \mathbb{T}$. The first estimator is the *constrained maximum likelihood* (CML) estimator which maximizes the likelihood function over $\mathbb{T}$. This estimator was e.g. used in [Van Trees, 2002; Zhu and Li, 2007] for linear equality/inequality and quadratic constraints and in [Babadi et al., 2009] for a 1-norm constraint. The second estimator we consider is the *affine minimax estimator* (AMX) [Pilz, 1986; Eldar, 2006a, 2008b] which has attracted a lot of interest in the past. The AMX minimizes the worst-case MSE for $\boldsymbol{\theta}_0(n) \in \mathbb{T}$. So far, the AMX has either only be considered for a time-invariant estimation problem or for the blind minimax tracking problem [Elron et al., 2008] and we will extend the AMX to the problem given by (3.1) and (3.2). The third estimator is the Bayesian *minimum mean squared error* (MMSE) estimator with a uniform prior on $\mathbb{T}$ [Mcleod and Quenneville, 2001; Jenkins and Watts, 1969; Robert, 2001] which we formulate in a recursive way.

The main challenge is to apply the three estimators (CML, AMX, and MMSE) to a time-varying estimation problem with an increasing number of observations $\mathbf{x}(0), \ldots, \mathbf{x}(n)$. A direct application of these estimators would lead to a growing computational complexity as the time n proceeds. The use of a suitable sufficient statistic transforms the original estimation problem to an equivalent one with a fixed number of "observations", thus resulting in a fixed computational complexity at each time step n. It will turn out that the recursive weighted least squares (RWLS) algorithm provides an efficient way to calculate this sufficient statistic. Hence, this chapter provides a uniform framework that allows the constrained tracking of $\boldsymbol{\theta}_0(n)$ from an increasing number of observations $\mathbf{x}(0), \ldots, \mathbf{x}(n)$. This unified framework, which has not been considered in the literature before, is the main contribution of this chapter.

The remainder of this chapter is organized as follows: Section 3.2 formulates the time-varying estimation problem which we will study. The sufficient statistic is derived and we show that the RWLS algorithm is an efficient way to compute the sufficient statistic. Section 3.3 is the main part of the chapter. There, we consider the CML estimator and also generalize the concepts of AMX and MMSE estimation to the time-variant signal model (3.2). We show that all three estimators can be efficiently calculated by using RWLS as a preprocessing step. After a comparison of the various approaches in Section 3.4, we show finally some simulation results in Section 3.5.

3.2 Signal Model and Sufficient Statistics

The aim of this chapter is to derive recursive estimators for the estimation of $\boldsymbol{\theta}_0(n)$ in (3.2) from the observations $\mathbf{x}(0), \ldots, \mathbf{x}(n)$ subject to the a priori constraint $\boldsymbol{\theta}_0(n) \in \mathbb{T}$. Assume that $\boldsymbol{\theta}_0(n)$ is slowly time-varying, such an adaptive estimation (tracking) is feasible. Given the instantaneous signal model (3.2) without any constraint on $\boldsymbol{\theta}_0(n)$, the following exponentially weighted

least squares (WLS) estimator is popular in adaptive filtering (see [Haykin, 2002; Farrell and Polycarpou, 2006])

$$\hat{\boldsymbol{\theta}}_{\text{WLS}}(n) = \arg\min_{\boldsymbol{\theta}(n)} \sum_{i=0}^{n} \beta^{n-i} \left(\mathbf{x}(i) - \mathbf{H}(i)\boldsymbol{\theta}(n)\right)^T \mathbf{C}(i)^{-1} \left(\mathbf{x}(i) - \mathbf{H}(i)\boldsymbol{\theta}(n)\right) \tag{3.3}$$

where $0 < \beta \leq 1$ is the exponential forgetting factor. This estimator can be efficiently calculated by the recursive weighted least squares (RWLS) algorithm as we will see later. The solution $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$ of (3.3) can also be interpreted as the ML estimator of the following stacked signal model

$$\mathbf{x}_n = \mathbf{H}_n\boldsymbol{\theta}(n) + \mathbf{z}_n. \tag{3.4}$$

$\mathbf{x}_n$, $\mathbf{H}_n$, and $\mathbf{z}_n$ are stacked versions of $\mathbf{x}(i)$, $\mathbf{H}(i)$, and $\mathbf{z}(i)$ for $i = 0, \ldots, n$, respectively:

$$\begin{aligned} \mathbf{x}_n &= \begin{bmatrix} \mathbf{x}(n)^T & \cdots & \mathbf{x}(0)^T \end{bmatrix}^T, \\ \mathbf{H}_n &= \begin{bmatrix} \mathbf{H}(n)^T & \cdots & \mathbf{H}(0)^T \end{bmatrix}^T, \\ \mathbf{z}_n &= \begin{bmatrix} \mathbf{z}(n)^T & \beta^{-1/2}\mathbf{z}(n-1)^T & \cdots & \beta^{-n/2}\mathbf{z}(0)^T \end{bmatrix}^T. \end{aligned} \tag{3.5}$$

Note that the noise term $\mathbf{z}_n$ is not merely the stacking of $\mathbf{z}(i)$, but also incorporates an intentional weighting of $\mathbf{z}(i)$ with $\beta^{-(n-i)/2}$. This results in a tendentiously larger noise (as it is in reality) for past measurements, marking them to be less valuable than recent observations for the estimation of $\boldsymbol{\theta}_0(n)$. The effect is exactly the same as the exponential down-weighting of error terms caused by past observations in (3.3). As the noise $\mathbf{z}(n)$ is temporally uncorrelated, we have $\mathbf{z}_n \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_n)$ where the covariance matrix $\mathbf{C}_n \in \mathbb{R}^{(n+1)K \times (n+1)K}$ is a block diagonal matrix

$$\mathbf{C}_n = \operatorname{diag}\left(\mathbf{C}(n), \beta^{-1}\mathbf{C}(n-1), \cdots, \beta^{-n}\mathbf{C}(0)\right). \tag{3.6}$$

The unconstrained maximum likelihood estimate to (3.4) is given by

$$\min_{\boldsymbol{\theta}(n)} \left(\mathbf{x}_n - \mathbf{H}_n\boldsymbol{\theta}(n)\right)^T \mathbf{C}_n^{-1} \left(\mathbf{x}_n - \mathbf{H}_n\boldsymbol{\theta}(n)\right). \tag{3.7}$$

It is identical to the WLS estimator (3.3). Using the stacked matrices $\mathbf{H}_n$, $\mathbf{C}_n$ and $\mathbf{x}_n$ defined in (3.5) and (3.6), the solution to (3.7) can be written as

$$\hat{\boldsymbol{\theta}}_{\text{WLS}}(n) = \mathbf{R}_n^{-1}\mathbf{H}_n^T\mathbf{C}_n^{-1}\mathbf{x}_n \quad \text{with} \quad \mathbf{R}_n = \mathbf{H}_n^T\mathbf{C}_n^{-1}\mathbf{H}_n \tag{3.8}$$

if $\mathbf{C}_n$ is invertible and $\mathbf{H}_n$ has a full column rank. It is obvious that $\boldsymbol{\theta}(n)$ in (3.4) is different from $\boldsymbol{\theta}_0(n)$ in (3.2) as (3.4) is an approximation of the original signal model (3.2) for a slowly varying parameter $\boldsymbol{\theta}_0(n)$. This approximation, however, has proved to be useful in many applications. If $\boldsymbol{\theta}_0(n) = \boldsymbol{\theta}_0$ is constant, $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$ will approach $\boldsymbol{\theta}_0$ with increasing n. For the case that $\boldsymbol{\theta}_0(n)$

changes slowly, $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$ will follow $\boldsymbol{\theta}_0(n)$. Therefore, we will use the stacked signal model (3.4) below and consider in addition the a priori constraint (3.1). We look for estimators for the signal model

$$\mathbf{x}_n = \mathbf{H}_n\boldsymbol{\theta}(n) + \mathbf{z}_n, \quad \boldsymbol{\theta}(n) \in \mathbb{T} \tag{3.9}$$

with a smaller MSE than the previous unconstrained WLS estimator in (3.8). Note that (3.9) describes a frequentist estimation problem as $\boldsymbol{\theta}(n)$ is deterministic. The length of $\mathbf{x}_n$, $\mathbf{H}_n$ and $\mathbf{C}_n$ grows linearly as the time n proceeds.

In order to avoid estimators with a growing computational complexity, we perform a signal model transform by using the concept of sufficient statistic as introduced in Section 2.5. Using (2.44) and (2.45) for the special choice of $\mathbf{Q} = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}$ yields the transformed signal model $\tilde{\mathbf{x}} = \boldsymbol{\theta} + \tilde{\mathbf{z}}$ with $\tilde{\mathbf{z}} \sim \mathcal{N}(\mathbf{0}, \tilde{\mathbf{C}})$ and the new covariance matrix $\tilde{\mathbf{C}} = \mathbf{T}^T\mathbf{C}\mathbf{T} = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}$. Therefore, instead of the stacked signal model (3.9), we can look for estimators based on the simpler model $\tilde{\mathbf{x}}_n = \boldsymbol{\theta}(n) + \tilde{\mathbf{z}}_n$ with $\tilde{\mathbf{z}}_n \sim \mathcal{N}(\mathbf{0}, \tilde{\mathbf{C}}_n)$. The transformed observation vector $\tilde{\mathbf{x}}_n$ and the new covariance matrix $\tilde{\mathbf{C}}_n$ are given by

$$\tilde{\mathbf{x}}_n = \mathbf{t}(\mathbf{x}) = \left(\mathbf{H}_n^T\mathbf{C}_n^{-1}\mathbf{H}_n\right)^{-1}\mathbf{H}_n^T\mathbf{C}_n^{-1}\mathbf{x}_n = \hat{\boldsymbol{\theta}}_{\text{WLS}}(n), \tag{3.10a}$$

$$\tilde{\mathbf{C}}_n = \left(\mathbf{H}_n^T\mathbf{C}_n^{-1}\mathbf{H}_n\right)^{-1} = \mathbf{R}_n^{-1}. \tag{3.10b}$$

where $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$ and $\mathbf{R}_n$ are defined in (3.8). The main advantage of this model transform is the fixed vector length M of $\tilde{\mathbf{x}}_n$ in comparison to the linearly growing vector length $(n+1)K$ of $\mathbf{x}_n$.

An efficient way to compute the sufficient statistic is to use the RWLS algorithm. According to the definition of $\mathbf{H}_n$ in (3.5) and $\mathbf{C}_n$ in (3.6), both matrices can be written recursively

$$\mathbf{H}_{n+1} = \begin{bmatrix} \mathbf{H}(n+1) \\ \mathbf{H}_n \end{bmatrix}, \qquad \mathbf{C}_{n+1} = \begin{bmatrix} \mathbf{C}(n+1) & \mathbf{0} \\ \mathbf{0} & \beta^{-1}\mathbf{C}_n \end{bmatrix},$$

which also implies

$$\mathbf{R}_{n+1} = \mathbf{H}_{n+1}^T\mathbf{C}_{n+1}^{-1}\mathbf{H}_{n+1} = \beta\mathbf{R}_n + \mathbf{H}(n+1)^T\mathbf{C}(n+1)^{-1}\mathbf{H}(n+1). \tag{3.11}$$

By applying the matrix inversion lemma

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}\left(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B}\right)^{-1}\mathbf{DA}^{-1}$$

we can update $\mathbf{R}_n^{-1}$ and the parameter estimate $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$ in (3.8) in a time recursive way. Introducing the gain matrix $\mathbf{G}_{n+1} = \mathbf{R}_{n+1}^{-1}\mathbf{H}(n+1)^T\mathbf{C}(n+1)^{-1}$, we finally obtain the RWLS

algorithm [Farrell and Polycarpou, 2006]

$$\hat{\boldsymbol{\theta}}_{\text{WLS}}(n+1) = \hat{\boldsymbol{\theta}}_{\text{WLS}}(n) + \mathbf{G}_{n+1}\left(\mathbf{x}(n+1) - \mathbf{H}(n+1)\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)\right), \tag{3.12a}$$

$$\mathbf{R}_{n+1}^{-1} = \frac{1}{\beta}\left(\mathbf{R}_n^{-1} - \mathbf{G}_{n+1}\mathbf{H}(n+1)\mathbf{R}_n^{-1}\right). \tag{3.12b}$$

At each time step, RWLS updates the sufficient statistic $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$ and the inverse correlation matrix $\mathbf{R}_n^{-1}$. Both values are needed by the three recursive estimators below subject to the constraint $\boldsymbol{\theta}(n) \in \mathbb{T}$.

3.3 Adaptive Estimation for a Linear Time-Variant Gaussian Model with General Constraints

3.3.1 Recursive Constrained ML Estimator

The first estimator we consider is the constrained maximum likelihood (CML) estimator [Bard, 1973]. Given the linear signal model $\mathbf{x} = \mathbf{H}\boldsymbol{\theta} + \mathbf{z}$, $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{C})$ with the constraint $\boldsymbol{\theta} \in \mathbb{T}$, the CML is given by the optimization problem

$$\hat{\boldsymbol{\theta}}_{\text{CML}} = \arg\max_{\boldsymbol{\theta}} p(\mathbf{x};\boldsymbol{\theta}) \;\text{ s.t. }\; \boldsymbol{\theta} \in \mathbb{T}. \tag{3.13}$$

Using the sufficient statistic $\hat{\boldsymbol{\theta}}_{\text{WLS}}$ in (2.44), we can rewrite (3.13) into

$$\hat{\boldsymbol{\theta}}_{\text{CML}} = \arg\min_{\boldsymbol{\theta}} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{WLS}})^T (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{WLS}}) \;\text{ s.t. }\; \boldsymbol{\theta} \in \mathbb{T}. \tag{3.14}$$

As $\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H} \succ \mathbf{0}$, we have to check whether $\hat{\boldsymbol{\theta}}_{\text{WLS}}$ satisfies the constraint or not. If $\hat{\boldsymbol{\theta}}_{\text{WLS}} \in \mathbb{T}$, then $\hat{\boldsymbol{\theta}}_{\text{CML}} = \hat{\boldsymbol{\theta}}_{\text{WLS}}$. Otherwise we have to find the minimum of $(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{WLS}})^T(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{WLS}})$ on the boundary of $\mathbb{T}$. For the special case that we can write $\boldsymbol{\theta} \in \mathbb{T}$ as a *linear matrix inequality* (LMI), we can use Schur's lemma[17] to rewrite (3.14) into

$$\hat{\boldsymbol{\theta}}_{\text{CML}} = \arg\min_{\tau,\boldsymbol{\theta}} \; \tau \;\text{ s.t. }\; \begin{bmatrix} (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1} & \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{WLS}} \\ (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{WLS}})^T & \tau \end{bmatrix} \succeq \mathbf{0} \;\text{ and }\; \boldsymbol{\theta} \in \mathbb{T}, \tag{3.15}$$

which is a *semidefinite program* (SDP) and can be solved efficiently [Boyd and Vandenberghe, 2007].

[17] Schur's lemma can be compactly stated as [Boyd and Vandenberghe, 2007]

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{B}^T & \mathbf{C} \end{bmatrix} \succeq \mathbf{0} \quad \text{if and only if} \quad \mathbf{A} \succ \mathbf{0} \text{ and } \mathbf{C} - \mathbf{B}^T\mathbf{A}^{-1}\mathbf{B} \succeq \mathbf{0}.$$

The extension of the CML to our time-varying signal model (3.9) is now straightforward. All we have to do is to replace $\hat{\boldsymbol{\theta}}_{\text{WLS}}$, $\mathbf{C}$ and $\mathbf{H}$ by $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$, $\mathbf{C}_n$ and $\mathbf{H}_n$ in (3.14). Therefore, for each time step we have to check $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n) \in \mathbb{T}$. If this is satisfied then $\hat{\boldsymbol{\theta}}_{\text{RCML}}(n) = \hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$. Otherwise, we have to find the minimum of $(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{WLS}}(n))^T(\mathbf{H}_n^T\mathbf{C}_n^{-1}\mathbf{H}_n)(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{WLS}}(n))$ on the boundary of $\mathbb{T}$. We call this estimator the *recursive constrained ML (RCML)* estimator.

3.3.2 Recursive Constrained Affine Minimax Estimator

The extension of the AMX (2.20) to our time-variant linear Gaussian model is straightforward. The unbiased efficient estimator $\hat{\boldsymbol{\theta}}_{\text{eff}}(n)$ for (3.9) is identical to $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$ in (3.8) and can be calculated efficiently by the RWLS algorithm. After each update of $\hat{\boldsymbol{\theta}}_{\text{eff}}(n)$, we multiply it with $\mathbf{I} + \mathbf{M}(n)$ and add $\mathbf{u}(n)$ to obtain the *recursive AMX* (RAMX) estimator

$$\hat{\boldsymbol{\theta}}_{\text{RAMX}}(n) = \big(\mathbf{I} + \mathbf{M}(n)\big)\,\hat{\boldsymbol{\theta}}_{\text{eff}}(n) + \mathbf{u}(n). \tag{3.16}$$

We have to solve the minimax problem (2.22) to obtain $\mathbf{M}(n)$ and $\mathbf{u}(n)$. The covariance matrix of $\hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$ in (2.22) is given by $\mathbf{R}_n^{-1} = \left(\mathbf{H}_n^T\mathbf{C}_n^{-1}\mathbf{H}_n\right)^{-1}$ which is also updated by the RWLS algorithm. Since $\mathbf{R}_n^{-1}$ is time-varying, the minimax problem (2.22) has to be solved repeatedly at each time step.

In the stationary case with $\mathbf{H}(n) = \mathbf{H}$ and $\mathbf{C}(n) = \mathbf{C}$, we can avoid the repeated solution of the minimax problem (2.22) by approximating $\mathbf{R}_n$ by its steady state value

$$\mathbf{R}_\infty = \lim_{n\to\infty} \mathbf{R}_n = \lim_{n\to\infty} \sum_{k=0}^{n} \beta^{n-k}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H} = \frac{1}{1-\beta}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H},\ (0 < \beta < 1). \tag{3.17}$$

In this case, we only need to compute $\mathbf{M}$ and $\mathbf{u}$ once from $\mathbf{R}_\infty$ and use them to change the efficient estimator $\hat{\boldsymbol{\theta}}_{\text{eff}}(n)$ to the RAMX estimator $\hat{\boldsymbol{\theta}}_{\text{RAMX}}(n)$ according to (3.16). This approach reduces the computational complexity of RAMX at the expense of an increased mean squared error for small n.

3.3.3 Recursive Constrained MMSE Estimator

The third approach we use after the calculation of the sufficient statistic is the MMSE estimation. In contrast to the two frequentist approaches (ML and AMX), we now assume $\boldsymbol{\theta}$ to be a random vector and recast the constraint $\boldsymbol{\theta} \in \mathbb{T}$ as an a priori PDF $p(\boldsymbol{\theta})$. In this way, we have reformulated the frequentist estimation problem ($\boldsymbol{\theta}$ deterministic) to a Bayesian one. We assume a uniform a priori PDF[18] (a natural choice as we have no information about $\boldsymbol{\theta}$ except for $\boldsymbol{\theta} \in \mathbb{T}$)

[18] Note that this prior PDF can be *improper* [Robert, 2001] if $\mathbb{T}$ is an unbounded set, i.e. it does not sum to one. Nevertheless, one can often compute (3.19) and therefore the RMMSE estimator still exists and is useful.

$$p(\boldsymbol{\theta}) = \begin{cases} \text{const} & \boldsymbol{\theta} \in \mathbb{T} \\ 0 & \text{otherwise} \end{cases}. \tag{3.18}$$

The MMSE estimator (2.27b) for this prior simplifies to

$$\hat{\boldsymbol{\theta}}_{\text{MMSE}} = \frac{\int\limits_{\mathbb{T}} \boldsymbol{\theta}\, p(\mathbf{x}|\boldsymbol{\theta}) d\boldsymbol{\theta}}{\int\limits_{\mathbb{T}} p(\mathbf{x}|\boldsymbol{\theta}) d\boldsymbol{\theta}}. \tag{3.19}$$

This estimator is sometimes called the *mean likelihood estimator* (MELE) [Jenkins and Watts, 1969; Mcleod and Quenneville, 2001] as (3.19) calculates the mean of the normalized likelihood function $p(\mathbf{x}|\boldsymbol{\theta}) / \int_{\mathbb{T}} p(\mathbf{x}|\boldsymbol{\theta}) d\boldsymbol{\theta}$.

Using the principle of Bayesian sufficient statistics as introduced in Section 2.5.2 significantly simplifies the adaptive MMSE estimation which is now split into two steps: First, we apply the RWLS algorithm to simplify the signal model and to compute the sufficient statistic $\tilde{\mathbf{x}}_n = \hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$. Then, we calculate the MMSE estimator from the simplified signal model $\tilde{\mathbf{x}}_n = \boldsymbol{\theta}(n) + \tilde{\mathbf{z}}_n$ with a uniform a priori distribution of $\boldsymbol{\theta}(n)$ in $\mathbb{T}$. We refer to this estimator as *recursive MMSE* (RMMSE) estimator.

Note that the MMSE estimation in the second step still involves two M-dimensional integrals which we in general cannot compute analytically. In 3.5.2, we will consider the case of ellipsoidal constraints on $\boldsymbol{\theta}$ and we will show that the computational complexity for this particular constraint can be reduced to the calculation of M one-dimensional integrals. Here, we apply a more general statistical approach which is not restricted to ellipsoidal constraints only. We use the Monte Carlo method to approximate the integrals. Since $\tilde{\mathbf{x}}_n|\boldsymbol{\theta} \sim \mathcal{N}(\boldsymbol{\theta}, \tilde{\mathbf{C}}_n)$ after the signal model transform and $\boldsymbol{\theta}$ is uniformly distributed in $\mathbb{T}$, we only need to generate random samples from a truncated multivariate Gaussian distribution, i.e. of a Gaussian distribution which is set to zero for $\boldsymbol{\theta} \notin \mathbb{T}$. This becomes clear if we consider the MMSE estimator for the simplified model

$$\hat{\boldsymbol{\theta}}_{\text{RMMSE}}(n) = \frac{\int\limits_{\mathbb{T}} \boldsymbol{\theta}\, \exp\{-\frac{1}{2}(\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)^T \tilde{\mathbf{C}}_n^{-1} (\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)\} d\boldsymbol{\theta}}{\int\limits_{\mathbb{T}} \exp\{-\frac{1}{2}(\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)^T \tilde{\mathbf{C}}_n^{-1} (\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)\} d\boldsymbol{\theta}}. \tag{3.20}$$

Using Monte Carlo integration [Liu, 2008], we have the approximation

$$\hat{\boldsymbol{\theta}}_{\text{RMMSE}}(n) \approx \frac{1}{I} \sum_{i=1}^{I} \boldsymbol{\theta}_i \tag{3.21}$$

where $\boldsymbol{\theta}_i$ is a random sample drawn from $\exp\{-\frac{1}{2}(\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)^T \tilde{\mathbf{C}}_n^{-1} (\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)\}$ satisfying $\boldsymbol{\theta}_i \in \mathbb{T}$. Therefore, $\boldsymbol{\theta}_i$ is a sample of a truncated Gaussian. To generate $\boldsymbol{\theta}_i$, we first draw a sample $\boldsymbol{\psi}$ from $\mathcal{N}(\mathbf{0}, \mathbf{I})$ and then transform it by $\boldsymbol{\varphi} = \mathbf{U}_n \boldsymbol{\psi} + \tilde{\mathbf{x}}_n$ with $\mathbf{U}_n \mathbf{U}_n^T = \tilde{\mathbf{C}}_n$. It is accepted as a

new value for $\boldsymbol{\theta}_i$ if $\boldsymbol{\varphi} \in \mathbb{T}$. Otherwise, it is rejected.

By doing so, we can include any constraints on $\boldsymbol{\theta}$ easily into the generation of samples from $p(\boldsymbol{\theta})p(\tilde{\mathbf{x}}_n|\boldsymbol{\theta})$. This rejection sampling works well when the truncated Gaussian PDF has enough "probability mass" in $\mathbb{T}$ as otherwise we would do many rejections before a sample $\boldsymbol{\varphi}$ is accepted. This happens, for example, for equality constraints. Such constraints cannot be handled by a rejection scheme, see Section 3.5.2 for a series expansion approach for this kind of constraints.

Note that there are also more sophisticated methods to draw multivariate samples from a truncated Gaussian, see e.g. [Robert, 1995]. They often yield a smaller rejection rate than the simple scheme presented here but have the drawback that they are only applicable if it is possible to easily compute the lower and upper truncation bounds of the PDF of the truncated Gaussian random variable.

3.4 Discussion of the Approaches

In this section, we compare all three approaches to solve our time-variant estimation problem (3.9). First, we introduce the concept of *hard* and *soft* constraints: If we need to ensure that the estimator satisfies $\hat{\boldsymbol{\theta}}(n) \in \mathbb{T}$ at any time instance n, we denote it as a hard constraint. If, however, the estimator is allowed to return a value $\hat{\boldsymbol{\theta}}(n)$ which does not lie in $\mathbb{T}$ but has a smaller MSE, we speak of a soft constraint[19].

From Section 3.3.1 and 3.3.2, we know that RCML and RAMX are frequentist approaches assuming $\boldsymbol{\theta}$ to be deterministic. On the other hand, Section 3.3.3 showed that RMMSE follows a different philosophy and assumes $\boldsymbol{\theta}(n)$ to be random with an a priori density and thus belongs to the class of Bayesian estimators. The native estimation problem in (3.9) is a frequentist problem with the additional information that $\boldsymbol{\theta}(n)$ has to lie in $\mathbb{T}$. We have four options:

1. Simply ignore the additional information $\boldsymbol{\theta}(n) \in \mathbb{T}$ and use the ML estimator for (3.9). This yields the WLS estimator (3.3) and the RWLS algorithm (3.12) for an efficient calculation.

2. A second option is to consider the constrained ML estimator for (3.9) which is given in (3.14). As $\mathbf{C}_n \succ \mathbf{0}$, we know that the solution is either the WLS in (3.8) if $\hat{\boldsymbol{\theta}}_{\text{WLS}} \in \mathbb{T}$ or otherwise it lies on the boundary of $\mathbb{T}$. The advantage of this method is that it will ensure $\hat{\boldsymbol{\theta}} \in \mathbb{T}$, i.e. the a priori information is taken into account as a hard constraint. However, the drawback of this method is that there are no efficient recursive algorithms for the case

[19] The notion of hard- and soft-constrained estimation was also used in [Benavoli et al., 2006] where hard-constrained estimation refers to a constrained LS estimator which is equivalent to a MAP with uniform prior on the feasible set. The soft-constrained estimator that Benavoli considers is the MAP with a Gaussian prior where the prior parameters are found via solving a convex optimization problem [Benavoli et al., 2007]. [Benavoli et al., 2006] provides an interesting discussion of both estimation approaches which also match with our observations.

that $\hat{\boldsymbol{\theta}}_{\text{WLS}} \notin \mathbb{T}$. Then, we have the original problem (3.14) at each time step. An exception is the case of linear equality/inequality, quadratic inequality constraints as considered in [Van Trees, 2002; Zhu and Li, 2007] or 1-norm constraints as in [Babadi et al., 2009]. In general, however, the constrained ML is only suited for practical online implementations where the constraints can be expressed such that the corresponding optimization problem can be solved efficiently. One example is to express the constraints as LMIs and the resulting problem is then a semidefinite program as illustrated in Section 3.5.3.

3. Restrict the attention to the class of affine estimators only and find the affine estimator that has the smallest worst-case MSE for all $\boldsymbol{\theta}(n)$ in $\mathbb{T}$. By doing this, the a priori information is taken into account as a soft constraint by using the minimax approach. Although this recursive AMX has in general an even higher computational complexity than RCML, it is in our view still to be favored due to the following three reasons: First, the AMX estimator often shows a smaller MSE compared to CML. Second, if $\mathbf{H}(n)$ and $\mathbf{C}(n)$ are constant over time, $\mathbf{M}(n)$ and $\mathbf{u}(n)$ reach their steady state value after some time steps and hence, no update of $\mathbf{M}(n)$ and $\mathbf{u}(n)$ is needed anymore. Third, we can restrict $\mathbf{M}(n)$ and $\mathbf{u}(n)$ to have a special structure, e.g. $\mathbf{M}(n)$ is only allowed to be diagonal. By this we can reduce the number of optimization variables and consequently the computational complexity.

4. Another possibility is to recast the frequentist problem to a Bayesian one by interpreting the a priori information $\boldsymbol{\theta} \in \mathbb{T}$ as a priori PDF $p(\boldsymbol{\theta})$ and thus $\boldsymbol{\theta} \in \mathbb{T}$ is used as a soft constraint. We choose $p(\boldsymbol{\theta})$ as uniform in $\mathbb{T}$ and zero otherwise as we do not have any other information about $\boldsymbol{\theta}$ except for $\boldsymbol{\theta} \in \mathbb{T}$. This is a natural choice as it does not prefer particular values of $\boldsymbol{\theta}$ and can be motivated by the concept of maximum entropy, see Section 4.2.3. In [Bard, 1973] it is discussed whether such a choice of the a priori density is useful or not. The main criticism against a uniform prior is that it is not invariant against reparameterization of the signal model. If we consider $\mathbf{g}(\boldsymbol{\theta})$ as a new parameter vector, still no a priori information is available about $\mathbf{g}(\boldsymbol{\theta})$ and thus the new a priori density of $\mathbf{g}(\boldsymbol{\theta})$ should also be uniform which, in general, is not the case. Therefore, the Jeffreys prior [Robert, 2001] which is invariant with respect to reparameterization is sometimes favored instead of a uniform prior. However, in this chapter we consider parameter estimation from real physical problems for which there is a natural choice of $\boldsymbol{\theta}$. Hence $p(\boldsymbol{\theta})$ has not to be invariant with respect to reparameterization [Robert, 2001]. As we use rejection sampling to incorporate the a priori information, the RMMSE estimator has the advantage compared to RCML and RAMX that it can handle quite complicated constraints.

In summary, all three approaches (RCML, RAMX and RMMSE with a uniform prior) can be viewed as different methods to model the a priori information $\boldsymbol{\theta} \in \mathbb{T}$ of the frequentist problem

(3.9). Especially the RAMX and RMMSE estimator allow to efficiently incorporate the a priori information into the estimation problem.

Note that the RAMX and the RMMSE estimator will in general not ensure $\hat{\boldsymbol{\theta}} \in \mathbb{T}$. The a priori knowledge is only exploited to achieve a smaller mean squared error of $\hat{\boldsymbol{\theta}}$. Consider, for example, the case that $\boldsymbol{\theta}$ has a uniform prior distribution in $\mathbb{T} = \{\boldsymbol{\theta} : R_1^2 \leq \|\boldsymbol{\theta}\|^2 \leq R_2^2\}$, a spherical shell with inner radius R_1 and outer radius R_2. Clearly, the a priori estimate (a priori mean) of $\boldsymbol{\theta}$ is $\hat{\boldsymbol{\theta}} = \mathbf{0}$ which is outside $\mathbb{T}$. For observations with a low SNR, both AMX and MMSE estimator still prefer an estimate that lies near the a priori estimate since the noisy observations have a smaller contribution to the estimate than the a priori information. This is advantageous in terms of a smaller MSE though $\hat{\boldsymbol{\theta}} \notin \mathbb{T}$. Relaxing the constraint as a soft constraint allows simpler recursive formulations of the estimators. Interestingly, the RMMSE estimator satisfies the constraint $\hat{\boldsymbol{\theta}} \in \mathbb{T}$ automatically if $\mathbb{T}$ is a convex set [Boyd and Vandenberghe, 2007]. This is not the case for the RAMX estimator, see [Pilz, 1986] for a discussion of this issue.

The final discussion whether hard or soft constraints should be used depends on the application. Sometimes, $\hat{\boldsymbol{\theta}} \in \mathbb{T}$ is more important than a smaller MSE value and then the RCML estimator should be used as this estimator will always ensure $\hat{\boldsymbol{\theta}} \in \mathbb{T}$. If, however, we only want to exploit the additional information to improve the estimation, then the RAMX and RMMSE approach should be favored.

3.5 Examples

In this section, we will now study three different examples to see the performance of the proposed time-varying estimators. The first example in Section 3.5.1 will consider the constraint that the parameter vector $\boldsymbol{\theta}$ has to lie inside an ellipsoid. The estimators are compared in terms of the MSE and it is also studied how many samples I are needed for the Monte-Carlo integration in (3.21). The second example in Section 3.5.2 will again study ellipsoidal constraints and it extends the first example by also considering the case that the parameter vector has to lie on the ellipsoid. Such constraints cannot be handled by the rejection method which was proposed in Section 3.3.3 and therefore we will show another way to compute both integrals in (3.20). The last example in Section 3.5.3 finally considers the problem of tracking a time-varying room impulse response where the estimated impulse response has to fulfill an energy conservation constraint that is used to improve the performance.

3.5.1 Example 1: Ellipsoidal Constraints

The following example compares the introduced estimators in terms of the mean squared error. We consider the estimation of $\boldsymbol{\theta}(n) \in \mathbb{T} \subset \mathbb{R}^2$, i.e. $M = 2$, under the ellipsoidal constraint $\mathbb{T} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^T \mathbf{A} \boldsymbol{\theta} \leq 1\}$ with

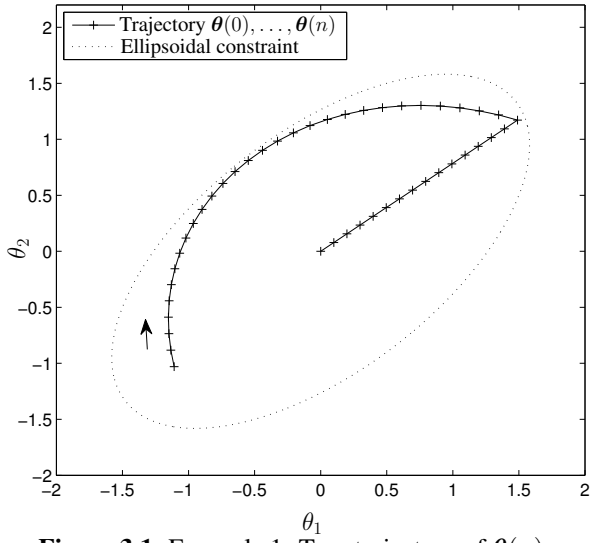


Figure 3.1: Example 1: True trajectory of $\boldsymbol{\theta}(n)$

$$\mathbf{A} = \begin{bmatrix} \cos(\frac{\pi}{4}) & \sin(\frac{\pi}{4}) \\ -\sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix}^T \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{4} \end{bmatrix} \begin{bmatrix} \cos(\frac{\pi}{4}) & \sin(\frac{\pi}{4}) \\ -\sin(\frac{\pi}{4}) & \cos(\frac{\pi}{4}) \end{bmatrix}.$$

Figure 3.1 shows the trajectory of $\boldsymbol{\theta}(n)$ (solid line) consisting of $N = 45$ discrete-time points and the boundary of $\mathbb{T}$ (dotted line). The signal model is given by $\mathbf{x}(n) = \boldsymbol{\theta}(n) + \mathbf{z}(n)$, with $\mathbf{z}(n) \sim \mathcal{N}(\mathbf{0}, 0.1\mathbf{I})$. This means $\mathbf{H}(n) = \mathbf{I}$.

We use the following five instantaneous estimators to estimate $\boldsymbol{\theta}(n)$ given only the observation $\mathbf{x}(n)$:

- *Least squares (LS)*: The LS estimator is given by

$$\hat{\boldsymbol{\theta}}_{\mathrm{LS}}(n) = \left(\mathbf{H}(n)^T\mathbf{H}(n)\right)^{-1}\mathbf{H}(n)^T\mathbf{x}(n) = \mathbf{x}(n).$$

  This means, LS accepts each observation $\mathbf{x}(n)$ as a new estimate and does not take the old observations $\mathbf{x}(n-1), \ldots, \mathbf{x}(0)$ into account.

- *Constrained maximum likelihood (CML)*: Consider the case of the quadratic constraint $\tilde{\mathbb{T}} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^T\mathbf{A}\boldsymbol{\theta} + 2\mathbf{b}^T\boldsymbol{\theta} + c \leq 0, \mathbf{A} \succ \mathbf{0}\}$ which also includes $\mathbb{T} = \{\boldsymbol{\theta} : \boldsymbol{\theta}^T\mathbf{A}\boldsymbol{\theta} \leq 1\}$ from this example. The problem in (3.14) is then a *quadratically constrained quadratic program* (QCQP), see Appendix B. In the following, we will use the *Karush-Kuhn-Tucker* (KKT) conditions [Boyd and Vandenberghe, 2007] which provide a set of sufficient conditions[20] for a solution of (3.14) with $\tilde{\mathbb{T}}$.

[20]The KKT conditions are in general only necessary. However, in our case we have a convex optimization problem and therefore they are also sufficient, see e.g. [Boyd and Vandenberghe, 2007]

The KKT conditions for the optimization problem (3.14) with constraint $\hat{\boldsymbol{\theta}}_{\text{CML}}(n) \in \tilde{\mathbb{T}}$ are given by

(a) Stationarity:

$$\left.\frac{\partial L(\boldsymbol{\theta},\lambda)}{\partial \boldsymbol{\theta}}\right|_{\substack{\boldsymbol{\theta}=\hat{\boldsymbol{\theta}}_{\text{CML}}(n)\\ \lambda=\lambda^*}} = 2(\hat{\boldsymbol{\theta}}_{\text{CML}}(n) - \hat{\boldsymbol{\theta}}_{\text{LS}}(n))^T\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H} + 2\lambda^*(\hat{\boldsymbol{\theta}}_{\text{CML}}(n)^T\mathbf{A} + \mathbf{b}^T) = \mathbf{0}^T,$$

(b) Primal feasibility:

$$\hat{\boldsymbol{\theta}}_{\text{CML}}(n)^T\mathbf{A}\hat{\boldsymbol{\theta}}_{\text{CML}}(n) + 2\mathbf{b}^T\hat{\boldsymbol{\theta}}_{\text{CML}}(n) + c \leq 0,$$

(c) Dual feasibility:

$$\lambda^* \geq 0,$$

(d) Complementary slackness:

$$\lambda^*(\hat{\boldsymbol{\theta}}_{\text{CML}}(n)^T\mathbf{A}\hat{\boldsymbol{\theta}}_{\text{CML}}(n) + 2\mathbf{b}^T\hat{\boldsymbol{\theta}}_{\text{CML}}(n) + c) = 0,$$

where $L(\boldsymbol{\theta})$ is the *Lagrangian function*, which is defined as

$$L(\boldsymbol{\theta},\lambda) = (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{LS}}(n))^T(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\text{LS}}(n)) + \lambda\left(\boldsymbol{\theta}^T\mathbf{A}\boldsymbol{\theta} + 2\mathbf{b}^T\boldsymbol{\theta} + c\right).$$

Now, we have to distinguish two cases:

(i) If $\lambda^* = 0$, then we immediately see from the stationarity condition (a) that $\hat{\boldsymbol{\theta}}_{\text{CML}}(n) = \hat{\boldsymbol{\theta}}_{\text{LS}}(n)$ and therefore, assuming that $\hat{\boldsymbol{\theta}}_{\text{LS}}(n) \in \tilde{\mathbb{T}}$, we see that all KKT conditions are fulfilled and the CML estimator is given by the LS estimator as we already discussed in Section 3.3.1.

(ii) If $\lambda^* > 0$, then the complementary condition (d) implies $\hat{\boldsymbol{\theta}}_{\text{CML}}(n)^T\mathbf{A}\hat{\boldsymbol{\theta}}_{\text{CML}}(n) + 2\mathbf{b}^T\hat{\boldsymbol{\theta}}_{\text{CML}}(n) + c = 0$, i.e. the CML estimator has to lie on the boundary of $\tilde{\mathbb{T}}$. Furthermore, the stationarity condition (a) results in

$$\begin{aligned}\hat{\boldsymbol{\theta}}_{\text{CML}}(n) &= \hat{\boldsymbol{\theta}}_{\text{CML}}(n,\lambda^*)\\ &= \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H} + \lambda^*\mathbf{A}\right)^{-1}\left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\hat{\boldsymbol{\theta}}_{\text{LS}}(n) - \lambda^*\mathbf{b}\right)\end{aligned}$$

and therefore, we can use Newton's algorithm to find a root of the function $\phi(\lambda^*) = \hat{\boldsymbol{\theta}}^T_{\text{CML}}(n,\lambda^*)\mathbf{A}\hat{\boldsymbol{\theta}}_{\text{CML}}(n,\lambda^*) + 2\mathbf{b}^T\hat{\boldsymbol{\theta}}_{\text{CML}}(n,\lambda^*) + c$. Note that it can be shown that $\phi(\lambda^*)$ has only one root for all $\lambda^* > 0$.

- *Affine minimax (AMX)*: Similar to LS, this estimator only uses the actual observation $\mathbf{x}(n)$. As the model matrix $\mathbf{H}(n)$ and the covariance matrix $\mathbf{C}(n)$ of the noise are time-invariant in our example, we have $\mathbf{M}(n) = \mathbf{M}$ and $\mathbf{u}(n) = \mathbf{u}$ for our affine estimator. Furthermore, as the center of the ellipsoid $\mathbb{T}$ is the origin, we have $\mathbf{u} = \mathbf{0}$. The optimization problem (2.22) for the special case of an ellipsoidal constraint $\boldsymbol{\theta}^T \mathbf{A} \boldsymbol{\theta} \leq 1$ was shown in [Eldar, 2006b] to be equivalent to

$$\min_{\tau, \mathbf{M}, \lambda \geq 0} \tau \tag{3.22a}$$

subject to

$$\begin{bmatrix} \tau - \lambda & \mathbf{b}^T \\ \mathbf{b} & \mathbf{I} \end{bmatrix} \succeq 0 \tag{3.22b}$$

$$\begin{bmatrix} \lambda \mathbf{I} & -\mathbf{A}^{-\frac{1}{2}} \mathbf{M}^T \\ -\mathbf{M} \mathbf{A}^{-\frac{1}{2}} & \mathbf{I} \end{bmatrix} \succeq 0 \tag{3.22c}$$

with $\mathbf{b} = \mathrm{vec}\{(\mathbf{H}^T \mathbf{C}^{-1} \mathbf{H})^{-\frac{1}{2}} \mathbf{M}^T\}$. These equations were used to find the optimal $\mathbf{M}$. Note that (3.22) defines a SDP which can be efficiently solved. We use the software packages SDPT3 [Toh et al.] and Yalmip [Löfberg, 2004] for solving (3.22).

- *Projected AMX*: This estimator is identical to the AMX but conditionally does a projection onto the boundary of $\mathbb{T}$ if $\hat{\boldsymbol{\theta}}_{\mathrm{AMX}}(n) \notin \mathbb{T}$. The projection problem is given by

$$\hat{\boldsymbol{\theta}}_{\mathrm{PAMX}}(n) = \arg\min_{\boldsymbol{\theta} \in \mathbb{T}} \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{AMX}}(n)\|^2$$

and using the KKT conditions, a similar algorithm as for the CML estimator can be derived as discussed in [Kiseliov, 1994].

- *MMSE*: The MMSE solution is calculated by means of Monte Carlo integration as described in Section 3.3.3 using $I = 200$ samples for approximating the integral.

Furthermore, we use the following five time-recursive estimators to estimate $\boldsymbol{\theta}(n)$ given the observations $\mathbf{x}(0), \ldots, \mathbf{x}(n)$:

- *Recursive weighted least squares (RWLS)*: We solve the weighted least squares problem (3.9) where old observations are weighted with a forgetting factor β. The RWLS does not care about the constraint $\boldsymbol{\theta}(n) \in \mathbb{T}$.

- *Recursive constrained ML (RCML)*: Similar to the CML, we have to solve the QCQP (3.14) where $\hat{\boldsymbol{\theta}}_{\mathrm{LS}}(n)$ is now replaced by $\hat{\boldsymbol{\theta}}_{\mathrm{WLS}}(n)$.

- *Recursive affine minimax (RAMX)*: As we have seen in Section 3.3.2, RAMX is given by a matrix multiplication of the RWLS solution with $\mathbf{I} + \mathbf{M}(n)$ plus an addition of $\mathbf{u}(n)$.

Thus, we have to solve the minimax optimization problem (2.22) repeatedly at each time step from $\mathbf{R}_n$. Similar to the AMX estimator, we use (3.22) to find $\mathbf{M}(n)$ at each step. Again, we have $\mathbf{u}(n) = \mathbf{0}$.

- *Projected RAMX*: Similar to the projected AMX, this estimator does a conditional projection if the RAMX estimate is not in $\mathbb{T}$.
- *Recursive MMSE (RMMSE)*: The RMMSE estimator combines RWLS and MMSE as described in Section 3.3.3. We use again Monte Carlo integration with $I = 200$ samples.

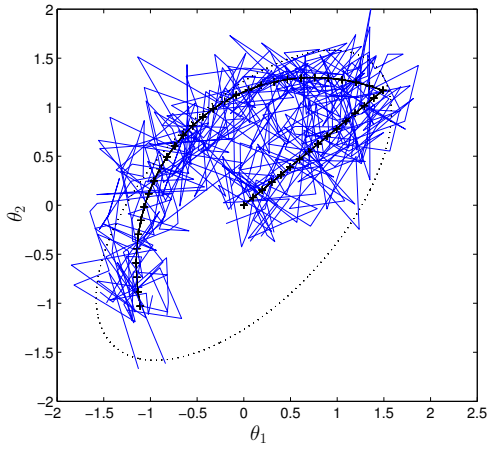
Figure 3.2 shows the overlay of ten realizations of the considered estimators with a forgetting factor of $\beta = 0.5$. Beside the projected AMX/RAMX and CML/RCML, also the MMSE and RMMSE estimators always ensure $\hat{\boldsymbol{\theta}}(n) \in \mathbb{T}$ as expected because of the convex set $\mathbb{T}$. This is not the case for the other estimators.

Figure 3.3 shows the mean squared error $\frac{1}{N}\sum_{n=1}^{N}\|\hat{\boldsymbol{\theta}}(n) - \boldsymbol{\theta}(n)\|^2$ averaged over $1\,500$ trials of this experiment for all estimators as a function of the forgetting factor β. Note that the five instantaneous estimators do not depend on β. The RMMSE estimator and the projected RAMX have the smallest mean squared error for all values of β among the recursive estimators. They also outperform the instantaneous estimators for $0 < \beta < 0.75$. The optimal forgetting factor for this time-variant problem is $\beta_{\text{opt}} \approx 0.5$. If the forgetting factor is smaller, the effective number $\frac{1}{1-\beta}$ of observations (degrees of freedom) is too small for a reliable estimation of $\boldsymbol{\theta}(n)$. If, however, the forgetting factor is too large, the estimator is not able to track the time-varying parameter $\boldsymbol{\theta}(n)$.

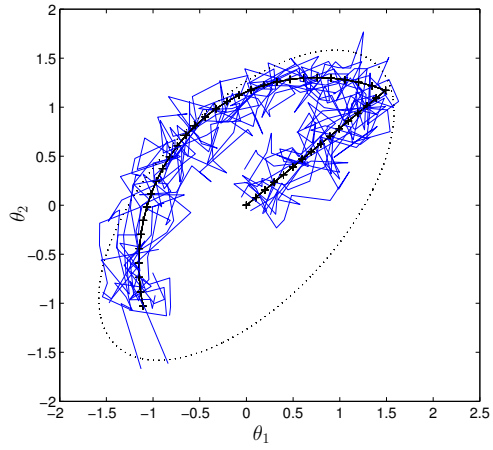
We rerun the first experiment with a slower change of $\boldsymbol{\theta}(n)$. The parameter $\boldsymbol{\theta}(n)$ is again assumed to follow the trajectory in Figure 3.1, however this time with three times more discrete points. Figure 3.4 shows the result. We see that the optimal choice of the forgetting factor β is now $\beta \approx 0.7$, indicating a slower decaying exponential window and hence a longer memory. As $\boldsymbol{\theta}(n)$ changes more slowly, the MSE can further be reduced compared to Figure 3.3.

The choice of the number of samples I in the Monte Carlo integration also plays an important role in the design of the estimator. To analyze the influence of I, we varied its value. Figure 3.5 shows the mean squared error of the RMMSE for different values of I and the noise variance σ^2. We see that it is sufficient to use $I \geq 100$ samples. To be on the safe side, we choose $I = 200$.

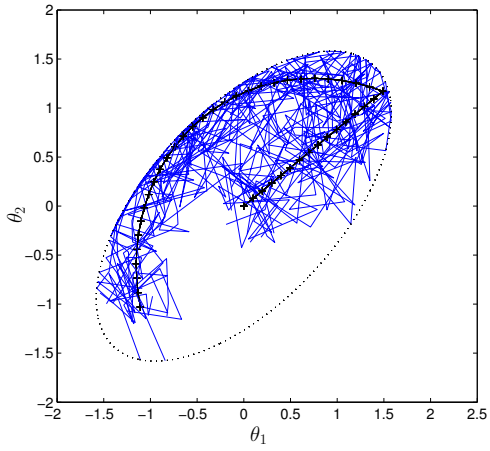
Finally, we compare the RAMX and its steady-state version. Instead of calculating $\mathbf{M}(n)$ at each time step, we calculate a fixed $\mathbf{M}$ from the steady-state correlation matrix $\mathbf{R}_\infty$ given in (3.17). Figure 3.6 shows the ratio of the mean squared errors of both estimators as a function of the time index n, i.e. $\kappa = \text{MSE(RAMX)}/\text{MSE(steady-state RAMX)}$. Clearly, the MSE values of the steady-state approximation are larger than the corresponding values of the RAMX. The difference between both estimators vanishes for increasing time indices n.



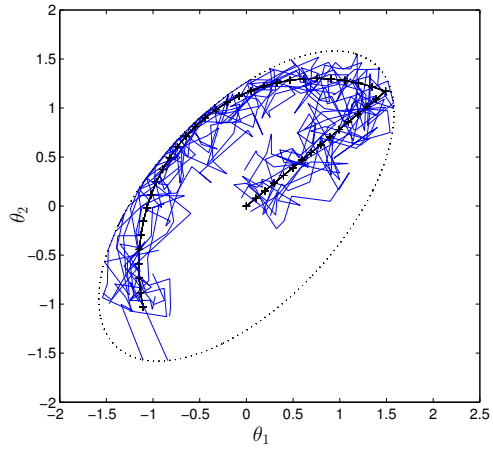
(a) Weighted least squares



(b) Recursive weighted least squares

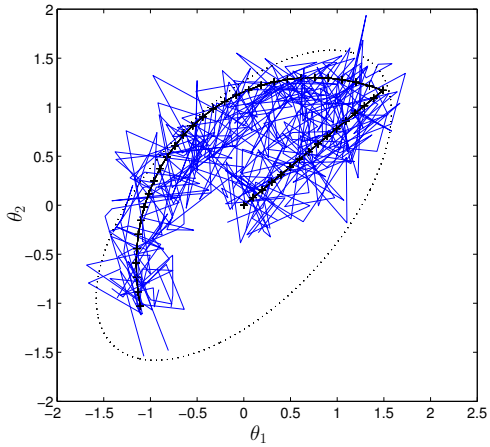


(c) Constrained ML

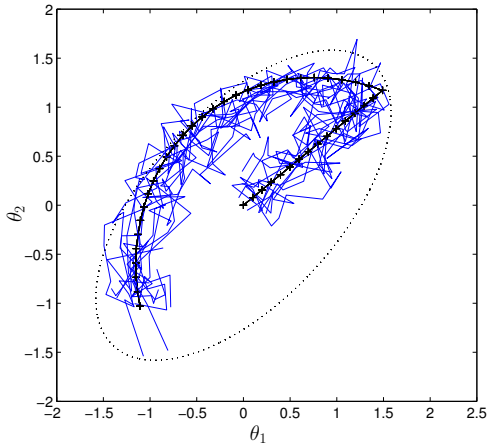


(d) Recursive Constrained ML

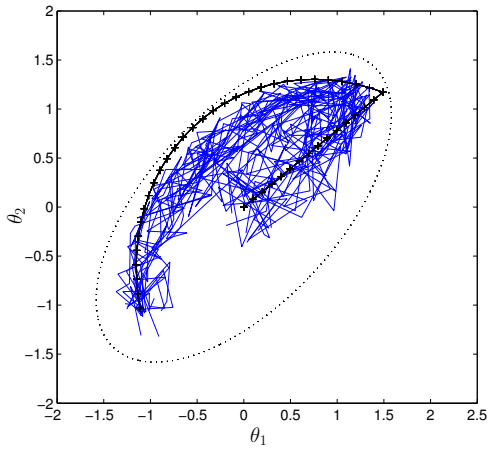
Figure 3.2: Example 1: Overlay of ten realizations for the first experiment with $\beta = 0.5$



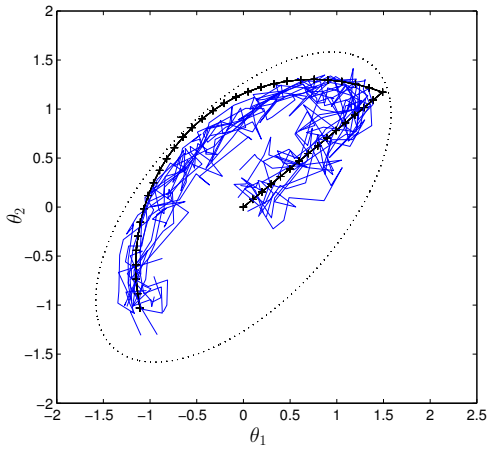
(e) Affine minimax



(f) Recursive affine minimax



(g) MMSE



(h) Recursive MMSE

Figure 3.2: Example 1: Overlay of ten realizations for the first experiment with $\beta = 0.5$
(continued from previous page)

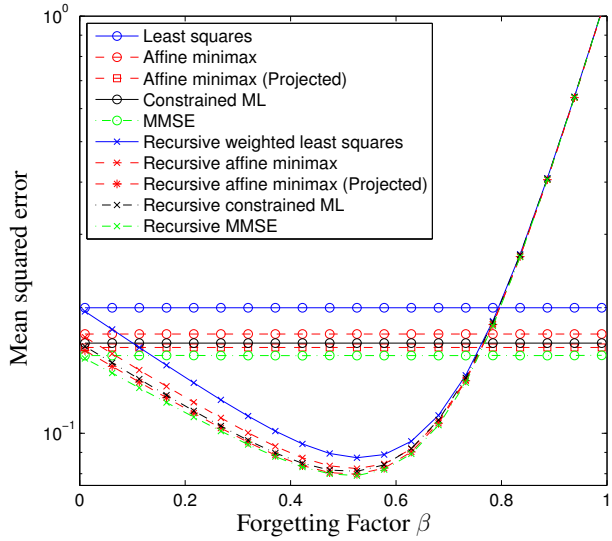


Figure 3.3: Example 1: Mean squared error vs. the forgetting factor β for $N = 45$

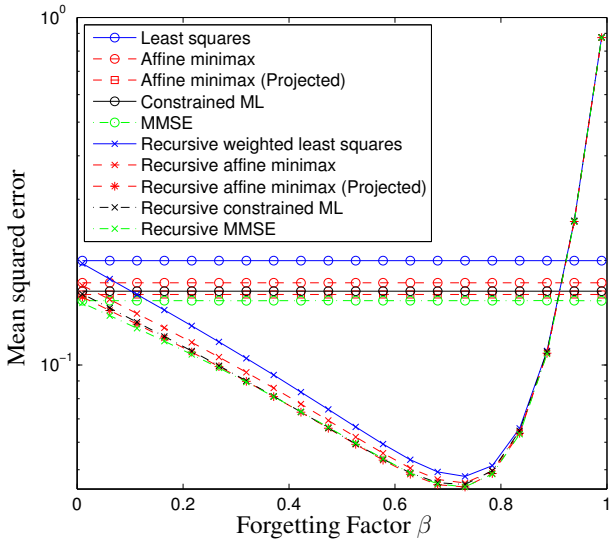


Figure 3.4: Example 1: Mean squared error vs. the forgetting factor β for $N = 137$

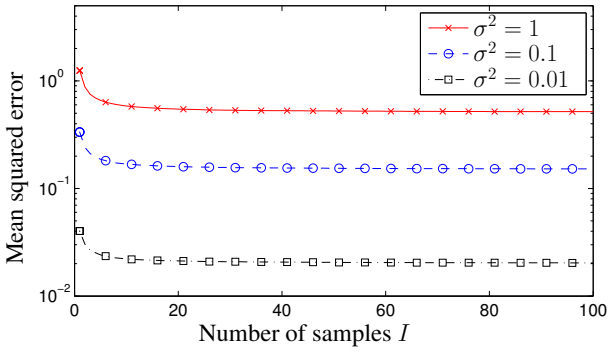


Figure 3.5: Example 1: Mean squared error for varying number of Monte Carlo samples

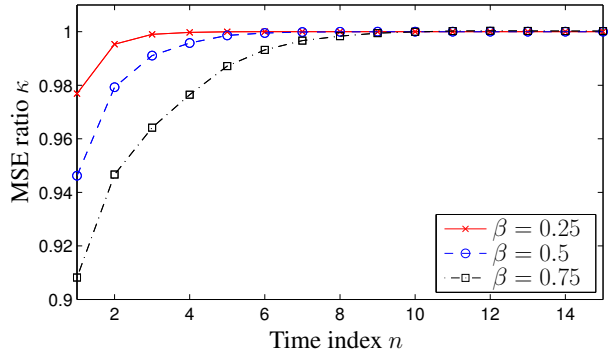


Figure 3.6: Example 1: Ratio of the mean squared errors for RAMX and its steady-state approximation

3.5.2 Example 2: Ellipsoidal Constraints

In this second example, we will show a different way to compute the RMMSE estimator than using rejection sampling as proposed in Section 3.3.3. This is necessary as the two constraints that we consider are given by

$$\text{Case A:} \quad \mathbb{T}_{\text{A}} = \{\boldsymbol{\theta} \in \mathbb{R}^M : (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \mathbf{A} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) \leq R^2, \mathbf{A} \succ \mathbf{0}\} \tag{3.23a}$$

$$\text{Case B:} \quad \mathbb{T}_{\text{B}} = \{\boldsymbol{\theta} \in \mathbb{R}^M : (\boldsymbol{\theta} - \boldsymbol{\theta}_0)^T \mathbf{A} (\boldsymbol{\theta} - \boldsymbol{\theta}_0) = R^2, \mathbf{A} \succ \mathbf{0}\}, \tag{3.23b}$$

i.e. $\boldsymbol{\theta}$ has to lie inside or on an ellipsoid. The constraint $\boldsymbol{\theta} \in \mathbb{T}_{\text{A}}$ can be handled by rejection sampling as shown in the first example (Section 3.5.1). The second constraint $\boldsymbol{\theta} \in \mathbb{T}_{\text{B}}$, however, is not suited for rejection sampling as the equality constraint for a random sample will not be fulfilled with probability one. Therefore, a different way to compute (3.20) is needed and we will show in this example that a series expansion approach of the occurring integrals for the RMMSE in (3.20) can be used.

Derivation of a Series Expansion for the RMMSE Estimators

We now derive $\hat{\boldsymbol{\theta}}_{\text{RMMSE}}(\mathbf{x})$ for the constraints in (3.23). From (3.20), we know that it is sufficient to consider

$$\hat{\boldsymbol{\theta}}_{\text{RMMSE}}(n) = \frac{\int\limits_{\mathbb{T}} \theta \, \exp\{-\frac{1}{2}(\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)^T \tilde{\mathbf{C}}_n^{-1} (\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)\} d\boldsymbol{\theta}}{\int\limits_{\mathbb{T}} \exp\{-\frac{1}{2}(\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)^T \tilde{\mathbf{C}}_n^{-1} (\boldsymbol{\theta} - \tilde{\mathbf{x}}_n)\} d\boldsymbol{\theta}} \tag{3.24}$$

which is obtained after the transformation using the sufficient statistic $\tilde{\mathbf{x}}_n = \hat{\boldsymbol{\theta}}_{\text{WLS}}(n)$. For convenience, we drop the dependence on the time instance n in the following and apply a two-step procedure to solve the problem:

(a) Whitening Transformation: First, we transform the problem to one of estimating a vector in white noise where the axes of the ellipsoid are parallel to the coordinate axes. Let $\mathbf{U}$ be one matrix square root of $\tilde{\mathbf{C}}$, i.e. $\tilde{\mathbf{C}} = \mathbf{U}^T \mathbf{U}$. One possibility is the Cholesky factor of $\tilde{\mathbf{C}}$. Furthermore, let $\mathbf{V}$ and $\mathbf{D} = \operatorname{diag}(d_1, \ldots, d_M)$ be the matrix of eigenvectors and eigenvalues of $\mathbf{U}\mathbf{A}\mathbf{U}^T$. Using the substitution $\boldsymbol{\theta} - \tilde{\mathbf{x}} = \mathbf{U}^T \mathbf{V} \mathbf{y}$, (3.24) reads

$$\begin{aligned}
\hat{\boldsymbol{\theta}}_{\text{RMMSE}} &= \frac{\int\limits_{\mathbb{T}^*} (\tilde{\mathbf{x}} + \mathbf{U}^T \mathbf{V} \mathbf{y}) \exp\{-\frac{1}{2}\|\mathbf{y}\|^2\} d\mathbf{y}}{\int\limits_{\mathbb{T}^*} \exp\{-\frac{1}{2}\|\mathbf{y}\|^2\} d\mathbf{y}} \\
&= \tilde{\mathbf{x}}_n + \mathbf{U}^T \mathbf{V} \frac{\int\limits_{\mathbb{T}^*} \mathbf{y} \exp\{-\frac{1}{2}\|\mathbf{y}\|^2\} d\mathbf{y}}{\int\limits_{\mathbb{T}^*} \exp\{-\frac{1}{2}\|\mathbf{y}\|^2\} d\mathbf{y}} =: \tilde{\mathbf{x}}_n + \mathbf{U}^T \mathbf{V} \frac{\mathbf{i}_1}{i_2}
\end{aligned} \tag{3.25}$$

where $\mathbb{T}^*$ is either $\mathbb{T}^*_{\mathrm{A}} = \{\mathbf{y} : (\mathbf{y}-\mathbf{y}_0)^T\mathbf{D}(\mathbf{y}-\mathbf{y}_0) \leq R^2\}$ or $\mathbb{T}^*_B = \{\mathbf{y} : (\mathbf{y}-\mathbf{y}_0)^T\mathbf{D}(\mathbf{y}-\mathbf{y}_0) = R^2\}$ with $\mathbf{y}_0 = (\mathbf{U}^T\mathbf{V})^{-1}(\boldsymbol{\theta}_0 - \tilde{\mathbf{x}}_n)$.

(b) Series Expansion: The second step is to represent both integrals in (3.25) in terms of a series of central χ^2 *cumulative distribution functions* (CDF)s or PDFs.

We first consider case A that $\boldsymbol{\theta}$ has to lie inside the ellipsoid, i.e. $\boldsymbol{\theta} \in \mathbb{T}_{\mathrm{A}}$. In Appendix C.2.2 it is shown that i_2 can be expressed as an infinite linear combination of χ^2-CDFs, i.e. we can write

$$i_2 = (2\pi)^{M/2} \sum_{k=0}^{\infty} a_k F_{\chi^2}(R^2/\beta; M+2k) \tag{3.26}$$

where $F_{\chi^2}(R^2/\beta; M+2k)$ is the central χ^2-CDF with $M+2k$ degrees of freedom evaluated at position R^2/β and β is an arbitrary constant. The coefficients a_k can be found by the recursive rule

$$a_0 = \prod_{m=1}^{M} \sqrt{\frac{\beta}{d_m}} e^{-\frac{1}{2}y_{0m}^2}, \qquad\qquad a_k = \frac{1}{2k}\sum_{l=0}^{k-1} b_{k-l}a_l,$$

where $b_k = \sum_{m=1}^{M}\left(1-\frac{\beta}{d_m}\right)^k + k\beta\sum_{m=1}^{M}\frac{y_{0m}^2}{d_m}\left(1-\frac{\beta}{d_m}\right)^{k-1}$, d_m is the mth diagonal element of $\mathbf{D}$ and y_{0m} is the mth component of $\mathbf{y}_0$. The calculation of i_2 is therefore straightforward where the summation is stopped if the relative error is small enough. An estimate of the truncation error is given in (C.29) and was used as a stopping criterion. It therefore remains to calculate $\mathbf{i}_1$. Its mth element reads

$$i_{1m} = \int_{\mathbb{T}^*_{\mathrm{A}}} y_m e^{-\frac{1}{2}\|\mathbf{y}\|^2} d\mathbf{y} = \int_{y_{0m}-\frac{R}{\sqrt{d_m}}}^{y_{0m}+\frac{R}{\sqrt{d_m}}} y_m e^{-\frac{1}{2}y_m^2} \left(\int_{\mathbb{T}^*_{\mathrm{A}m}} e^{-\frac{1}{2}\|\tilde{\mathbf{y}}_m\|^2} d\tilde{\mathbf{y}}_m \right) dy_m \tag{3.27}$$

where $\tilde{\mathbf{y}}_m = \mathbf{P}_m\mathbf{y}$ is identical to $\mathbf{y}$ except for the removed mth element. The selection matrix $\mathbf{P}_m \in \mathbb{R}^{(M-1)\times M}$ is the identity matrix with the mth row erased. Note that the inner integral is of the same type as i_2 and can thus also be expressed as an infinite series of χ^2-CDFs. The integration area $\mathbb{T}^*_{\mathrm{A}m}$ is given by $\mathbb{T}^*_{\mathrm{A}m} = \{\tilde{\mathbf{y}}_m : (\tilde{\mathbf{y}}_m - \mathbf{P}_m\mathbf{y}_0)^T\mathbf{P}_m\mathbf{D}\mathbf{P}_m^T(\tilde{\mathbf{y}}_m - \mathbf{P}_m\mathbf{y}_0) \leq R^2 - d_m(y_m - y_{0m})^2\}$ and is a function of y_m. Hence, (3.27) can be efficiently solved by an one-dimensional numerical integration.

For case B we can use the same approach as before, i.e.

$$i_{1m} = \int_{\mathbb{T}^*_B} y_m e^{-\frac{1}{2}\|\mathbf{y}\|^2} d\mathbf{y} = \int_{y_{0m}-\frac{R}{\sqrt{d_m}}}^{y_{0m}+\frac{R}{\sqrt{d_m}}} y_m e^{-\frac{1}{2}y_m^2} \left(\int_{\mathbb{T}^*_{\mathrm{B}m}} e^{-\frac{1}{2}\|\tilde{\mathbf{y}}_m\|^2} d\tilde{\mathbf{y}}_m \right) dy_m \tag{3.28a}$$

$$i_2 = \frac{(2\pi)^{M/2}}{\beta} \sum_{k=0}^{\infty} a_k p_{\chi^2}(R^2/\beta; M+2k) \tag{3.28b}$$

where $p_{\chi^2}(R^2/\beta; M+2k)$ is the central χ^2-PDF with $M+2k$ degrees of freedom evaluated at position R^2/β and $\mathbb{T}^*_{\mathrm{B}m} = \{\tilde{\mathbf{y}}_m : (\tilde{\mathbf{y}}_m - \mathbf{P}_m\mathbf{y}_0)^T\mathbf{P}_m\mathbf{D}\mathbf{P}_m^T(\tilde{\mathbf{y}}_m - \mathbf{P}_m\mathbf{y}_0) = R^2 - d_m(y_m - y_{0m})^2\}$. Similar to before, we can express i_2 as an infinite series of χ^2-PDFs and $\mathbf{i}_1$ is found again by M one-dimensional numerical integrations.

Simulation Results

In the following we show some simulation results for both cases where $M = 3$ unknowns have to be estimated from $K = 10$ observations in a stationary setup. The model matrix $\mathbf{H}$ is randomly chosen for each trial and in total we average over $10\,000$ trials, except for Figure 3.8 where we average over $100\,000$ trials. The noise $\mathbf{z}$ has the distribution $\mathcal{N}(\mathbf{0}, \sigma^2\mathbf{I})$ and the SNR is defined as $\mathrm{SNR} = 10\log_{10}\|\mathbf{H}\boldsymbol{\theta}\|^2/(K\sigma^2)$. For case A, we assume that the parameter vector $\boldsymbol{\theta}$ is restricted to lie in the unit sphere $\mathbb{T}_{\mathrm{A}} = \{\boldsymbol{\theta} : \|\boldsymbol{\theta}\|^2 \leq 1\}$ and is uniformly distributed for the simulation. For case B, we assume that $\boldsymbol{\theta}$ is uniformly distributed on the unit sphere $\mathbb{T}_{\mathrm{B}} = \{\boldsymbol{\theta} : \|\boldsymbol{\theta}\|^2 = 1\}$.

(a) Case A: We compare the derived MMSE estimator with the following four estimators for a spherical constraint:

- *Least squares (LS) estimator*: According to (2.16), the LS is given by [Scharf, 1990]

$$\hat{\boldsymbol{\theta}}_{\mathrm{LS}} = (\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\mathbf{H}^T\mathbf{C}^{-1}\mathbf{x} = \tilde{\mathbf{x}} \tag{3.29a}$$

- *Projected LS (PLS) estimator*: The PLS does a conditional scaling if the constraint $\hat{\boldsymbol{\theta}}_{\mathrm{LS}} \in \mathbb{T}_{\mathrm{A}}$ is not fulfilled [Stoica and Ganesan, 2000], i.e.

$$\hat{\theta}_{\mathrm{PLS}} = \begin{cases} \hat{\boldsymbol{\theta}}_{\mathrm{LS}} & \|\hat{\boldsymbol{\theta}}_{\mathrm{LS}}\| \leq R \\ \frac{R}{\|\hat{\boldsymbol{\theta}}_{\mathrm{LS}}\|}\hat{\boldsymbol{\theta}}_{\mathrm{LS}} & \|\hat{\boldsymbol{\theta}}_{\mathrm{LS}}\| > R \end{cases} \tag{3.29b}$$

- *Affine minimax (AMX) estimator*: The AMX for this problem can be given analytically and has the form [Pilz, 1986; Eldar et al., 2005]

$$\hat{\boldsymbol{\theta}}_{\mathrm{AMX}} = \frac{R^2}{R^2 + \mathrm{tr}\{(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H})^{-1}\}}\hat{\boldsymbol{\theta}}_{\mathrm{LS}} \tag{3.29c}$$

- *Projected affine minimax (PAMX) estimator*: The PAMX is [Stoica and Ganesan, 2000]

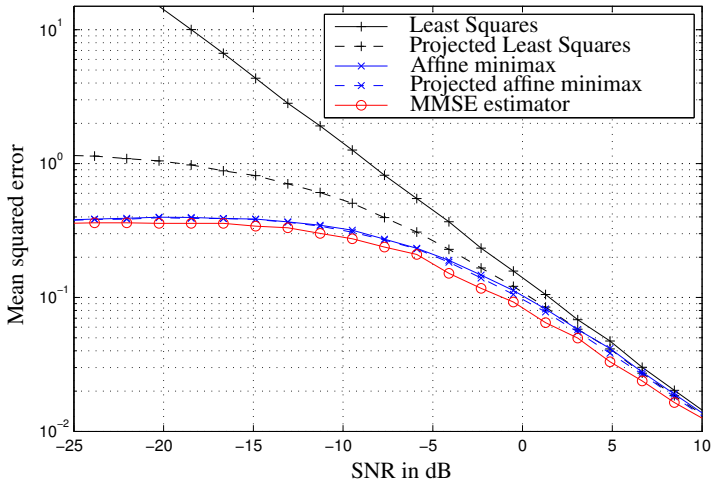


Figure 3.7: Example 2, Case A: Comparison of averaged squared error

$$\hat{\boldsymbol{\theta}}_{\text{PAMX}} = \begin{cases} \hat{\boldsymbol{\theta}}_{\text{AMX}} & \|\hat{\boldsymbol{\theta}}_{\text{AMX}}\| \leq R \\ \frac{R}{\|\hat{\boldsymbol{\theta}}_{\text{AMX}}\|}\hat{\boldsymbol{\theta}}_{\text{AMX}} & \|\hat{\boldsymbol{\theta}}_{\text{MX}}\| > R \end{cases} \tag{3.29d}$$

Figure 3.7 shows the simulation results for the squared error $\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|^2$ averaged over all $\boldsymbol{\theta} \in \mathbb{T}_{\text{A}}$. The derived MMSE estimator has clearly the minimal averaged squared error and is therefore superior to the other estimators as expected. It has, however, a higher computational complexity because of the M one-dimensional numerical integrations.

The second comparison is in terms of the risk of the affine minimax and the MMSE estimator.[21] The risk of an estimator that corresponds to a quadratic loss function is defined as [Scharf, 1990]

$$R(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \mathrm{E}_{\mathbf{x}|\boldsymbol{\theta}}\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}(\mathbf{x})\|^2 = \int \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}(\mathbf{x})\|^2 p(\mathbf{x}|\boldsymbol{\theta}) d\mathbf{x}. \tag{3.30}$$

It is the quadratic loss $\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|^2$ averaged over the distribution of the measurements with $\boldsymbol{\theta}$ fixed. This comparison is interesting as the affine minimax estimator minimizes the worst case mean squared error for each deterministic $\boldsymbol{\theta} \in \mathbb{T}_{\text{A}}$ opposed to the MMSE estimator which only considers the overall mean squared error. Note that the estimator risk is rotational invariant in the parameter space of $\boldsymbol{\theta}$ as the model matrix $\mathbf{H}$ is chosen randomly. Thus, it is sufficient to plot the estimator risk as a function of the norm of $\boldsymbol{\theta}$ only and the risk in (3.30) is calculated by averaging the squared error over all $\boldsymbol{\theta}$ with $\|\boldsymbol{\theta}\| = \text{const}$. Figure 3.8 depicts the simulation results for a SNR of 0dB. It shows that the MMSE estimator is, except for a small region around $\|\boldsymbol{\theta}\| = 0.4$, better than the affine minimax estimator. Thus, although we derived the estimator

[21] Note that the definition of the risk $R(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ in (3.30) can be used in the frequentist and Bayesian case and therefore we can use it to compare the affine minimax with the MMSE estimator.

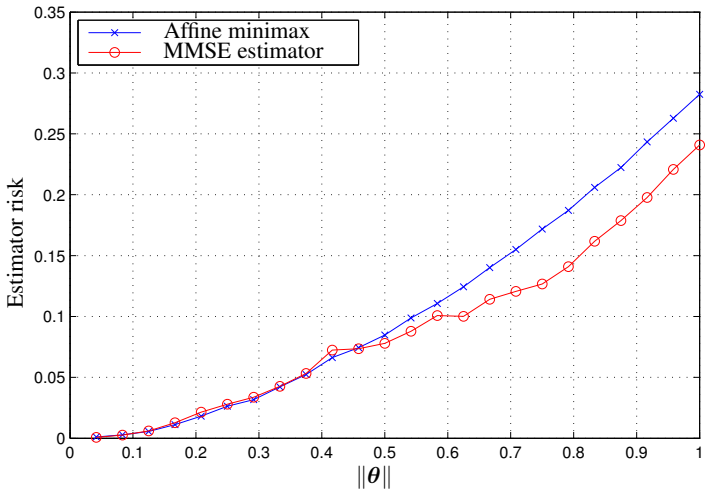


Figure 3.8: Example 2, Case A: Comparison of estimator risk

which minimizes the mean squared error averaged over all possible parameter vectors in $\mathbb{T}_{\mathrm{A}}$, it is still almost always better in terms of the risk than the affine minimax estimator which was designed to minimize the worst case error for the deterministic case.

(a) Case B: We compare the derived MMSE estimator with the least squares estimator (3.29a) and the following three estimators:

- *Scaled LS (ScLS) estimator*: The ScLS is inspired by [Stoica and Ganesan, 2000] and defined by

$$\hat{\boldsymbol{\theta}}_{\mathrm{ScLS}} = \frac{R}{\|\hat{\boldsymbol{\theta}}_{\mathrm{LS}}\|}\hat{\boldsymbol{\theta}}_{\mathrm{LS}} \tag{3.31}$$

- *Spherical LS (SLS) estimator*: The SLS corresponds to the class of constrained LS estimators [Moon and Stirling, 2000]

$$\hat{\boldsymbol{\theta}}_{\mathrm{SLS}} = \min_{\boldsymbol{\theta}} \|\mathbf{x} - \mathbf{H}\boldsymbol{\theta}\|^2 \text{ s.t. } \|\boldsymbol{\theta}\| = R \tag{3.32}$$

  which we solved by expressing $\boldsymbol{\theta}$ in spherical coordinates and using a nonlinear least squares optimization procedure.

- *Affine minimax (AMX) estimator*: The AMX for case B is equal to the estimator given in (3.29c) as the worst case MSE to be minimized is always located on the boundary of the ellipsoid. Hence, calculating the affine minimax estimator with respect to $\mathbb{T}_{\mathrm{B}}$ is equal to calculating it with respect to $\mathbb{T}_{\mathrm{A}}$.

Figure 3.9 and 3.10 show simulation results for case B with respect to the averaged squared error $\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|^2$ and the averaged angle between $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$ in degrees $\cos^{-1}\left(\frac{\boldsymbol{\theta}^T\hat{\boldsymbol{\theta}}}{\|\boldsymbol{\theta}\|\|\hat{\boldsymbol{\theta}}\|}\right)$. The derived

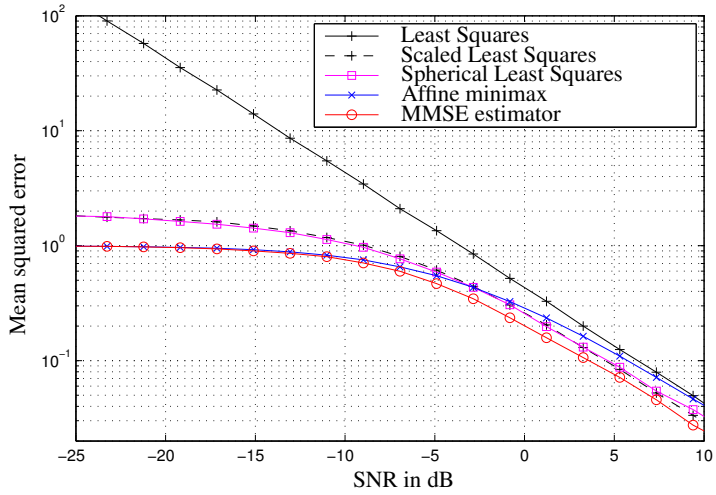


Figure 3.9: Example 2, Case B: Comparison of averaged squared error

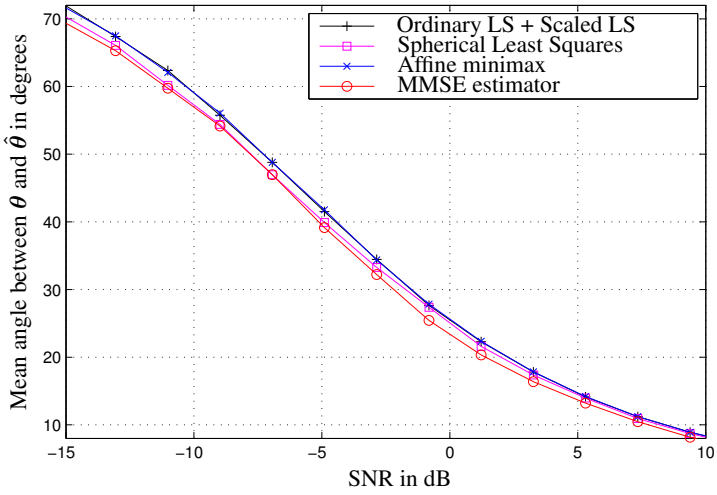


Figure 3.10: Example 2, Case B: Comparison of averaged angle in degrees

MMSE estimator is superior to the other estimators for both error measures. Especially the averaged angle error is interesting for DOA applications as in [Mukai et al., 2005] where it is important that $\boldsymbol{\theta}$ and $\hat{\boldsymbol{\theta}}$ point to the same direction.

3.5.3 Example 3: Room Impulse Response Estimation

In the last example, we study the tracking of a room impulse response (RIR) which is time-varying. In particular, we consider the following setup: Let $\boldsymbol{\theta}(n) = \begin{bmatrix}\theta_0(n) & \cdots & \theta_{M-1}(n)\end{bmatrix}^T$ denote the RIR which we want to estimate and let $\mathbf{s}(n) = \begin{bmatrix}s_0(n) & \cdots & s_{S-1}(n)\end{bmatrix}^T$ be the known

signal which is transmitted by the source at time instance n. The corresponding microphone signal $\mathbf{x}(n) = \begin{bmatrix} x_0(n) & \cdots & x_{K-1}(n) \end{bmatrix}^T$ with length $K = M + S - 1$ is given by

$$\mathbf{x}(n) = \boldsymbol{\Theta}(n)\mathbf{s}(n) + \mathbf{z}(n) = \mathbf{S}(n)\boldsymbol{\theta}(n) + \mathbf{z}(n) \tag{3.33}$$

where $\boldsymbol{\Theta}(n) \in \mathbb{R}^{K \times S}$ and $\mathbf{S}(n) \in \mathbb{R}^{K \times M}$ are Toeplitz matrices that consist of shifted column vectors $\boldsymbol{\theta}(n)$ and $\mathbf{s}(n)$, i.e.

$$\boldsymbol{\Theta}(n) = \begin{bmatrix} \boldsymbol{\theta}(n) & & & \mathbf{0} \\ & \boldsymbol{\theta}(n) & & \\ & & \ddots & \\ \mathbf{0} & & & \boldsymbol{\theta}(n) \end{bmatrix}, \qquad \mathbf{S}(n) = \begin{bmatrix} \mathbf{s}(n) & & & \mathbf{0} \\ & \mathbf{s}(n) & & \\ & & \ddots & \\ \mathbf{0} & & & \mathbf{s}(n) \end{bmatrix}.$$

Eq. (3.33) describes a "burst model" where we assume that the RIR is stationary for one input burst $\mathbf{s}(n)$. The noise $\mathbf{z}(n)$ in (3.33) is Gaussian with $\mathbf{z}(n) \sim \mathcal{N}(\mathbf{0}, \mathbf{C}(n))$ and temporally uncorrelated, i.e. $\mathrm{E}[\mathbf{z}(n_1)\mathbf{z}(n_2)^T] = \mathbf{0}$ for all $n_1 \neq n_2$.

To improve the recursive estimation of the RIR $\boldsymbol{\theta}(n)$, we study the problem (3.33) with the additional energy conservation constraint

$$\|\boldsymbol{\Theta}(n)\mathbf{s}(n)\|^2 \leq \|\mathbf{s}(n)\|^2, \tag{3.34}$$

which states that the energy of the received signal does not exceed the energy of the transmitted signal. The basic idea of (3.34) is to avoid "outliers" of $\hat{\boldsymbol{\theta}}(n)$ which violate the physical constraint that the impulse response is energy conserving.

Beside the many techniques to estimate RIR, which differ in the different choices of excitation signals [Stan et al., 2002], Lin proposed in [Lin and Lee, 2006] two constraints to improve the RIR estimation for the stationary case. He introduced the nonnegativity constraints $\theta_m \geq 0$, $m = 0, \ldots, M-1$ and added an 1-norm penalty which controls the sparsity of the solution. It is shown that adding such constraints to the RIR estimation improves the robustness to different noise distributions and decreases the mean squared error. Adding such additional constraints to the RCML, RAMX and RMMSE is also possible and can be done in a straightforward way. Therefore, we will focus in this example on the energy conservation constraint (3.34) to improve the tracking performance.

Energy Conservation — Mathematical Formulation

To improve the estimation performance, we make use of the energy conservation (3.34) which must hold for real-world impulse responses. We first give two equivalent mathematical representations of (3.34) and then show how to conveniently approximate them using the *discrete Fourier transform* (DFT).

Expressing the energy conservation mathematically, we need to ensure that[22]

$$\|\mathbf{\Theta}(n)\mathbf{s}(n)\|^2 = \mathbf{s}(n)^T\mathbf{\Theta}(n)^T\mathbf{\Theta}(n)\mathbf{s}(n) \leq \mathbf{s}(n)^T\mathbf{s}(n) = \|\mathbf{s}(n)\|^2$$

which implies

$$\mathbf{\Theta}(n)^T\mathbf{\Theta}(n) - \mathbf{I} \preceq \mathbf{0}. \tag{3.35}$$

Note that (3.35) has to hold for all signal lengths S. If the condition is fulfilled for a particular signal length S, it is automatically fulfilled for all smaller signal lengths since an upper-left submatrix of a negative semidefinite matrix is again negative semidefinite. In the following, we will now give two equivalent representations of the set that is described by (3.35).

(a) LMI representation: Using Schur's lemma [Boyd and Vandenberghe, 2007], we can rewrite (3.35) into

$$\begin{bmatrix} \mathbf{I} & \mathbf{\Theta}(n)^T \\ \mathbf{\Theta}(n) & \mathbf{I} \end{bmatrix} \succeq \mathbf{0} \tag{3.36}$$

which is a LMI. Note that (3.36) immediately implies that (3.35) describes a convex set.

(b) Frequency domain representation: From above, we know that it is sufficient to consider the case $S \to \infty$. This case can be efficiently computed using the theory of bandlimited Toeplitz matrices [Moon and Stirling, 2000; Alkire, 2003] as we will now show.

Eq. (3.35) is equivalent to requiring that all eigenvalues of $\mathbf{\Theta}(n)^T\mathbf{\Theta}(n)$ are smaller than or equal to 1. We will therefore now show how the eigenvalues of $\mathbf{\Theta}(n)^T\mathbf{\Theta}(n)$ can be computed for $S \to \infty$. Let

$$r(n) = \begin{cases} \sum_{m=0}^{M-1-|n|} \theta_m(n)\theta_{|n|+m}(n) & |n| \leq M-1 \\ 0 & \text{otherwise} \end{cases} \tag{3.37}$$

be the (unnormalized) auto-correlation function of the RIR $\boldsymbol{\theta}(n)$. Then, the matrix $\mathbf{\Theta}(n)^T\mathbf{\Theta}(n) \in \mathbb{R}^{S\times S}$ can be written as a symmetric Toeplitz matrix where the first row is given by $r(n)$ for $n = 0, \ldots, S-1$. Note that $\mathbf{\Theta}(n)^T\mathbf{\Theta}(n)$ is bandlimited as only the first $M-1$ off-diagonals are unequal to zero and $S > M$ as we consider the case $S \to \infty$. Using the asymptotic equivalence of the eigenvalues of a bandlimited Toeplitz matrix and the corresponding circulant matrix, it immediately follows that the condition (3.35) can be transformed into

$$\left|\Theta(\omega, n)\right|^2 = \left|\sum_{m=0}^{M-1} \theta_m(n)e^{-j\omega m}\right|^2 \leq 1 \;\forall\, \omega \in [0, 2\pi) \tag{3.38}$$

[22] We assume in this paper that the source and sensor gains are known, i.e. we have no ambiguity in terms of a scaling of $\mathbf{\Theta}(n)\mathbf{s}(n)$. This can e.g. be achieved by using a calibration step before estimating the RIRs.

where $\Theta(\omega, n)$ denotes the room frequency response at time instance n. To derive (3.38), we used the fact that the eigenvalues of a circulant matrix are given by the DFT of its first column as shown in Appendix C.3.6 and therefore (3.35) results in (3.38). As we consider only real-valued impulse responses $\boldsymbol{\theta}(n)$, it is sufficient to restrict ω to $[0, \pi]$ in (3.38).

Energy Conservation — Approximation using DFT

All three estimators will be based on (3.38) where we evaluate $\Theta(\omega, n)$ at discrete frequencies ω_l using the DFT. Let $\omega_l = \frac{2\pi l}{L}$ with $l = 0, \ldots, L-1$ be the equidistant frequency bins which we consider. $L \geq M$ denotes the DFT length where $L > M$ corresponds to the case of zero-padding. As we only need to evaluate $\Theta(\omega, n)$ in $[0, \pi]$, we have $l = 0, \ldots, \tilde{L}$ with $\tilde{L} = \lfloor L/2 \rfloor$ where $\lfloor x \rfloor$ is the largest integer not greater than x. Using vector notation, we can therefore approximate (3.38) by

$$\boldsymbol{\theta}^T \mathbf{f}_l \mathbf{f}_l^H \boldsymbol{\theta} \leq 1 \quad \forall\, l = 0, \ldots, \tilde{L} \tag{3.39}$$

where $\mathbf{f}_l \in \mathbb{C}^M$ is composed of the first M elements of the lth column of the DFT matrix $\mathbf{U} \in \mathbb{C}^{L \times L}$, i.e. $\mathbf{f}_l^H = [\, 1 \; e^{-j\omega_l} \; \cdots \; e^{-j\omega_l(M-1)} \,]$. Using the DFT, we can therefore approximate $\mathbb{T} = \{\boldsymbol{\theta}(n) : \boldsymbol{\theta}(n) \text{ fulfills (3.38)}\}$ by the new constraint $\tilde{\mathbb{T}} = \{\boldsymbol{\theta}(n) : \boldsymbol{\theta}(n) \text{ fulfills (3.39)}\}$ with arbitrary precision if L is chosen large enough.

In the following, we will see that it is convenient to consider the zero-padded vector $\tilde{\boldsymbol{\theta}} \in \mathbb{R}^L$ instead of $\boldsymbol{\theta} \in \mathbb{R}^M$. Let $\mathbf{P} \in \mathbb{R}^{M \times L}$ be the matrix that consists of the first M rows of the $L \times L$ identity matrix. Then, introducing the zero-padded vector $\tilde{\boldsymbol{\theta}} = \mathbf{P}^T \boldsymbol{\theta}$, condition (3.39) transforms into

$$\boldsymbol{\theta}^T \mathbf{P} \tilde{\mathbf{f}}_l \tilde{\mathbf{f}}_l^H \mathbf{P}^T \boldsymbol{\theta} = \tilde{\boldsymbol{\theta}}^T \tilde{\mathbf{f}}_l \tilde{\mathbf{f}}_l^H \tilde{\boldsymbol{\theta}} \leq 1 \quad \forall\, l = 0, \ldots, \tilde{L} \tag{3.40}$$

where $\tilde{\mathbf{f}}_l \in \mathbb{C}^L$ is the lth column of the DFT matrix $\mathbf{U} \in \mathbb{C}^{L \times L}$, i.e. $\tilde{\mathbf{f}}_l^H = [\, 1 \; e^{-j\omega_l} \; \cdots \; e^{-j\omega_l(L-1)} \,]$, and $\mathbf{f}_l = \mathbf{P} \tilde{\mathbf{f}}_l$ holds.

Recursive Constrained Estimators

In the following, we will describe in more detail the three recursive tracking algorithms for the estimation of a time-varying RIR.

(a) Recursive CML: The RCML estimator $\hat{\boldsymbol{\theta}}_{\text{RCML}}(n)$ for our problem is given by (3.14) with the constraint (3.39). This is a quadratically constrained quadratic program which can be solved by standard convex solvers [Boyd and Vandenberghe, 2007].

For completeness, we would like to mention the transform $\bar{\boldsymbol{\theta}} = \frac{1}{\sqrt{L}} \mathbf{V}^T \tilde{\boldsymbol{\theta}}$ where $\frac{1}{\sqrt{L}} \mathbf{V} \in \mathbb{R}^{L \times L}$ is an orthogonal matrix and for an even L is given by

$$\mathbf{V} = \begin{bmatrix} 1 & \sqrt{2} & \cdots & \sqrt{2} & 1 & \sqrt{2} & \cdots & \sqrt{2} \\ 1 & \sqrt{2}\cos(\omega_1) & \cdots & \sqrt{2}\cos(\omega_{\tilde{L}-1}) & \cos(\omega_{\tilde{L}}) & \sqrt{2}\sin(\omega_1) & \cdots & \sqrt{2}\sin(\omega_{\tilde{L}-1}) \\ \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 1 & \sqrt{2}\cos(\omega_1(L-1)) & \cdots & \sqrt{2}\cos(\omega_{\tilde{L}-1}(L-1)) & \cos(\omega_{\tilde{L}}(L-1)) & \sqrt{2}\sin(\omega_1(L-1)) & \cdots & \sqrt{2}\sin(\omega_{\tilde{L}-1}(L-1)) \end{bmatrix}$$

This transform can be used to simplify the constraints to have the form $\bar{\theta}_0^2 \leq \frac{1}{L}$, $\bar{\theta}_l^2 + \bar{\theta}_{\tilde{L}+l}^2 \leq \frac{2}{L}$ for all $1 \leq l < \tilde{L}$ and $\bar{\theta}_{\tilde{L}}^2 \leq \frac{1}{L}$. Note that the transformation with $\mathbf{V}$ can be efficiently calculated using the *fast Fourier transform* (FFT).

(b) Recursive AMX: To reduce the computational complexity of RAMX, we consider the special case of $\mathbf{u} = \mathbf{0}$ and $\mathbf{M} = \alpha\mathbf{I}$, i.e. we only allow a shrinkage by the factor $1 + \alpha$. First, we transform the optimization problem (2.22) into the epigraphic[23] form

$$\min_{\mathbf{M},\mathbf{u},\tau} \ \tau \tag{3.41a}$$

subject to

$$\begin{bmatrix} \boldsymbol{\theta} \\ 1 \end{bmatrix}^T \begin{bmatrix} -\mathbf{M}^T\mathbf{M} & -\mathbf{M}^T\mathbf{u} \\ -\mathbf{u}^T\mathbf{M} & \tau - \mathbf{u}^T\mathbf{u} - \operatorname{tr}\{(\mathbf{I}+\mathbf{M})\mathbf{R}_n^{-1}(\mathbf{I}+\mathbf{M})^T\} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta} \\ 1 \end{bmatrix} \geq 0 \ \forall\, \boldsymbol{\theta} \in \mathbb{T}, \tag{3.41b}$$

which is better suited to deal with the constraint (3.40). Using $\mathbf{u} = \mathbf{0}$ and $\mathbf{M} = \alpha\mathbf{I}$, (3.41) simplifies to

$$\min_{\alpha,\tau} \ \tau \tag{3.42a}$$

subject to

$$\begin{bmatrix} \boldsymbol{\theta} \\ 1 \end{bmatrix}^T \begin{bmatrix} -\alpha^2\mathbf{I} & \mathbf{0} \\ \mathbf{0}^T & \tau - \operatorname{tr}\{(1+\alpha)^2\mathbf{R}_n^{-1}\} \end{bmatrix} \begin{bmatrix} \boldsymbol{\theta} \\ 1 \end{bmatrix} \geq 0 \ \forall\, \boldsymbol{\theta} \in \mathbb{T}. \tag{3.42b}$$

Condition (3.40) is equivalent to

$$\begin{bmatrix} \tilde{\boldsymbol{\theta}}(n) \\ 1 \end{bmatrix}^T \begin{bmatrix} -\tilde{\mathbf{f}}_l\tilde{\mathbf{f}}_l^H & \mathbf{0} \\ \mathbf{0}^T & 1 \end{bmatrix} \begin{bmatrix} \tilde{\boldsymbol{\theta}}(n) \\ 1 \end{bmatrix} \geq 0 \quad \forall\, l = 0, \ldots, \tilde{L} \tag{3.43}$$

which has to be fulfilled for all $l = 0, \ldots, \tilde{L}$. Now, the S-procedure can be used to reformulate the optimization problem (3.42) into a SDP [Boyd and Vandenberghe, 2007]. The S-procedure shows that a sufficient condition for the statement

$$\text{for all } \mathbf{z}\text{: } \mathbf{z}^T\mathbf{F}_0\mathbf{z} \geq 0, \ldots, \mathbf{z}^T\mathbf{F}_{\tilde{L}}\mathbf{z} \geq 0 \Rightarrow \mathbf{z}^T\mathbf{G}\mathbf{z} \geq 0$$

to be true is the existence of $\lambda_0, \ldots, \lambda_{\tilde{L}} \geq 0$ such that $\mathbf{G} \succeq \lambda_0\mathbf{F}_0 + \cdots + \lambda_{\tilde{L}}\mathbf{F}_{\tilde{L}}$. Please see

[23] The *epigraph* $\operatorname{epi}\{f\}$ of a function $f : \mathbb{R}^n \to \mathbb{R}$ is the set of all points lying on or above its graph, i.e. $\operatorname{epi}\{f\} = \{(\mathbf{x}, \tau) \in \mathbb{R}^{n+1} : \tau \geq f(\mathbf{x})\}$.

Appendix C.2.1 for a detailed discussion of the S-procedure. The optimization problem can therefore be rewritten as

$$\min_{\substack{\lambda_0 \geq 0, \ldots, \lambda_{\tilde{L}} \geq 0 \\ \tau, \alpha}} \tau \tag{3.44a}$$

subject to

$$\lambda_0 \tilde{\mathbf{f}}_0 \tilde{\mathbf{f}}_0^H + \cdots + \lambda_{\tilde{L}} \tilde{\mathbf{f}}_{\tilde{L}} \mathbf{f}_{\tilde{L}}^H \succeq \alpha^2 \mathbf{I}, \tag{3.44b}$$

$$\tau - (1+\alpha)^2 \operatorname{tr}\{\mathbf{R}_n^{-1}\} \geq \lambda_0 + \cdots + \lambda_{\tilde{L}}. \tag{3.44c}$$

Note that (3.44) is still not a SDP as (3.44b) and (3.44c) are not linear in α. However, using the same idea as in [Eldar, 2008b] and introducing a new variable x and the constraint $x \geq \alpha^2$, we finally obtain a SDP which has the same solution as (3.44). Furthermore, the constraint (3.44b) can be simplified by exploiting the fact that $\mathbf{U}^H \tilde{\mathbf{f}}_l \tilde{\mathbf{f}}_l^H \mathbf{U} = L\mathbf{J}^{l,l}$ where $\mathbf{J}^{i,j}$ is the single-entry matrix which is all zero except for a one at position i, j. Hence, we can rewrite (3.44b) as the new set of constraints $\lambda_l \geq \alpha^2 / L$ for all $l = 0, \ldots, \tilde{L}$ and the following SDP is equivalent to (3.44)

$$\min_{\substack{\lambda_0 \geq 0, \ldots, \lambda_{\tilde{L}} \geq 0 \\ \tau, \alpha, x}} \tau \tag{3.45a}$$

subject to

$$\lambda_l \geq x/L \quad (l = 0, \ldots, \tilde{L}), \tag{3.45b}$$

$$\begin{bmatrix} x & \alpha \\ \alpha & 1 \end{bmatrix} \succeq \mathbf{0}, \tag{3.45c}$$

$$\tau - (1 + 2\alpha + x) \operatorname{tr}\{\mathbf{R}_n^{-1}\} \geq \lambda_0 + \cdots + \lambda_{\tilde{L}}. \tag{3.45d}$$

(c) Recursive MMSE: For the RMMSE, we use directly (3.38) for the rejection sampling. A FFT of length L is used to calculate $\Theta(\omega, n)$ at discrete frequency points ω_l. Note that (3.38) describes a convex set. Therefore, we know that the estimate $\hat{\boldsymbol{\theta}}_{\text{RMMSE}}(n)$ of the RMMSE will lie in $\mathbb{T}$, i.e. it fulfills the energy conservation constraint as was pointed out in Section 3.4.

Simulation Results

To compare the estimators, we consider the problem of estimating the time-varying RIR of a moving source to a fixed microphone. The room size is $3 \times 3 \times 2.3$m. The source moves along a straight line with an increment of 10cm between neighbouring positions. There are $N = 11$ positions which implies that the source moves one meter in total. At each discrete source position n, the RIR $\boldsymbol{\theta}(n)$ is calculated using an image source model [Lehmann and Johansson, 2008]

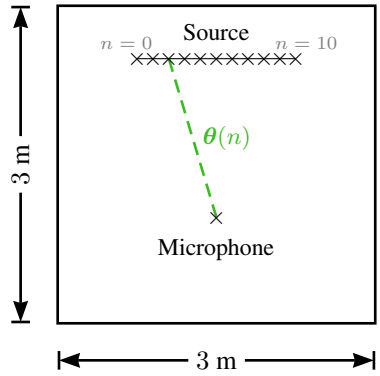


Figure 3.11: Time-varying RIR estimation setup

with the help of the Matlab implementation from [Lehmann, 2008]. Figure 3.11 summarizes the simulation setup.

The reverberation time T_{60} of the room, defined as the time for the reflections to fall 60 dB below the direct sound, is 120ms and the sampling frequency is $f_s = 12$kHz. The image source model returned RIRs with $M = 1241$ taps. The noise covariance is $\mathbf{C}(n) = \sigma^2\mathbf{I}$ and the corresponding SNR is defined as $\text{SNR} = 10\log_{10}\|\boldsymbol{\Theta}(n)\mathbf{s}(n)\|^2/((M+S-1)\sigma^2)$. To compare the estimators, the normalized error measure $E = 1/N\sum_{n=1}^{N}\|\hat{\boldsymbol{\theta}}(n) - \boldsymbol{\theta}(n)\|^2/\|\boldsymbol{\theta}(n)\|^2$ is used. The DFT length is chosen as $L = 16\,384$ and the signal $\mathbf{s}(n)$ is a white Gaussian noise of length $S = 100$ such that the total number of observations is $K = M + S - 1 = 1\,340$. The RMMSE estimator is calculated from $I = 3\,000$ samples to approximate the MMSE integrals.

Table 3.1 shows the simulation results for the four instantaneous estimators WLS, CML, AMX and MMSE, which rely only on $\mathbf{x}(n)$ and do not take past measurements into account (i.e. $\beta = 0$) and the recursive estimators RWLS, RCML, RAMX and RMMSE. The results are averaged over 200 trials. The optimal forgetting factors are $\beta_{\text{opt}} \approx 0.55$ for a SNR of 5 dB and $\beta_{\text{opt}} \approx 0.28$ for a SNR of 10 dB. The results show that the energy conservation constraint improves the estimation performance, especially for a SNR of 5 dB as the additional information (3.34) is more important if the SNR is small. If the SNR is higher then all estimators are equivalent, e.g. the RCML does not have to solve for the minimum on the boundary of $\mathbb{T}$ as $\hat{\boldsymbol{\theta}}_{\text{RWLS}}(n) \in \mathbb{T}$. The MMSE/RMMSE with a uniform prior shows the best results among all estimators.

SNR	WLS	CML	AMX	MMSE
5 dB	1.23×10^{0}	1.20×10^{0}	1.21×10^{0}	1.02×10^{0}
10 dB	3.88×10^{-1}	3.88×10^{-1}	3.87×10^{-1}	3.72×10^{-1}
SNR	**RWLS**	**RCML**	**RAMX**	**RMMSE**
5 dB	7.02×10^{-1}	6.97×10^{-1}	6.99×10^{-1}	6.52×10^{-1}
10 dB	2.98×10^{-1}	2.98×10^{-1}	2.98×10^{-1}	2.94×10^{-1}

Table 3.1: Example 3: Normalized error of the estimated RIRs

3.6 Observations and Conclusions

To conclude this chapter, we want two point out two observations which are worthwhile to mention:

First, the computational complexity of the recursive affine minimax estimator can be simplified if $\mathbf{M}(n)$ and $\mathbf{u}(n)$ are modified such that they exhibit a special structure. This was e.g. used in the third example, where we used $\mathbf{M}(n) = \alpha(n)\mathbf{I}$ and then only $\alpha(n)$ had to be found during the optimization step. This possibility together with the option to use a steady-state approximation of $\alpha(n)$, i.e. $\alpha(n) = \alpha$, renders the RAMX suitable for practical applications as α can be computed offline.

Second, the RMMSE estimator showed for all three examples the best results. This was expected in the second example as the true a priori density was a uniform distribution. However, this could not be expected for the other examples. Therefore, we can conclude that using a uniform prior to model the constraints is a useful technique and the RMMSE estimator should be further investigated. Especially the task of sampling from the a posteriori density is an interesting topic which should be studied in a future work.

Chapter 4

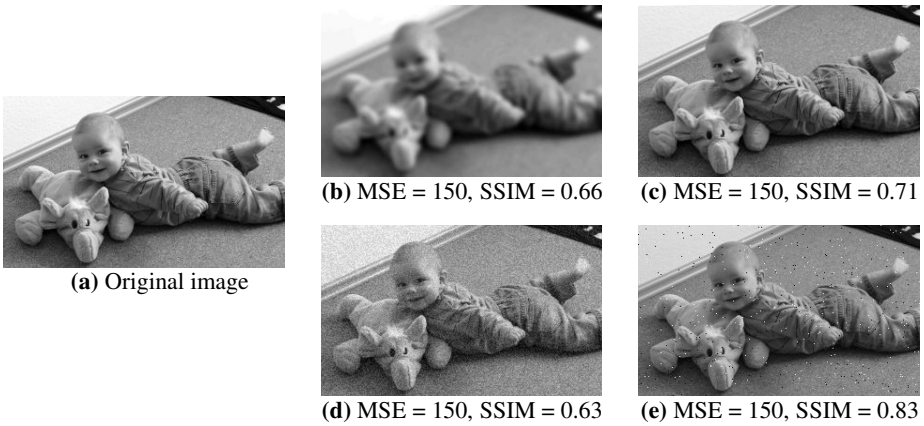
Incorporation of Additional Information using Nonstandard Loss Functions

4.1 Overview

Most often in Bayesian estimation, the MMSE estimator (2.27b) or MAP estimator (2.28b) are used to estimate an unknown parameter $\boldsymbol{\theta} \in \mathbb{T} \subset \mathbb{R}^M$ from the data $\mathbf{x} \in \mathbb{R}^K$. As we have seen in Section 2.4.2, the underlying loss functions $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))$ are the squared loss (2.27a) and the hit-or-miss loss (2.28a). The reason that they are used so widely is often not their suitability to the problem at hand but that the corresponding OBEs are well known and, at least for the MAP estimator, are often computable as they are the mean and maximum of the a posteriori density $p(\boldsymbol{\theta}|\mathbf{x})$ and powerful methods are available to calculate the estimate $\hat{\boldsymbol{\theta}}(\mathbf{x})$, see Section 2.4.3.

In this chapter, we will consider Bayesian estimation with other than those loss functions. This problem is very important for practical applications as the loss function should reflect the cost that is connected with a certain estimation error, see. e.g. [Gelman et al., 2003; Norstrom, 1996]. The following two examples illustrate this more clearly:

- Consider the problem of constructing a dam [Zellner, 1986]. Underestimating the peak water level from older measurements is clearly more serious than overestimating it and this fact should be reflected in the choice of the loss function $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))$. This example motivates the use of an asymmetric loss function, i.e. $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \neq L(-\boldsymbol{\theta}, -\hat{\boldsymbol{\theta}})$ and it is obvious that the hit-or-miss loss and the squared loss are not suited for such an estimation problem.
- Another example that gives rise to other loss functions can be found in the field of image processing. Traditionally, the mean squared error is used to compare images and therefore many algorithms are optimized for this loss function [Wang and Bovik, 2009]. The problem with the MSE is that it does not well represent human perception as is shown in Figure 4.1. Images which have a small mean squared error may still look very



(a) Original image
(b) MSE = 150, SSIM = 0.66
(c) MSE = 150, SSIM = 0.71
(d) MSE = 150, SSIM = 0.63
(e) MSE = 150, SSIM = 0.83

Figure 4.1: Comparison of MSE with SSIM for different image operations
(b) = Gaussian blurring with $\sigma = 2.27$, (c) = image rotation by $\beta = 0.7°$,
(d) = additive uniform noise, (e) = salt and pepper noise

different and therefore in [Wang and Bovik, 2009] it is suggested to use other distance measures. One is the structural similarity (SSIM) index, which was introduced by Wang in [Wang et al., 2004] and e.g. used in [Channappayya et al., 2008] for the design of linear equalizers. Figure 4.1 clearly shows that the SSIM index better correlates to the human perception than the MSE as the SSIM index for (b) and (d) are smaller compared to the SSIM index for (c) and (e). A related example that discusses the design of loss functions for the reconstruction of images is given by Rue in [Rue, 1995]. He shows how information about the image structure can be used to find a suitable loss function and he proposes the use of MCMC and simulated annealing methods to calculate the Bayesian estimates.

More examples of Bayesian estimation with non-standard loss functions can also be found in cluster analysis [Binder, 1978, 1981; Lau and Green, 2007] and mixture modeling [Celeux et al., 2000; Hurn et al., 2003].

Calculating the OBE for many nonstandard loss functions is nontrivial and can often only be stated in terms of the optimization problem (2.24) which has to be solved for each new data $\mathbf{x}$. Therefore, we propose in this chapter a parametric family $\mathbb{F}$ of estimators which are suited for a large variety of loss functions but still have a computational complexity comparable to the MMSE estimator for the same problem. Thus, using the best estimator in $\mathbb{F}$ that has the smallest Bayes risk for a given loss function will be a good approximation of the OBE. Our parametric family of estimators can be viewed as a compromise between the perfect OBE on one side and a (nonlinear) regression approach on the other. It trades off performance against computational complexity as it will have a larger Bayes risk than the OBE but will be easier to learn due to the small and fixed number of parameters compared to a regression approach.

In the following, we will introduce our parametric family of estimators and show its relation to many well-known estimators, including the MAP, MMSE and OBE for the LinEx loss.

4.2 Parametric Family of Estimators

4.2.1 Basic Family of Estimators

Let $\mathbb{F}_{\mathrm{B}}$ be the set of estimators that have the form

$$\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda) = \frac{\int \boldsymbol{\theta}\, p(\boldsymbol{\theta},\mathbf{x})^{\lambda} d\boldsymbol{\theta}}{\int p(\boldsymbol{\theta},\mathbf{x})^{\lambda} d\boldsymbol{\theta}} \tag{4.1}$$

and are parameterized by λ. We call $\mathbb{F}_{\mathrm{B}}$ the *basic family of estimators*. Thinking of $p(\boldsymbol{\theta},\mathbf{x})^{\lambda}$ as a new (unnormalized) density, we see that (4.1) calculates the mean of the conditional density $p(\boldsymbol{\theta},\mathbf{x})^{\lambda}/\int p(\boldsymbol{\theta},\mathbf{x})^{\lambda} d\boldsymbol{\theta}$ and therefore looks similar to the MMSE estimator except for the modified PDF.

Note that it is reasonable to restrict λ to positive values, i.e. $\lambda \in [0,\infty)$. Otherwise we average over a new density $p(\boldsymbol{\theta},\mathbf{x})^{\lambda}/\int p(\boldsymbol{\theta},\mathbf{x})^{\lambda} d\boldsymbol{\theta}$ which is inverted in the sense that it has large values at positions where $p(\boldsymbol{\theta},\mathbf{x})$ is small, i.e. it emphasizes points $(\boldsymbol{\theta},\mathbf{x}) \in \mathbb{R}^{M+K}$ that are not likely to occur and we can expect therefore a poor performance for $\lambda < 0$.[24]

We will now show that $\mathbb{F}_{\mathrm{B}}$ includes three important estimators, namely the uniform a priori MMSE estimator, the MMSE estimator and the MAP estimator. By uniform a priori MMSE estimator, we refer to the estimator where we have no data $\mathbf{x}$ about $\boldsymbol{\theta} \in \mathbb{T} \subset \mathbb{R}^M$ and the a priori distribution $p(\boldsymbol{\theta})$ is assumed to be uniform in $\mathbb{T}$. The estimator with the minimum MSE is then the "center of gravity" of $\mathbb{T}$, i.e. $\hat{\boldsymbol{\theta}} = E[\boldsymbol{\theta}] = \int \boldsymbol{\theta}\, p(\boldsymbol{\theta}) d\boldsymbol{\theta} = \int_{\mathbb{T}} \boldsymbol{\theta} d\boldsymbol{\theta} / \int_{\mathbb{T}} 1 d\boldsymbol{\theta}$ which is well defined if $\mathbb{T}$ is bounded. The following theorem shows that all three estimators are in $\mathbb{F}_{\mathrm{B}}$. The proof can be found in Appendix C.3.1.

Theorem 2 (Relationship to other estimators)**.** *The estimator family $\mathbb{F}_B$ defined in* (4.1) *includes the following special cases:*

(a) If $\mathbb{T} \subset \mathbb{R}^M$ is bounded and $p(\boldsymbol{\theta},\mathbf{x}) \neq 0$ then $\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda)$ for $\lambda \to 0$ exists and is equivalent to the uniform a priori MMSE estimator.

(b) The case $\lambda = 1$ corresponds to the MMSE estimator.

(c) The case $\lambda \to \infty$ corresponds to the MAP estimator.

[24] For example the loss $L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}}) = 1 - L_{\mathrm{MAP}}(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})$ results in seeking the minimum of $p(\boldsymbol{\theta}|\mathbf{x})$ which is related (but in general not identical) to $\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda)$ for $\lambda \to -\infty$.

Although it is interesting to see the relationship of this basic family of estimators to other estimators, we also see that the loss functions associated with $\lambda \in \{0, 1, \infty\}$ are all symmetric as they are the squared loss (2.27a) and the hit-or-miss loss (2.28a). In the following, we will prove in Theorem 3 that if there is a continuously differentiable loss function that results in $\hat{\boldsymbol{\theta}}(\mathbf{x}; \lambda)$, then the loss function has to be symmetric.[25] For the proof of Theorem 3, we need the following Lemma. The proofs can be found in Appendix C.3.2.

Lemma 1. *The estimator $\hat{\boldsymbol{\theta}}(\mathbf{x}; \lambda)$ for the PDFs $p(\boldsymbol{\theta}, \mathbf{x}) = \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$ and $p(\boldsymbol{\theta}, \mathbf{x}) = P\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0) + (1-P)\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_1)$ is given by $\hat{\boldsymbol{\theta}}(\mathbf{x}; \lambda) = \boldsymbol{\theta}_0$ and $\hat{\boldsymbol{\theta}}(\mathbf{x}; \lambda) = (P^\lambda \boldsymbol{\theta}_0 + (1-P)^\lambda \boldsymbol{\theta}_1)/(P^\lambda + (1-P)^\lambda)$, respectively.*

Theorem 3 (Corresponding loss is symmetric)**.** *Let $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ be a continuously differentiable loss function that results in the optimal Bayesian estimator $\hat{\boldsymbol{\theta}}(\mathbf{x}; \lambda)$ for an arbitrary PDF $p(\boldsymbol{\theta}, \mathbf{x})$. Then $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ is symmetric, i.e. $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = L(-\boldsymbol{\theta}, -\hat{\boldsymbol{\theta}})$.*

From this Theorem, we see that no estimator resulting from an asymmetric, continuously differentiable loss function is included in $\mathbb{F}_{\mathrm{B}}$. However, we would like to use such asymmetric loss functions due to their practical relevance and hence we have to extend $\mathbb{F}_{\mathrm{B}}$. This is done in the next subsection.

4.2.2 A Generalized Family of Estimators

In order to extend the basic family of estimators $\mathbb{F}_{\mathrm{B}}$ given in (4.1), we will now modify its parametric form such that the OBE for LinEx loss is also included. By doing this, we obtain a new family of estimators $\mathbb{F}$ which can deal with the important case of asymmetric loss functions.

From (2.30b) we know that the OBE for LinEx loss has the form

$$\hat{\theta}_{\mathrm{LinEx},m} = -\frac{1}{a_m} \ln \int e^{-a_m \theta_m} p(\theta_m|\mathbf{x}) d\theta_m, \quad m = 1, \ldots, M. \tag{4.2}$$

Knowing the OBE for LinEx loss, we can now extend our basic family of estimators $\mathbb{F}_{\mathrm{B}}$. This will be done in such a way that the new family of estimators $\mathbb{F}$ is a kind of "superposition" of both $\mathbb{F}_{\mathrm{B}}$ and the OBE (4.2). We define this new family of estimators in the following way: Let $\mathbb{F}$ be the set of estimators where each estimator has the form

$$\hat{\boldsymbol{\theta}}(\mathbf{x}; \mathbf{p}) = \mathbf{f}_1 \left(\frac{\int \mathbf{f}_2(\boldsymbol{\theta}; \mathbf{p}_2)\, p(\boldsymbol{\theta}, \mathbf{x})^\lambda d\boldsymbol{\theta}}{\int p(\boldsymbol{\theta}, \mathbf{x})^\lambda d\boldsymbol{\theta}}; \mathbf{p}_1 \right) \tag{4.3a}$$

[25] Note that it is difficult to prove the existence of such a loss function for an arbitrary λ and the corresponding estimator $\hat{\boldsymbol{\theta}}(\mathbf{x}; \lambda)$.

and depends on the $2M + 4$ parameters $\mathbf{p} = \left[\lambda, \mathbf{p}_1^T, \mathbf{p}_2^T\right]^T$ with $\mathbf{p}_1 = \left[\xi_1, \phi_1, \ldots, \phi_M\right]^T$ and $\mathbf{p}_2 = \left[\xi_2, \xi_3, \psi_1, \ldots, \psi_M\right]^T$. $\mathbf{f}_1$ and $\mathbf{f}_2$ are defined as

$$\mathbf{f}_1(\mathbf{z}; \mathbf{p}_1) = \xi_1 \mathbf{z} + \boldsymbol{\phi} \circ \ln|\mathbf{z}|, \tag{4.3b}$$

$$\mathbf{f}_2(\mathbf{z}; \mathbf{p}_2) = \xi_2 \mathbf{z} + \xi_3 e^{\boldsymbol{\psi} \circ \mathbf{z}} \tag{4.3c}$$

with $\boldsymbol{\phi} = \left[\phi_1, \ldots, \phi_M\right]^T$ and $\boldsymbol{\psi} = \left[\psi_1, \ldots, \psi_M\right]^T$. Note that $e^{\mathbf{z}}$, $\ln \mathbf{z}$ and $|\mathbf{z}|$ are understood elementwise and the operator "$\circ$" denotes the elementwise multiplication of two vectors or matrices. λ is again chosen such that $\lambda \in [0, \infty)$ as discussed in Section 4.2.1.

First, we would like to note that $\mathbb{F}_{\mathrm{B}} \subset \mathbb{F}$ as all estimators $\hat{\boldsymbol{\theta}}(\mathbf{x}; \lambda)$ from (4.1) are included in $\hat{\boldsymbol{\theta}}(\mathbf{x}; \mathbf{p})$ for $\xi_1 = \xi_2 = 1$, $\xi_3 = 0$ and $\phi_1 = \cdots = \phi_M = 0$. Therefore, we already know from Theorem 2 that the uniform a priori MMSE, the MMSE and the MAP estimator are included in this family. Furthermore, it is straightforward to see that $\mathbb{F}$ also includes the OBE for LinEx loss as plugging in the values $\xi_1 = \xi_2 = 0$, $\xi_3 = 1$, $\lambda = 1$ and $\psi_m = 1/\phi_m = -a_m$ for $m = 1, \ldots, M$ into (4.3a) results in (4.2). Thus, we see that the new estimator family $\mathbb{F}$ is more general than $\mathbb{F}_{\mathrm{B}}$ and also includes estimators with asymmetric loss functions.

4.2.3 Practical Considerations

This section explains the general approach how to obtain the estimator for a given signal model and loss function and also shows how the estimate $\hat{\boldsymbol{\theta}}(\mathbf{x}; \mathbf{p})$ can be calculated efficiently for the given data $\mathbf{x}$. In the sequel, we will make the following two assumptions:

- The generation of samples $(\boldsymbol{\theta}_i, \mathbf{x}_i) \sim p(\boldsymbol{\theta}, \mathbf{x})$ is manageable, where $p(\boldsymbol{\theta}, \mathbf{x})$ is the joint PDF of $\boldsymbol{\theta}$ and $\mathbf{x}$. This is often the case as $p(\boldsymbol{\theta}, \mathbf{x})$ can be written as $p(\boldsymbol{\theta}, \mathbf{x}) = p(\mathbf{x}|\boldsymbol{\theta})p(\boldsymbol{\theta})$, where $p(\boldsymbol{\theta})$ is the a priori PDF of $\boldsymbol{\theta}$ and $p(\mathbf{x}|\boldsymbol{\theta})$ is the likelihood PDF. Very often, both are known: $p(\boldsymbol{\theta})$ from expert knowledge and $p(\mathbf{x}|\boldsymbol{\theta})$ through the signal model.
- The generation of samples $\boldsymbol{\theta}_i \sim p(\boldsymbol{\theta}|\mathbf{x})$ is manageable. This is not a hard restriction as the MMSE estimator is often calculated using Markov chain Monte Carlo methods [Robert, 2001; Liu, 2008]. MCMC allows the approximate generation of correlated samples from the a posteriori distribution and the MMSE estimator is then simply the average over all samples. Here, we will use importance sampling where the conditional distribution $p(\boldsymbol{\theta}|\mathbf{x})$ is the importance function.

Given the loss function and the signal model, the use of our estimator family for a general estimation problem consists of two steps:

Step 1 – Find the optimal estimator in $\mathbb{F}$

In a first step, we have to find the estimator $\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p}_0) \in \mathbb{F}$ that has the smallest Bayes risk for the particular loss function and joint PDF $p(\boldsymbol{\theta},\mathbf{x})$, i.e. we have to solve the optimization problem

$$\mathbf{p}_0 = \arg\min_{\mathbf{p}} \iint L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p}))p(\boldsymbol{\theta},\mathbf{x})d\boldsymbol{\theta}d\mathbf{x}. \tag{4.4}$$

This optimization has only to be carried out once to learn the optimal values of the parameters $\mathbf{p}$. In the Appendix C.3.3, we give the gradient vector of the Bayes risk in (4.4) with respect to the parameters in $\mathbf{p}$. The knowledge of the gradient vector allows to use a gradient descent method to find the locally[26] optimal parameter values. Note that the integration with respect to $\boldsymbol{\theta}$ and $\mathbf{x}$ can be carried out by a plain MC integration using samples $(\boldsymbol{\theta}_i, \mathbf{x}_i) \sim p(\boldsymbol{\theta},\mathbf{x})$. The optimization problem (4.4) becomes then

$$\mathbf{p}_0 = \arg\min_{\mathbf{p}} \frac{1}{I_1}\sum_{i=1}^{I_1} L(\boldsymbol{\theta}_i, \hat{\boldsymbol{\theta}}(\mathbf{x}_i;\mathbf{p})). \tag{4.5}$$

If the generation of samples from $p(\boldsymbol{\theta},\mathbf{x})$ is not directly possible, then importance sampling as discussed in Section 2.4.3 is another possibility to obtain an accurate approximation of the integral.

Step 2 – Calculate the estimate $\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p}_0)$

In a second step, we calculate the estimate for the given data $\mathbf{x}$. Therefore, we need an efficient method to compute both integrals in (4.3a). Note that (4.3a) can be written as

$$\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p}) = \mathbf{f}_1\left(\int \mathbf{f}_2(\boldsymbol{\theta};\mathbf{p}_2)\frac{p(\boldsymbol{\theta},\mathbf{x})^\lambda}{\int p(\boldsymbol{\theta},\mathbf{x})^\lambda d\boldsymbol{\theta}}d\boldsymbol{\theta};\mathbf{p}_1\right) = \mathbf{f}_1\left(E_{p_\lambda}\left[\mathbf{f}_2(\boldsymbol{\theta};\mathbf{p}_2)\right];\mathbf{p}_1\right). \tag{4.6}$$

We see that we can write the integrals as the expectation of $\mathbf{f}_2(\boldsymbol{\theta};\mathbf{p})$ with respect to a new conditional density $p_\lambda(\boldsymbol{\theta}|\mathbf{x}) = p(\boldsymbol{\theta},\mathbf{x})^\lambda / \int p(\boldsymbol{\theta},\mathbf{x})^\lambda d\boldsymbol{\theta}$. Assuming that we can generate samples from the a posteriori distribution $\boldsymbol{\theta}_k \sim p(\boldsymbol{\theta}|\mathbf{x}) = p(\boldsymbol{\theta},\mathbf{x}) / \int p(\boldsymbol{\theta},\mathbf{x})d\boldsymbol{\theta}$, we can use importance sampling as shown in Section 2.4.3 for (4.6). Generalizing (2.42) from Section 2.4.3 to the more general case of calculating $\mathrm{E}[h(\boldsymbol{\theta})] = \int h(\boldsymbol{\theta})p(\boldsymbol{\theta})d\boldsymbol{\theta}$, the importance sampling approximation is

$$\mathrm{E}\left[h(\boldsymbol{\theta})\right] \approx \frac{\sum_{i=1}^{I} w_i h(\boldsymbol{\theta}_i)}{\sum_{i=1}^{I} w_i} \tag{4.7}$$

[26] In general, the gradient descent method will only converge to a local minimum of (4.4) and therefore it is often advantageous to restart the descent method with a different start value.

where $\boldsymbol{\theta}_i$ are drawn from a trial distribution $g(\boldsymbol{\theta})$ and the importance weights w_i are defined as $w_i = p(\boldsymbol{\theta}_i)/g(\boldsymbol{\theta}_i)$. Note that w_i has only to be known up to a multiplicative constant. Using importance sampling for our problem, we finally obtain the approximation

$$\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p}) \approx \mathbf{f}_1 \left(\frac{\sum_{i=1}^{I_2} w_i \; \mathbf{f}_2(\boldsymbol{\theta}_i;\mathbf{p}_2)}{\sum_{i=1}^{I_2} w_i} ;\mathbf{p}_1 \right) \tag{4.8}$$

with $g(\boldsymbol{\theta}) = p(\boldsymbol{\theta},\mathbf{x})$ and thus $w_i = p(\boldsymbol{\theta}_i,\mathbf{x})^{\lambda-1}$. The computational complexity is hence comparable to that of a MMSE estimation if the MMSE estimator also uses MC integration.

4.3 Examples

In the following, we will give three examples that compare the OBE for a nonstandard loss function to our estimator family $\mathbb{F}$.

4.3.1 Example 1: BLinEx Loss

The first example is as follows: Given the signal model $x = \theta + z$, estimate θ which is uniformly distributed in $[0,1]$ from the observation x where we know that the observation is perturbed by additive Gaussian noise $z \sim \mathcal{N}(0,\sigma^2)$. Furthermore, z and θ are independently distributed. The considered loss function is the bounded LinEx (BLinEx) loss introduced in [Wen and Levy, 2001]. The univariate BLinEx loss function is given by

$$L_{\text{BLinEx}}(\theta,\hat{\theta}) = \frac{L_{\text{LinEx}}(\theta,\hat{\theta})}{1+\rho L_{\text{LinEx}}(\theta,\hat{\theta})}, \; \rho > 0. \tag{4.9}$$

Plugging $L_{\text{LinEx}}(\theta,\hat{\theta})$ from (2.29) into (4.9), we obtain

$$L_{\text{BLinEx}}(\theta,\hat{\theta}) = \frac{1}{\rho}\left(1 - \frac{1}{1+c(e^{a(\hat{\theta}-\theta)} - a(\hat{\theta}-\theta) - 1)}\right) \tag{4.10}$$

with $c = \rho b$. It differs from the usually used loss functions (2.27a) and (2.28a) in two main properties, namely it is (a) asymmetric and (b) bounded:

(a) If $a > 0$ then the positive error $\hat{\theta} > \theta$ results in a larger loss than the corresponding negative error of the same magnitude. If $a < 0$ then negative errors $\hat{\theta} < \theta$ have a larger loss. A case where such an emphasis of negative errors is useful is the dam construction example given in Section 4.1 as underestimating the peak water level is more severe than overestimating it.

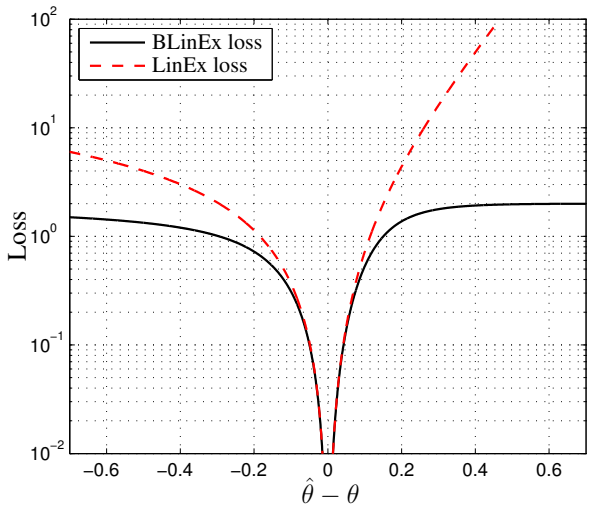


Figure 4.2: Example 1: LinEx and BLinEx loss ($\rho = 0.5$, $a = 10$ and $b = 1$)

(b) $L_{\text{BLinEx}}(\theta, \hat{\theta})$ is bounded by 0 and $1/\rho$. Such a requirement for a loss function may occur naturally out of the considered problem or may be artificially introduced to improve the robustness of the estimator in the case of outliers.

In our example, we choose $\rho = 0.5$, $a = 10$ and $b = 1$[27]. Figure 4.2 shows the graph of the BLinEx loss function for this choice of parameters. Furthermore, the noise variance is $\sigma^2 = 0.25$.

We compare the following five estimators with respect to the squared error loss (2.27a) and the BLinEx loss (4.10):

- *MAP estimator*: The MAP estimator is in general given by $\hat{\theta} = \arg\max_\theta p(\theta|x)$ with $p(\theta|x) \sim p(\theta, x) = e^{-(x-\theta)^2/(2\sigma^2)} u_{[0,1]}(\theta)$ and $u_{[0,1]}(\theta)$ is the a priori PDF of θ which is uniformly distributed in $[0, 1]$. This yields

$$\hat{\theta}_{\text{MAP}} = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & x > 1 \end{cases} \tag{4.11}$$

- *MMSE estimator*: The MMSE estimator is given by $\hat{\theta}_{\text{MMSE}} = E[\theta|x]$. For our signal model, the conditional mean can be calculated analytically and one obtains

[27] We choose these parameter values in order to achieve the following two effects: First, we want to study an asymmetric loss function and therefore a has to be large. Second, we want a loss function which is bounded and therefore different from the LinEx loss. To see this effect, we choose $\rho = 0.5$.

Estimator	Mean squared error loss	Mean BLinEx loss
MAP estimator	1.21×10^{-1}	1.02×10^{0}
MMSE estimator	6.28×10^{-2}	9.03×10^{-1}
OBE for LinEx loss	1.23×10^{-1}	8.70×10^{-1}
Optimal estimator $\in \mathbb{F}$	8.16×10^{-2}	8.21×10^{-1}
OBE for BLinEx loss	8.70×10^{-2}	8.12×10^{-1}

Table 4.1: Example 1: Comparison of the Bayes risks

$$\hat{\theta}_{\text{MMSE}} = x + \sqrt{\frac{2}{\pi}}\sigma \frac{e^{-\frac{x^2}{2\sigma^2}} - e^{-\frac{(x-1)^2}{2\sigma^2}}}{\operatorname{erf}\left(\frac{x}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{x-1}{\sqrt{2}\sigma}\right)} \tag{4.12}$$

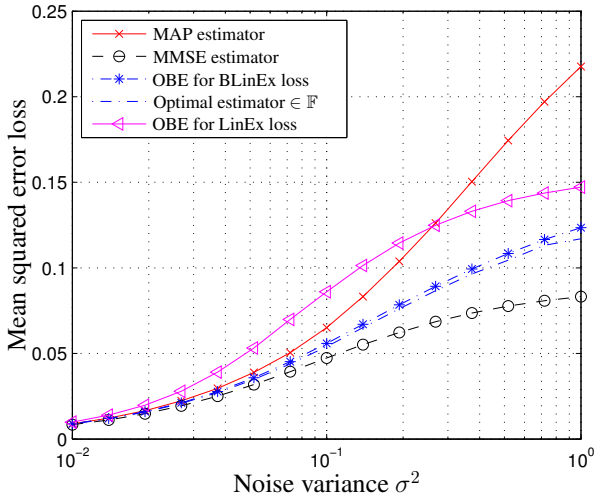
- *OBE for LinEx loss*: The OBE for LinEx loss is given by (2.30b) which can be calculated analytically. It is given by

$$\hat{\theta}_{\text{OBE,LinEx}} = x - \frac{a\sigma^2}{2} - \frac{1}{a}\ln\left(\frac{\operatorname{erf}\left(\frac{1+a\sigma^2-x}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{a\sigma^2-x}{\sqrt{2}\sigma}\right)}{\operatorname{erf}\left(\frac{x}{\sqrt{2}\sigma}\right) - \operatorname{erf}\left(\frac{x-1}{\sqrt{2}\sigma}\right)}\right) \tag{4.13}$$

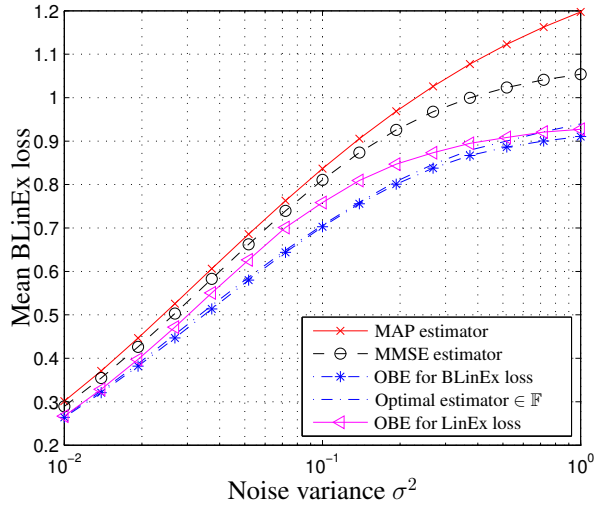
- *OBE for BLinEx loss*: The optimization problem (2.24) for this example cannot be carried out analytically and thus (2.24) has to be solved for each new observation x individually, either by Monte Carlo integration or numerical quadrature. For our simulations, we used the Matlab functions `fminunc` and `quad` to solve (2.24).

- *Estimator family* (4.3) *with optimal parameters*: The optimal parameters are found via the Matlab function `fmincon` using 50 random start points for the gradient descent. The found parameters are $\xi_1 \approx 6.77\times10^{-1}$, $\xi_2 \approx 4.03\times10^{-1}$, $\xi_3 = 1.33\times10^{-1}$, $\lambda \approx 8.31$, $\phi \approx 4.02\times10^{-3}$ and $\psi \approx 1.91$. $I_1 = 5\,000$ samples are used for the Monte Carlo approximation in (4.5) and $I_2 = 5\,000$ samples are drawn from the a posteriori density $p(\theta|x)$ for (4.8) using the sampling method proposed in Robert [1995]. The values for I_1 and I_2 were found by simulations to ensure statistical stable results of the MC integral approximations.

Table 4.1 shows the results averaged over $10\,000$ trials. Clearly, the MMSE estimator is optimal in terms of the squared error loss as expected. Similarly, the OBE for the BLinEx loss gives the smallest Bayes risk if the BLinEx loss function is used. The optimal estimator $\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p}_0)$ from the set $\mathbb{F}$ is a good approximation of the OBE for the BLinEx loss as it has a similar Bayes risk. Thus, although the OBE for the BLinEx loss itself is not an element of $\mathbb{F}$, there is an estimator $\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p}_0)$ in $\mathbb{F}$ which gives nearly the same performance.



(a) Squared error



(b) BLinEx loss

Figure 4.3: Example 1: Squared error and BLinEx loss for a varying noise variance σ^2

In order to study the influence of the noise variance on the simulation results, we rerun the first experiment with varying σ^2 values. Figure 4.3 shows the simulation results and it can be concluded that the relative performance of $\hat{\boldsymbol{\theta}}(\mathbf{x}; \mathbf{p}_0)$ with respect to the OBE for BLinEx loss is almost constant.

Finally, the run times to compute the estimates on a standard desktop computer are given in Table 4.2 in order to compare the computational costs of the different approaches. It can be observed that the run time of the OBE for BLinEx loss is roughly ten times larger as for our estimator family which justifies to use the approximation given by (4.8) rather than the OBE

Estimator	Run time
MAP estimator	$< 1\times10^{-3}$ sec.
MMSE estimator	$< 1\times10^{-1}$ sec.
OBE for LinEx loss	$< 1\times10^{-1}$ sec.
Optimal estimator $\in \mathbb{F}$	2.3×10^{1} sec.
OBE for BLinEx loss	3.2×10^{2} sec.

Table 4.2: Example 1: Comparison of the run times for $10\,000$ trials

itself. Note that the computation of the estimator family according to (4.8) consists of two steps: First, we have to sample from the a posteriori distribution which in our case is a truncated Gaussian density. We used the sampling algorithm proposed by Robert in [Robert, 1995] for this step. Second, we have to use importance sampling as shown in (4.8) to find the estimate. The run time for the first step is 21 seconds and for the second step 2 seconds which results in the 23 seconds that are given in Table 4.2. These numbers show that most of the run time is spent on computing samples from the a posteriori density.

4.3.2 Example 2: Speech Enhancement

The second example which we consider is the enhancement of a distorted speech signal. The goal is to suppress an unwanted noise signal while leaving the speech as undistorted as possible, see e.g. [Benesty et al., 2005; Loizou, 2011].

In the time domain, the speech enhancement problem can be written as

$$x(n) = s(n) + z(n), \tag{4.14}$$

where $s(n)$ is the original (clean) speech signal at time instance n which is distorted by noise $z(n)$ to result in the observed signal $x(n)$. One solution for this problem is the traditional approach of *short-time spectral attenuation* (STSA) which was introduced by [Berouti et al., 1979; Ephraim and Malah, 1984] and extended in later work [Ephraim and Malah, 1985; Loizou, 2005; You et al., 2005]. While [Berouti et al., 1979] is based on the method of *spectral subtraction*, the other papers use a more statistically motivated approach by introducing a suitable loss function and signal model for each frequency bin. The corresponding OBE is then used to perform the STSA operation.

In the following, we will state the speech enhancement problem in the frequency domain where we assume a Gaussian signal model. All necessary elements to use our family of estimators from (4.3) are derived and this estimator is then compared to the OBE.

Problem Formulation and Solution Approach

Using the short-time Fourier transform of (4.14), the signal model can be written in the frequency domain as

$$X_{l,i} = S_{l,i} + Z_{l,i}, \tag{4.15}$$

where $X_{l,i} = |X_{l,i}|e^{j\theta_{l,i}}$, $S_{l,i} = |S_{l,i}|e^{j\phi_{l,i}}$ and $Z_{l,i}$ are the lth spectral component of the noisy signal $x(n)$, clean speech $s(n)$ and noise $z(n)$ in the ith frame. The frequency index l ranges from 0 to $L-1$ where L is the FFT length. In STSA, the speech enhancement problem is solved by using

$$\hat{S}_{l,i} = |\hat{S}_{l,i}|e^{j\theta_{l,i}}, \tag{4.16}$$

i.e. the amplitude $|X_{l,i}|$ of the noisy spectral component $X_{l,i}$ is replaced by the estimate $|\hat{S}_{l,i}| = |\hat{S}_{l,i}|(X_{l,i})$. For convenience, we will drop the dependence of the spectral components on the frame index i in the following.

Using the Gaussian model $S_l = |S_l|e^{j\phi_l} \sim \mathcal{CN}(0, \sigma_s^2(l))$, i.e. S_l is *circular complex Gaussian*, we know that the PDF of $|S_l|$ and ϕ_l is given by

$$p(|S_l|, \phi_l) = \begin{cases} \frac{1}{2\pi}\frac{|S_l|}{\sigma_s^2(l)/2}e^{-\frac{|S_l|^2}{\sigma_s^2(l)}} & |S_l| \geq 0, 0 \leq \phi_l < 2\pi \\ 0 & \text{otherwise} \end{cases}, \tag{4.17}$$

i.e. $|S_l|$ follows a *Rayleigh* distribution, ϕ_l is uniformly distributed on $[0, 2\pi)$ and both are independent of each other. Assuming furthermore $Z_l \sim \mathcal{CN}(0, \sigma_z^2(l))$ and Z_l is independent of S_l, then we have clearly $X_l \sim \mathcal{CN}(0, \sigma_s^2(l) + \sigma_z^2(l))$. The a posteriori density $p\left(|S_l|\Big|X_l\right)$ for $|S_l| \geq 0$ is given by

$$\begin{aligned} p\left(|S_l|\Big|X_l\right) &= \frac{1}{p(X_l)}\int_0^{2\pi} p\left(X_l\Big||S_l|, \phi_l\right)p(|S_l|, \phi_l)d\phi_l \\ &= \frac{|S_l|(\sigma_z^2(l) + \sigma_s^2(l))}{\pi\sigma_z^2(l)\sigma_s^2(l)}\exp\left\{-\frac{\sigma_z^2(l) + \sigma_s^2(l)}{\sigma_z^2(l)\sigma_s^2(l)}|S_l|^2 - \frac{\sigma_s^2(l)}{\sigma_z^2(l)(\sigma_z^2(l) + \sigma_s^2(l))}|X_l|^2\right\} \\ &\quad \times \int_0^{2\pi}\exp\left\{\frac{2|S_l||X_l|}{\sigma_z^2(l)}\cos\left(\phi_l - \theta_l\right)\right\}d\phi_l. \end{aligned} \tag{4.18}$$

Introducing the *modified Bessel function of the first kind and nth order* $I_n(z)$ which has the integral representation

$$I_n(z) = \frac{1}{2\pi}\int_0^{2\pi}\cos(\beta n)\exp\left\{z\cos(\beta)\right\}d\beta, \tag{4.19}$$

and using the shorthand notations $v_l = \frac{\sigma_s^2(l)}{\sigma_z^2(l)(\sigma_z^2(l)+\sigma_s^2(l))}|X_l|^2$ and $\lambda_l^{-1} = \frac{\sigma_z^2(l)+\sigma_s^2(l)}{\sigma_z^2(l)\sigma_s^2(l)}$, we can finally write the a posteriori density $p\left(|S_l|\Big|X_l\right)$ as

$$p\left(|S_l|\Big|X_l\right) = 2\frac{|S_l|}{\lambda_l}\exp\left\{-\frac{|S_l|^2}{\lambda_l} - v_l\right\} I_0\left(2|S_l|\sqrt{v_l\lambda_l^{-1}}\right). \tag{4.20}$$

This density is well known in the literature and shows that $|S_l|$ given the observation X_l follows a *Rice distribution* [Papoulis and Pillai, 2002]. It is interesting to note that $p(|S_l|\Big|X_l)$ only depends on $|X_l|$ and therefore, $|\hat{S}_l| = |\hat{S}_l|(|X_l|)$. To derive the OBE in the next Section, we will need to calculate the moments $E\left[|S_l|^m\Big|X_l\right]$. Interestingly, they can be given analytically using *Kummer's function* $M(a,b,z)$ as defined in Appendix C.3.4 and they are

$$E\left[|S_l|^m\Big|X_l\right] = \lambda_l^{m/2}\Gamma\left(\frac{m}{2}+1\right)M\left(-\frac{m}{2},1,-v_l\right) \tag{4.21}$$

for all $m > -2$ where $\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt$ is the *Gamma function*. Eq. (4.21) results from the identities [Abramowitz and Stegun, 1964, 11.4.28] and [Abramowitz and Stegun, 1964, 13.1.27].

Loss Function and Corresponding OBE

In the literature, many different loss functions were proposed to perform STSA speech enhancement. The first approach in [Ephraim and Malah, 1984] was to use the squared loss function $L(|S_l|,|\hat{S}_l|) = (|S_l| - |\hat{S}_l|)^2$ which results in the *MMSE-STSA* algorithm. Later, other loss functions were proposed in [Ephraim and Malah, 1985; Loizou, 2005; You et al., 2005] which show a better performance with respect to perceptual motivated quality measures, e.g. the *perceptual evaluation of speech quality* (PESQ) measure [ITU-T Recommendation P.862; Möller et al., 2011]. In [Plourde and Champagne, 2008], these loss functions were combined into a family of loss functions of the form

$$L(|S_l|,|\hat{S}_l|) = \left(\frac{|S_l|^\beta - |\hat{S}_l|^\beta}{|S_l|^\alpha}\right)^2. \tag{4.22}$$

This loss function was later generalized in [Plourde and Champagne, 2009] to include even more proposed loss functions. The corresponding OBE for (4.22) can easily be found by using (2.25) together with $\partial L/\partial|\hat{S}_l| = -2\beta|\hat{S}_l|^{\beta-1}(|S_l|^\beta - |\hat{S}_l|^\beta)/|S_l|^{2\alpha}$ and is given by

$$|\hat{S}_l| = \left(\frac{\int_0^\infty |S_l|^{\beta-2\alpha}p\left(|S_l|\Big||X_l|\right)d|S_l|}{\int_0^\infty |S_l|^{-2\alpha}p\left(|S_l|\Big||X_l|\right)d|S_l|}\right)^{\frac{1}{\beta}} = \left(\frac{\mathrm{E}\left[|S_l|^{\beta-2\alpha}\Big||X_l|\right]}{\mathrm{E}\left[|S_l|^{-2\alpha}\Big||X_l|\right]}\right)^{\frac{1}{\beta}}. \tag{4.23}$$

Simulation Results

In the following, we will compare the OBE for the loss function (4.22) with the best estimator from the generalized family (4.3). Two experiments are conducted: In the first experiment, we find the best estimator in $\mathbb{F}$ with respect to the loss function (4.22) for $\alpha = 0.5$ and $\beta = 1$. This parameter setup was shown in [Plourde and Champagne, 2008] to result in an STSA algorithm with the best PESQ value, which is called *Weighted Euclidean STSA (WE-STSA)*. In contrast, the second experiment optimizes directly on the PESQ measure.

Experiment 1: Fitting of the estimator family to WE-STSA
The following three estimators are considered:

- *Minimum Mean-Squared Error STSA (MMSE-STSA)*: The MMSE-STSA estimator results from the special choice $\alpha = 0$ and $\beta = 1$ in (4.22). The corresponding OBE is given by [Ephraim and Malah, 1984]

$$|\hat{S}_l| = \mathrm{E}\left[|S_l| \middle| |X_l|\right] = \frac{\sqrt{\pi\lambda_l}}{2} e^{-\frac{v_l}{2}} \left[(1+v_l) I_0\left(\frac{v_l}{2}\right) + v_l I_1\left(\frac{v_l}{2}\right)\right] \tag{4.24}$$

  where we used the identities [Abramowitz and Stegun, 1964, 13.1.27] and [Abramowitz and Stegun, 1964, 13.3.6] in (4.21) for $m = 1$.

- *Weighted Euclidean (WE-STSA)*: The WE-STSA estimator is the OBE that corresponds to the choice $\alpha = 0.5$ and $\beta = 1$. It is given by [Loizou, 2005]

$$|\hat{S}_l| = \left(\mathrm{E}\left[|S_l|^{-1} \middle| |X_l|\right]\right)^{-1} = \sqrt{\frac{\lambda_l}{\pi}} \frac{e^{v_l/2}}{I_0\left(\frac{v_l}{2}\right)} \tag{4.25}$$

  where we used the identity $M(\frac{1}{2}, 1, z) = e^{z/2} I_0\left(\frac{z}{2}\right)$ [Abramowitz and Stegun, 1964, 9.6.47] in (4.21).

- *Estimator Family*: To learn the optimal parameters $\mathbf{p}_0$, $I_1 = 5\,000$ samples from the joint PDF $p(|S|, |X|)$ and $I_2 = 5\,000$ samples from the a posteriori PDF $p(|S_l| \big| |X_l|)$ are drawn using a uniform (hyper-)prior distribution for $\sigma_z^2(l)$ and $\sigma_s^2(l)$. They were chosen to be $\sigma_z^2(l) \sim \mathcal{U}(10^{-2}, 10^{0})$ and $\sigma_s^2(l) \sim \mathcal{U}(10^{-12}, 10^{3})$.

We used ten female and ten male speakers from the TIMIT database which resulted in a total of 144 utterances. The noise was assumed to be white Gaussian with a SNR of 10dB. The short-time Fourier transform was computed using a Hamming window of length 32ms and an overlap of 50% as in [Plourde and Champagne, 2008]. The noise variance $\sigma_z^2(l)$ was estimated from noise-only segments where those segments were found by a voice activity detector (VAD). $\sigma_s^2(l)$ is estimated from the decision-directed approach as proposed in [Ephraim and Malah, 1984].

	MMSE loss ($\alpha = 0, \beta = 1$)	**WE loss** ($\alpha = 0.5, \beta = 1$)	**PESQ**
Noisy speech signal	4.13×10^{-2}	8.34×10^{0}	2.26
OBE for $\alpha = 0, \beta = 1$ (MMSE-STSA)	1.47×10^{-2}	1.02×10^{0}	2.65
OBE for $\alpha = 0.5, \beta = 1$ (WE-STSA)	2.02×10^{-2}	2.01×10^{-1}	2.86
Optimal estimator in $\mathbb{F}$ (WE Loss)	2.10×10^{-2}	1.53×10^{-1}	2.80

Table 4.3: Example 2, Exp. 1: Performance of the STSA estimators

The results are shown in Table 4.3. Beside the MMSE loss and the WE loss, we also give the results with respect to the PESQ measure. It can take on values between "1" (bad) and "4.5" (excellent) and was shown to be a good objective quality measure for speech enhancement [Hu and Loizou, 2008]. From the results we see that WE-STSA gives the best results with respect to the PESQ measure which was already observed in [Plourde and Champagne, 2008]. Furthermore, we also see that the best estimator from $\mathbb{F}$ is a good approximation of the OBE for WE loss. It gives a better PESQ measure than the MMSE-STSA and therefore we could adapt the parametric family to the WE loss function. It is interesting to note that the best estimator from $\mathbb{F}$ has a smaller WE loss than the OBE for this loss function. This stems from the fact that estimates of $\sigma_z^2(l)$ and $\sigma_s^2(l)$ were used during the speech enhancement which influences the performance of the estimators.

Experiment 2: Fitting of the estimator family to PESQ
Instead of using the WE loss as for Experiment 1, we also studied the performance of the estimator family $\mathbb{F}$ if the PESQ measure is used as loss function, i.e. we rerun the first experiment with the same setup but this time we search the best estimator in $\mathbb{F}$ that yields the maximum PESQ value. We splitted the 144 files into two sets, a training set consisting of one male and one female speaker, and a disjoint test set which contains the remaining 142 files. The optimization problem (4.5) was solved using Matlab's `fminsearch` procedure from 50 different randomly chosen starting points.

Table 4.4 and Table 4.5 show the training and generalization performance for this new optimized estimator. It can be seen that the estimator which is adapted to the PESQ loss has an improved mean PESQ value of 2.87 compared to the estimator we found in the first experiment which had a PESQ loss of 2.80. A difference of 0.07 in the PESQ measure corresponds roughly to a 1dB difference in SNR and hence, we can conclude that the found estimator is capable of fitting to the PESQ loss function. Furthermore, it performs also slightly better than WE-STSA on the 142 utterances of the test set.

	MMSE loss ($\alpha = 0, \beta = 1$)	**WE loss** ($\alpha = 0.5, \beta = 1$)	**PESQ**
Noisy speech signal	2.89×10^{-2}	9.71×10^{0}	2.14
OBE for $\alpha = 0, \beta = 1$ (MMSE-STSA)	1.03×10^{-2}	1.22×10^{0}	2.58
OBE for $\alpha = 0.5, \beta = 1$ (WE-STSA)	1.35×10^{-2}	2.20×10^{-1}	2.81
Optimal estimator in $\mathbb{F}$ (for WE Loss)	1.33×10^{-2}	1.60×10^{-1}	2.73
Optimal estimator in $\mathbb{F}$ (for PESQ)	1.08×10^{-2}	1.30×10^{-1}	2.82

Table 4.4: Example 2, Exp. 2: Performance of the STSA estimators (Training error)

	MMSE loss ($\alpha = 0, \beta = 1$)	**WE loss** ($\alpha = 0.5, \beta = 1$)	**PESQ**
Noisy speech signal	4.15×10^{-2}	8.28×10^{0}	2.26
OBE for $\alpha = 0, \beta = 1$ (MMSE-STSA)	1.48×10^{-2}	1.01×10^{0}	2.65
OBE for $\alpha = 0.5, \beta = 1$ (WE-STSA)	2.04×10^{-2}	1.99×10^{-1}	2.85
Optimal estimator in $\mathbb{F}$ (for WE Loss)	2.09×10^{-2}	1.52×10^{-1}	2.80
Optimal estimator in $\mathbb{F}$ (for PESQ)	1.57×10^{-1}	1.36×10^{-1}	2.87

Table 4.5: Example 2, Exp. 2: Performance of the STSA estimators (Generalization error)

4.3.3 Example 3: Image Denoising

The second example we consider is the denoising of an image which is corrupted by additive Gaussian noise. We do not intend to design in this section a "state-of-the-art" denoising algorithm but we would like to show that there is an estimator in the proposed family that yields better results with respect to a perceptional motivated image distance measure than the ordinary MMSE estimator. This shows that the proposed family is also applicable to such estimation problems.

Let $\boldsymbol{\theta} \in \mathbb{R}^{M\times 1}$ be the vector containing the pixel luminance values of a grayscale image of size $M_1 \times M_2$ which is obtained after column stacking it, i.e. $M = M_1 M_2$. Assuming the signal model

$$\mathbf{x} = \boldsymbol{\theta} + \mathbf{z} \text{ with } \mathbf{z} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}), \tag{4.26}$$

the task is to recover the original image $\boldsymbol{\theta}$ from the distorted image $\mathbf{x}$. To use a Bayesian esti-

mator for this problem, we have to answer two questions: First, which loss function is adequate for this problem, i.e. is perceptual relevant. Second, we have to choose an a priori density which can handle the inter-pixel correlation of neighbouring pixels efficiently. Both issues will now be addressed.

Loss Function

A good loss function for the comparison of images is the structural similarity (SSIM) index [Wang et al., 2004]. It was shown that this index performs very well compared with other metrics, see [Sheikh et al., 2006], and that it is a perceptional measure of image similarity [Wang and Bovik, 2009].

The SSIM works by locally calculating the similarity of two image blocks $\boldsymbol{\theta}^{(b)}$ and $\hat{\boldsymbol{\theta}}^{(b)}$ in terms of the luminance (brightness values) $l(\boldsymbol{\theta}^{(b)}, \hat{\boldsymbol{\theta}}^{(b)})$, contrast $c(\boldsymbol{\theta}^{(b)}, \hat{\boldsymbol{\theta}}^{(b)})$ and structure $s(\boldsymbol{\theta}^{(b)}, \hat{\boldsymbol{\theta}}^{(b)})$ and is defined as [Wang and Bovik, 2009]

$$\begin{aligned} S(\boldsymbol{\theta}^{(b)}, \hat{\boldsymbol{\theta}}^{(b)}) &= l(\boldsymbol{\theta}^{(b)}, \hat{\boldsymbol{\theta}}^{(b)}) c(\boldsymbol{\theta}^{(b)}, \hat{\boldsymbol{\theta}}^{(b)}) s(\boldsymbol{\theta}^{(b)}, \hat{\boldsymbol{\theta}}^{(b)}) \\ &= \frac{2\mu_\theta^{(b)}\mu_{\hat{\theta}}^{(b)} + C_1}{\mu_\theta^{2(b)} + \mu_{\hat{\theta}}^{2(b)} + C_1} \cdot \frac{2\sigma_\theta^{(b)}\sigma_{\hat{\theta}}^{(b)} + C_2}{\sigma_\theta^{2(b)} + \sigma_{\hat{\theta}}^{2(b)} + C_2} \cdot \frac{\sigma_{\theta\hat{\theta}}^{(b)} + C_3}{\sigma_\theta^{(b)}\sigma_{\hat{\theta}}^{(b)} + C_3}. \end{aligned} \tag{4.27}$$

$\mu_\theta^{(b)}$, $\mu_{\hat{\theta}}^{(b)}$, $\sigma_\theta^{2(b)}$, $\sigma_{\hat{\theta}}^{2(b)}$ and $\sigma_{\theta\hat{\theta}}^{(b)}$ are given by

$$\begin{aligned} \mu_\theta^{(b)} &= \sum_{q=1}^{Q} w_q \theta_q^{(b)}, & \mu_{\hat{\theta}}^{(b)} &= \sum_{q=1}^{Q} w_q \hat{\theta}_q^{(b)}, \\ \sigma_\theta^{2(b)} &= \sum_{q=1}^{Q} w_q (\theta_q^{(b)} - \mu_\theta^{(b)})^2, & \sigma_{\hat{\theta}}^{2(b)} &= \sum_{q=1}^{Q} w_q (\hat{\theta}_q^{(b)} - \mu_\theta^{(b)})^2, \\ \sigma_{\theta\hat{\theta}}^{(b)} &= \sum_{q=1}^{Q} w_q (\theta_q^{(b)} - \mu_\theta^{(b)})(\hat{\theta}_q^{(b)} - \mu_{\hat{\theta}}^{(b)}), \end{aligned}$$

where Q is the number of pixels in the bth image block. The weights w_q are the coefficients of a suitable window function, e.g. a two-dimensional Gaussian window. The constants C_1, C_2 and C_3 in (4.27) have small positive values to avoid numerical instabilities. The SSIM index is then computed by averaging the local similarity values, i.e.

$$\mathrm{SSIM}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \frac{1}{B} \sum_{b=1}^{B} S(\boldsymbol{\theta}^{(b)}, \hat{\boldsymbol{\theta}}^{(b)}) \tag{4.28}$$

where B is the total number of image blocks. It is easy to show that the SSIM index is symmetric, i.e. $\mathrm{SSIM}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \mathrm{SSIM}(\hat{\boldsymbol{\theta}}, \boldsymbol{\theta})$, and that it is bounded, i.e. $-1 \leq \mathrm{SSIM}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \leq 1$, which are two properties which we would expect from a good image comparison metric. The SSIM index

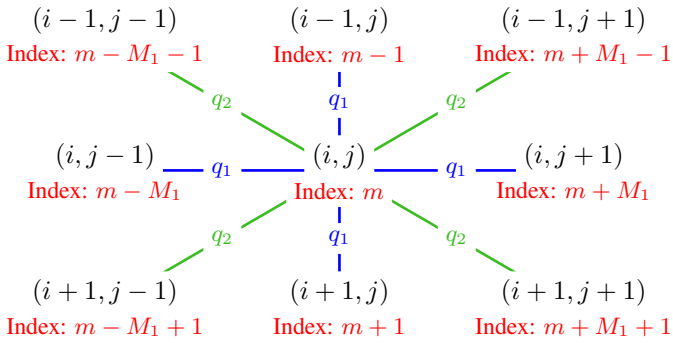


Figure 4.4: Example 3: Nearest neighbour dependency graph at position (i, j). The corresponding linear index of the pixels with respect to the whole image are given in red.

has been successfully used in many areas, see [Wang and Bovik, 2009] for a list of applications to which the SSIM index has already been applied to in the past. For our example, we will use the special choice $C_1 = (0.01L)^2$, $C_2 = (0.03L)^2 = 2C_3$ where L is the dynamical range of the pixel values. The used window is a Gaussian window with size 11×11 and therefore $Q = 11^2 = 121$. These values were proposed in [Wang et al., 2004] and the SSIM index (4.28) simplifies to

$$\mathrm{SSIM}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = \frac{1}{B} \sum_{b=1}^{B} \frac{(2\mu_\theta^{(b)} \mu_{\hat{\theta}}^{(b)} + C_1)(2\sigma_{\theta\hat{\theta}}^{(b)} + C_2)}{(\mu_\theta^{2\,(b)} + \mu_{\hat{\theta}}^{2\,(b)} + C_1)(\sigma_\theta^{2\,(b)} + \sigma_{\hat{\theta}}^{2\,(b)} + C_2)}. \tag{4.29}$$

The negative SSIM index from (4.29) will be used as loss function for our minimization (4.4), i.e. $L_{\mathrm{SSIM}}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = -\mathrm{SSIM}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$. Note that due to its simple structure, the derivative of (4.29) with respect to $\hat{\boldsymbol{\theta}}$ can be obtained after some calculations.

A Priori Density

As a priori density $p(\boldsymbol{\theta})$, we use a Gaussian Markov Random Field (GMRF) [Rue and Held, 2005; Chellappa, 1985]. It will enable us to efficiently model the correlation of neighbouring pixels by specifying the sparse precision matrix $\tilde{\mathbf{Q}}$ which is the inverse of the covariance matrix. For this example, we will use a GMRF model of order 2, i.e. the neighbour set is given by the eight neighbour pixels

$$\mathbb{S} = \{(1,0), (-1,0), (0,1), (0,-1), (1,1), (1,-1), (-1,1), (-1,-1)\}. \tag{4.30}$$

Figure 4.4 shows the corresponding nearest neighbour dependency graph for an arbitrary pixel at the position (i, j) where $1 < i < M_1$ and $1 < j < M_2$, i.e. it does not lie on the image boundary.

Using this neighbour set, it is well known that $\tilde{\mathbf{Q}}$ will have a sparse and symmetric Toeplitz

structure [Rue and Held, 2005]. It has the form $\tilde{\mathbf{Q}} = d\mathbf{Q}(q_1, q_2)$ where $\mathbf{Q}$ is the $M \times M$ Toeplitz matrix[28]

$$\mathbf{Q} = \begin{bmatrix} 1 & q_1 & 0 & \cdots & 0 & q_2 & q_1 & q_2 & 0 & & \cdots & & 0 \\ q_1 & 1 & q_1 & 0 & \cdots & 0 & q_2 & q_1 & q_2 & 0 & & \cdots & 0 \\ 0 & q_1 & 1 & q_1 & 0 & \cdots & 0 & q_2 & q_1 & q_2 & 0 & \cdots & 0 \\ & & \ddots & \ddots & \ddots & & & & \ddots & \ddots & \ddots & & \end{bmatrix}, \mathbf{Q} = \mathbf{Q}^T$$

with $M = M_1 M_2$. q_1 and q_2 are the precision factors with respect to all neighbours of order 1 and 2, respectively. The a priori density is therefore given by a Gaussian PDF with mean $\mu\mathbf{1}$, where $\mathbf{1}$ denotes a vector of all ones, and covariance matrix $\tilde{\mathbf{Q}}^{-1}$, i.e.

$$p(\boldsymbol{\theta}; \mathbb{U}) = \frac{d^{K/2} \det\{\mathbf{Q}\}^{1/2}}{(2\pi)^{K/2}} e^{-\frac{d}{2}(\boldsymbol{\theta}-\mu\mathbf{1})^T \mathbf{Q}(\boldsymbol{\theta}-\mu\mathbf{1})}. \tag{4.31}$$

The set $\mathbb{U} = \{\mu, d, q_1, q_2\}$ denotes all unknown parameters. For a practical denoising algorithm, we need to estimate $\mathbb{U}$ from a noisy image $\mathbf{x} \in \mathbb{R}^K \sim \mathcal{N}(\mu\mathbf{1}, d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I})$ with $K = M$. This will be done by using the ML method. The log-likelihood function is

$$L(\mathbb{U}) = -\frac{K}{2}\ln(2\pi) - \frac{1}{2}\ln\det\left\{d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I}\right\} - \frac{1}{2}(\mathbf{x}-\mu\mathbf{1})^T\left(d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I}\right)^{-1}(\mathbf{x}-\mu\mathbf{1}). \tag{4.32}$$

The first and second-order derivatives with respect to the elements of $\mathbb{U}$ can be calculated in a straightforward way and are given in Appendix C.3.5. Thus, we can use a Newton algorithm to solve the optimization problem

$$\hat{\mathbb{U}} = \arg\max_{\mathbb{U}} L(\mathbb{U}). \tag{4.33}$$

To further reduce the computational complexity of the problem, we approximate the Toeplitz matrix $\mathbf{Q}$ by a circulant matrix $\mathbf{C}$. This can be done as $\mathbf{Q}$ is a bandlimited Toeplitz matrix where only the first $M_1 + 1$ off-diagonals are unequal to zero [Gray, 2006]. Using the circulant matrix $\mathbf{C}$ instead of $\mathbf{Q}$ has the advantage that all matrix-matrix and matrix-vector operations can be carried out in "FFT-speed" and thus we can efficiently work with large images [Golub and Van Loan, 1996]. Appendix C.3.6 gives a brief overview of the basic properties of circulant matrices which are useful for our calculations.

[28] Note that this Toeplitz structure is only correct for all inner pixels, i.e. pixels that do not lie on the image boundary. However, for computational convenience, it is advantageous to use a global Toeplitz structure for the precision matrix as the induced error will be negligible.

Denoising Algorithms and Setup

The following four denoising algorithms are used in this example:

- *Median filter*: A simple denoising algorithm is to use a two-dimensional median filter. We use a neighbourhood of size 5×5.

- *MMSE estimator*: As we have a Gaussian a priori PDF and additive Gaussian noise, the MMSE estimator is equivalent to the LMMSE estimator

$$\hat{\boldsymbol{\theta}}_{\text{LMMSE}}(\mathbf{x}) = \boldsymbol{\mu} + \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \boldsymbol{\mu}) \,. \tag{4.34}$$

  Note that the estimate $\hat{\boldsymbol{\theta}}_{\text{LMMSE}}(\mathbf{x})$ can be efficiently computed using the circulant matrix approximation of the precision matrix as discussed before.

- *Estimator Family*: Finding the estimator which gives the largest averaged SSIM index was done as suggested in (4.4). Instead of learning $2M+3$ parameters in (4.4), we do the simplification to use a fixed ϕ and ψ value for each pixel, i.e. $\phi_1 = \phi_2 = \ldots = \phi_M$ and $\psi_1 = \psi_2 = \ldots = \psi_M$. Using such a uniform parameter setting assumes that the image is spatial homogeneous and also has the advantage that it should yield a parameter setting that is as general as possible and not fitted to a particular image. The value $I_1 = 32$ was used in (4.5) and $I_2 = 32$ a posteriori samples are used for the importance sampling (4.8).

- *Bayes Least Squares - Gaussian Scale Mixture (BLS-GSM) algorithm*: The BLS-GSM algorithm was introduced in [Portilla et al., 2003]. It performs the following three steps: First, the image is transformed into subbands using a *steerable pyramid transform* which provides an overcomplete set of coefficients [Simoncelli et al., 1992]. The basis functions of such a multiscale decomposition are spatially localized, oriented and bandwidth limited. Second, each subband is denoised, except for the lowpass residual band. Third, the inverse pyramid transform is applied to obtain the denoised image.

  For each neighbourhood, a *Gaussian scale mixture* (GSM) is assumed which yields the signal model

$$\mathbf{x} = \boldsymbol{\theta} + \mathbf{z} = \sqrt{c}\mathbf{u} + \mathbf{z}, \tag{4.35}$$

  where $\mathbf{x}$, $\boldsymbol{\theta}$ and $\mathbf{z}$ correspond to the observed coefficients, original (noiseless) coefficients and the noise. Each set of coefficients for one neighbourhood are assumed to have the GSM form $\sqrt{c}\mathbf{u}$ where $\mathbf{u}$ is zero-mean Gaussian with covariance matrix $\mathbf{C}_\theta$ and $\sqrt{c}$ is a positive scalar random variable with a noninformative Jeffreys prior [Robert, 2001] which controls the local variance for each neighbourhood. Finally, an MMSE estimator is applied to estimate the original coefficient value from $\mathbf{x}$. The interested reader is referred to [Portilla et al., 2003] for more details.

To learn the a priori parameters $\mathbb{U}$ which are needed by the MMSE estimator and the estimator family, we used samples from the same image which ought to be denoised. Furthermore, instead of using samples from the joint PDF $p(\boldsymbol{\theta}, \mathbf{x})$, we used image patches $(\boldsymbol{\theta}_k, \mathbf{x}_k)$ from the image in (4.5). Therefore, we consider in this example a *supervised* denoising problem, where information about the image is already present in the learning phase. We will discuss in the next section, what would change if we had chosen an unsupervised scheme.

Simulation Results

Figure 4.5 shows the simulation results for $\sigma = 30$. The best overall performance is shown by the BLS-GSM algorithm.

Comparing the MMSE estimator with the estimator family, we see that $\hat{\boldsymbol{\theta}}(\mathbf{x}; \mathbf{p}_0)$ shows clearly a better result than the MMSE estimator, as

$$\text{SSIM}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}; \mathbf{p}_0)) = 7.35 \times 10^{-1} \qquad \text{and} \qquad \text{SSIM}(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}_{\text{MMSE}}(\mathbf{x})) = 7.00 \times 10^{-1}$$

though the MSE value of $\hat{\boldsymbol{\theta}}(\mathbf{x}; \mathbf{p}_0)$ is larger than for $\hat{\boldsymbol{\theta}}_{\text{MMSE}}(\mathbf{x})$. These results are comparable to the results that were obtained in [Channappayya et al., 2008] where the best linear estimator with respect to the SSIM was considered. The difference between the MMSE and the estimator family can also be observed in Figure 4.5. However, $\hat{\boldsymbol{\theta}}(\mathbf{x}; \mathbf{p}_0)$ does not perform as well as BLS-GSM. This can be due to one of the following two reasons:

- Currently, the signal model that is used, i.e. the GMRF (4.31), makes the assumption of spatial homogeneity and Gaussian distributions. In [Portilla et al., 2003], these assumptions are discussed and it is argued that especially the spatial homogeneity is too restrictive for a good image representation.
- The estimator family operates until now in the pixel domain. However, it is well known that multiscale representations like the wavelet transform allow a substantially better description of an image [Gonzalez and Woods, 2002] which should result in a better denoising algorithm.

As already pointed out, the previous results were obtained in a supervised setup. We also rerun the simulations where we used different images for the learning of the parameters and the image denoising and thus considered the generalization error of the denoising algorithms. We started by using only one image for the training and another image for the test of the learned parameters and we could observe that the estimator family for this scenario has a lower SSIM index as the MMSE estimator. This is due to overfitting and the result of training the estimator family with just one image. Therefore, we also considered the case of using several images for the training (in our case four different images). However, for this scenario, the improvement with respect to the MMSE estimator becomes negligible which points out another problem: The proposed

(a) Original image

(b) Noisy image
MSE: 9.61×10^2
SSIM: 2.98×10^{-1}

(c) Median
MSE: 2.78×10^2
SSIM: 6.19×10^{-1}

(d) MMSE
MSE: 1.82×10^2
SSIM: 7.00×10^{-1}

(e) Estimator family
MSE: 3.03×10^2
SSIM: 7.35×10^{-1}

(f) BLS-GSM
MSE: 1.00×10^1
SSIM: 8.21×10^{-1}

Figure 4.5: Example 3: Results of the various reconstruction methods ($\sigma = 30$)

family in (4.3) is currently not versatile enough to adequately learn the OBE for the SSIM loss that can be used for new, i.e. before unseen, images.

4.4 Observations and Conclusions

From the simulation results of all three examples in the last section, we see that the estimator family could be successfully used to approximate the OBE for a given loss function and signal model.

However, especially the last example also showed the limitations of our current approach: There are algorithms like BLS-GSM that perform better with respect to the SSIM index. We can conclude that using the estimator family out of the box does not work for this application and more a priori knowledge, e.g. working in the wavelet domain than in the pixel domain, has to be taken into account as wavelets are a more universal representation of an image. Therefore, carefully

adapting the estimator family to the problem at hand is necessary. Beside using a different representation of the image data, also the a priori density can be changed or the definition of the estimator family in (4.3) could be generalized.

Chapter 5

Loss Reconstruction for a Given Estimator

5.1 Introduction

As we have seen in Section 2.4, the key idea of Bayesian estimation is to minimize the posterior Bayes risk for a given loss function $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$. In this chapter, we will now consider the inverse problem of finding a good approximation of the loss function $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ for a given estimator $\hat{\boldsymbol{\theta}}(\mathbf{x})$. Such a reconstruction is interesting for the parametric family of estimators $\hat{\boldsymbol{\theta}}(\mathbf{x}; \mathbf{p})$ which we introduced in the last chapter. For most parameter values $\mathbf{p}$, the corresponding loss function that would result in such an estimator is not known and hence, solving the inverse problem as considered in this chapter is interesting as it can be used to study the influence of a particular parameter γ of $\mathbf{p}$.

Another application, where the reconstruction of a loss function is interesting and was already used, arises in economics and finance for testing the rationality of expectations [Muth, 1961]. For example [Elliott et al., 2005] proposed to use a flexible loss function to perform such a hypothesis test and thus allowing a better judging of the forecast rationality. Elliott's loss function depends on two unknown parameters which are learned from time-series data and allow the loss to become asymmetric. This is in contrast to most works which rely on the squared error loss. Elliott's approach differs in two aspects to our work: First, he considers a specific parameterized family of loss functions whereas we deal with the more general aspect of fitting a sum of squares of polynomials. Second, he derived the estimators for the two loss parameters relying on a linear forecast and orthogonality conditions underlying the theory of rationality of expectations. Our reconstruction approach only relies on using a sufficient number of parameter realizations $\boldsymbol{\theta}_i$, corresponding observations $\mathbf{x}_i$ and estimator outputs $\hat{\boldsymbol{\theta}}(\mathbf{x}_i)$ and is thus applicable to a broader class of problems.

This chapter is organized as follows: First, we formulate in Section 5.2 and Section 5.3 the problem of loss reconstruction as an optimization problem to find the optimal coefficients of a sum of squares of polynomials. We will show that additional constraints are needed to avoid

the trivial solution of all zero coefficients. Two possible constraints are therefore introduced in Section 5.4 to circumvent this problem. Furthermore, we discuss in Section 5.5 how additional a priori information can be incorporated using results from algebraic geometry, in particular the *Positivstellensatz*. Therefore, we show similar to Chapter 3 and Chapter 4 how specific application knowledge can be exploited to solve the inverse problem. Section 5.6 finally provides three examples and discusses the reconstruction results.

5.2 Problem Formulation

Instead of minimizing (2.24) to find the OBE $\hat{\boldsymbol{\theta}}(.)$ for a given loss $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$, we consider the inverse problem of reconstructing[29] the loss $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ by

$$\min_{L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})} \iint L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}(\mathbf{x}))p(\boldsymbol{\theta}, \mathbf{x})d\boldsymbol{\theta}d\mathbf{x} \tag{5.1a}$$

subject to

$$L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \geq 0 \quad \forall\, \boldsymbol{\theta}, \hat{\boldsymbol{\theta}} \tag{5.1b}$$

for given $\hat{\boldsymbol{\theta}}(.)$ and given joint PDF $p(\boldsymbol{\theta}, \mathbf{x})$. This is an ill-posed problem as we may have different loss functions that result in the same OBE for a particular signal model. One example is the linear Bayesian model where the OBEs corresponding to the hit-or-miss error, squared error and absolute error are all identical as the maximum, mean and median of the Gaussian a posteriori PDF coincide. In this case, additional requirements, i.e. constraints, can be formulated to obtain a unique solution.

The optimization problem (5.1) has the trivial solution $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = 0$ which we have to avoid. This will be done in Section 5.4 by introducing suitable regularizations. Furthermore, we want to point out that we can reconstruct the loss function $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ in (5.1) only up to a multiplicative scaling. This stems from the fact that the OBE is also invariant with respect to a scaling of the loss.

We will now proceed with explaining our solution approach.

5.3 Solution Approach

We will make the following two simplifications to (5.1) in order to find efficient semidefinite programming (SDP) formulations:

[29] Note that there might be cases, where no loss function $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ exists, e.g. if the considered estimator results from a different statistical model. However, our approach can still be used and (5.1a) will yield the loss that best fits to the given estimator and statistical model.

(S1) First, we do not consider all possible loss functions $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) \geq 0$ but restrict our attention to nonnegative polynomials of degree $2D$ which can be expressed as a sum of squares [Lasserre, 2010; Parrilo, 2003]. Introducing the notation $\mathbf{z} = [\boldsymbol{\theta}; \hat{\boldsymbol{\theta}}] \in \mathbb{R}^{2M}$, this implies $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = L(\mathbf{z}) = \sum_i f_i(\mathbf{z})^2$ where $f_i(\mathbf{z})$ are multivariate polynomials of degree $\leq D$. Denote by $\mathbb{R}[\mathbf{z}]$ the set of all polynomials with $2M$ variables and let

$$\mathbb{SOS} = \{f \in \mathbb{R}[\mathbf{z}] : f(\mathbf{z}) = \sum_i (f_i(\mathbf{z}))^2, \deg\{f_i\} \leq D\}$$

be the set of all sum of squares polynomials. Then it is obvious that $f(\mathbf{z}) \in \mathbb{SOS}$ implies that $f(\mathbf{z})$ is a nonnegative polynomial, i.e. $f(\mathbf{z}) \geq 0$ for all $\mathbf{z} \in \mathbb{R}^{2M}$. The converse, however, is not true as there are nonnegative polynomials which are not sum of squares, e.g. the famous Motzkin example which is studied in Appendix C.4.1. However, there are denseness results that show that sum of squares polynomials allow a good approximation of any nonnegative polynomial, e.g. [Lasserre, 2007].

Introducing the monomial vector $\mathbf{m}(\mathbf{z}) \in \mathbb{R}^{\binom{D+2M}{D}}$, an equivalent representation of every sum of squares polynomial $f(\mathbf{z}) \in \mathbb{SOS}$ is $f(\mathbf{z}) = \mathbf{m}(\mathbf{z})^T \mathbf{Q} \mathbf{m}(\mathbf{z})$ with $\mathbf{Q} \succeq \mathbf{0}$ as is shown in Appendix C.4.2[30]. This representation is well suited for SDP formulations as we will see in the following [Parrilo, 2003].

For completeness, we would like to point out that SOS polynomials play a central role in *polynomial optimization* where the goal is to optimize a polynomial cost function with respect to constraints given by polynomial equalities/inequalities. Appendix C.4.3 contains a brief introduction into the topic of polynomial optimization.

(S2) The second simplification is to approximate the integral in (5.1) by Monte Carlo integration [Liu, 2008] using samples $\mathbf{z}_1, \ldots, \mathbf{z}_I$ which are tuples of $\boldsymbol{\theta}_i$ and $\hat{\boldsymbol{\theta}}(\mathbf{x}_i)$ generated according to $p(\boldsymbol{\theta}, \mathbf{x})$.

Using these two simplifications, we obtain the following new optimization problem which builds the basis for our further discussion:

$$\min_{L(\mathbf{z}) \in \mathbb{SOS}} \frac{1}{I} \sum_{i=1}^{I} L(\mathbf{z}_i) = \min_{\mathbf{Q} \succeq \mathbf{0}} \operatorname{tr}\left\{\mathbf{M}^T \mathbf{Q} \mathbf{M}\right\} = \min_{\mathbf{Q} \succeq \mathbf{0}} \operatorname{tr}\left\{\mathbf{Q} \tilde{\mathbf{M}}\right\} \tag{5.2}$$

where $\mathbf{M} = \frac{1}{\sqrt{I}} \begin{bmatrix} \mathbf{m}(\mathbf{z}_1) & \cdots & \mathbf{m}(\mathbf{z}_I) \end{bmatrix}$ and $\tilde{\mathbf{M}} = \mathbf{M}\mathbf{M}^T$. The complete information about the signal model (i.e. $p(\boldsymbol{\theta}, \mathbf{x})$) and the estimator $\hat{\boldsymbol{\theta}}$ are incorporated into (5.2) in form of the

[30] For example, consider the case $M = 1$, $D = 2$ which results in the monomial vector $\mathbf{m}(\mathbf{z}) = \begin{bmatrix} 1 & \theta & \hat{\theta} & \theta^2 & \theta\hat{\theta} & \hat{\theta}^2 \end{bmatrix}^T$. Using $f_i(\mathbf{z}) = \mathbf{c}_i^T \mathbf{m}(\mathbf{z})$ we can write every $f(\mathbf{z}) \in \mathbb{SOS}$ as $f(\mathbf{z}) = \sum_i f_i(\mathbf{z})^2 = \mathbf{m}^T(\mathbf{z}) \sum_i \mathbf{c}_i \mathbf{c}_i^T \mathbf{m}(\mathbf{z}) = \mathbf{m}(\mathbf{z})^T \mathbf{Q} \mathbf{m}(\mathbf{z})$, i.e. all coefficients of $f(\mathbf{z})$ are contained in $\mathbf{Q} \succeq \mathbf{0}$.

estimated moment matrix $\tilde{\mathbf{M}}$. If $p(\boldsymbol{\theta}, \mathbf{x})$ and $\hat{\boldsymbol{\theta}}$ allow an analytic calculation of $\tilde{\mathbf{M}}$, then this result could be used in (5.2) instead.

Note that (5.2) has similar to (5.1) the trivial solution $\mathbf{Q} = \mathbf{0}$ which corresponds to $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}}) = 0$. Section 5.4 and 5.5 will therefore introduce now suitable additional constraints to (5.2) to obtain a reasonable estimate of the underlying loss function.

5.4 SDP Formulations

5.4.1 SDP Formulation with Trace Regularization

The first SDP we introduce uses the simple trace regularization $\mathrm{tr}\{\mathbf{Q}\} \geq \delta$ with $\delta > 0$ to avoid the trivial solution. It is given by

$$\min_{\mathbf{Q} \succeq \mathbf{0}} \mathrm{tr}\left\{\mathbf{Q}\tilde{\mathbf{M}}\right\} \quad \text{s.t.} \quad \mathrm{tr}\{\mathbf{Q}\} \geq \delta. \tag{5.3}$$

The particular value of $\delta > 0$ is not important as any scaling of $\mathbf{Q}$ does not change the solution.

The simple constraint $\mathrm{tr}\{\mathbf{Q}\} \geq \delta$ has the drawback that an increasing degree D of the monomial vector will yield a reconstructed loss function which is almost zero everywhere and thus this regularization is only suited for small degrees D. The last example in Section 5.6 will give a scenario where the trace regularization fails.

5.4.2 SDP Formulation with Unit Volume Regularization

A better constraint is given by requiring $L(\mathbf{z})$ to have a unit volume in a specific hyperrectangle $\mathbb{H}$ in $\mathbb{R}^{2M}$ with $\mathbb{H} = \{\mathbf{z} : z_{i,l} \leq z_i \leq z_{i,u}\}$ where $z_{i,l}$ and $z_{i,u}$ are given lower and upper bounds for the ith component of $\mathbf{z}$. $\mathbb{H}$ defines the region of interest where we want to approximately reconstruct our loss function. The modified optimization problem is

$$\min_{\mathbf{Q} \succeq \mathbf{0}} \mathrm{tr}\left\{\mathbf{Q}\tilde{\mathbf{M}}\right\} \quad \text{s.t.} \quad \frac{1}{V(\mathbb{H})} \int_{\mathbb{H}} L(\mathbf{z}) d\mathbf{z} \geq 1 \tag{5.4}$$

with $V(\mathbb{H}) = \int_{\mathbb{H}} d\mathbf{z} = \prod_{i=1}^{2M} \left(z_{i,u} - z_{i,l}\right)$. Using $L(\mathbf{z}) = \mathrm{tr}\{\mathbf{Q}\mathbf{m}(\mathbf{z})\mathbf{m}(\mathbf{z})^T\}$, we obtain

$$\frac{1}{V(\mathbb{H})} \int_{\mathbb{H}} L(\mathbf{z}) d\mathbf{z} = \mathrm{tr}\left\{\mathbf{Q} \frac{1}{V(\mathbb{H})} \int_{\mathbb{H}} \mathbf{m}(\mathbf{z})\mathbf{m}(\mathbf{z})^T d\mathbf{z}\right\}. \tag{5.5}$$

Hence, we can replace the integral constraint by $\mathrm{tr}\left\{\mathbf{Q}\tilde{\mathbf{U}}\right\} \geq 1$ where we can think of $\tilde{\mathbf{U}}$ as the moment matrix of a random vector which is uniformly distributed on $\mathbb{H}$, i.e. $\tilde{\mathbf{U}} = \mathrm{E}\left[\mathbf{m}(\mathbf{z})\mathbf{m}(\mathbf{z})^T\right]$. The elements of $\tilde{\mathbf{U}}$ are given by

$$\mathrm{E}[z_1^{\alpha_1} \cdots z_{2M}^{\alpha_{2M}}] = \prod_{i=1}^{2M} \frac{z_{i,u}^{\alpha_i+1} - z_{i,l}^{\alpha_i+1}}{(\alpha_i+1)(z_{i,u} - z_{i,l})}.$$

The unit volume constraint in $\mathbb{H}$ forces $L(\mathbf{z})$ to take only small positive values in regions where the samples $\mathbf{z}_i$ are located. By this, it prevents the trivial solution and also gives more reasonable results than the simple trace regularization $\mathrm{tr}\{\mathbf{Q}\} \geq \delta$ as we will see in Section 5.6.

5.5 Incorporation of A Priori Knowledge

In this section, we will discuss the possibility of incorporating a priori knowledge into the loss reconstruction. Using polynomial optimization this can easily be achieved. Before, however, we have to briefly introduce some notions from algebraic geometry. The interested reader is referred to [Jarvis-Wloszek et al., 2005; Parrilo, 2003].

5.5.1 Positivstellensatz

The Positivstellensatz allows to give certificates that a system of polynomial constraints has no solution. The following three algebraic structures are needed to state the Positivstellensatz:

(D1) The *multiplicative monoid* $\mathbb{M}$ generated by the polynomials $a_1, \ldots, a_s \in \mathbb{R}[\mathbf{z}]$ is defined as the set of all finite products of the $a_i(\mathbf{z})$'s including the empty product, the identity.

(D2) The *cone* $\mathbb{K}$ generated by the polynomials $b_1, \ldots, b_t \in \mathbb{R}[\mathbf{z}]$ is defined as

$$\mathbb{K}(b_1, \ldots, b_t) = \left\{ s_0 + \sum_i s_i \tilde{b}_i \middle| s_i \in \mathbb{SOS}, \tilde{b}_i \in \mathbb{M}(b_1, \ldots, b_t) \right\}.$$

(D3) The *ideal* $\mathbb{I}$ generated by the polynomials $c_1, \ldots, c_u \in \mathbb{R}[\mathbf{z}]$ is defined as

$$\mathbb{I}(c_1, \ldots, c_u) = \left\{ \sum_{i=1}^{u} c_i p_i \middle| p_i \in \mathbb{R}[\mathbf{z}] \right\}.$$

After introducing these three definitions, we can now state the Positivstellensatz (see e.g. [Parrilo, 2003]):

Given the polynomials $\{a_1, \ldots, a_s\}$, $\{b_1, \ldots, b_t\}$ and $\{c_1, \ldots, c_u\}$ in $\mathbb{R}[\mathbf{z}]$, the following two statements are equivalent:

1. $$\left\{ \mathbf{z} \in \mathbb{R}^{2M} \middle| \begin{matrix} a_1(\mathbf{z}) \neq 0, \ldots, a_s(\mathbf{z}) \neq 0 \\ b_1(\mathbf{z}) \geq 0, \ldots, b_t(\mathbf{z}) \geq 0 \\ c_1(\mathbf{z}) = 0, \ldots, c_u(\mathbf{z}) = 0 \end{matrix} \right\} = \emptyset$$

2. There exist polynomials $a \in \mathbb{M}(a_1, \ldots, a_s)$, $b \in \mathbb{K}(b_1, \ldots, b_t)$ and $c \in \mathbb{I}(c_1, \ldots, c_u)$ such that
$$a^2 + b + c = 0.$$

We will now show that the second statement implies the first one and by this also motivate the definitions given above. The full prove can be found in [Stengle, 1974]. Assume that the set of the first statement is not empty, i.e. there is a point $\mathbf{z}_0$ such that all conditions are fulfilled. Then, we know that $a^2(\mathbf{z}_0) > 0$, $b(\mathbf{z}_0) \geq 0$ and $c(\mathbf{z}_0) = 0$ as a is an element of $\mathbb{M}(a_1, \ldots, a_s)$, b an element of $\mathbb{K}(b_1, \ldots, b_t)$ and c an element of $\mathbb{I}(c_1, \ldots, c_u)$. However, this contradicts the second statement $a^2 + b + c = 0$ and therefore we have proven that the second statement implies the first one.

5.5.2 Incorporation of Additional Constraints

After we have introduced the necessary concepts, we can now show how to incorporate additional a priori information. We will do this by the following example: Assume that we know that $L(\boldsymbol{\theta}, \hat{\boldsymbol{\theta}})$ is unequal to zero for all $\|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|^2 \geq \delta$. Thus, our original optimization problem (5.2) is now

$$\min_{\mathbf{Q} \succeq \mathbf{0}} \operatorname{tr}\left\{ \mathbf{Q}\tilde{\mathbf{M}} \right\} \tag{5.6a}$$

subject to

$$L(\mathbf{z}) \neq 0 \ \ \forall \ \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|^2 \geq \delta. \tag{5.6b}$$

The constraint (5.6b) is equivalent to the statement that the set

$$\left\{ \mathbf{z} = [\boldsymbol{\theta}; \hat{\boldsymbol{\theta}}] \in \mathbb{R}^{2M} \middle| \|\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}\|^2 \geq \delta, L(\mathbf{z}) = 0 \right\} = \left\{ \mathbf{z} \in \mathbb{R}^{2M} \middle| b_1(\mathbf{z}) \geq 0, c_1(\mathbf{z}) = L(\mathbf{z}) = 0 \right\} \tag{5.7}$$

is empty where $b_1(\mathbf{z}) = \mathbf{z}^T \begin{bmatrix} \mathbf{I}_M & -\mathbf{I}_M \\ -\mathbf{I}_M & \mathbf{I}_M \end{bmatrix} \mathbf{z} - \delta$ and $\mathbf{I}_M \in \mathbb{R}^{M \times M}$ the identity matrix. Using the Positivstellensatz, we can reformulate (5.7) into the condition that there are polynomials

$s_0, s_1 \in \mathbb{SOS}$ and $p_1 \in \mathbb{R}[\mathbf{z}]$ such that $s_0 + s_1 b_1 + c_1 p + 1 = 0$. Defining $s_i(\mathbf{z}) = \mathbf{m}(\mathbf{z})^T \mathbf{Q}_i \mathbf{m}(\mathbf{z})$ and $p(\mathbf{z}) = \mathbf{m}(\mathbf{z})^T \mathbf{P} \mathbf{m}(\mathbf{z})$ we obtain

$$\min_{\mathbf{Q} \succeq \mathbf{0}, \mathbf{Q}_0 \succeq \mathbf{0}, \mathbf{Q}_1 \succeq \mathbf{0}, \mathbf{P}} \operatorname{tr} \left\{ \mathbf{Q} \tilde{\mathbf{M}} \right\} \tag{5.8a}$$

subject to

$$s_0 + s_1 b_1 + c_1 p_1 + 1 = 0 \tag{5.8b}$$

where the constraint (5.8b) requires all coefficients of the lefthand side to be zero. To obtain the corresponding set of equality constraints, a computer algebra system like Maxima or Maple can be used. Note that the constraint (5.8b) is bilinear due to the term $c_1(\mathbf{z}) p_1(\mathbf{z})$. One possibility to solve such bilinear problems is the V-K iteration [El Ghaoui and Balakrishnan, 1994] which cyclically fixes either $\mathbf{Q}$ or $\mathbf{P}$ and solves the SDP (5.8) several times until it converges to a solution.

As the Positivstellensatz allows constraints involving "$=$", "$\leq$" and "$\neq$", quite general a priori information can be incorporated into the optimization to reconstruct the loss function. Note that there are also simpler certificates than the Positivstellensatz, e.g. the certificates of Schmüdgen and Putinar [Lasserre, 2010] for the case that we only have inequality constraints.

5.6 Examples

In the following we will consider three examples involving different loss functions. The common signal model for all examples is

$$x = \theta + z, \tag{5.9}$$

i.e. $M = N = 1$ with $\theta \sim \mathcal{N}(\mu_\theta, \sigma_\theta^2)$, $z \sim \mathcal{N}(0, \sigma_z^2)$ and both are independent of each other.

The values μ_θ, σ_θ^2 and σ_z^2 determine the valid reconstruction area of the loss function. They should be chosen such that the following two requirements are fulfilled:

- To get a valid reconstruction of the loss function, we have to make sure that we have samples $\mathbf{z}_i$ that are spread over the whole space where we are interested in $L(\mathbf{z})$. Therefore, the noise variances σ_z^2 and σ_θ^2 have to be large enough to "cover" the interesting area.
- The influence of the a priori information has to be small. This can be achieved by using a large enough σ_θ^2 and different values for μ_θ. This has the effect that the a priori information is averaged out.

In the following examples, we are interested in a reconstruction of the loss function in the rectangle defined by $-1 \leq \theta, \hat{\theta} \leq 1$. Therefore, we choose μ_θ equidistant in 20 steps from -2

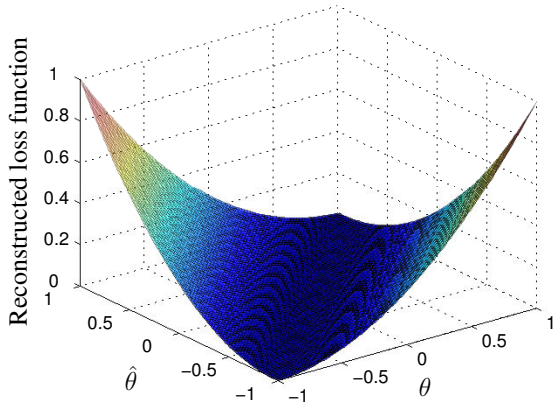


Figure 5.1: Example 1: Reconstructed loss ($D = 1$, trace and unit volume regularization)

to 2. Furthermore, $\sigma_\theta^2 = \sigma_z^2 = 1$ for all examples except $\sigma_\theta^2 = 0.3$ for the third example. We use $I = 10^7$ samples for the Monte Carlo integration in (5.2) and the hyperrectangle $\mathbb{H}$ for the unit volume regularization is set to $z_{1,l} = z_{2,l} = -1$ and $z_{1,u} = z_{2,u} = 1$.

5.6.1 Example 1: Squared Error Loss

Using the squared error loss $L_{\text{SE}}(\theta, \hat{\theta}) = (\theta - \hat{\theta})^2$, it is well known that the corresponding OBE for our signal model (5.9) is given by [Kay, 1993]

$$\hat{\theta}_{\text{SE}} = \mu_\theta + \frac{\sigma_\theta^2}{\sigma_\theta^2 + \sigma_z^2}(x - \mu_\theta). \tag{5.10}$$

Figure 5.1 shows the result of the reconstructed loss for $D = 1$, i.e. we consider quadratic polynomials. The transparency of the graph in this figure and all following figures is given by the bivariate histogram of $\mathbf{z}_1, \ldots, \mathbf{z}_I$ where the graph is more opaque for larger values in the bivariate histogram. Both proposed semidefinite programs from Section 5.4 show the same result for $D = 1$ and yield a good reconstruction of the original loss function $L_{\text{SE}}(\theta, \hat{\theta})$. If the degree D is increased then the reconstructed loss function is more flat around $\theta - \hat{\theta} = 0$ as shown by Figure 5.2 and 5.3. This flattening result is especially severe for the trace regularization. The unit volume regularization is more robust with respect to choosing the degree of the polynomial $L(\mathbf{z})$ too large. This effect can be further reduced if the size of the hyperrectangle $\mathbb{H}$ is reduced.

5.6.2 Example 2: LinEx Loss

The LINEX loss is an asymmetric loss function, which was defined in (2.29) and has the form

$$L_{\text{LINEX}}(\theta, \hat{\theta}) = b\left(e^{a(\hat{\theta}-\theta)} - a(\hat{\theta} - \theta) - 1\right) \tag{5.11}$$

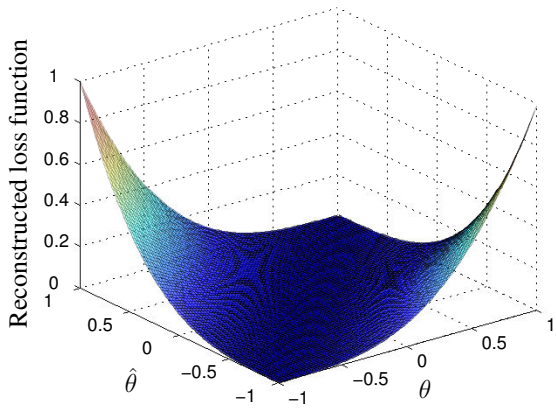


Figure 5.2: Example 1: Reconstructed loss ($D = 3$, unit volume regularization)

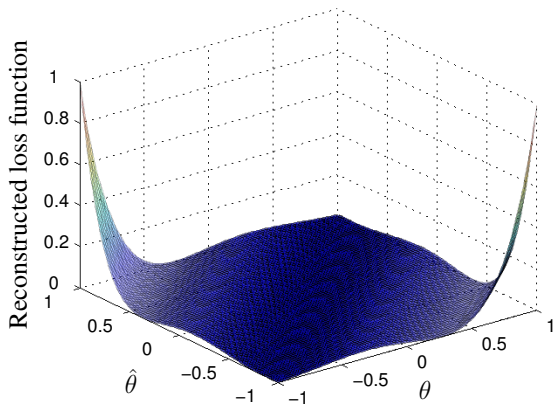


Figure 5.3: Example 1: Reconstructed loss ($D = 3$, trace regularization)

with $b > 0$ and $a \neq 0$. Asymmetric loss functions are interesting for applications where under-/overestimating θ is more severe than over-/underestimating it. The corresponding OBE for our signal model is according to (2.30b) given by

$$\hat{\theta}_{\text{LINEX}} = \hat{\theta}_{\text{SE}} - \frac{a}{2}\left(\sigma_\theta^2 - \frac{\sigma_\theta^4}{\sigma_\theta^2 + \sigma_z^2}\right). \tag{5.12}$$

Figure 5.4 shows the reconstructed loss function for the unit volume regularization with $D = 1$ and $b = 1$, $a = 2$. Both regularizations from Section 5.4 show almost the same reconstruction result which is asymmetric with respect to $\theta - \hat{\theta} = 0$ as expected from the original LINEX loss. Similar as before, we observe that for increasing D the reconstructed loss functions are more flat around $\theta - \hat{\theta} = 0$. Again, the unit volume regularization can be made more stable by adjusting

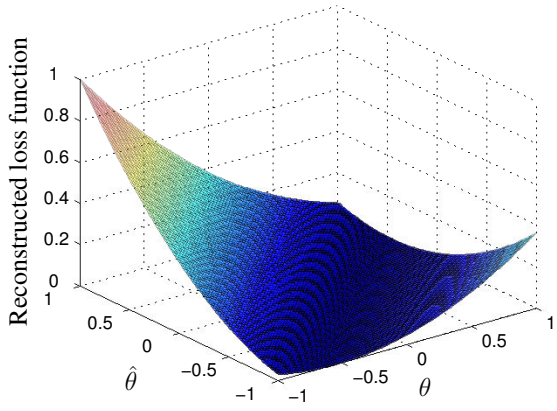


Figure 5.4: Example 2: Reconstructed loss ($D = 1$, unit volume regularization)

the hyperrectangle $\mathbb{H}$.

5.6.3 Example 3: Normalized Squared Error Loss

For this example, we use the normalized squared error loss

$$L_{\text{NSE}}(\theta, \hat{\theta}) = \frac{(\theta - \hat{\theta})^2}{\hat{\theta}^2} \tag{5.13}$$

which was introduced in [Norstrom, 1996] for risk analysis. Figure 5.5 shows the shape of the normalized squared error loss. It has the characteristic that the loss $L_{\text{NSE}}(\theta, \hat{\theta})$ approaches infinity if $\hat{\theta} \to 0$ and thus is "conservative" if small values $\hat{\theta}$ are estimated. The corresponding OBE is [Norstrom, 1996]

$$\hat{\theta}_{\text{NSE}} = \hat{\theta}_{\text{SE}} + \frac{1}{\hat{\theta}_{\text{SE}}} \frac{\sigma_\theta^2 \sigma_z^2}{\sigma_\theta^2 + \sigma_z^2}. \tag{5.14}$$

Figure 5.6 show the reconstruction results for the unit volume regularization with $D = 5$. The normalized squared error loss is an example where the trace regularization does not work as we need a higher polynomial order $2D$ to approximate the original loss function which results in a very flat loss reconstruction for the trace regularization. The unit volume regularization, however, gives a good reconstruction result.

5.7 Observations and Conclusions

The example in Section 5.6.3 revealed that we need a quite high polynomial degree D in order to approximate the normalized squared error loss function $L_{\text{NSE}}(\theta, \hat{\theta})$. The reason is that $L_{\text{NSE}}(\theta, \hat{\theta})$ is not differentiable for $\hat{\theta} = 0$ in contrast to our sum of squares approximations. To

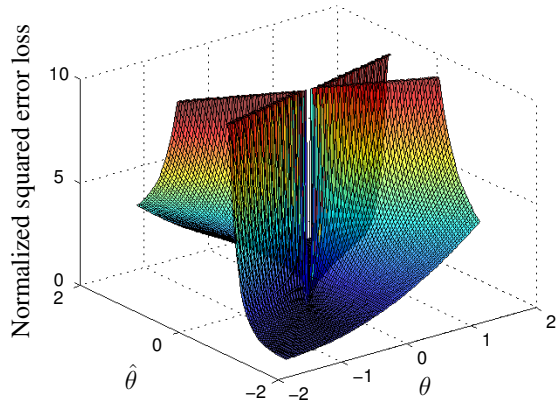


Figure 5.5: Example 3: Normalized squared error loss

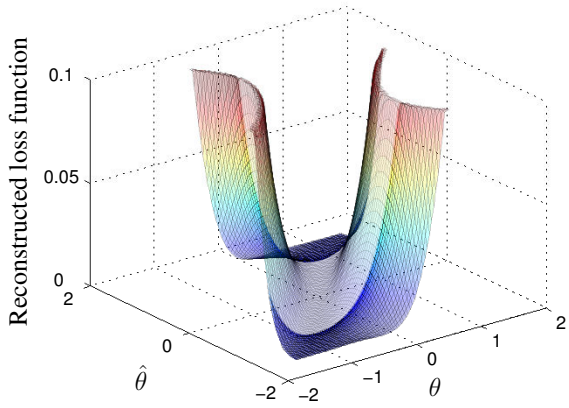


Figure 5.6: Example 3: Reconstructed loss ($D = 5$, unit volume regularization)

compensate this, the polynomial degree D has to be increased. This is an issue that should be further investigated as solving this problem is important if we want to apply it to our parametric family of estimators from Chapter 4. For the speech enhancement problem in Section 4.3.2 for example, we expect that the underlying loss function is discontinuous (similar to the WE loss) and therefore using the SOS approach is not optimal. One idea could be to use an approach similar to the one that is used with *splines* [Unser, 1999] and to introduce piecewise-polynomial functions.

Chapter 6

Conclusions and Future Work

6.1 Conclusions

Chapter 3 and Chapter 4 showed two different ways to incorporate a priori information into the estimation procedure. In Chapter 3, we considered three recursive estimators – the recursive constrained ML, the recursive affine minimax and the recursive MMSE – for the adaptive estimation of a time-varying parameter vector $\boldsymbol{\theta}(n)$ in a linear Gaussian signal model with the a priori knowledge $\boldsymbol{\theta}(n) \in \mathbb{T}$. All three estimators use the RWLS algorithm in a preprocessing step to transform a dimensionally growing estimation problem to a fixed-dimension task. As the estimators take care of the history of measurements and the a priori knowledge $\boldsymbol{\theta}(n) \in \mathbb{T}$, they outperform the instantaneous estimators (least squares, constrained ML, affine minimax, MMSE) as well as the unconstrained recursive estimator RWLS. Especially the derived RMMSE estimator is interesting as the proposed rejection sampling allows a simple handling of general constraints. Furthermore, the computational complexity is moderate as we only have to generate samples from a truncated Gaussian density. In Section 3.5, we studied three different examples and could conclude that the RMMSE estimator showed the best performance among the three proposed estimators.

In Chapter 4, a family of estimators was proposed for the Bayesian estimation with nonstandard loss functions. This family has the advantage that it is parameterized by a small number of variables which can be determined offline for a particular loss function and signal model. We proved that the family includes many important estimators known from the literature, namely MMSE, MAP, and OBE under LINEX loss which shows that it is quite versatile. The computational complexity of our approach is comparable to that of an MMSE estimation for the same signal model if we assume that Monte Carlo integration is used for the calculation of the MMSE estimator. The examples showed how the parametric family of estimators can be applied to practical problems and we saw that using nonstandard loss functions gives better results compared to the traditional approaches.

Finally, we considered in Chapter 5 the problem of approximately reconstructing the underlying loss function for a given Bayesian estimator $\hat{\boldsymbol{\theta}}(\mathbf{x})$. Two SDP formulations that allow an efficient reconstruction were proposed. Furthermore, we showed how to incorporate additional a priori information using the Positivstellensatz. From the three considered examples in 5.6, it turned out that the SDP using a unit volume regularization allows a good reconstruction of the loss function.

6.2 Other Work

Beside the work that was presented in this thesis, also the following unrelated work was done which we want to briefly mention.

In [Uhlich and Yang, 2008], we studied a generalized and optimum *multiple description coding* (MDC) prefilter scheme. It allows redundant descriptions to be transmitted which offer an improved robustness against erasure channels. Several properties of the distortion measure could be proved which help to understand the function of the prefilter. Especially the knowledge of the Hessian of the distortion measure allows to efficiently find a solution of the optimization problem. This idea was further extended to the problem of Kalman tracking with lossy channels using a MDC precoding in [Blind et al., 2009].

In [Uhlich et al., 2009], we investigated the problem of estimating an unknown parameter using a polynomial LMMSE estimator. There, it was shown in a case study of frequency estimation that only a sufficiently large polynomial order allows a good estimation performance if we do not know the signal model at all. A high polynomial order, however, results in two problems: high computational complexity due to a large length of the augmented observation vector and ill condition of the autocorrelation matrix due to strong correlation of some elements. In order to combat these problems, the potential of the *sequential floating forward selection* (SFFS) algorithm was considered. It is shown that the SFFS yields almost the same performance with only a small percentage of selected elements from the augmented observation vector.

6.3 Future Work

In this last section, we would like to point the interested reader to five possible directions for future work.

A first issue is the rejection sampling used by the RMMSE estimator in Section 3.3.3. Using rejection sampling allowed us to incorporate a wide variety of constraints. However, this simple scheme is not always optimal. Some work can be found in the literature which aims at designing more efficient sampling schemes for truncated Gaussian densities, see e.g. [Robert, 1995; Lu, 2008]. Using this techniques and extending them to more sophisticated constraints, e.g. equality

constraints which we cannot handle with rejection sampling, could yield fast random variable generators and would be therefore worthwhile for practical signal processing applications.

A second issue which needs further investigation is the efficient solution of the minimax problem (2.19) for a more general class of constraints as was considered in this thesis. Up to now, we only studied the RAMX for a superposition of quadratic constraints using the S-procedure and from the standard literature, it is also well known how to deal with linear constraints (see *Finsler*'s lemma [Boyd et al., 1994]). However, more advanced results are known, especially from robust control [Ben-Tal et al., 2009], which could be transferred to our problems and hence allowing affine minimax estimation for more general constraints.

As was already pointed out in Section 3.4, another possibility to include a priori information beside our three considered estimators, is to use the approach of [Benavoli et al., 2007]. He studied a MAP estimator where the constraints are incorporated by choosing adequate values of the Gaussian prior density, i.e. its mean and covariance matrix. Extending this approach to our time-varying scenario by using the sufficient statistic (3.10) and comparing it with the other three recursive estimators from Chapter 3 would be interesting.

Another direction for further work is the study of lower bounds for estimation problems with additional information. The additional information that we considered in Chapter 3 was given by constraints which the estimator has to fulfill. Therefore, it is worthwile to compare the found estimators of Chapter 3 to the constrained Cramér-Rao bound, which was e.g. derived in [Gorman and Hero, 1990; Ben-Haim and Eldar, 2009]. [Gorman and Hero, 1990] uses the Bobrovsky-Zakai bound where the test points are chosen such that the constraints are fulfilled, i.e. $\mathbf{h}$ is chosen such that $\boldsymbol{\theta} + \mathbf{h} \in \mathbb{T}$. Besides looking at lower bounds for constrained problems it would also be interesting to look at lower bounds for other loss functions than the traditional squared error. Interestingly, the Ziv-Zakai bound (2.12) can be modified to be a lower bound for other symmetric loss functions as discussed in [Bell et al., 1997]. These bounds would naturally complement the estimators from Chapter 4 and would be therefore worthwile to study.

A last issue is related to the parametric family from Section 4.2. From the examples we could conclude that there are applications where the parametric family needs to be extended, either by generalizing it using more parameters or by adequately changing the signal model. The first approach would allow us to incorporate the Bayesian estimators for other loss functions which would render the family more flexible – this however at the cost of solving an optimization problem with increased complexity as more parameters have to be determined. The other approach, i.e. altering the signal model, is motivated by the image denoising example from Section 4.3.3. This example revealed that it would be better to work in the wavelet domain instead of the original pixel domain. Hence, finding suitable representations of our signal model, which allow the direct usage of the parametric family of estimators, are also interesting.

Appendix A

Definition of Derivatives for Vectors and Matrices

Matrix calculus extends the notion of derivatives for scalars to vectors and matrices and is a frequently used tool in signal processing, see e.g. [Minka, 2000; Moon and Stirling, 2000; Magnus and Neudecker, 2007]. Table A.1 summarizes the definitions for the real-valued case that are used throughout this thesis. Note that $\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}}$ results in a row vector[31] and therefore the gradient is given by the transpose, i.e.

$$\boldsymbol{\nabla} f(\mathbf{x}) = \left(\frac{\partial f(\mathbf{x})}{\partial \mathbf{x}} \right)^T .$$

Using the above definitions of matrix derivatives, the following useful properties can be shown to be true [Moon and Stirling, 2000]

$$\begin{aligned}
\partial \mathbf{A} &= \mathbf{0}, \qquad \mathbf{A} \text{ constant}, \\
\partial(\alpha \mathbf{X}) &= \alpha \partial \mathbf{X}, \\
\partial(\mathbf{X} + \mathbf{Y}) &= \partial \mathbf{X} + \partial \mathbf{Y}, \\
\partial(\mathbf{X}\mathbf{Y}) &= (\partial \mathbf{X})\mathbf{Y} + \mathbf{X}(\partial \mathbf{Y}), \\
\\
\partial \operatorname{tr}(\mathbf{X}) &= \operatorname{tr}(\partial \mathbf{X}), \\
\partial \det(\mathbf{X}) &= \det(\mathbf{X}) \operatorname{tr}(\mathbf{X}^{-1} \partial \mathbf{X}), \\
\partial(\mathbf{X}^{-1}) &= -\mathbf{X}^{-1}(\partial \mathbf{X})\mathbf{X}^{-1}, \\
\\
\partial(\mathbf{X}^T) &= (\partial \mathbf{X})^T .
\end{aligned}$$

Note that some properties hold in general only if we compute the derivative with respect to a scalar, i.e. $\partial = \frac{\partial}{\partial x}$. Otherwise, problems with the matrix dimensions occur. For example,

[31] Some authors use a different (transposed) convention which results in a column vector and therefore their results are also transposed.

	$f \in \mathbb{R}$ **Scalar**	$\mathbf{f} \in \mathbb{R}^N$ **Vector**	$\mathbf{F} \in \mathbb{R}^{N_1 \times N_2}$ **Matrix**
$x \in \mathbb{R}$ **Scalar**	$\frac{\partial f}{\partial x}$ 1×1	$\frac{\partial \mathbf{f}}{\partial x} = \left[\frac{\partial f_i}{\partial x}\right]_i$ $N \times 1$	$\frac{\partial \mathbf{F}}{\partial x} = \left[\frac{\partial F_{ij}}{\partial x}\right]_{ij}$ $N_1 \times N_2$
$\mathbf{x} \in \mathbb{R}^M$ **Vector**	$\frac{\partial f}{\partial \mathbf{x}} = \left[\frac{\partial f}{\partial x_j}\right]_j$ $1 \times M$	$\frac{\partial \mathbf{f}}{\partial \mathbf{x}} = \left[\frac{\partial f_i}{\partial x_j}\right]_{ij}$ $N \times M$	$N_1 \cdot N_2 \cdot M$ derivatives
$\mathbf{X} \in \mathbb{R}^{M_1 \times M_2}$ **Matrix**	$\frac{\partial f}{\partial \mathbf{X}} = \left[\frac{\partial f}{\partial x_{ij}}\right]_{ij}$ $M_1 \times M_2$	$N \cdot M_1 \cdot M_2$ derivatives	$N_1 \cdot N_2 \cdot M_1 \cdot M_2$ derivatives

Table A.1: Definitions of the derivatives for vectors and matrices

$$\frac{\partial}{\partial \mathbf{x}} \mathbf{x}^T \mathbf{x} = 2\mathbf{x}^T \neq \left(\frac{\partial}{\partial \mathbf{x}} \mathbf{x}\right)^T \mathbf{x} + \mathbf{x}^T \frac{\partial}{\partial \mathbf{x}} \mathbf{x} = \mathbf{x} + \mathbf{x}^T. \quad \lightning$$

Two extensions exist in the literature that we would like to mention for completeness. The first extension deals with the problem of defining a proper derivative for matrix-valued functions with respect to a matrix argument, i.e. $\partial \mathbf{F}(\mathbf{X}) / \partial \mathbf{X}$. Such problems e.g. occur if the Newton method for a scalar cost function of the form $f(\mathbf{X})$ is derived as this requires the second-order derivative of $f(\mathbf{X})$. One possibility to solve this problem is to use the *column-stacking operator* $\mathrm{vec}\{.\}$ and to define $\partial \mathbf{F}(\mathbf{X}) / \partial \mathbf{X}$ as

$$\frac{\partial \mathbf{F}(\mathbf{X})}{\partial \mathbf{X}} = \frac{\partial \,\mathrm{vec}\{\mathbf{F}(\mathrm{vec}\{\mathbf{X}\})\}}{\partial \,\mathrm{vec}\{\mathbf{X}\}},$$

which was e.g. done in [Magnus and Neudecker, 2007]. Another possibility is to use the concept of *directional derivatives* as it allows to solve many problems, e.g. deriving the Newton method, without explicitly computing the second-order derivative $\partial \mathbf{F}(\mathbf{X}) / \partial \mathbf{X}$. This technique was e.g. used in [Dattorro, 2011] to prove convexity of a matrix-valued function.

The second extension is with respect to the domain of a function. Very often, it is natural to use complex-valued arguments instead of real-valued ones. A convenient way to compute the derivatives in this case is given by *Wirtinger calculus*, see e.g. [Gentle, 2007; Schreier and Scharf, 2010; Hjørungnes, 2011].

Appendix B

A Brief Introduction to Convex Optimization

In this Chapter, we will briefly review the basic concepts from optimization theory, which are needed in this thesis. We will in particular focus on the special case of convex optimization. A more detailed introduction and discussion can e.g. be found in [Luo and Yu, 2006; Boyd and Vandenberghe, 2007].

Consider the general optimization problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} f_0(\mathbf{x}) \tag{B.1a}$$

subject to

$$f_i(\mathbf{x}) \leq 0 \quad \text{for all } i = 1, \ldots, m, \tag{B.1b}$$

$$h_j(\mathbf{x}) = 0 \quad \text{for all } j = 1, \ldots, r, \tag{B.1c}$$

where $f_0(\mathbf{x})$ denotes the *objective function* and $\{f_i(\mathbf{x})\}_{i=1}^m$, $\{h_j(\mathbf{x})\}_{j=1}^r$ denote the *inequality* and *equality constraint functions*, respectively. The set which is described by the inequality and equality constraint functions is called the *feasible set* and defined as

$$\text{FS} = \{\mathbf{x} : f_i(\mathbf{x}) \leq 0, h_j(\mathbf{x}) = 0\}.$$

A feasible solution $\mathbf{x}^* \in$ FS is called *globally optimal*, if $f_0(\mathbf{x}^*) \leq f_0(\mathbf{x})$ for all $\mathbf{x} \in$ FS. In contrast, a feasible vector $\mathbf{x}^*$ is called *locally optimal*, if there exists an $\epsilon > 0$ such that $f_0(\mathbf{x}^*) \leq f_0(\mathbf{x})$ for all feasible $\mathbf{x}$ in the ϵ-neighbourhood of $\mathbf{x}^*$, i.e. all points $\mathbf{x}$ satisfying $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon$.

Interestingly, there exists a set of local optimality conditions, known as *Karush-Kuhn-Tucker* (KKT) conditions, which have to be fulfilled for a local optimum $\mathbf{x}^*$, given that the optimization

problem (B.1) satisfies some regularity conditions, the so-called *constraint qualifications*. Let $L(\mathbf{x}, \boldsymbol{\lambda}, \boldsymbol{\nu}) = f_0(\mathbf{x}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{x}) + \sum_{j=1}^{r} \nu_j h_j(\mathbf{x})$ denote the *Lagrangian function*. Then, for a locally optimal point $\mathbf{x}^*$, there exist *dual variables* $\boldsymbol{\lambda}^*$ and $\boldsymbol{\nu}^*$ such that the following conditions are satisified:

(a) *Stationarity:*

$$\boldsymbol{\nabla}_{\mathbf{x}} L(\mathbf{x}^*, \boldsymbol{\lambda}^*, \boldsymbol{\nu}^*) = \mathbf{0},$$

(b) *Primal feasibility:*

$$\begin{aligned} f_i(\mathbf{x}^*) &\leq 0 \quad \text{for all } i = 1, \ldots, m, \\ h_j(\mathbf{x}^*) &= 0 \quad \text{for all } j = 1, \ldots, r, \end{aligned}$$

(c) *Dual feasibility:*

$$\lambda_i^* \geq 0 \quad \text{for all } i = 1, \ldots, m,$$

(d) *Complementary slackness:*

$$\lambda_i^* f_i(\mathbf{x}^*) = 0 \quad \text{for all } i = 1, \ldots, m.$$

An important special case of the general optimization problem (B.1) are *convex optimization problems*. To introduce them, we have to first define the notions of a convex set and function.

- *Convex Set:*

  A set $\mathbb{S} \subset \mathbb{R}^n$ is called *convex* if for every point pair the line segment joining these two points is also an element of $\mathbb{S}$, i.e.

  $$\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2 \in \mathbb{S} \text{ for all } \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S} \text{ and } \lambda \in (0, 1).$$

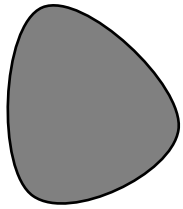
  Figure B.1 gives an example of a convex and non-convex set.

- *Convex Function:*

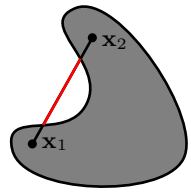
  A function $f : \mathbb{R}^n \to \mathbb{R}$ is called *convex* if the function always lies below the straight line connecting two points on the curve, i.e.

  $$f(\lambda \mathbf{x}_1 + (1 - \lambda)\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda) f(\mathbf{x}_2) \text{ for all } \lambda \in (0, 1).$$

  Figure B.2 gives an example of a convex and non-convex function.

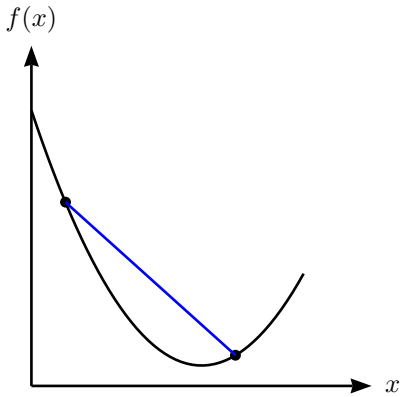


(a) Convex set

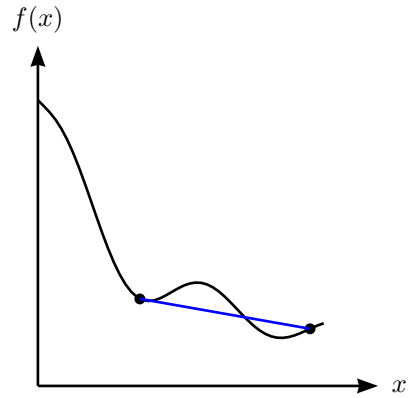


(b) Non-convex set

Figure B.1: Example of a convex and non-convex set



(a) Convex function



(b) Non-convex function

Figure B.2: Example of a convex and non-convex function

This allows us now to define convex optimization problems: A *convex optimization problem* denotes the special case that the cost function $f_0(\mathbf{x})$ and the feasible set FS are convex. FS is e.g. convex if the inequality constraint functions $\{f_i(\mathbf{x})\}_{i=1}^m$ are convex and the equality constraint functions $\{h_j(\mathbf{x})\}_{j=1}^r$ are affine in $\mathbf{x}$, i.e. they can be written as $h_j(\mathbf{x}) = \mathbf{a}_j^T\mathbf{x} + b_j$ for all $j = 1, \ldots, r$.

Convex optimization problems have the important property that each local minimum is also globally optimal and hence each local optimization technique, e.g. gradient descent, will yield the global optimum. This fact can easily be proved by contradiction: Let $\mathbf{x}^* \in$ FS be a local minimum, i.e. $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in$ FS with $\|\mathbf{x} - \mathbf{x}^*\| \leq \epsilon$. Furthermore, assume that there is another $\mathbf{x}^{**} \in$ FS with $f(\mathbf{x}^{**}) < f(\mathbf{x}^*)$. As the feasible set FS and the objective function $f_0(\mathbf{x})$ are convex, we know that

(i) $\lambda\mathbf{x}^* + (1-\lambda)\mathbf{x}^{**} \in$ FS for all $\lambda \in (0, 1)$,

(ii) $f(\lambda\mathbf{x}^* + (1-\lambda)\mathbf{x}^{**}) \leq \lambda f(\mathbf{x}^*) + (1-\lambda)f(\mathbf{x}^{**})$ for all $\lambda \in (0, 1)$.

Using $f(\mathbf{x}^{**}) < f(\mathbf{x}^{*})$, we obtain $f(\lambda\mathbf{x}^{*} + (1-\lambda)\mathbf{x}^{**}) < f(\mathbf{x}^{*})$, which contradicts our assumption that $f(\mathbf{x}^{*})$ is locally optimal.

The class of convex optimization problems includes many well known special cases which we will now briefly list:

- *Linear program (LP)*:
  - linear objective function, i.e. $f_0(\mathbf{x}) = \mathbf{c}^T\mathbf{x}$
  - linear constraints, i.e. $\mathbf{A}_{\text{ineq}}\mathbf{x} \leq \mathbf{b}_{\text{ineq}}, \mathbf{A}_{\text{eq}}\mathbf{x} = \mathbf{b}_{\text{eq}}$
- *Quadratic program (QP)*:
  - quadratic objective function, i.e. $f_0(\mathbf{x}) = \mathbf{x}^T\mathbf{Q}_0\mathbf{x} + \mathbf{c}^T\mathbf{x}$ with $\mathbf{Q}_0 \succeq \mathbf{0}$
  - linear constraints, i.e. $\mathbf{A}_{\text{ineq}}\mathbf{x} \leq \mathbf{b}_{\text{ineq}}, \mathbf{A}_{\text{eq}}\mathbf{x} = \mathbf{b}_{\text{eq}}$
- *Quadratically constrained quadratic program (QCQP)*:
  - quadratic objective function, i.e. $f_0(\mathbf{x}) = \mathbf{x}^T\mathbf{Q}_0\mathbf{x} + \mathbf{c}^T\mathbf{x}$ with $\mathbf{Q}_0 \succeq \mathbf{0}$
  - linear equality and quadratic inequality constraints, i.e.

$$\begin{aligned} \mathbf{x}^T\mathbf{Q}_i\mathbf{x} + \mathbf{a}_i^T\mathbf{x} + b_i \leq 0 \;\text{ with }\; \mathbf{Q}_i \succeq \mathbf{0} \text{ and } i = 1, \ldots, m \\ \mathbf{A}_{\text{eq}}\mathbf{x} = \mathbf{b}_{\text{eq}} \end{aligned}$$

- *Semidefinite program (SDP)*[32]:
  - linear objective function, i.e. $f_0(\mathbf{x}) = \mathbf{c}^T\mathbf{x}$
  - *linear matrix inequality* (LMI) constraints[33], i.e. $\sum_i \mathbf{A}_i x_i \succeq \mathbf{A}_0$

It can be shown that LP $\subset$ QP $\subset$ QCQP $\subset$ SDP, i.e. an SDP is the most general form of the four optimization problems. Many efficient solvers exist for these special cases, for example the *interior-point algorithm* [Boyd and Vandenberghe, 2007], which allow to find the optimal solution even for many optimization variables.

[32] Please note that we give here only an SDP in standard inequality form. The interested reader should consult [Boyd and Vandenberghe, 2007] for SDPs with equality constraints.

[33] A good overview of LMIs and their applications can be found in [Boyd et al., 1994; VanAntwerp and Braatz, 2000]. It is for example easy to show that the feasible set that is described by a LMI is convex.

Appendix C

Details, Derivations and Proofs

C.1 Details, Derivations and Proofs for Chapter 2

C.1.1 Optimal choice of A for the Covariance-Inequality Family

As (2.2) has to hold for all matrices $\mathbf{A}$, we can choose the optimal $\mathbf{A}_{\text{opt}}$ which results in a good lower bound [Van Trees and Bell, 2007]. Consider therefore the function $J(\mathbf{A})$ defined as

$$J(\mathbf{A}) = \mathbf{v}^T \left(\mathbf{T}\mathbf{A}^T + \mathbf{A}\mathbf{T}^T - \mathbf{A}\mathbf{G}\mathbf{A}^T \right) \mathbf{v}, \qquad \mathbf{v} \neq \mathbf{0}, \tag{C.1}$$

which measures the "size" of the right-hand side in (2.2). The first-order derivative is given by

$$\frac{\partial J(\mathbf{A})}{\partial \mathbf{A}} = 2\mathbf{v}\mathbf{v}^T \left(\mathbf{T} - \mathbf{A}\mathbf{G} \right) \tag{C.2}$$

which, after equating it to zero, yields $\mathbf{A}_{\text{opt}} = \mathbf{T}\mathbf{G}^{-1}$. Note that $J(\mathbf{A})$ is concave, i.e.

$$J(\lambda \mathbf{A}_1 + (1-\lambda)\mathbf{A}_2) \geq \lambda J(\mathbf{A}_1) + (1-\lambda)\mathbf{A}_2 \quad \forall\, 0 \leq \lambda \leq 1$$

as $J(\lambda \mathbf{A}_1 + (1-\lambda)\mathbf{A}_2) = \lambda J(\mathbf{A}_1) + (1-\lambda)\mathbf{A}_2 + \lambda(1-\lambda)\mathbf{v}^T(\mathbf{A}_1 - \mathbf{A}_2)\mathbf{G}(\mathbf{A}_1 - \mathbf{A}_2)^T\mathbf{v}$ and therefore there is only one (global) maximum at $\mathbf{A}_{\text{opt}} = \mathbf{T}\mathbf{G}^{-1}$.

C.1.2 Proof of the Ziv-Zakai Identity (2.11)

Let X be an arbitrary scalar random variable with PDF $p(x)$ and CDF $P(x)$. Then $\mathrm{E}[X^2] = \frac{1}{2}\int_0^\infty x \Pr\left\{|X| \geq \frac{x}{2}\right\} dx$ holds as

$$\mathrm{E}[X^2] = \int_{-\infty}^{\infty} x^2 p(x) dx = \lim_{c\to\infty} \int_0^c x^2 [p(x) + p(-x)] dx$$

$$
\begin{aligned}
&= \lim_{c\to\infty} \left\{ \Big[x^2\big(P(x)-P(-x)\big)\Big]_0^c - 2\int_0^c x[P(x)-P(-x)]dx \right\} \\
&= \lim_{c\to\infty} \left\{ c^2\big(P(c)-P(-c)\big) - 2\int_0^c x\Big[1-\Pr\big\{|X|\geq x\big\}\Big]\,dx \right\} \\
&= \lim_{c\to\infty} \left\{ c^2\big(P(c)-P(-c)\big) - c^2 + 2\int_0^c x\Pr\big\{|X|\geq x\big\}\,dx \right\} \\
&= \frac{1}{2}\int_0^\infty x\Pr\left\{|X|\geq\frac{x}{2}\right\}dx.
\end{aligned}
$$

C.1.3 Derivation of the Ziv-Zakai Bound

In the following, we will briefly review the basic ideas that are used to develop the Ziv-Zakai bound. The interested reader is referred to [Ziv and Zakai, 1969; Bellini and Tartara, 1974; Bell et al., 1997] for the Bayesian case and [Gu and Wong, 1991] for the deterministic case.

Bayesian Bound

Using the identity (2.11), we can write

$$
\begin{aligned}
\mathbf{a}^T\mathbf{S}_{\hat{\boldsymbol{\theta}}}\mathbf{a} = \mathrm{E}\left[\Big(\mathbf{a}^T(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\Big)^2\right] &= \frac{1}{2}\int_0^\infty h\,\Pr\left\{\Big|\mathbf{a}^T(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\Big|\geq\frac{h}{2}\right\}dh \\
&= \frac{1}{2}\int_0^\infty h\left(\Pr\left\{\mathbf{a}^T(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})>\frac{h}{2}\right\}+\Pr\left\{\mathbf{a}^T(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\leq-\frac{h}{2}\right\}\right)dh \\
&= \frac{1}{2}\int_0^\infty h\int\left(p(\boldsymbol{\theta}_0)\Pr\left\{\mathbf{a}^T(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})>\frac{h}{2}\Big|\boldsymbol{\theta}=\boldsymbol{\theta}_0\right\}\right. \\
&\qquad \left.+p(\boldsymbol{\theta}_0+\boldsymbol{\delta})\Pr\left\{\mathbf{a}^T(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta})\leq-\frac{h}{2}\Big|\boldsymbol{\theta}_0+\boldsymbol{\delta}\right\}\right)d\boldsymbol{\theta}_0\,dh, \qquad \text{(C.3)}
\end{aligned}
$$

where $p(\boldsymbol{\theta})$ denotes the a priori PDF of $\boldsymbol{\theta}$. Choosing $\boldsymbol{\delta}$ as $\mathbf{a}^T\boldsymbol{\delta}=h$, we obtain

$$
\begin{aligned}
\mathbf{a}^T\mathbf{S}_{\hat{\boldsymbol{\theta}}}\mathbf{a} = \frac{1}{2}\int_0^\infty h\int\big(p(\boldsymbol{\theta}_0)+p(\boldsymbol{\theta}_0+\boldsymbol{\delta})\big)\left[\frac{p(\boldsymbol{\theta}_0)}{p(\boldsymbol{\theta}_0)+p(\boldsymbol{\theta}_0+\boldsymbol{\delta})}\Pr\left\{\mathbf{a}^T\hat{\boldsymbol{\theta}}>\mathbf{a}^T\boldsymbol{\theta}+\frac{h}{2}\Big|\boldsymbol{\theta}=\boldsymbol{\theta}_0\right\}\right. \\
\left.+\frac{p(\boldsymbol{\theta}_0+\boldsymbol{\delta})}{p(\boldsymbol{\theta}_0)+p(\boldsymbol{\theta}_0+\boldsymbol{\delta})}\Pr\left\{\mathbf{a}^T\hat{\boldsymbol{\theta}}\leq\mathbf{a}^T\boldsymbol{\theta}+\frac{h}{2}\Big|\boldsymbol{\theta}_0+\boldsymbol{\delta}\right\}\right]d\boldsymbol{\theta}_0\,dh. \qquad \text{(C.4)}
\end{aligned}
$$

Comparing this expression with results from detection theory, it can be seen that the term in brackets can be interpreted as the error probability of the binary hypothesis test

$$\begin{aligned} \mathcal{H}_0 &: \boldsymbol{\theta} = \boldsymbol{\theta}_0 \;\text{ where }\; \mathbf{x} \sim p(\mathbf{x}|\boldsymbol{\theta}_0) \\ \mathcal{H}_1 &: \boldsymbol{\theta} = \boldsymbol{\theta}_0 + \boldsymbol{\delta} \;\text{ where }\; \mathbf{x} \sim p(\mathbf{x}|\boldsymbol{\theta}_0 + \boldsymbol{\delta}) \end{aligned} \tag{C.5}$$

with $\Pr(\mathcal{H}_0) = p(\boldsymbol{\theta}_0)/(p(\boldsymbol{\theta}_0) + p(\boldsymbol{\theta}_0 + \boldsymbol{\delta}))$, $\Pr(\mathcal{H}_1) = p(\boldsymbol{\theta}_0 + \boldsymbol{\delta})/(p(\boldsymbol{\theta}_0) + p(\boldsymbol{\theta}_0 + \boldsymbol{\delta}))$ and the suboptimal decision rule

$$\begin{aligned} &\text{Decide } \mathcal{H}_0 \;\text{ if } \mathbf{a}^T\hat{\boldsymbol{\theta}} \leq \mathbf{a}^T\boldsymbol{\theta}_0 + \frac{h}{2} \\ &\text{Decide } \mathcal{H}_1 \;\text{ if } \mathbf{a}^T\hat{\boldsymbol{\theta}} > \mathbf{a}^T\boldsymbol{\theta}_0 + \frac{h}{2}. \end{aligned} \tag{C.6}$$

Thus, the term in brackets can be lower bounded by the minimum error probability $P_{\min}(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0 + \boldsymbol{\delta})$ obtained from the likelihood ratio test. This results in

$$\mathbf{a}^T\mathbf{S}_{\hat{\boldsymbol{\theta}}}\mathbf{a} \leq \frac{1}{2}\int_0^\infty h \int \big(p(\boldsymbol{\theta}_0) + p(\boldsymbol{\theta}_0 + \boldsymbol{\delta})\big)\, P_{\min}(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0 + \boldsymbol{\delta})\, dh, \tag{C.7}$$

which is valid for all $\boldsymbol{\delta}$ with $\mathbf{a}^T\boldsymbol{\delta} = h$. As we are interested in a tight lower bound, we can therefore replace (C.7) by

$$\mathbf{a}^T\mathbf{S}_{\hat{\boldsymbol{\theta}}}\mathbf{a} \leq \frac{1}{2}\int_0^\infty h \max_{\mathbf{a}^T\boldsymbol{\delta}=\mathbf{h}} \int \big(p(\boldsymbol{\theta}_0) + p(\boldsymbol{\theta}_0 + \boldsymbol{\delta})\big)\, P_{\min}(\boldsymbol{\theta}_0, \boldsymbol{\theta}_0 + \boldsymbol{\delta})\, dh. \tag{C.8}$$

Furthermore, a valley-filling function $\mathcal{V}(.)$ can be introduced as we know that the original $\Pr\left\{\left|\mathbf{a}^T(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})\right| \geq \frac{h}{2}\right\}$ is monotonically decreasing. This finally results in (2.12).

Deterministic Bound

Analog to the Bayesian case, we can derive the Ziv-Zakai bound for the deterministic case. The only difference is that we have to require that the estimator $\hat{\boldsymbol{\theta}}(\mathbf{x})$ is translation invariant with respect to a translation of $\boldsymbol{\theta}$, see [Gu and Wong, 1991]. The final result is

$$\mathbf{a}^T\mathbf{S}_{\hat{\boldsymbol{\theta}}}(\boldsymbol{\theta})\mathbf{a} \geq \int_0^\infty h\, \mathcal{V}\left(\max_{\mathbf{a}^T\boldsymbol{\delta}=h} P_{\min}(\boldsymbol{\theta}, \boldsymbol{\theta} + \boldsymbol{\delta})\right) dh. \tag{C.9}$$

C.1.4 Derivation of the MMSE Estimator

The following derivation is adapted from [Scharf, 1990][34] and uses a geometric argument to prove that the estimator with the minimum MSE is the mean of the a posteriori density. Another possibility is to apply (2.25) which was e.g. done in [Kay, 1993].

Let $\mathbf{g}(\mathbf{x}) : \mathbb{R}^K \rightarrow \mathbb{R}^Q$ be an arbitrary statistic and $\mathbf{W} \in \mathbb{R}^{M \times Q}$ be an arbitary weighting matrix. Then it holds

$$\begin{aligned} \mathrm{E}\left[\boldsymbol{\theta}^T \mathbf{W} \mathbf{g}(\mathbf{x})\right] &= \iint \boldsymbol{\theta}^T \mathbf{W} \mathbf{g}(\mathbf{x}) p(\boldsymbol{\theta}, \mathbf{x}) d\boldsymbol{\theta} d\mathbf{x} = \iint \boldsymbol{\theta}^T \mathbf{W} \mathbf{g}(\mathbf{x}) p(\boldsymbol{\theta}|\mathbf{x}) p(\mathbf{x}) d\boldsymbol{\theta} d\mathbf{x} \\ &= \int \left(\int \boldsymbol{\theta}\, p(\boldsymbol{\theta}|\mathbf{x}) d\boldsymbol{\theta} \right)^T \mathbf{W} \mathbf{g}(\mathbf{x}) p(\mathbf{x}) d\mathbf{x} = \mathrm{E}\left[\hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x})^T \mathbf{W} \mathbf{g}(\mathbf{x})\right] \end{aligned}$$

which results in

$$\mathrm{E}\left[\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x}) \right)^T \mathbf{W} \mathbf{g}(\mathbf{x}) \right] = \mathbf{0}, \tag{C.10}$$

i.e. the error $\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x})$ is orthogonal to any statistic $\mathbf{g}(\mathbf{x})$. This allows us to easily verify the following three facts:

1. *Unbiasedness:* Using $\mathbf{g}(\mathbf{x}) = 1$, it follows that

$$\mathrm{E}\left[\hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x})\right] = \mathrm{E}\left[\boldsymbol{\theta}\right], \tag{C.11}$$

   i.e. the MMSE estimator is unbiased.

2. *Orthogonality of error to observations:* Using $\mathbf{g}(\mathbf{x}) = \mathbf{x}$, we obtain

$$\mathrm{E}\left[\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x}) \right)^T \mathbf{W} \mathbf{x} \right] = \mathbf{0}, \tag{C.12}$$

   which implies that the error $\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x})$ is orthogonal to the observations.

3. *Orthogonality of error to MMSE estimator:* Using $\mathbf{g}(\mathbf{x}) = \hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x})$, we obtain

$$\mathrm{E}\left[\left(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x}) \right)^T \mathbf{W} \hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x}) \right] = \mathbf{0}, \tag{C.13}$$

   which implies that the error $\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x})$ is orthogonal to the MMSE estimator.

Figure C.1 shows the orthogonality property (C.10) of the MMSE estimator. The estimation error $\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}_{\mathrm{MMSE}}(\mathbf{x})$ is orthogonal to any statistic $\mathbf{g}(\mathbf{x})$. Using this property allows us to easily

[34] Note that [Scharf, 1990] considers the more general case of outer products, i.e. he considers expectations of the form $\mathrm{E}\left[\boldsymbol{\theta}\mathbf{g}(\mathbf{x})^T\right]$.

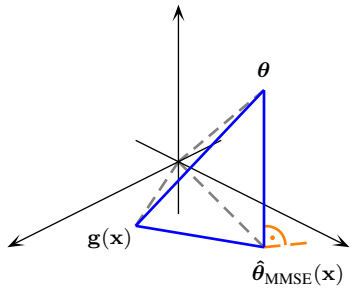


Figure C.1: Orthogonality property of the MMSE estimator

verify that $\hat{\boldsymbol{\theta}}_{\text{MMSE}}(\mathbf{x})$ is really the estimator with the minimum MSE as we can use Pythagoras' theorem and write

$$\begin{aligned}\mathrm{E}\left[\left\|\boldsymbol{\theta}-\mathbf{g}(\mathbf{x})\right\|_{\mathbf{W}}^2\right] &= \mathrm{E}\left[\left\|\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{\text{MMSE}}(\mathbf{x})\right\|_{\mathbf{W}}^2\right] + \mathrm{E}\left[\left\|\hat{\boldsymbol{\theta}}_{\text{MMSE}}(\mathbf{x})-\mathbf{g}(\mathbf{x})\right\|_{\mathbf{W}}^2\right] \\ &\geq \mathrm{E}\left[\left\|\boldsymbol{\theta}-\hat{\boldsymbol{\theta}}_{\text{MMSE}}(\mathbf{x})\right\|_{\mathbf{W}}^2\right]\end{aligned}$$

which proves this statement. Note that we require here $\mathbf{g}(\mathbf{x}) \in \mathbb{R}^M$, i.e. $Q = M$, and $\mathbf{W} \succ \mathbf{0}$.

C.1.5 Jeffreys Prior

The Jeffreys prior is defined as [Robert, 2001]

$$p(\boldsymbol{\theta}) = \sqrt{\det\left\{\mathbf{G}_{\mathrm{D}}(\boldsymbol{\theta})\right\}} \tag{C.14}$$

where $\mathbf{G}_{\mathrm{D}}(\boldsymbol{\theta}) = \mathrm{E}\left[\frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}^T \frac{\partial \ln p(\mathbf{x};\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\right]$ is the Fisher information matrix from (2.5). The Jeffreys prior has the nice property that it is invariant against reparameterization: Let $\boldsymbol{\vartheta} = \mathbf{h}(\boldsymbol{\theta})$ denote a one-to-one transform of $\boldsymbol{\theta}$, then the new prior according to Jeffreys rule is $p(\boldsymbol{\vartheta}) = \sqrt{\det\left\{\mathbf{G}_{\mathrm{D}}(\boldsymbol{\vartheta})\right\}}$. Using the transformation property of the Fisher information matrix [Kay, 1993], this can be reformulated into

$$\begin{aligned}p(\boldsymbol{\vartheta}) &= \sqrt{\det\left\{\mathbf{G}_{\mathrm{D}}(\boldsymbol{\vartheta})\right\}} = \sqrt{\det\left\{\mathbf{J}(\boldsymbol{\theta})^{-T}\mathbf{G}_{\mathrm{D}}(\boldsymbol{\theta})\mathbf{J}(\boldsymbol{\theta})^{-1}\right\}}\Bigg|_{\boldsymbol{\theta}=\mathbf{h}^{-1}(\boldsymbol{\vartheta})} \\ &= \frac{1}{\det\left\{\mathbf{J}(\boldsymbol{\theta})\right\}}\sqrt{\det\left\{\mathbf{G}_{\mathrm{D}}(\boldsymbol{\theta})\right\}}\Bigg|_{\boldsymbol{\theta}=\mathbf{h}^{-1}(\boldsymbol{\vartheta})}\end{aligned}$$

with $\mathbf{J}(\boldsymbol{\theta}) = \frac{\partial \mathbf{h}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}$. The same result would have been obtained using the change of variables law for random vectors and hence, we can conclude that the Jeffreys prior is invariant with respect to a reparameterization of $\boldsymbol{\theta}$.

C.1.6 Numerical Quadrature

In this section, we will briefly review numerical integration for $M = 1$, i.e. the one-dimensional case. The idea is to use

$$\int_a^b f(\theta)d\theta \approx \sum_{i=1}^{I} w_i f(\theta_i) \qquad a \leq \theta_1, \ldots, \theta_I \leq b \tag{C.15}$$

with the requirement that this approach is exact for polynomials up to degree P, i.e.

$$\int_a^b \theta^p d\theta \stackrel{!}{=} \sum_{i=1}^{I} w_i \theta_i^p \quad \text{for all} \quad 0 \leq p \leq P. \tag{C.16}$$

Using equispaced nodes θ_i yields the well known *Newton-Cotes* formulas with $P = I - 1$. If the nodes are chosen carefully but not equispaced, then the *Gaussian quadrature* rules with $P = 2I - 1$ are obtained.

For example, using equispaced nodes with $I = 2$ yields the two requirements

$$p = 0: \qquad \int_a^b 1 d\theta = b - a \stackrel{!}{=} w_1 + w_2$$

$$p = 1: \qquad \int_a^b \theta \, d\theta = \frac{b^2}{2} - \frac{a^2}{2} \stackrel{!}{=} w_1 a + w_2 b$$

and, solving for w_1, w_2, we obtain the *trapezoidal rule*

$$\int_a^b f(\theta)d\theta \approx \frac{b-a}{2}\left(f(a) + f(b)\right). \tag{C.17}$$

Instead of increasing the quadrature order I, which may yield numerical problems e.g. due to loss of significant bits, *composite rules* are used which split $[a, b]$ into L subintervals where for each subinterval a quadrature rule is used. This allows to obtain better numerical approximations to the integral. For example, if we use the trapezoidal rule, we obtain

$$\int_a^b f(\theta)d\theta \approx \frac{b-a}{L}\left(\frac{f(a)}{2} + \sum_{l=1}^{L-1} f\left(a + \frac{b-a}{L}l\right) + \frac{f(b)}{2}\right). \tag{C.18}$$

Table C.1 summarizes the integration error for the Newton-Cotes and the Gaussian quadrature rules. The shorthand notations $\bar{D}_j = \max_{\theta\in[a,b]}|f^{(j)}(\theta)|$ and $S = IL$, which denotes the

Name		Integration error
Newton-Cotes, Trapezoidal rule	$I = 2$	$\frac{(b-a)^3}{12L^2}\bar{D}_2 \sim \mathcal{O}\left(S^{-2}\right)$
Newton-Cotes, Simpson's rule	$I = 3$	$\frac{(b-a)^5}{2880L^4}\bar{D}_4 \sim \mathcal{O}\left(S^{-4}\right)$
Newton-Cotes, Simpson's $3/8$ rule	$I = 4$	$\frac{(b-a)^5}{6480L^4}\bar{D}_4 \sim \mathcal{O}\left(S^{-4}\right)$
Newton-Cotes, Boole's rule	$I = 5$	$\frac{(b-a)^7}{1\,935\,360L^6}\bar{D}_6 \sim \mathcal{O}\left(S^{-6}\right)$
Gaussian Quadrature	I	$\mathcal{O}\left(S^{-2I}\right)$

Table C.1: Overview of Newton-Cotes and Gaussian quadrature

number of nodes in $[a, b]$, are used. For a derivation of the integration error, see e.g. [Dahlquist and Björck, 2008].

C.2 Details, Derivations and Proofs for Chapter 3

C.2.1 S-Procedure

The S-procedure is a tool which is frequently used in robust optimization to deal with constraints of the form

$$f_0(\mathbf{x}) \geq 0 \;\; \textit{for all} \;\; \mathbf{x} \;\; \textit{such that} \;\; f_1(\mathbf{x}) \geq 0, \ldots, f_L(\mathbf{x}) \geq 0 \tag{C.19}$$

where $f_l(\mathbf{x}) = \mathbf{x}^T\mathbf{F}_l\mathbf{x} + 2\mathbf{s}_l^T\mathbf{x} + r_l$ are quadratic functions with $\mathbf{F}_l = \mathbf{F}_l^T$ [Boyd et al., 1994]. Condition (C.19) corresponds in general to checking the nonnegativity of a nonconvex quadratic form on a nonconvex set as $\mathbf{F}_l$ with $l = 0, \ldots, L$ does not have to be positive definite. Therefore, the S-procedure allows us to relax nonconvex problems into convex ones and is therefore very popular. In the following, we will outline the basic result of the S-procedure. For more details, the interested reader should consult e.g. [Boyd and Vandenberghe, 2007; Jönsson, 2006].

A sufficient condition to check (C.19) is the existence of $\lambda_1, \ldots, \lambda_L \geq 0$ such that

$$f_0(\mathbf{x}) - \lambda_1 f_1(\mathbf{x}) - \ldots - \lambda_L f_L(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \tag{C.20}$$

is fulfilled. This can be seen from the fact that (C.20) ensures $f_0(\mathbf{x}) \geq \lambda_1 f_1(\mathbf{x}) + \ldots + \lambda_L f_L(\mathbf{x})$ which implies (C.19). Furthermore, (C.20) corresponds to a *linear matrix inequality* (LMI) as

$$\begin{aligned}
\text{(C.20)} \;&\Leftrightarrow \exists \lambda_l \geq 0 \;\text{ s.t. }\; f_0(\mathbf{x}) - \lambda_1 f_1(\mathbf{x}) - \ldots - \lambda_L f_L(\mathbf{x}) \geq 0 \quad \forall \mathbf{x} \\
&\Leftrightarrow \exists \lambda_l \geq 0 \;\text{ s.t. }\; \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix}^T \begin{bmatrix} \mathbf{F}_0 - \sum_l \lambda_l \mathbf{F}_l & \mathbf{s}_0 - \sum_l \lambda_l \mathbf{s}_l \\ \mathbf{s}_0^T - \sum_l \lambda_l \mathbf{s}_l^T & r_0 - \sum_l \lambda_l r_l \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \geq 0 \quad \forall \mathbf{x}
\end{aligned}$$

$$\Leftrightarrow \exists \lambda_l \geq 0 \text{ s.t. } \begin{bmatrix} \mathbf{F}_0 - \sum_l \lambda_l \mathbf{F}_l & \mathbf{s}_0 - \sum_l \lambda_l \mathbf{s}_l \\ \mathbf{s}_0^T - \sum_l \lambda_l \mathbf{s}_l^T & r_0 - \sum_l \lambda_l r_l \end{bmatrix} \succeq \mathbf{0} \tag{C.21}$$

which is a LMI.[35] Note that the converse, i.e. (C.19) $\Rightarrow$ (C.20), is in general not true and the S-procedure is hence often called *lossy*. An exception for a *lossless* S-procedure is the case for $L = 1$, see e.g. [Boyd and Vandenberghe, 2007].

C.2.2 Series Expansions for the Distribution of Quadratic Forms of Gaussian Random Variables on Elliptic Regions

In this section, we will give a series expansion of the CDF $P[(\mathbf{x} - \mathbf{x}_0)^T \mathbf{A}(\mathbf{x} - \mathbf{x}_0) \leq t]$ with $\mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \mathbf{C})$ and of its corresponding PDF. This problem has been intensively studied in the past. One of the first results was given by [Ruben, 1962] who established a series expansion in terms of central and noncentral χ^2-CDFs/PDFs. This result was further generalized by [Kotz et al., 1967a,b; Gideon and Gurland, 1976], who derived other series expansions, namely a power series, Laguerre series and χ^2-series representation using a unifying framework and thus combining many results that were published before.

The following derivation follows the same lines as in [Johnson and Kotz, 1970] and gives the expansion in terms of central χ^2-CDFs/PDFs. It has the advantage compared to the other expansions that it allows to develop a *mixture representation* which does not show the effect of loss of significance. Moreover, a simple yet good approximation of the truncation error can be established.

Transformation to Standard Form

Similar to Section 3.5.2, we can apply the transform $\mathbf{x} - \boldsymbol{\mu} = \mathbf{U}^T \mathbf{V} \mathbf{y}$ where $\mathbf{U}$ is a Cholesky factor of $\mathbf{C}$, i.e. $\mathbf{C} = \mathbf{U}^T \mathbf{U}$ and $\mathbf{V}$, $\mathbf{D}$ are the matrix of eigenvectors/eigenvalues of $\mathbf{U}\mathbf{A}\mathbf{U}^T$. Applying this transform turns the problem

$$P[(\mathbf{x} - \mathbf{x}_0)^T \mathbf{A}(\mathbf{x} - \mathbf{x}_0) \leq t] = \int\limits_{(\mathbf{x}-\mathbf{x}_0)^T \mathbf{A}(\mathbf{x}-\mathbf{x}_0) \leq t} \frac{\exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \mathbf{C}^{-1}(\mathbf{x} - \boldsymbol{\mu})\right\}}{(2\pi)^{n/2} \det\{\mathbf{C}\}^{1/2}} d\mathbf{x}$$

into the standard form

$$P\left[(\mathbf{y} - \mathbf{y}_0)^T \mathbf{D}(\mathbf{y} - \mathbf{y}_0) \leq t\right] = P\left[\sum_{m=1}^{M} d_m (y_m - y_{m0})^2 \leq t\right] = F(t; \mathbf{d}, \mathbf{y}_0) \tag{C.22}$$

[35] The last equivalence stems from the fact that each quadratic polynomial can be written as a sum of squares polynomial, see e.g. [Lasserre, 2010].

where $\mathbf{y} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ and $\mathbf{y}_0 = (\mathbf{U}^T\mathbf{V})^{-1}(\mathbf{x}_0 - \boldsymbol{\mu})$. $F(t; \mathbf{d}, \mathbf{y}_0)$ gives the probability of M independent unit normal variables to lie in an ellipsoid where the principal axes are parallel to the coordinate axes. The corresponding density will be denoted by $p(t; \mathbf{d}, \mathbf{y}_0) = \frac{dF(t;\mathbf{d},\mathbf{y}_0)}{dt}$.

Expansion using Central χ^2-Distributions/Densities

The idea is now to find expansions of the form[36]

$$p(t; \mathbf{d}, \mathbf{y}_0) = \frac{1}{\beta} \sum_{k=0}^{\infty} a_k p_{\chi^2}(t/\beta; M + 2k) \tag{C.23a}$$

$$F(t; \mathbf{d}, \mathbf{y}_0) = \sum_{k=0}^{\infty} a_k F_{\chi^2}(t/\beta; M + 2k) \tag{C.23b}$$

where $F_{\chi^2}(t/\beta; M + 2k)$ and $p_{\chi^2}(t/\beta; M + 2k)$ denote the CDF and PDF of a central χ^2-distribution with $M + 2k$ degrees of freedom evaluated at t/β. $\beta > 0$ is a suitable chosen constant which will be discussed later. Note that (C.23b) follows directly from (C.23a) by termwise integration with respect to t and hence it suffices to proof (C.23a).

In order to find the coefficients a_k in (C.23), we calculate the characteristic function of the density $p(t; \mathbf{d}, \mathbf{y}_0)$ in (C.23a) and obtain

$$\begin{aligned}
\Phi(\omega; \mathbf{d}, \mathbf{y}_0) &= \mathrm{E}\left[\exp\left\{j\omega \sum_{m=1}^{M} d_m (y_m - y_{0m})^2\right\}\right] \\
&= \frac{1}{(2\pi)^{M/2}} \int_{\mathbb{R}^M} \exp\left\{j\omega \sum_{m=1}^{M} d_m (y_m - y_{m0})^2 - \frac{1}{2}\sum_{m=1}^{M} y_m^2\right\} d\mathbf{y} \\
&= \exp\left\{-\frac{1}{2}\sum_{m=1}^{M} y_{0m}^2 + \frac{1}{2}\sum_{m=1}^{M} \frac{y_{0m}^2}{1 - 2j\omega d_m}\right\} \prod_{m=1}^{M} (1 - 2j\omega d_m)^{-1/2}
\end{aligned} \tag{C.24}$$

as $\int_{\mathbb{R}^M} \exp\left\{\mathbf{b}^T\mathbf{y} - \frac{1}{2}\mathbf{y}^T\mathbf{A}\mathbf{y}\right\} d\mathbf{y} = \frac{(2\pi)^{M/2}}{\det\{\mathbf{A}\}^{1/2}} \exp\left\{\frac{1}{2}\mathbf{b}^T\mathbf{A}^{-1}\mathbf{b}\right\}$ for $\mathbf{A} \succ \mathbf{0}$. The characteristic function of a central χ^2-CDF with $2k$ degrees of freedom is $\Phi_{\chi^2}(\omega; 2k) = (1 - 2j\omega)^{-k}$, and thus we need to expand (C.24) in powers of $\beta(1 - 2j\omega\beta)^{-1}$ in order to obtain the series expansion as in (C.23a). Using the shorthand notation $\theta = \beta(1 - 2j\omega\beta)^{-1}$ in (C.24) yields

$$\Phi(\omega; \mathbf{d}, \mathbf{y}_0) = \exp\left\{\frac{\theta - \beta}{2} \sum_{m=1}^{M} \frac{y_{0m}^2}{\beta - (1 - \frac{\beta}{d_m})\theta}\right\} \prod_{m=1}^{M} \left(\frac{\beta\theta}{d_m}\right)^{1/2} \left(\beta - \left(1 - \frac{\beta}{d_m}\right)\theta\right)^{-1/2}$$

[36] Note that some care is needed if working with infinite series as in (C.23). However, Ruben proved in [Ruben, 1962] that such an expansion exists and that it is uniformly convergent on any finite interval of t. Furthermore, he also proved the uniform convergence of the derivative and therefore, we can e.g. safely differentiate each element of the sum w.r.t. t to obtain a series expansion of the derivative, see also [Kotz et al., 1967a].

$$= \frac{\theta^{M/2}}{\beta} \sum_{k=0}^{\infty} a_k \theta^k,$$

i.e. the coefficients a_k can be obtained by the Taylor series expansion

$$\exp\left\{ \frac{\theta-\beta}{2} \sum_{m=1}^{M} \frac{y_{0m}^2}{\beta - (1-\frac{\beta}{d_m})\theta} \right\} \prod_{m=1}^{M} \left(\frac{\beta}{d_m}\right)^{1/2} \left(\beta - \left(1-\frac{\beta}{d_m}\right)\theta\right)^{-1/2} = \frac{1}{\beta} \sum_{k=0}^{\infty} a_k \theta^k. \tag{C.25}$$

In order to obtain a recursive formula for the coefficients a_k, we consider the Taylor series expansion of the logarithm of the lefthand side in (C.25) [37]

$$\frac{\theta-\beta}{2} \sum_{m=1}^{M} \frac{y_{0m}^2}{\beta - (1-\frac{\beta}{d_m})\theta} + \frac{1}{2} \sum_{m=1}^{M} \ln\left(\frac{\beta}{d_m}\right) - \frac{1}{2} \sum_{m=1}^{M} \ln\left(\beta - \left(1-\frac{\beta}{d_m}\right)\theta\right) = \sum_{k=0}^{\infty} b_k \theta^k. \tag{C.26}$$

Using $(1-\alpha x)^{-1} = \sum_{k=0}^{\infty} \alpha^k x^k$ and $\ln(1-\alpha x) = -\sum_{k=1}^{\infty} k^{-1} \alpha^k x^k$ in (C.26) and some further manipulations yields

$$b_k = \sum_{m=1}^{M} \left(1-\frac{\beta}{d_m}\right)^k + k\beta \sum_{m=1}^{M} \frac{y_{0m}^2}{d_m} \left(1-\frac{\beta}{d_m}\right)^{k-1} \quad \text{for all } k \geq 1$$

and we finally obtain the following recursive representation of the series coefficients a_k

$$a_k = \begin{cases} \exp\left\{ -\frac{1}{2} \sum_{m=1}^{M} y_{0m}^2 \right\} \prod_{m=1}^{M} \left(\frac{\beta}{d_m}\right)^{\frac{1}{2}} & k = 0 \\ \frac{1}{2k} \sum_{l=0}^{k-1} d_{k-l} c_l & k \geq 1 \end{cases}. \tag{C.27}$$

Practical Considerations: Choice of β and Truncation Error Estimate

In [Ruben, 1962], the choice of β was discussed. Ruben showed that if $0 < \beta \leq \min\{d_1, \ldots, d_M\}$ is chosen, then the series (C.23) are mixture representations, i.e. $a_k \geq 0$ and $\sum_{k=0}^{\infty} a_k = 1$. This

[37] Consider the Taylor series expansion $f(x) = \sum_{k=0}^{\infty} c_k x^k$ and the related Taylor series $\ln(f(x)) = \sum_{k=0}^{\infty} d_k x^k$ (assuming that it exists). Then, the coefficients show the relationship $c_k = \begin{cases} f(0) & k = 0 \\ \frac{1}{k} \sum_{l=0}^{k-1} (k-l) d_{k-l} c_l & k \geq 1 \end{cases}$. This can be seen from the identity

$$\frac{d}{dx} \ln(f(x)) = \sum_{k=1}^{\infty} k d_k x^{k-1} = \frac{f'(x)}{f(x)} = \frac{\sum_{k=1}^{\infty} k c_k x^{k-1}}{\sum_{k=0}^{\infty} c_k x^k} \Leftrightarrow \sum_{k=0}^{\infty} \sum_{j=1}^{\infty} c_k j d_j x^{k+j-1} = \sum_{k=1}^{\infty} k c_k x^{k-1}$$

which implies $c_k = \frac{1}{k} \sum_{l=0}^{k-1} (k-l) d_{k-l} c_l$ for all $k \geq 1$.

renders the χ^2-expansion especially useful for practical implementations as we can now easily obtain approximations of the truncation error if only the first K terms are used in (C.23). Let the truncation error be defined as

$$e_K = \frac{1}{\beta} \sum_{k=K+1}^{\infty} a_k p_{\chi^2}(t/\beta; M + 2k) \tag{C.28a}$$

$$E_K = \sum_{k=K+1}^{\infty} a_k F_{\chi^2}(t/\beta; M + 2k) \tag{C.28b}$$

then an upper bound on e_K and E_K is given by

$$e_K \leq \frac{1}{\beta} \left(1 - \sum_{k=0}^{K} a_k \right) p_{\chi^2}(t/\beta; M + 2(K+1)) \tag{C.29a}$$

$$E_K \leq \left(1 - \sum_{k=0}^{K} a_k \right) F_{\chi^2}(t/\beta; M + 2(K+1)) \tag{C.29b}$$

as $p_{\chi^2}(t/\beta; M + 2k)$ and $F_{\chi^2}(t/\beta; M + 2k)$ are monotonically decreasing in k for fixed t/β.[38] These upper bounds on the error can be used as stopping criterion in a practical implementation.

C.3 Details, Derivations and Proofs for Chapter 4

C.3.1 Proof of Theorem 2

The proof of Theorem 2 is straightforward and is given below.

(a) Assuming $\mathbb{T}$ to be a bounded set on $\mathbb{R}^M$, we immediately see that

$$\lim_{\lambda \to 0} \frac{p(\boldsymbol{\theta}, \mathbf{x})^\lambda}{\int p(\boldsymbol{\theta}, \mathbf{x})^\lambda d\boldsymbol{\theta}} = \text{const.},$$

i.e. it converges pointwise to a uniform distribution on $\mathbb{T}$. Therefore, $\hat{\boldsymbol{\theta}}(\mathbf{x}; 0)$ calculates the center of gravity of $\mathbb{T}$ which is equivalent to the a priori MMSE estimator.

(b) Setting $\lambda = 1$ in (4.1), we obtain

$$\frac{p(\boldsymbol{\theta}, \mathbf{x})}{\int p(\boldsymbol{\theta}, \mathbf{x}) d\boldsymbol{\theta}} = p(\boldsymbol{\theta}|\mathbf{x})$$

[38] Note that the PDF $p_{\chi^2}(t/\beta; M + 2k)$ is only monotonically decreasing for all $k > K_0$ where K_0 is the solution to $\ln(t/\beta) - \ln(2) = \Psi(K_0/2)$ and $\Psi(.)$ is the digamma function which is defined as $\Psi(x) = \Gamma'(x)/\Gamma(x)$. However, K_0 is in our case often very small and thus (C.29a) is valid.

and thus $\hat{\boldsymbol{\theta}}(\mathbf{x};1) = \int \boldsymbol{\theta}\, p(\boldsymbol{\theta}|\mathbf{x}) d\boldsymbol{\theta} = \mathrm{E}[\boldsymbol{\theta}|\mathbf{x}]$, which is the MMSE estimator.

(c) To prove this part, we use a result from [Pincus, 1968]: Given a continuous function $f(\boldsymbol{\theta})$, which attains a global maximum at exactly one point in $\mathbb{T}$, then Pincus showed

$$\arg\max_{\boldsymbol{\theta}} f(\boldsymbol{\theta}) = \lim_{\lambda\to\infty} \frac{\int_{\mathbb{T}} \boldsymbol{\theta}\, f(\boldsymbol{\theta})^\lambda d\boldsymbol{\theta}}{\int_{\mathbb{T}} f(\boldsymbol{\theta})^\lambda d\boldsymbol{\theta}}.$$

Using this theorem, we conclude that $\lim_{\lambda\to\infty} \hat{\boldsymbol{\theta}}(\mathbf{x};\lambda)$ is the MAP estimator.

C.3.2 Proof of Theorem 3

To prove Theorem 3, we have to first show that Lemma 1 is true.

First of all, we would like to point out that the delta function can be expressed as a limit of the normal distribution, i.e.

$$g(\boldsymbol{\theta}; a^2) = \frac{1}{a^M \pi^{M/2}} e^{-\|\boldsymbol{\theta}\|^2/a^2} \xrightarrow{a\to 0} \delta(\boldsymbol{\theta}).$$

They are equivalent in the sense that $f(\mathbf{0}) = \int f(\boldsymbol{\theta})\delta(\boldsymbol{\theta}) d\boldsymbol{\theta} = \lim_{a\to 0} \int f(\boldsymbol{\theta}) g(\boldsymbol{\theta}; a^2) d\boldsymbol{\theta}$.

(a) $p(\boldsymbol{\theta}, \mathbf{x}) = \delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0)$:

$$\begin{aligned}
\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda) &= \lim_{a\to 0} \frac{\int \boldsymbol{\theta}\, g(\boldsymbol{\theta} - \boldsymbol{\theta}_0; a^2)^\lambda d\boldsymbol{\theta}}{\int g(\boldsymbol{\theta} - \boldsymbol{\theta}_0; a^2)^\lambda d\boldsymbol{\theta}} = \lim_{a\to 0} \frac{\int \boldsymbol{\theta}\, g(\boldsymbol{\theta} - \boldsymbol{\theta}_0; \frac{a^2}{\lambda}) d\boldsymbol{\theta}}{\int g(\boldsymbol{\theta} - \boldsymbol{\theta}_0; \frac{a^2}{\lambda}) d\boldsymbol{\theta}} \\
&= \lim_{a\to 0} \int \boldsymbol{\theta}\, g(\boldsymbol{\theta} - \boldsymbol{\theta}_0; \frac{a^2}{\lambda}) d\boldsymbol{\theta} = \boldsymbol{\theta}_0
\end{aligned}$$

(b) $p(\boldsymbol{\theta}, \mathbf{x}) = P\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_0) + (1-P)\delta(\boldsymbol{\theta} - \boldsymbol{\theta}_1)$:

$$\begin{aligned}
\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda) &= \lim_{a\to 0} \frac{\int \boldsymbol{\theta}\, [P g(\boldsymbol{\theta} - \boldsymbol{\theta}_0; a^2) + (1-P) g(\boldsymbol{\theta} - \boldsymbol{\theta}_1, a^2)]^\lambda d\boldsymbol{\theta}}{\int [P g(\boldsymbol{\theta} - \boldsymbol{\theta}_0; a^2) + (1-P) g(\boldsymbol{\theta} - \boldsymbol{\theta}_1; a^2)]^\lambda d\boldsymbol{\theta}} \\
&= \lim_{a\to 0} \frac{P^\lambda}{P^\lambda + (1-P)^\lambda} \int \boldsymbol{\theta}\, g(\boldsymbol{\theta} - \boldsymbol{\theta}_0; \frac{a^2}{\lambda}) d\boldsymbol{\theta} \\
&\quad + \lim_{a\to 0} \frac{(1-P)^\lambda}{P^\lambda + (1-P)^\lambda} \int \boldsymbol{\theta}\, g(\boldsymbol{\theta} - \boldsymbol{\theta}_1; \frac{a^2}{\lambda}) d\boldsymbol{\theta} \\
&= \frac{P^\lambda \boldsymbol{\theta}_0 + (1-P)^\lambda \boldsymbol{\theta}_1}{P^\lambda + (1-P)^\lambda}
\end{aligned}$$

where we used the fact that

$$[P g(\boldsymbol{\theta} - \boldsymbol{\theta}_0; a^2) + (1-P) g(\boldsymbol{\theta} - \boldsymbol{\theta}_1; a^2)]^\lambda$$

$$\rightarrow P^\lambda g(\boldsymbol{\theta}-\boldsymbol{\theta}_0;a^2)^\lambda + (1-P)^\lambda g(\boldsymbol{\theta}-\boldsymbol{\theta}_1;a^2)^\lambda$$

for $\boldsymbol{\theta}_0 \neq \boldsymbol{\theta}_1$ and $a \rightarrow 0$.

Now, we are in the position to prove Theorem 3 by contradiction. Suppose $\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda)$ has a corresponding loss function $L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})$ which is continuously differentiable but not symmetric. Then at least one of the following two cases has to be true:

(a) There is a $\boldsymbol{\theta}_0$ such that

$$\left|\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}}\right|\Bigg|_{\substack{\boldsymbol{\theta}=\boldsymbol{\theta}_0\\ \hat{\boldsymbol{\theta}}=\boldsymbol{\theta}_0}} \neq \left|\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}}\right|\Bigg|_{\substack{\boldsymbol{\theta}=-\boldsymbol{\theta}_0\\ \hat{\boldsymbol{\theta}}=-\boldsymbol{\theta}_0}} \tag{C.30}$$

Now, consider the special PDF $p(\boldsymbol{\theta},\mathbf{x}) = \delta(\boldsymbol{\theta}-\boldsymbol{\theta}_0)$. As $\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda)$ from (4.1) holds for all densities, we can directly use the result of the Lemma and obtain $\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda) = \boldsymbol{\theta}_0$. A necessary condition that $\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda)$ is the OBE for the loss function $L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})$ is (2.25)

$$\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}}\Bigg|_{\substack{\boldsymbol{\theta}=\boldsymbol{\theta}_0\\ \hat{\boldsymbol{\theta}}=\boldsymbol{\theta}_0}} = \mathbf{0}.$$

Furthermore, consider the special PDF $p(\boldsymbol{\theta},\mathbf{x}) = \delta(\boldsymbol{\theta}+\boldsymbol{\theta}_0)$ which has the OBE $\hat{\boldsymbol{\theta}}(\mathbf{x};\lambda) = -\boldsymbol{\theta}_0$. Using again (2.25), we obtain the necessary condition

$$\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}}\Bigg|_{\substack{\boldsymbol{\theta}=-\boldsymbol{\theta}_0\\ \hat{\boldsymbol{\theta}}=-\boldsymbol{\theta}_0}} = \mathbf{0}$$

which cannot be true as we assumed (C.30).

(b) There is a $\boldsymbol{\theta}_0$ and $\boldsymbol{\theta}_1$ such that

$$\left|\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}}\right|\Bigg|_{\substack{\boldsymbol{\theta}=\boldsymbol{\theta}_0\\ \hat{\boldsymbol{\theta}}=\boldsymbol{\theta}_1}} \neq \left|\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}}\right|\Bigg|_{\substack{\boldsymbol{\theta}=-\boldsymbol{\theta}_0\\ \hat{\boldsymbol{\theta}}=-\boldsymbol{\theta}_1}} \tag{C.31}$$

Now consider the special PDF $p(\boldsymbol{\theta},\mathbf{x}) = P\delta(\boldsymbol{\theta}-\boldsymbol{\theta}_0)+(1-P)\delta(\boldsymbol{\theta}-\boldsymbol{\theta}_1)$ which, according to the above Lemma, has the OBE $\mathbf{u} = \hat{\boldsymbol{\theta}}(\mathbf{x};\lambda) = (P^\lambda\boldsymbol{\theta}_0+(1-P)^\lambda\boldsymbol{\theta}_1)/(P^\lambda+(1-P)^\lambda)$. A necessary condition that has to be fulfilled is (2.25) which yields

$$P\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}}\Bigg|_{\substack{\boldsymbol{\theta}=\boldsymbol{\theta}_0\\ \hat{\boldsymbol{\theta}}=\mathbf{u}}} + (1-P)\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial \hat{\boldsymbol{\theta}}}\Bigg|_{\substack{\boldsymbol{\theta}=\boldsymbol{\theta}_1\\ \hat{\boldsymbol{\theta}}=\mathbf{u}}} = \mathbf{0}.$$

Furthermore, the PDF $p(\boldsymbol{\theta},\mathbf{x}) = P\delta(\boldsymbol{\theta}+\boldsymbol{\theta}_0)+(1-P)\delta(\boldsymbol{\theta}+\boldsymbol{\theta}_1)$ results in the OBE $-\mathbf{u}$

and the necessary condition (2.25) is

$$P\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial\hat{\boldsymbol{\theta}}}\bigg|_{\substack{\boldsymbol{\theta}=-\boldsymbol{\theta}_0\\ \hat{\boldsymbol{\theta}}=-\mathbf{u}}}+(1-P)\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial\hat{\boldsymbol{\theta}}}\bigg|_{\substack{\boldsymbol{\theta}=-\boldsymbol{\theta}_1\\ \hat{\boldsymbol{\theta}}=-\mathbf{u}}}=\mathbf{0}.$$

Without loss of generality, we can assume

$$\left|\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial\hat{\boldsymbol{\theta}}}\right|\bigg|_{\substack{\boldsymbol{\theta}=\boldsymbol{\theta}_1\\ \hat{\boldsymbol{\theta}}=\boldsymbol{\theta}_1}}=\left|\frac{\partial L(\boldsymbol{\theta},\hat{\boldsymbol{\theta}})}{\partial\hat{\boldsymbol{\theta}}}\right|\bigg|_{\substack{\boldsymbol{\theta}=-\boldsymbol{\theta}_1\\ \hat{\boldsymbol{\theta}}=-\boldsymbol{\theta}_1}}$$

as we can otherwise use (a) and show that the loss is asymmetric. Taking the limit $P\to 0$ ($P>0$), we see that both necessary conditions contradict the assumption (C.31).

C.3.3 Gradient of (4.4)

In this section, we derive the gradient of the Bayes risk with respect to an element γ of $\mathbf{p}$. Using the gradient is advantageous to solve the optimization problem (4.4) as gradient descent methods can be used. Taking the derivative of BR in (4.4) with respect to γ, we obtain for the first-order derivative

$$\frac{\partial\mathrm{BR}}{\partial\gamma}=\iint\left(\frac{\partial L(\boldsymbol{\theta},\mathbf{u})}{\partial\mathbf{u}}\bigg|_{\mathbf{u}=\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p})}\right)^T\frac{\partial\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p})}{\partial\gamma}p(\boldsymbol{\theta},\mathbf{x})d\boldsymbol{\theta}d\mathbf{x}.$$

Using the shorthand notations $p_\lambda(\boldsymbol{\theta}|\mathbf{x})=p(\boldsymbol{\theta},\mathbf{x})^\lambda/\int p(\boldsymbol{\theta},\mathbf{x})^\lambda d\boldsymbol{\theta}$ and $\mathbf{D}=\frac{\partial\mathbf{f}_1}{\partial\mathbf{z}}=\xi_1\mathbf{I}+\mathrm{diag}\{\phi_1/z_1,\ldots,\phi_M/z_M\}$ evaluated at $\mathbf{z}=\int\mathbf{f}_2(\boldsymbol{\theta},\mathbf{p}_2)p_\lambda(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\neq\mathbf{0}$, we obtain

$$\begin{aligned}
\frac{\partial\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p})}{\partial\xi_1}&=\int\mathbf{f}_2(\boldsymbol{\theta},\mathbf{p}_2)p_\lambda(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\\
\frac{\partial\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p})}{\partial\xi_2}&=\mathbf{D}\cdot\int\boldsymbol{\theta}\,p_\lambda(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\\
\frac{\partial\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p})}{\partial\xi_3}&=\mathbf{D}\cdot\int e^{\boldsymbol{\psi}\circ\boldsymbol{\theta}}\,p_\lambda(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\\
\frac{\partial\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p})}{\partial\lambda}&=\mathbf{D}\cdot\bigg(\int\mathbf{f}_2(\boldsymbol{\theta};\mathbf{p}_2)\,\ln\big(p(\boldsymbol{\theta},\mathbf{x})\big)\,p_\lambda(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\\
&\qquad-\int\mathbf{f}_2(\boldsymbol{\theta};\mathbf{p}_2)\,p_\lambda(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\int\ln\big(p(\boldsymbol{\theta},\mathbf{x})\big)\,p_\lambda(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\bigg)\\
\frac{\partial\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p})}{\partial\boldsymbol{\phi}}&=\mathrm{diag}\left\{\ln\left|\int\mathbf{f}_2(\boldsymbol{\theta};\mathbf{p}_2)\,p_\lambda(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\right|\right\}\\
\frac{\partial\hat{\boldsymbol{\theta}}(\mathbf{x};\mathbf{p})}{\partial\boldsymbol{\psi}}&=\xi_3\mathbf{D}\cdot\mathrm{diag}\left\{\int\boldsymbol{\theta}\circ e^{\boldsymbol{\psi}\circ\boldsymbol{\theta}}\,p_\lambda(\boldsymbol{\theta}|\mathbf{x})d\boldsymbol{\theta}\right\}.
\end{aligned}$$

Note that all integrals can again be calculated using Monte Carlo integration, especially importance sampling as discussed in Section 4.2.3.

C.3.4 Definition of Kummer's Function $M(a, b, z)$

Kummer's function $M(a, b, z)$ is one of the two linearly independent solutions of *Kummer's equation* which is an *ordinary differential equation* (ODE) and defined as [Abramowitz and Stegun, 1964]

$$z\frac{d^2w(z)}{dz^2} + (b - z)\frac{dw(z)}{dz} - aw(z) = 0. \tag{C.33}$$

The two solutions are *Kummer's function* $M(a, b, z)$ and *Tricomi's function* $U(a, b, z)$. Using special choices for the arguments of $M(a, b, z)$ and $U(a, b, z)$, many well known functions can be expressed by them, e.g. the Bessel functions.

For completeness, we would like to mention that Kummer's function $M(a, b, z)$ can be expressed as a generalized hypergeometric series, i.e.

$$M(a, b, z) = \sum_{n=0}^{\infty} \frac{(a)_n}{(b)_n} \frac{z^n}{n!}, \tag{C.34}$$

where $(a)_n$ and $(b)_n$ are *rising factorials* given by $(a)_n = a(a + 1)(a + 2) \cdots (a + n - 1)$ and in particularly $(a)_0 = 1$. Plugging (C.34) into (C.33), it is easy to prove that $M(a, b, z)$ is a solution of Kummer's equation (C.33). The interested reader should consult [National Institute of Standards and Technology, 2012] for more details and relations.

C.3.5 Gradient and Hessian of the Log-Likelihood (4.32)

In the following, we will give the elements of the gradient and Hessian of the Log-likelihood function $L(\mathbb{U})$ in order to use efficient optimization algorithms. The first-order derivatives are

$$\begin{aligned}
\frac{\partial L(\mathbb{U})}{\partial \mu} &= \mathbf{1}^T \left(d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) = \mathbf{1}^T d\mathbf{Q} \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
\frac{\partial L(\mathbb{U})}{\partial d} &= \frac{1}{2} \operatorname{tr}\left\{\left(d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I}\right)^{-1} d^{-2}\mathbf{Q}^{-1}\right\} - \frac{1}{2} (\mathbf{x} - \mu\mathbf{1})^T \left(d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I}\right)^{-1} \\
&\qquad \times\ d^{-2}\mathbf{Q}^{-1} \left(d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
&= \frac{1}{2d} \operatorname{tr}\left\{(\mathbf{I} + \sigma^2 d\mathbf{Q})^{-1}\right\} - \frac{1}{2} (\mathbf{x} - \mu\mathbf{1})^T \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} \mathbf{Q} \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
\frac{\partial L(\mathbb{U})}{\partial q_i} &= \frac{1}{2} \operatorname{tr}\left\{\left(d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I}\right)^{-1} d^{-1}\mathbf{Q}^{-1}(\partial_{q_i}\mathbf{Q})\mathbf{Q}^{-1}\right\} - \frac{1}{2} (\mathbf{x} - \mu\mathbf{1})^T \left(d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I}\right)^{-1} \\
&\qquad \times\ d^{-1}\mathbf{Q}^{-1}(\partial_{q_i}\mathbf{Q})\mathbf{Q}^{-1} \left(d^{-1}\mathbf{Q}^{-1} + \sigma^2\mathbf{I}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
&= \frac{1}{2} \operatorname{tr}\left\{(\mathbf{I} + \sigma^2 d\mathbf{Q})^{-1}(\partial_{q_i}\mathbf{Q})\mathbf{Q}^{-1}\right\} \\
&\quad - \frac{d}{2} (\mathbf{x} - \mu\mathbf{1})^T \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\partial_{q_i}\mathbf{Q}) \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}),
\end{aligned}$$

where we used $\partial \ln(\det\{\mathbf{X}\}) = \mathrm{tr}\{\mathbf{X}^{-1}\partial\mathbf{X}\}$, $\partial(\mathbf{X}^{-1}) = -\mathbf{X}^{-1}(\partial\mathbf{X})\mathbf{X}^{-1}$ and $(\mathbf{X}^{-1} + \mathbf{Y}^{-1})^{-1} = \mathbf{X}(\mathbf{X}+\mathbf{Y})^{-1}\mathbf{Y}$. Similarly, the second-order derivatives can be computed. They are given by

$$
\begin{aligned}
\frac{\partial^2 L(\mathbb{U})}{\partial \mu^2} &= -\mathbf{1}^T d\mathbf{Q} \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} \mathbf{1} \\
\frac{\partial^2 L(\mathbb{U})}{\partial d^2} &= -\frac{1}{2d^2}\,\mathrm{tr}\left\{(\mathbf{I} + \sigma^2 d\mathbf{Q})^{-1}\right\} - \frac{\sigma^2}{2d}\,\mathrm{tr}\left\{(\mathbf{I} + \sigma^2 d\mathbf{Q})^{-2}\mathbf{Q}\right\} \\
&\quad + \sigma^2 (\mathbf{x} - \mu\mathbf{1})^T \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} \mathbf{Q} \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} \mathbf{Q} \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
\frac{\partial^2 L(\mathbb{U})}{\partial q_i^2} &= -\frac{\sigma^2 d}{2}\,\mathrm{tr}\left\{(\mathbf{I} + \sigma^2 d\mathbf{Q})^{-2}(\partial_{q_i}\mathbf{Q})^2\mathbf{Q}^{-1}\right\} - \frac{1}{2}\,\mathrm{tr}\left\{(\mathbf{I} + \sigma^2 d\mathbf{Q})^{-1}\mathbf{Q}^{-2}(\partial_{q_i}\mathbf{Q})^{-2}\right\} \\
&\quad + \sigma^2 d^2 (\mathbf{x} - \mu\mathbf{1})^T \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\partial_{q_i}\mathbf{Q}) \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\partial_{q_i}\mathbf{Q}) \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
\frac{\partial^2 L(\mathbb{U})}{\partial \mu \partial d} &= \mathbf{1}^T \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} \mathbf{Q} \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
\frac{\partial^2 L(\mathbb{U})}{\partial \mu \partial q_i} &= d\mathbf{1}^T \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\partial_{q_i}\mathbf{Q}) \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
\frac{\partial^2 L(\mathbb{U})}{\partial d \partial q_i} &= \frac{\sigma^2}{2}\,\mathrm{tr}\left\{\left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-2} (\partial_{q_i}\mathbf{Q})\right\} \\
&\quad + \sigma^2 d\,(\mathbf{x} - \mu\mathbf{1})^T \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\partial_{q_i}\mathbf{Q}) \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} \mathbf{Q} \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
&\quad - \frac{1}{2}\,(\mathbf{x} - \mu\mathbf{1})^T \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\partial_{q_i}\mathbf{Q}) \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1}) \\
\frac{\partial^2 L(\mathbb{U})}{\partial q_1 \partial q_2} &= -\frac{\sigma^2 d}{2}\,\mathrm{tr}\left\{(\mathbf{I} + \sigma^2 d\mathbf{Q})^{-2}(\partial_{q_1}\mathbf{Q})(\partial_{q_2}\mathbf{Q})\mathbf{Q}^{-1}\right\} - \frac{1}{2}\,\mathrm{tr}\left\{(\mathbf{I} + \sigma^2 d\mathbf{Q})^{-1}\mathbf{Q}^{-2}(\partial_{q_1}\mathbf{Q})(\partial_{q_2}\mathbf{Q})\right\} \\
&\quad + \sigma^2 d^2 (\mathbf{x} - \mu\mathbf{1})^T \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\partial_{q_1}\mathbf{Q}) \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\partial_{q_2}\mathbf{Q}) \left(\mathbf{I} + \sigma^2 d\mathbf{Q}\right)^{-1} (\mathbf{x} - \mu\mathbf{1})\,.
\end{aligned}
$$

C.3.6 Circulant Matrices

In this section, we will briefly review the basic properties of circulant matrices. For a more detailed treatment, the interested reader should consult [Gray, 2006; Moon and Stirling, 2000].

A circulant matrix $\mathbf{C} \in \mathbb{C}^{L\times L}$ has the form

$$
\mathbf{C} = \begin{bmatrix}
c_0 & c_{L-1} & \cdots & c_2 & c_1 \\
c_1 & c_0 & c_{L-1} & & c_2 \\
\vdots & c_1 & c_0 & \ddots & \vdots \\
c_{L-2} & & \ddots & \ddots & c_{L-1} \\
c_{L-1} & c_{L-2} & \cdots & c_1 & c_0
\end{bmatrix}, \tag{C.35}
$$

i.e. each column is a cyclically downshifted version of its predecessor. To find the eigenvectors and eigenvalues of $\mathbf{C}$, we have to consider the equation $\mathbf{C}\mathbf{v} = \lambda\mathbf{v}$ where λ is the eigenvalue that corresponds to the eigenvector $\mathbf{v} = \begin{bmatrix} v_0 & v_1 & \dots & v_{L-1} \end{bmatrix}^T$. More specifically, we have

$$\mathbf{C}\mathbf{v} = \mathbf{c} \circledast \mathbf{v} = \lambda\mathbf{v}, \tag{C.36}$$

where $\mathbf{c}$ denotes the first column of $\mathbf{C}$ and $\circledast$ denotes the circular convolution. For (C.36), we used the fact that the matrix-vector multiplication $\mathbf{C}\mathbf{v}$ corresponds to a circular convolution of $\mathbf{c}$ with $\mathbf{v}$. Introducing the *discrete Fourier transform* (DFT) operator $\mathrm{DFT}\{.\}$, (C.36) can be rewritten as

$$\mathrm{DFT}\{\mathbf{c}\} \circ \mathrm{DFT}\{\mathbf{v}\} = \lambda\,\mathrm{DFT}\{\mathbf{v}\} \quad \Leftrightarrow \quad \big(\mathrm{DFT}\{\mathbf{c}\} - \lambda\mathbf{1}\big) \circ \mathrm{DFT}\{\mathbf{v}\} = \mathbf{0}, \tag{C.37}$$

where $\circ$ denotes the elementwise (Hadamard) product. Using the special choice

$$\mathbf{v}_l = \begin{bmatrix} 1 & e^{j\frac{2\pi}{L}l} & \dots & e^{j\frac{2\pi}{L}(L-1)l} \end{bmatrix}^T,$$

we obtain $\mathrm{DFT}\{\mathbf{v}_l\} = L\mathbf{e}_l$ and therefore the corresponding eigenvalue λ has to be the lth element of $\mathrm{DFT}\{\mathbf{c}\}$. Thus, each circulant matrix $\mathbf{C}$ can be decomposed as $\mathbf{C} = \mathbf{U}\mathbf{D}\mathbf{U}^{-1}$ where $\mathbf{D}$ is a diagonal matrix which contains the DFT values of the first column of $\mathbf{C}$ on its diagonal and $\mathbf{U} \in \mathbb{C}^{L\times L}$ is the well known DFT matrix with $[\mathbf{U}]_{l,k} = e^{j\frac{2\pi}{L}(l-1)(k-1)}$ and $\mathbf{U}^{-1} = \frac{1}{L}\mathbf{U}^H$.

Using this eigenvalue decomposition, the following nice properties can be proved (see e.g. [Gray, 2006]):

1. Let $\mathbf{C}_1$ and $\mathbf{C}_2$ be two circulant matrices then the matrix-product commutes, i.e. $\mathbf{C}_1\mathbf{C}_2 = \mathbf{C}_2\mathbf{C}_1$. Furthermore, $\mathbf{C}_1\mathbf{C}_2$ is again circulant.

2. Let $\mathbf{C}_1$ and $\mathbf{C}_2$ be two circulant matrices then $\mathbf{C}_1 + \mathbf{C}_2$ is again circulant.

3. If all eigenvalues $\boldsymbol{\lambda} = \mathrm{DFT}\{\mathbf{c}\}$ of a circulant matrix $\mathbf{C}$ are unequal to zero then the inverse exists and is given by $\mathbf{U}\mathbf{D}^{-1}\mathbf{U}^{-1}$, i.e. it is again circulant.

These properties allow us to do the basic matrix operations that we need in Section 4.3.2 in “FFT speed”.

C.4 Details, Derivations and Proofs for Chapter 5

C.4.1 Motzkin's Example

A famous example of a polynomial that is nonnegative but which cannot be represented as a sum of squares is the Motzkin polynomial [Reznick, 2000]

$$f(\mathbf{z}) = z_1^4 z_2^2 + z_1^2 z_2^4 - 3z_1^2 z_2^2 + 1. \tag{C.38}$$

We will first establish that it is nonnegative: Using the inequality of the arithmetic and geometric mean (*AM-GM inequality*)

$$\frac{a+b+c}{3} \geq \sqrt[3]{abc}$$

with $a = 1$, $b = z_1^4 z_2^2$ and $c = z_1^2 z_2^4$ proves this fact. To show that it cannot be written as a sum of squares, we use the general Ansatz $f(\mathbf{z}) = \sum_i (f_i(\mathbf{z}))^2$ with

$$f_i(\mathbf{z}) = c_{1,i} z_1^2 z_2 + c_{2,i} z_1 z_2^2 + c_{3,i} z_1 z_2 + c_{4,i} z_1^2 + c_{5,i} z_2^2 + c_{6,i} z_1 + c_{7,i} z_2 + c_{8,i}. \tag{C.39}$$

Comparing (C.38) with (C.39), we can immediately conclude that $c_{4,i} = c_{5,i} = c_{6,i} = c_{7,i} = 0$ and thus $f_i(\mathbf{z})$ from (C.39) reduces to

$$f_i(\mathbf{z}) = c_{1,i} z_1^2 z_2 + c_{2,i} z_1 z_2^2 + c_{3,i} z_1 z_2 + c_{8,i}.$$

In order to obtain the monomial $-3z_1^2 z_2^2$ in (C.38) we see that $\sum_i c_{3,i}^2 = -3$ has to hold which cannot be fulfilled for $c_{3,i} \in \mathbb{R}$ and therefore we have proved that the Motzkin polynomial (C.38) is not SOS.

C.4.2 Equivalence between SOS Polynomials and Quadratic Forms

In this section, we will show that an equivalent representation of $f(\mathbf{z}) = \sum_i (f_i(\mathbf{z}))^2$ of degree $2D$ with $\mathbf{z} \in \mathbb{R}^{2M}$ is given by $f(\mathbf{z}) = \mathbf{m}(\mathbf{z})^T \mathbf{Q} \mathbf{m}(\mathbf{z})$ with $\mathbf{Q} \succeq \mathbf{0}$ where $\mathbf{m}(\mathbf{z}) \in \mathbb{R}^{\binom{D+2M}{D}}$ is a column vector containing all monomials up to order D, i.e. the elements have the form $\mathbf{z}^{\boldsymbol{\alpha}} = \prod_{i=1}^{2M} z_i^{\alpha_i}$ where $\alpha_1 + \ldots + \alpha_{2M} \leq D$.

First, we will show that each sum of squares polynomial can be written as $f(\mathbf{z}) = \mathbf{m}(\mathbf{z})^T \mathbf{Q} \mathbf{m}(\mathbf{z})$ with $\mathbf{Q} \succeq \mathbf{0}$ which can be easily established. It simply stems from the fact that $\mathbf{m}(\mathbf{z})$ contains all monomials of degree less or equal to D and therefore we can represent each $f_i(\mathbf{z})$ as $f_i(\mathbf{z}) = \mathbf{t}_i^T \mathbf{m}(\mathbf{z})$ which implies $f(\mathbf{z}) = \mathbf{m}(\mathbf{z})^T \mathbf{T}^T \mathbf{T} \mathbf{m}(\mathbf{z})$ where the ith row of $\mathbf{T}$ is given by $\mathbf{t}_i^T$ and $\mathbf{T}^T \mathbf{T} \succeq \mathbf{0}$.

To prove the converse, i.e. that each polynomial $f(\mathbf{z}) = \mathbf{m}(\mathbf{z})^T \mathbf{Q} \mathbf{m}(\mathbf{z})$ with $\mathbf{Q} \succeq \mathbf{0}$ is a sum of squares polynomial, we have to use the *Cholesky decomposition* of $\mathbf{Q}$ [Golub and Van Loan,

1996]. For a positive semi-definite matrix $\mathbf{Q}$, the Cholesky factorization is given by $\mathbf{Q} = \mathbf{L}^T\mathbf{D}\mathbf{L}$ where $\mathbf{L}$ is a lower triangular matrix and $\mathbf{D}$ is diagonal with elements larger or equal to zero. Therefore, introducing $\mathbf{T} = \mathbf{D}^{1/2}\mathbf{L}$ we see that $f(\mathbf{z})$ is a SOS polynomial.

C.4.3 A Brief Introduction to Polynomial Optimization

Polynomial optimization deals with finding a good approximation of the global minimum or maximum of a multivariate polynomial $f_0(\mathbf{z})$ with polynomial equality and/or inequality constraints [Parrilo, 2003; Lasserre, 2010]. In general, the following two families of problems can be distinguished:

- *Unconstrained problems* which have the form

$$\min_{\mathbf{z}\in\mathbb{R}^n} \; f_0(\mathbf{z}) \tag{C.40}$$

  where $f_0(\mathbf{z})$ is a multivariate polynomial in $\mathbf{z}$, i.e. $f_0 \in \mathbb{R}[\mathbf{z}]$.

- *Constrained problems* which are given by

$$\min_{\mathbf{z}\in\text{FS}} \; f_0(\mathbf{z}) \tag{C.41a}$$

  where the feasible set FS is given by

$$\begin{aligned} a_i(\mathbf{z}) &\neq 0, \; i = 1, \ldots, s \\ b_j(\mathbf{z}) &\geq 0, \; j = 1, \ldots, t \\ c_k(\mathbf{z}) &= 0, \; k = 1, \ldots, u \end{aligned} \tag{C.41b}$$

  and $f_0, a_i, b_j, c_k \in \mathbb{R}[\mathbf{z}]$. The feasible set is therefore the set of solutions fulfilling polynomial equality/inequality constraints.[39]

Both optimization problems (C.40) and (C.41) are in general nonconvex and therefore difficult to solve. Using however the SOS relaxation of a nonnegative polynomial, we can come up with efficient SDP formulations to obtain approximate solutions to (C.40) and (C.41).

Nonnegative Polynomials and Sum of Squares Approximation

As we will see later, nonnegative polynomials play a central role in polynomial optimization. A polynomial $f(\mathbf{z})$ is called *nonnegative* if $f(\mathbf{z}) \geq 0$ for all $\mathbf{z} \in \mathbb{R}^n$ holds. Furthermore, a polynomial $f(\mathbf{z})$ is called *sum of squares* (SOS) if $f(\mathbf{z})$ can be written as $f(\mathbf{z}) = \sum_i (f_i(\mathbf{z}))^2$.

[39] Such a set is called a *semialgebraic set*.

$$
\begin{array}{ccc}
f(\mathbf{z}) \geq 0 \quad \forall\, \mathbf{z} \in \mathbb{R}^n & \not\Leftrightarrow & f(\mathbf{z}) = \textstyle\sum_i f_i^2(\mathbf{z}) \\
\Updownarrow & & \Updownarrow \\
f(\mathbf{x}) = \frac{p(\mathbf{z})}{q(\mathbf{z})}, \quad p, q \text{ SOS} & & f(\mathbf{z}) = \mathbf{m}(\mathbf{z})^T \mathbf{Q} \mathbf{m}(\mathbf{z}), \quad \mathbf{Q} \succeq \mathbf{0} \\
\textit{(Hilbert's 17th problem)} & & \textit{(SDP formulation)}
\end{array}
$$

Figure C.2: Relationship between nonnegative polynomials and SOS

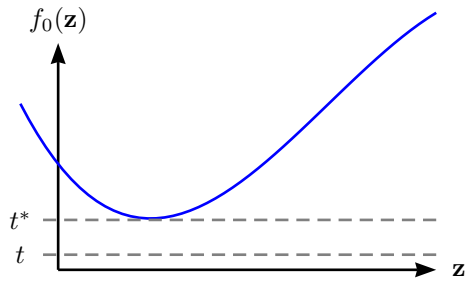


Figure C.3: Visual representation of the optimization problem (C.42)

It is obvious that if $f(\mathbf{z})$ is SOS then it is also nonnegative. The converse, however, does not hold in general as was proven by Hilbert in 1888 [Reznick, 2000].[40] The diagram in Figure C.2 summarizes the relationship between nonnegativity and SOS where $\mathbf{m}(\mathbf{z})$ is the monomial vector as introduced in Section 5.3.

SOS Relaxation of Polynomial Optimization Problems

The optimization problem (C.40) can be rewritten in epigraphic form as

$$\max_{t} \; t \qquad \text{s.t.} \quad f_0(\mathbf{z}) - t \geq 0 \quad \forall\, \mathbf{z} \in \mathbb{R}^n \tag{C.42}$$

i.e. t is increased such that the polynomial $f_0(\mathbf{z}) - t$ is nonnegative for all $\mathbf{z} \in \mathbb{R}^n$ as is shown in Figure C.3. Relaxing the constraint $f_0(\mathbf{z}) - t \geq 0$ in (C.42) with the constraint $f_0(\mathbf{z}) - t$ is SOS, we obtain the SDP

$$\begin{aligned} \max_{t, \mathbf{Q}_0} \; t \qquad \text{s.t.} \quad & \mathbf{Q}_0 - t\mathbf{J}^{1,1} \succeq 0 \\ & \mathbf{m}(\mathbf{z})^T \mathbf{Q}_0 \mathbf{m}(\mathbf{z}) = f_0(\mathbf{z}) \end{aligned} \tag{C.43}$$

[40] Exceptions are given by the cases $\{n = 1, d \text{ arbitrary}\}$ or $\{n \text{ arbitrary}, d = 2\}$ or $\{n = 2, d = 4\}$ where d denotes the degree of the polynomial.

which can be solved efficiently. $\mathbf{J}^{1,1}$ denotes the single-entry matrix which is zero everwhere except a one at $(1,1)$.[41] The optimal solution $\mathbf{z}^*$ can be extracted from

$$f_0(\mathbf{z}^*) = t^*_{\text{SDP}} \quad \Longleftrightarrow \quad \mathbf{m}(\mathbf{z}^*)^T \left(\mathbf{Q}_0 - t^*_{\text{SDP}} \mathbf{J}^{1,1} \right) \mathbf{m}(\mathbf{z}^*) = 0. \tag{C.44}$$

Further details about the extraction of solutions can e.g. be found in [Lasserre, 2010].

Solving constrained polynomial optimization problems as in (C.41) can be done in the same manner as for the unconstrained case. The only difference is that we have $f_0(\mathbf{z}) - t \geq 0$ for all $\mathbf{z} \in \text{FS}$, i.e. we have to ensure nonnegativity of $f_0(\mathbf{z}) - t$ on the feasible set. The Positivstellensatz as given in Section 5.5 can be used to handle such constraints and therefore again, we can give efficient SDP formulations which give good approximations to (C.41).

[41] We assume here that the monomials in $\mathbf{m}(\mathbf{z})$ are sorted in increasing order and therefore the monomial of order zero is at the first position.

List of Figures

List of Tables

Bibliography

J. S. Abel. A bound on mean-square-estimate error. *IEEE Transactions on Information Theory*, 39(5):1675–1680, 1993.

M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. U.S. Department of Commerce, National Bureau of Standards, 1964.

B. F. Alkire. *Convex Optimization Problems Involving Autocorrelation Sequences*. PhD thesis, University of California, 2003.

F. Athley. Threshold region performance of maximum likelihood direction of arrival estimators. *IEEE Transactions on Signal Processing*, 53(4):1359–1373, April 2005.

B. Babadi, N. Kalouptsidis, and V. Tarokh. Comparison of SPARLS and RLS algorithms for adaptive filtering. *Proceedings of the 2009 IEEE Sarnoff Symposium*, 2009.

Y. Bard. *Nonlinear Parameter Estimation*. Academic Press, 1973.

K. Bell, Y. Ephraim, and H. L. Van Trees. Comparison of the Chazan-Zakai-Ziv, Weiss-Weinstein, and Cramer-Rao bounds for bearing estimation. *Proc. of Conf. Info. Sci and Syst., Baltimore*, 1993.

K. L. Bell, Y. Steinberg, Y. Ephraim, and H. L. Van Trees. Extended Ziv-Zakai lower bound for vector parameter estimation. *IEEE Transactions on Information Theory*, 43(2):624–637, 1997.

S. Bellini. An upper bound on the performance of pulse modulation systems. *IEEE Transactions on Communications*, 26(9):1352–1355, 1978.

S. Bellini and G. Tartara. Bounds on error in signal parameter estimation. *IEEE Transactions on Communications*, 22(3):340–342, 1974.

Z. Ben-Haim and Y. C. Eldar. On the constrained Cramer-Rao bound with a singular fisher information matrix. *IEEE Signal Processing Letters*, 16(6):453–456, 2009.

A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, 2009.

A. Benavoli, L. Chisci, A. Farina, L. Ortenzi, and G. Zappa. Hard-constrained versus soft-constrained parameter estimation. *IEEE Transactions on Aerospace and Electronic Systems*, 42(4):1224–1239, 2006.

A. Benavoli, L. Chisci, and A. Farina. Estimation of constrained parameters with guaranteed MSE improvement. *IEEE Transactions on Signal Processing*, 55(4):1264–1274, 2007.

J. Benesty, S. Makino, and J. Chen. *Speech Enhancement*. Springer, 2005.

M. Berouti, R. Schwartz, and J. Makhoul. Enhancement of speech corrupted by acoustic noise. *Proc. IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 1979.

A. Bhattacharyya. On some analogues of the amount of information and their use in statistical estimation. *Sankhya Indian Journal of Statistics*, 8:1–14, 201–218, 315–328, 1946.

D. A. Binder. Bayesian cluster analysis. *Biometrika*, 65(1):31–38, 1978.

D. A. Binder. Approximations to Bayesian clustering rules. *Biometrika*, 68(1):275–285, 1981.

R. Blind, S. Uhlich, B. Yang, and F. Allgöwer. Robustification and optimization of a Kalman filter with measurement loss using linear precoding. *Proc. of the American Control Conference*, 2009.

B. Z. Bobrovsky and M. Zakai. A lower bound on the estimation error for Markov processes. *IEEE Transactions on Automatic Control*, 20(6):785–788, 1975.

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge Univ. Pr., 2007. ISBN 0-521-83378-7.

S. Boyd, L. E. Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. Society for Industrial and Applied Mathematics (SIAM), 1994. ISBN 0-89871-334-X.

L. D. Brown. Inadmissibility of the usual estimators of scale parameters in problems with unknown location and scale parameters. *Annals of Mathematical Statistics*, 39(1):29–48, 1968.

E. Candes and M. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, pages 21–30, 2008.

G. Celeux, M. Hurn, and C. P. Robert. Compuational and inferential difficulties with mixture posterior distributions. *Journal of the American Statistical Association*, 95(451):957–970, 2000.

S. S. Channappayya, A. C. Bovik, C. Caramanis, and R. W. Heath Jr. Design of linear equalizers optimized for the structural similarity index. *IEEE Transactions on Image Processing*, 17(6): 857–872, 2008.

R. Chellappa. Two-dimensional discrete Gaussian Markov Random Fields models for image processing. *Progress in Pattern Recognition 2*, 1985.

T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley, 2006.

H. Cramer. *Mathematical Methods of Statistics*. Princeton University Press, 1946.

R. Cranley and T. N. L. Patterson. Randomization of number theoretic methods for multiple integration. *SIAM Journal on Numerical Analysis*, 13(6):904–914, 1976.

G. Dahlquist and A. Björck. *Numerical Methods in Scientific Computing - Volume I*. Society for Industrial and Applied Mathematics, 2008.

J. Dattorro. *Convex Optimization and Euclidean Distance Geometry*. Meboo Publishing USA, 2011.

N. G. de Bruijn. *Asymptotic Methods in Analysis*. Amsterdam : North-Holland Publ., 1961.

A. Dempster, N. Laird, and D. Rubin. Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, 39(1):1–38, 1977.

A. Doucet and X. Wang. Monte Carlo methods for signal processing: A review in the statistical signal processing context. *IEEE Signal Processing Magazine*, 22(6):152–170, 2005.

M. Duarte, M. A. Davenport, D. Takhar, J. N. Laska, T. Sun, K. F. Kelly, and R. G. Baraniuk. Single-pixel imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2): 83–91, 2008.

L. El Ghaoui and V. Balakrishnan. Synthesis of fixed-structure controllers via numerical optimization. *Proc. IEEE Conf. on Decision and Control*, pages 2678–2683, 1994.

Y. C. Eldar. Uniformly improving the Cramer-Rao bound and maximum-likelihood estimation. *IEEE Transactions on Signal Processing*, 54(8):2943–2956, 2006a.

Y. C. Eldar. Comparing between estimation approaches: Admissible and dominating linear estimators. *IEEE Transactions on Signal Processing*, 54(5):1689– 1702, 2006b.

Y. C. Eldar. Universal weighted MSE improvement of the least-squares estimator. *IEEE Transactions on Signal Processing*, 56(5):1788–1800, 2008a.

Y. C. Eldar. MSE bounds with affine bias dominating the Cramer-Rao bound. *IEEE Transactions on Signal Processing*, 56(8):3824–3836, 2008b.

Y. C. Eldar, A. Ben-Tal, and A. Nemirovski. Robust mean-squared error estimation in the presence of model uncertainties. *IEEE Transactions on Signal Processing*, 53(1):168–181, 2005.

Y. C. Eldar. *Rethinking biased estimation: Improving Maximum Likelihood and the Cramer-Rao Bound*. Now, Boston [u.a.], 2008c. ISBN 1-60198-130-1.

G. Elliott, I. Komunjer, and A. Timmermann. Estimation and testing of forecast rationality und flexible loss. *Review of Economic Studies*, 72(4):1107–1125, 2005.

A. Elron, G. Leibovitz, Z. Ben-Haim, and Y. C. Eldar. Recursive blind minimax estimation: Improving MSE over recursive least squares. *IEEE 25th Convention of Electrical and Electronics Engineers in Israel (IEEEI)*, pages 548–552, 2008.

Y. Ephraim and D. Malah. Speech enhancement using a minimum mean square error short-time spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 32:1109–1121, 1984.

Y. Ephraim and D. Malah. Speech enhancement using a minimum mean square error log-spectral amplitude estimator. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 33: 443–445, 1985.

M. Evans and T. Swartz. Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems. *Statistical Science*, 10(3):254–272, 1995.

J. A. Farrell and M. M. Polycarpou. *Adaptive Approximation Based Control*. Wiley, 2006.

R. A. Fisher. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A*, 222:309–368, 1922.

P. Forster and P. Larzabal. On lower bounds for deterministic parameter estimation. *Proc. IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2:1137–1140, 2002.

A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin. *Bayesian Data Analysis*. Chapman & Hall, 2003.

J. E. Gentle. *Matrix Algebra - Theory, Computations, and Applications in Statistics*. Springer, 2007.

J. K. Ghosh, M. Delampady, and T. Samanta. *An Introduction to Bayesian Analysis: Theory and Methods*. Springer, 2006.

R. A. Gideon and J. Gurland. Series expansions for quadratic forms in normal variables. *Journal of the American Statistical Association*, 71(353):227–232, 1976.

G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins Press, Baltimore, 3rd edition, 1996.

R. C. Gonzalez and R. E. Woods. *Digital Image Processing*. Prentice-Hall, 2002.

J. D. Gorman and A. O. Hero. Lower bounds for parametric estimation with constraints. *IEEE Transactions on Information Theory*, 36(6):1285–1301, 1990.

R. M. Gray. *Toeplitz and Circulant Matrices: A Review*. Foundations and Trends in Communications and Information Theory. Now Publishers, 2006.

X. Gu and M. Wong. A modified ziv-zakai lower bound and its application in array processing. *Proc. IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 1477–1480, 1991.

J. Halton. On the efficiency of certain quasi-random sequences of points in evaluating multi-dimensional integrals. *Numerische Mathematik*, 2:84–90, 1960.

S. Haykin. *Adaptive Filter Theory*. Prentice-Hall, 4th edition, 2002.

H. Heyer. *Theory of Statistical Experiments*. Springer, 1982.

A. Hjørungnes. *Complex-valued matrix derivatives : with applications in signal processing and communications*. Univ. Press, Cambridge, 2011. ISBN 978-0-521-19264-4.

Y. Hu and P. C. Loizou. Evaluation of objective quality measures for speech enhancement. *IEEE Transactions on Audio, Speech and Language Processing*, 16(1):229–238, 2008.

M. Hurn, A. Justel, and C. P. Robert. Estimating mixtures of regressions. *Journal of Computational and Graphical Statistics*, 12(1):1, 25 2003.

ITU-T Recommendation P.862. Perceptual Evaluation of Speech Quality (PESQ): An objective method for end-to-end speech quality assessment of narrow-band telephone networks and speech codecs. (02/2001).

Z. Jarvis-Wloszek, R. Feeley, W. Tan, K. Sun, and A. Packard. Control applications of sum of squares programming. *Lecture Notes in Control and Information Sciences*, pages 3–22, 2005.

G. M. Jenkins and D. G. Watts. *Spectral Analysis and its Applications*. Emerson-Adams Press, 1969.

N. L. Johnson and S. Kotz. *Distributions in Statistics: Continuous Univariate Distributions 2*. John Wiley & Sons, 1970.

U. T. Jönsson. A lecture on the S-procedure. *Royal Institute of Technology (KTH), Lecture Notes*, 2006.

A. Kaminska and Z. Porosinski. On robust Bayesian estimation under some asymmetric and bounded loss function. *Statistics*, 43(3):253–265, 2009.

S. M. Kay. *Fundamentals of Statistical Signal Processing, Volume 1: Estimation Theory*. Prentice-Hall, 1993.

Y. N. Kiseliov. Algorithms of projection of a point onto an ellipsoid. *Lithuanian Mathematical Journal*, 34(2):141–159, 1994.

L. Knockaert. The Barankin bound and threshold behavior in frequency estimation. *IEEE Transactions on Signal Processing*, 45(9):2398–2401, 1997.

A. N. Kolmogorov. Definition of the center of scattering and the measure of accuracy by a limited number of observations. *Izvestia Akademii Nauk SSSR*, 6:3–32, 1942.

S. Kotz, N. L. Johnson, and D. W. Boyd. Series representations of distributions of quadratic forms in normal variables. I. central case. *Ann. of Math. Statist.*, 38(3):823–837, 1967a.

S. Kotz, N. L. Johnson, and D. W. Boyd. Series representations of distributions of quadratic forms in normal variables II. Non-central case. *Ann. of Math. Statist.*, 38(3):838–848, 1967b.

J.-B. Lasserre. A sum of squares approximation of nonnegative polynomials. *SIAM Review*, 49 (4):651–669, 2007.

J.-B. Lasserre. *Moments, Positive Polynomials and their Applications*. Imperial College Press, London, 2010. ISBN 978-1-84816-445-1.

J. W. Lau and P. J. Green. Bayesian model based clustering procedures. *Journal of Computational and Graphical Statistics*, 16:526–558, 2007.

O. Lee, J. Min Kim, Y. Bresler, and J. Chul Ye. Compressive diffuse optical tomography: Noniterative exact reconstruction using joint sparsity. *IEEE Transactions on Medical Imaging*, 30 (5):1129–1142, 2011.

E. Lehmann. Image-source method: Matlab implementation. Available: http://www.eric-lehmann.com/, 2008.

E. Lehmann and A. Johansson. Prediction of energy decay in room impulse responses simulated with an image-source model. *The Journal of the Acoustical Society of America*, 124(1):269–277, 2008.

E. L. Lehmann and G. Casella. *Theory of Point Estimation*. Springer, 1998.

R. D. Levine and M. Tribus. *Maximum Entropy Formalism*. MIT Press, 1978.

Y. Lin and D. D. Lee. Bayesian regularization and nonnegative deconvolution for room impulse response estimation. *IEEE Transactions on Signal Processing*, 54(3):839–847, 2006.

J. S. Liu. *Monte Carlo Strategies in Scientific Computing*. Springer, 2008.

J. Löfberg. YALMIP: A toolbox for modeling and optimization in MATLAB. In *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

P. Loizou. Speech enhancement algorithms: A survey. *in Recent Advances in Robust Speech Recognition Technology, (J. Ramírez and J. M. Górriz, eds.)*, pages 60–102, 2011.

P. C. Loizou. Speech enhancement based on perceptually motivated Bayesian estimators of the magnitude spectrum. *IEEE Transactions on Speech and Audio Processing*, 13(5):857–869, 2005.

J. Lu. *Multivariate Slice Sampling*. PhD thesis, Drexel University, 2008.

Z.-Q. Luo and W. Yu. An introduction to convex optimization for communications and signal processing. *IEEE Journal on Selected Topics in Communications*, 24(8):1426–1438, 2006.

M. Lustig, D. L. Donoho, J. M. Santos, and J. M. Pauly. Compressed sensing MRI. *IEEE Signal Processing Magazine*, 25(2):72–82, 2008.

J. R. Magnus and H. Neudecker. *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Wiley, third edition, 2007.

A. I. Mcleod and B. Quenneville. Mean likelihood estimators. *Statistics and Computing*, 11(1): 57–65, 2001.

J. M. Mendel. *Lessons in Estimation Theory for Signal Processing, Communications, and Control*. Prentice-Hall, 2 edition, 1995.

T. P. Minka. Old and new matrix algebra useful for statistics, 2000.

S. Möller, W.-Y. Chan, N. Cote, T. H. Falk, A. Raake, and M. Waltermann. Speech quality estimation: Models and trends. *IEEE Signal Processing Magazine*, 28(6):18–28, 2011.

T. K. Moon and W. C. Stirling. *Mathematical Methods and Algorithms for Signal Processing*. Prentice-Hall, 2000.

R. Mukai, H. Sawada, S. Araki, and S. Makino. Real-time blind source separation and DOA estimation using small 3-D microphone array. *Proc. International Workshop for Acoustic Echo and Noise Control (IWAENC)*, pages 45–48, 2005.

J. F. Muth. Rational expectations and the theory of price movements. *Econometrica*, 29(3): 315–335, 1961.

National Institute of Standards and Technology. Digital Library of Mathematical Functions, available online: http://dlmf.nist.gov/, 2012.

J. Neyman. Sur un teorema concernente le cosidette statistiche sufficienti. *Giorn. Ist. Ital. Att.*, 6:320–334, 1935.

H. Nguyen and H. L. Van Trees. Comparison of performance bounds for DOA estimation. *IEEE Workshop on Statistical Signal and Array Processing*, pages 313–316, 1994.

J. G. Norstrom. The use of precautionary loss functions in risk analysis. *IEEE Trans. on Reliability*, 45(3):400–403, 1996.

A. Papoulis and S. U. Pillai. *Probability, Random Variables and Stochastic Processes*. McGraw-Hill, 2002.

P. A. Parrilo. Semidefinite programming relaxations for semialgebraic problems. *Mathematical Programming Ser. B*, 96(2):293–320, 2003.

J. Pilz. Minimax linear regression estimation with symmetric parameter restrictions. *Journal of Statistical Planning and Inference*, 13:297–318, 1986.

M. Pincus. A closed form solution of certain programming problems. *Operations Research*, 16 (3):690–694, 1968.

M. S. Pinsker. Optimal filtration of square-integrable signals in Gaussian noise. *Problems of Information Transmission*, 16:120–133, 1980.

E. Plourde and B. Champagne. Auditory-based spectral amplitude estimators for speech enhancement. *IEEE Transactions on Audio, Speech and Language Processing*, 16(8):1614–1623, 2008.

E. Plourde and B. Champagne. Generalized Bayesian estimators of the spectral amplitude for speech enhancement. *IEEE Signal Processing Letters*, 2009.

J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12 (11):1338–1351, 2003.

F. Pukelsheim. *Optimal Design of Experiments*. Classics in Applied Mathematics. Society for Industrial and Applied Mathematics, 2006.

C. R. Rao. Information and the accuracy attainable in the estimation of statistical parameters. *Bulletin of the Calcutta Mathematical Society*, 37:81–89, 1945.

I. Reuven and H. Messer. A Barankin-type lower bound on the estimation error of a hybrid parameter vector. *IEEE Transactions on Information Theory*, 43(3):1084–1093, 1997.

B. Reznick. Some concrete aspects of Hilbert's 17th problem. *Contemporary Mathematics*, 253: 251–272, 2000.

C. P. Robert. Simulation of truncated normal variables. *Statistics and Computing*, 5(2):121–125, 1995.

C. P. Robert. *The Bayesian Choice: From Decision-Theoretic Foundations to Computational Implementation*. Springer, 2001.

T. Routtenberg and J. Tabrikian. General classes of Bayesian lower bounds for outage error probability and MSE. *Proc. Workshop on Statistical Signal Processing*, pages 69–72, 2009.

H. Ruben. Probability content of regions under spherical normal distributions, IV: The distribution of homogeneous and non-homogeneous quadratic functions of normal variables. *Ann. of Math. Statist.*, 33(2):542–570, 1962.

H. Rue. New loss functions in Bayesian imaging. *Journal of the American Statistical Association*, 90(431):900–908, 1995.

H. Rue and L. Held. *Gaussian Markov Random Fields*. Chapman & Hall/CRC, 2005.

F. J. Samaniego. *A Comparison of the Bayesian and Frequentist Approaches to Estimation*. Springer, 2010.

L. L. Scharf. *Statistical Signal Processing: Detection, Estimation and Time Series Analysis*. Addison-Wesley, 1990.

P. J. Schreier and L. L. Scharf. *Statistical signal processing of complex-valued data: The theory of improper and noncircular signals*. Univ. Press, Cambridge, 2010. ISBN 978-0-521-89772-3.

J. E. H. Shaw, A. Genz, J. Monahan, M. J. Schervish, L. Wasserman, and R. Wolfinger. Comments on: "Methods for approximating integrals in statistics with special emphasis on Bayesian integration problems" by M. Evans and T. Swartz. *Statistical Science*, 11(1):54–64, 1996.

H. Sheikh, M. Sabir, and A. Bovik. A statistical evaluation of recent full reference image quality assessment algorithms. *IEEE Transactions on Image Processing*, 15(11):3440–3451, 2006.

E. P. Simoncelli, W. T. Freeman, E. H. Adelson, and D. J. Heeger. Shiftable multiscale transforms. *IEEE Transactions on Information Theory*, 38(2):587–607, 1992.

I. M. Sobol. On the systematic search in a hypercube. *SIAM Journal on Numerical Analysis*, 16 (5):790–793, 1979.

J. C. Spall. Estimation via Markov chain Monte Carlo. *IEEE Control Systems Magazine*, 23(2): 34– 45, 2003.

G.-B. Stan, J.-J. Embrechts, and D. Archambeau. Comparison of different impulse response measurement techniques. *Journal of the Audio Engineering Society*, 50(4):249–262, 2002.

C. Stein. Inadmissibility of the usual estimator for the variance of a normal distribution with unknown mean. *Annals of the Institute of Statistical Mathematics*, 16(1):155–160, 1964.

G. Stengle. A Nullstellensatz and a Positivstellensatz in semialgebraic geometry. *Mathematische Annalen*, 207(2):87–97, 1974.

P. Stoica and G. Ganesan. Linear regression constrained to a ball. *Digital Signal Processing*, 11 (1):80–90, 2000.

R. D. Thompson and A. P. Basu. Asymmetric loss functions for estimating system reliability. *In D.A. Berry, K.M. Chaloner, and J.K. Geweke, editors, Bayesian Analysis in Statistics and Econometrics*, pages 471–482, 1996.

R. Tibshirani. Regression shrinkage and selection via the lasso. *J. Royal. Statist. Soc., Series B.*, 58(1):267–288, 1996.

L. Tierney and J. B. Kadane. Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, 81(393):82–86, 1986.

K. Todros and J. Tabrikian. General classes of performance lower bounds for parameter estimation – Part I: Non-Bayesian bounds for unbiased estimators. *IEEE Transactions on Information Theory*, 56(10):5045–5063, 2010a.

K. Todros and J. Tabrikian. General classes of performance lower bounds for parameter estimation – Part II: Bayesian bounds. *IEEE Transactions on Information Theory*, 56(10):5064–5082, 2010b.

K. C. Toh, M. J. Todd, and R. H. Tutuncu. SDPT3 – a MATLAB software for semidefinite-quadratic-linear programming. Available: http://www.math.nus.edu.sg/ mattohkc/sdpt3.html.

S. Uhlich and B. Yang. A generalized optimal correlating transform for multiple description coding and its theoretical analysis. *Proc. IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2008.

S. Uhlich and B. Yang. MMSE estimation in a linear signal model with ellipsoidal constraints. *Proc. IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 3249–3252, 2009.

S. Uhlich and B. Yang. A parametric family of Bayesian estimators for non-standard loss functions. *Proc. European Signal Processing Conference (EUSIPCO)*, 2010a.

S. Uhlich and B. Yang. Efficient recursive estimators for a linear, time-varying Gaussian model with general constraints. *IEEE Transactions on Signal Processing*, 58(9):4910–4915, 2010b.

S. Uhlich and B. Yang. Recursive estimation of room impulse responses with energy conservation constraints. *Proc. IEEE Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pages 3648–3651, 2011.

S. Uhlich and B. Yang. Bayesian estimation for non-standard loss functions using a parametric family of estimators. *IEEE Transactions on Signal Processing*, 60(3):1022–1031, 2012.

S. Uhlich, B. Loesch, and B. Yang. Polynomial LMMSE estimation: A case study. *Proc. Workshop on Statistical Signal Processing*, 2009.

M. Unser. Splines: A perfect fit for signal and image processing. *IEEE Signal Processing Magazine*, 16(6):22–38, 1999.

H. L. Van Trees. *Detection, Estimation, and Modulation Theory*. Wiley, 1968.

H. L. Van Trees and K. L. Bell, editors. *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Wiley-Interscience, 2007. ISBN 978-0-470-12095-8.

H. L. Van Trees. *Optimum Array Processing*. Part IV of Detection, Estimation and Modulation Theory. Wiley, 2002.

J. G. VanAntwerp and R. D. Braatz. A tutorial on linear and bilinear matrix inequalities. *Journal of Process Control*, 10(4):363–385, 2000.

H. R. Varian. A Bayesian approach to real estate assessment. In S. E. Fienberg and A. Zellner, editors, *Studies in Bayesian Econometrics and Statistics in Honor of Leonard J. Savage*, pages 195–208. North Holland Press, 1974.

Z. Wang and A. C. Bovik. Mean squared error: Love it or leave it? A new look at signal fidelity measures. *IEEE Signal Processing Magazine*, 26(1):98–117, 2009.

Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 2004.

E. Weinstein and A. J. Weiss. A general class of lower bounds in parameter estimation. *IEEE Transactions on Information Theory*, 34(2):338–342, 1988.

A. Weiss and E. Weinstein. A lower bound on the mean-square error in random parameter estimation. *IEEE Transactions on Information Theory*, 31(5):680–682, 1985.

D. Wen and M. S. Levy. BLINEX: A bounded asymmetric loss function with application to Bayesian estimation. *Communications in Statistics - Theory and Methods*, 30(1):147–153, 2001.

C. H. You, S. N. Koh, and S. Rahardja. Beta-order MMSE spectral amplitude estimation for speech enhancement. *IEEE Transactions on Speech and Audio Processing*, 13(4):475–486, 2005.

A. Zellner. Bayesian estimation and prediction using asymmetric loss functions. *Journal of the American Statistical Association*, 81(394):446–451, 1986.

Y. Zhu and X. R. Li. Recursive least squares with linear constraints. *Communications in Information and Systems*, 7(3):287–312, 2007.

J. Ziv and M. Zakai. Some lower bounds on signal parameter estimation. *IEEE Transactions on Information Theory*, 15(3):386–391, 1969.