Regularized Estimation of High-dimensional Covariance Matrices

by

Yilun Chen

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Electrical Engineering: Systems)
in The University of Michigan
2011

Doctoral Committee:

Professor Alfred O. Hero III, Chair
Assistant Professor Clayton D. Scott
Assistant Professor Raj Rao Nadakuditi
Associate Professor Ji Zhu
Senior Lecturer Ami Wiesel, Hebrew University of Jerusalem

Dedicated to :
my wife Mrs. Yingjie Qi
my parents Mr. Xin Chen and Mrs. Xuean Zhan
whom I can never thank enough

ACKNOWLEDGEMENTS

Through the years since I started the Ph.D. journey I have been very fortunate to receive guidance, encouragement, patience and trust from my mentors, colleagues, friends and family. This dissertation could not have been possible without their support.

First, I would like to express my most sincere gratitude to my advisor Professor Alfred Hero for his invaluable support and guidance. I have always been admiring his vast knowledge, deep insight and passion to research. His trust of my abilities has provided me complete academic freedom to pursue my research interest and to develop my professional skills. His encouragement has given me invaluable opportunities to interact with world-class scholars. There has never been any doubt in my mind that I was granted an exceptional opportunity that every doctoral student would ever dream of.

I am grateful to Dr. Ami Wiesel for his tremendous help throughout my research experiences. His sharp insight always provided me the biggest support at the most challenging time and I am so appreciative for his genuine care of my professional development. I consider myself extremely lucky to have such a brilliant person as a mentor and life long friend. I am sure he will have great success in his academic career. I also would like to thank Professor Clayton Scott, Professor Rajesh Rao Nadakuditi and Professor Ji Zhu for being my committee members and their invaluable input and suggestions.

During my graduate study I have also been very fortunate to have opportunities to collaborate with Professor Yonina Eldar and Dr. Moshe Mishali on the exciting sampling project. They generously shared with me their valuable knowledge on analog compressive sensing. I am also thankful to the members of Hero group, particularly Dr. Mark Kliger, Dr. Yongsheng Huang, Mr. Kevin Xu, Mr. Arnau Tibau Puig, Mr. Zhaoshi Meng, and Dr. Xu Chen who have honored me with their friendship and helped me in so many ways.

Finally and most of all, I would like to thank my dear parents and wife, for their constant love and firm support throughout the years, which provides a never ending source of energy to my doctoral studies and future endeavors.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Regularized Estimation of High-dimensional Covariance Matrices

by
Yilun Chen

Chair: Alfred O. Hero III

Many signal processing methods are fundamentally related to the estimation of covariance matrices. In cases where there are a large number of covariates the dimension of covariance matrices is much larger than the number of available data samples. This is especially true in applications where data acquisition is constrained by limited resources such as time, energy, storage and bandwidth. This dissertation attempts to develop necessary components for covariance estimation in the high-dimensional setting. The dissertation makes contributions in two main areas of covariance estimation: (1) high dimensional shrinkage regularized covariance estimation and (2) recursive online complexity regularized estimation with applications of anomaly detection, graph tracking, and compressive sensing.

New shrinkage covariance estimation methods are proposed that significantly outperform previous approaches in terms of mean squared error. Two multivariate data scenarios are considered: (1) independently Gaussian distributed data; and (2) heavy tailed elliptically contoured data. For the former scenario we improve on the Ledoit-Wolf (LW) shrinkage estimator using the principle of Rao-Blackwell conditioning and

iterative approximation of the clairvoyant estimator. In the latter scenario, we apply a variance normalizing transformation and propose an iterative robust LW shrinkage estimator that is distribution-free within the elliptical family. The proposed robustified estimator is implemented via fixed point iterations with provable convergence and unique limit.

A recursive online covariance estimator is proposed for tracking changes in an underlying time-varying graphical model. Covariance estimation is decomposed into multiple decoupled adaptive regression problems. A recursive recursive group lasso is derived using a homotopy approach that generalizes online lasso methods to group sparse system identification. By reducing the memory of the objective function this leads to a group lasso regularized LMS that provably dominates standard LMS. Finally, we introduce a state-of-the-art sampling system, the Modulated Wideband Converter (MWC) which is based on recently developed analog compressive sensing theory. By inferring the block-sparse structures of the high-dimensional covariance matrix from a set of random projections, the MWC is capable of achieving sub-Nyquist sampling for multiband signals with arbitrary carrier frequency over a wide bandwidth.

CHAPTER I

Introduction

1.1 Covariance matrix estimation

Estimating the covariance matrix or its inverse of a random vector is one of the most fundamental problems in signal processing and related fields. Covariance matrices are important statistics that describe the pairwise correlation between random variables. Indeed, when the random variables are jointly Gaussian distributed and their mean values are pre-processed to zeros, the covariance matrix is the minimal sufficient statistics and encodes all the necessary statistical information for estimation and detection. As several typical examples, accurate covariance estimation plays an essential role in the following scenarios.

- *Regression problems.* The interplay between covariance estimation and vector regression has long been noticed. There are numerous examples how covariance matrices naturally enter into filtering or prediction problems. On the other hand, covariance estimation methods are usually connected to regularized regression algorithms. For example, it is well known that diagonal loading techniques, which is popular in covariance estimation of array signals, is equivalent to the ridge regression where the ℓ_2 norm is used as the penalty function for the coefficient vector to be estimated. In [3], the authors considers covariance esti-

mation with an unknown sparse graphical model and uses a set of ℓ_1- penalized regression problems to determine the graph structure.

- *Statistical inference in graphical models.*

  Inverse covariance matrices reveal the conditional independence of Gaussian distributed random variables, which is usually referred to as the covariance selection problem [4]. When there is an underlying graphical model associating variables of interest, the inverse covariance matrix determines the structure of the graph.

- *Subspace methods.*

  Covariance estimation is the initial stage of principal components analysis (PCA), linear discriminant analysis (LDA) and other factor analysis algorithms. In [5], efficient covariance estimators have been demonstrated to achieve less classification error in LDA based classification methods. Similar examples also widely exist in array signal processing applications [6].

1.2 High-dimensional challenges

As we have entered the era of data, the fast-developing sensing technologies have enabled the generation of vast amounts of information, often in the form of high-dimensional signals. This can be easily viewed within sensor/agent networks [7, 8], imaging [9], spectrum sensing [2], gene microarrays [10] and many others applications. While the dimension of signals has been significantly expanded, data acquisition processes are generally constrained in various kinds of resources and result in limited sample size, including but not limited to:

- *Budget.* Despite of the expanded data dimension, data acquisition can be expensive and limited in size. For example, the obtaining Affymetrix Human Genome

U133Plus2.0 DNA micro-arrays, which has 54,675 variables, costs as much as 500$ per sample, and limited budgets then yield limited sample size.

- *Energy.* Most sensors are powered by batteries and each sample is acquired at the cost of shortening the battery life. Take the Mica2 wireless sensor as an example, a standard nine-volt battery can support about 4000 times of sampling and communication cycles, while the data dimension of a sensor networks increases quadratically with the number of sensors.
- *Time.* In many signal processing and machine learning applications, the system of interest is time-varying that yields non-stationary samples. One example is the beamforming problem of moving targets, where only a small amount of samples are locally stationary.

Other resource constraints may also include limited storage and bandwidth. Those factors may occur simultaneously and cause the high-dimensional challenge in traditional covariance estimation methods.

Take the most used sample covariance as an example. While it is a consistent estimator for low-dimensional matrices with sufficient amount of samples, its poor performance in the high-dimensional setting has been well addressed. In fact, it has been demonstrated that the sample covariance suffers from a systematical distortion of its eigen-structures [11]. The eigenvalues of the sample covariance tends to over-spread which makes the sample covariance ill-conditioned unless the sample size n is significantly larger than the dimension p. In [11, 12], it has also been demonstrated that estimation of principal eigenvectors of the sample covariance breaks down if n/p is less than a threshold. Another drawback of the sample covariance based estimators is the sensitivity to the heavy-tails or outliers. This becomes even more severe for high-dimensional matrices as the sample size is usually not sufficient to

perform an outlier rejection. Such sensitivity may also deteriorate performance of other covariance estimators that are based on the sample covariance.

1.3 Previous works

General covariance estimation. Early work to improve upon the sample covariance can be traced back to 1960s when Stein demonstrated that superior performance can be obtained by shrinking the sample covariance [13, 14]. Since then, many shrinkage estimators have been proposed under different performance measures. For example, Haff [15] introduced an estimator inspired by the empirical Bayes approach. Dey and Srinivasan [16] derived a minimax estimator under Stein's entropy loss function. These works addressed the case of invertible sample covariance when $n \geq p$. Recently, Ledoit and Wolf (LW) proposed a shrinkage estimator for the case $n < p$ which asymptotically minimizes the MSE [17]. The LW estimator is well conditioned for small sample sizes and can thus be applied to high dimensional problems. In contrast to previous approaches, they show that performance advantages are distribution-free and not restricted to Gaussian assumptions. Those Steinian shrinkage methods are fundamentally related to ridge regression and diagonal loading, where the ℓ_2 norm regularization is used.

Model based covariance estimation. The Steinian shrinkage based estimators deal with the general covariance estimation problems and can be applied to any covariance matrices. Those methods are usually simple and computationally efficient; yet they do not fully exploit prior covariance structures for specific applications. Prior structures, if they exist, usually imply that the true model complexity is much lower than the number of unknown parameters. By using a proper regularization, consistency of covariance estimators can be obtained under "large p small n" settings. For

example, in [18, 19, 20, 21], the ℓ_1 norm is adopted as the regularization function to promote sparsity on the inverse covariance matrix. This is closely related to the problem of covariance selection [4] when there is an underlying undirected graphical model associating jointly Gaussian distributed variables. For variables that have a natural ordering and those far apart in the ordering have small partial correlations, [22, 23] have been proposed to exploit the banded structure on the modified Cholesky decomposition of the inverse covariance matrix. In [24, 25, 26, 27, 28], covariance tapering, banding or thresholding methods have been proposed to explore the sparse or banded structures on the covariance rather than its inverse. [29] considers a multi-factor model and uses it to reduce dimensionality and to estimate the covariance matrix.

Robust covariance estimation for non-Gaussian samples. As the sample covariance is sensitive to heavy tails or outliers, covariance estimation methods that operate on the sample covariance (explicitly or implicitly) may suffer from such sensitivity as well. In fact, robust covariance estimation has been investigated for years. Typical approaches include the M-estimators [30, 31], the minimum volume ellipsoid estimator (MVE) and the minimum covariance determinant (MCD) [32, 33] estimator, and the S-estimators [34]. These methods are robust to non-Gaussian samples but generally require $n \gg p$, which are not suitable for high-dimensional problems. Robust covariance estimation is even more challenging in the "large p small n" setting, as the insufficient number of samples makes it difficult to perform a Gaussianity test. Furthermore, traditional robust estimators designed for low dimensional problems may not be scalable with the dimensionality p. For covariance estimation in graphical models, the authors in [35] presented a robust estimator for multivariate student-T distributed samples, which is based on a Markov chain Monte Carlo

Expectation Maximization (EM) usings a Gibbs sampler. They describe a simple variational approximation to make the resulting method computationally feasible in high-dimensional problems.

Online covariance estimation for time-varying systems. The majority of literatures for covariance estimation are based on off-line processing of stationary samples. In the contrast, much less has been done on time-varying covariance estimation via online estimators. Recent developments in signal processing and machine learning has attracted growing interests in statistical inference problems of time-varying models [7, 8], as time-varying systems and non-stationary data are one of the main causes of "large p small n" problems. In the context of radar signal processing, Abramovich [36, 37, 38] proposes a series of work on covariance estimation for time-varying autoregressive (TVAR) models. While the AR coefficients are non-stationary, the banded structure of the inverse covariance is assumed to be time invariant. In [7] and [39], inference problems on time-varying stochastic networks are considered, where both covariance structures (or network topologies) and model parameters are temporally dynamic. Those methods are implemented in the offline mode and the optimization algorithms are computationally intensive for online implementation.

1.4 Dissertation contributions

The focus of the work presented in this dissertation is to develop high-dimensional covariance estimation methodologies in "large p small n" settings. The contents for each chapter are specified as follows:

- Chapter II addresses the general covariance estimation problem in the sense of minimum mean squared error (MMSE). The samples are assumed to be Gaussian distributed and Steinian shrinkage methods are considered. We begin by

improving the Ledoit-Wolf (LW) method by conditioning on a sufficient statistic. By the Rao-Blackwell theorem, we obtain a new estimator, called RBLW, which dominates LW in MSE. Then, to further reduce the estimation error, we propose an iterative approach which approximates the clairvoyant shrinkage estimator. Convergence of this iterative method is established and a closed-form expression for the limit is determined, which is referred to as the oracle approximating shrinkage (OAS) estimator. Both RBLW and OAS estimators have simple expressions and are easily implemented. Although the two methods are developed from different persepctives, their structure is identical up to specified constants. The RBLW estimator provably dominates the LW method for Gaussian samples. Numerical simulations demonstrate that the OAS approach can perform even better than RBLW, especially when n is much less than p. We also demonstrate the performance of these techniques in the context of adaptive beamforming.

- Chapter III presents a robustfied generalization of the work proposed in Chapter II, where we extend the Gaussian sample distribution to the family of elliptical distributions. We start from a classical robust covariance estimator [31], which is firstly proposed by Tyler and is distribution-free within the family of elliptical distribution but inapplicable when $n < p$. Using a shrinkage coefficient, we regularize Tyler's fixed point iterations. We prove that, for all n and p, the proposed fixed point iterations converge to a unique limit regardless of the initial condition. Next, we propose a simple, closed-form and data dependent choice for the shrinkage coefficient, which is based on a minimum mean squared error framework. Simulations demonstrate that the proposed method achieves low estimation error and is robust to heavy-tailed samples. As a real world

application we demonstrate the performance of the proposed technique in the context of activity/intrusion detection using a wireless sensor network.

- In Chapter IV we address covariance estimation of time-varying systems using lasso type penalties. Specifically we are interested in online algorithms for streaming data. Instead of solving the covariance matrix directly, we propose to split the covariance estimation into a set of adaptive regularized regression problems and develop recursive algorithms for efficient online implementation. We introduce a recursive adaptive group lasso, as a generalization to the standard lasso, for real-time penalized least squares regression. At each time index the proposed algorithm computes an exact update of the optimal $\ell_{1,\infty}$-penalized recursive least squares (RLS) predictor. We develop an online homotopy method to reduce the computational complexity. Numerical simulations demonstrate that the online homotopy method has lower implementation complexity than direct path following algorithms. We finally integrate the adaptive recursive lasso into online covariance estimation and demonstrate its performance in anomaly detection for the Abilene Network.

- Chapter V develops stochastic gradient algorithms for adaptive regularized regression problems, as a low cost approximation of the work presented in Chapter IV. We consider time-varying adaptive regression problems in the stochastic gradient framework and propose a family of Least-Mean-Square (LMS) algorithms regularized by general convex penalties. We show that with a properly selected regularization parameter the regularized LMS provably dominates its conventional counterpart in terms of MSE and establish simple and closed-form expressions for choosing this regularization parameter. The proposed regularized LMS can be applied to online regression problems with any convex constraints.

As special examples, we propose sparse and group sparse LMS algorithms for adaptive sparse regression and demonstrate their performances using numerical simulations.

- Chapter VI[1] introduces an innovative sensing system developed for sub-Nyquist sampling of sparse spectrum wide-band signals, where structure estimation of sparse high-dimensional covariance matrices plays an important role. The system is designed based on recently developed theories of compressive sensing [40, 41] and the board-level prototype supports input signals with 2 GHz Nyquist rate and 120 MHz spectrum occupancy with arbitrary transmission frequencies, while the sampling rate is as low as 280 MHz. The key to achieve sub-Nyquist sampling is the capability to recover a high-dimensional covariance matrix, where only a small portion of columns and rows are non-zero, from its low-dimensional random projections. Unlike previous chapters which focus on methodologies, Chapter VI covers implementation details that are crucial to the success of the practical system.

1.5 List of relevant publications

The following publications were produced based on the research presented in this dissertation:

1. Y. Chen , A. Wiesel, Y. C. Eldar, A. O. Hero, "Shrinkage Algorithms for MMSE Covariance Estimation ," *IEEE Transactions on Signal Processing*, October 2010

2. Y. Chen, A. Wiesel, A. O. Hero, "Robust Shrinkage Estimation for High-dimensional Covariance Matrices,"accepted by *IEEE Transactions on Signal*

[1]This is based on the joint work between University of Michigan (Yilun Chen, Prof. Alfred Hero) and Technion – Israel Institute of Technology (Dr. Moshe Mishali, Prof. Yonina Eldar).

Processing

3. Y. Chen, A. O. Hero, "Recursive $\ell_{1,\infty}$ Group lasso," *IEEE Transactions on Signal Processing*, in review

4. Y. Chen, Y. Gu, A. O. Hero, "Regularized Least-Mean-Square Algorithms," *IEEE Transactions on Signal Processing*, in preparation to *EURASIP Signal Processing*

5. Y. Chen , A. Wiesel, A. O. Hero, "Robust Shrinkage Estimation of High-dimensional Covariance Matrices," *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, Oct. 2010

6. Y. Chen, M. Mishali, Y. C. Eldar, A. O. Hero, "Modulated Wideband Converter with Non-ideal Lowpass Filters," *IEEE Intl Conf. on Acoust., Speech, and Signal Processing (ICASSP)*, Dallas, TX, USA, Mar. 2010

7. Y. Chen, A. Wiesel, A. O. Hero, "Shrinkage estimation of high dimensional covariance matrices," *IEEE Intl Conf. on Acoust., Speech, and Signal Processing (ICASSP),* Taiwan, Mar. 2009

8. Y. Chen , Y. Gu, A. O. Hero, "Sparse LMS for System Identification," *IEEE Intl Conf. on Acoust., Speech, and Signal Processing, Taiwan (ICASSP)*, Mar. 2009

9. K. S. Xu, M. Kliger, Y. Chen, P. Woolf, A. O. Hero, "Social Networks of Spammers Through Spectral Clustering," *IEEE Intl. Conf. on Communications (ICC)*, Jun. 2009

CHAPTER II

Shrinkage Algorithms for Covariance Estimation: Gaussian Samples

In this chapter we address covariance estimation in the sense of minimum mean-squared error (MMSE) when the samples are Gaussian distributed. Specifically, we consider shrinkage methods which are suitable for high dimensional problems with a small number of samples (large p small n). First, we improve on the Ledoit-Wolf (LW) method by conditioning on a sufficient statistic. By the Rao-Blackwell theorem, this yields a new estimator called RBLW, whose mean-squared error dominates that of LW for Gaussian variables. Second, to further reduce the estimation error, we propose an iterative approach which approximates the clairvoyant shrinkage estimator. Convergence of this iterative method is established and a closed form expression for the limit is determined, which is referred to as the oracle approximating shrinkage (OAS) estimator. Both RBLW and OAS estimators have simple expressions and are easily implemented. Although the two methods are developed from different perspectives, their structure is identical up to specified constants. The RBLW estimator provably dominates the LW method for Gaussian samples. Numerical simulations demonstrate that the OAS approach can perform even better than RBLW, especially when n is much less than p. We also demonstrate the performance of these techniques in the context of adaptive beamforming.

2.1 Introduction

Covariance matrix estimation is a fundamental problem in signal processing and related fields. Many applications varying from array processing [42] to functional genomics [10] rely on accurately estimated covariance matrices. In recent years, estimation of high dimensional $p \times p$ covariance matrices under small sample size n has attracted considerable interest. Examples include classification on gene expression from microarray data [10], financial forecasting [43, 44], spectroscopic imaging [45], brain activation mapping from fMRI [9] and many others. Standard estimation methods perform poorly in these large p small n settings. This is the main motivation for this work.

The sample covariance is a common estimate for the unknown covariance matrix. When it is invertible, the sample covariance coincides with the classical maximum likelihood estimate. However, while it is an unbiased estimator, it does not minimize the mean-squared error (MSE). Indeed, Stein demonstrated that superior performance may be obtained by shrinking the sample covariance [13, 14]. Since then, many shrinkage estimators have been proposed under different performance measures. For example, Haff [15] introduced an estimator inspired by the empirical Bayes approach. Dey and Srinivasan [16] derived a minimax estimator under Stein's entropy loss function. These works addressed the case of invertible sample covariance when $n \geq p$. Recently, Ledoit and Wolf (LW) proposed a shrinkage estimator for the case $n < p$ which asymptotically minimizes the MSE [17]. The LW estimator is well conditioned for small sample sizes and can thus be applied to high dimensional problems. In contrast to previous approaches, they show that performance advantages are distribution-free and not restricted to Gaussian assumptions.

In this chapter, we show that the LW estimator can be significantly improved when the samples are in fact Gaussian. Specifically, we develop two new estimation techniques that result from different considerations. The first follows from the Rao-Blackwell theorem, while the second is an application of the ideas of [46] to covariance estimation.

We begin by providing a closed form expression for the optimal clairvoyant shrinkage estimator under an MSE loss criteria. This estimator is an explicit function of the unknown covariance matrix that can be used as an oracle performance bound. Our first estimator is obtained by applying the well-known Rao-Blackwell theorem [47] to the LW method, and is therefore denoted by RBLW. Using several nontrivial Haar integral computations, we obtain a simple closed form solution which provably dominates the LW method in terms of MSE. We then introduce an iterative shrinkage estimator which tries to approximate the oracle. This approach follows the methodology developed in [46] for the case of linear regression. Beginning with an initial naive choice, each iteration is defined as the oracle solution when the unknown covariance is replaced by its estimate obtained in the previous iteration. Remarkably, a closed form expression can be determined for the limit of these iterations. We refer to the limit as the oracle approximating shrinkage (OAS) estimator.

The OAS and RBLW solutions have similar structure that is related to a sphericity test as discussed in [48, 49, 50]. Both OAS and RBLW estimators are intuitive, easy to compute and perform well with finite sample size. The RBLW technique provably dominates LW. Numerical results demonstrate that for small sample sizes, the OAS estimator is superior to both the RBLW and the LW methods.

To illustrate the proposed covariance estimators we apply them to problems of time series analysis and array signal processing. Specifically, in the context of time

series analysis we establish performance advantages of OAS and RBLW to LW for covariance estimation in autoregressive models and in fractional Brownian motion models, respectively. In the context of beamforming, we show that RBLW and OAS can be used to significantly improve the Capon beamformer. In [42] a multitude of covariance matrix estimators were implemented in Capon beamformers, and the authors reported that the LW approach substantially improves performance as compared to other methods. We show here that even better performance can be achieved by using the techniques introduced in this work.

Finally, we would like to point out that the performance gain achieved in the proposed estimators is only guaranteed when the samples are approximately Gaussian distributed. Although this is a common assumption in signal processing, when the sample distribution is far from Gaussian the proposed estimators may not perform as well as the LW estimator. The discussion of non-Gaussian shrinkage estimator will be continued in the following chapter.

The chapter is organized as follows. Section 2.2 formulates the problem. Section 2.3 introduces the oracle estimator together with the RBLW and OAS methods. Section 2.4 represents numerical simulation results and applications in adaptive beamforming. Section 2.5 summarizes our principal conclusions. The proofs of theorems and lemmas are provided in the Appendix.

Notation: In the following, we depict vectors in lowercase boldface letters and matrices in uppercase boldface letters. $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and the conjugate transpose, respectively. $\mathrm{Tr}\,(\cdot)$, $\|\cdot\|_F$, and $\det(\cdot)$ are the trace, the Frobenius norm, and the determinant of a matrix, respectively. Finally, $\mathbf{A} \prec \mathbf{B}$ means that the matrix $\mathbf{B} - \mathbf{A}$ is positive definite, and $\mathbf{A} \succ \mathbf{B}$ means that the matrix $\mathbf{A} - \mathbf{B}$ is positive definite.

2.2 Problem formulation

Let $\{\mathbf{x}_i\}_{i=1}^n$ be a sample of independent identical distributed (i.i.d.) p-dimensional Gaussian vectors with zero mean and covariance $\mathbf{\Sigma}$. We do not assume $n \geq p$. Our goal is to find an estimator $\widehat{\mathbf{\Sigma}}\left(\{\mathbf{x}_i\}_{i=1}^n\right)$ which minimizes the MSE:

$$E\left\{\left\|\widehat{\mathbf{\Sigma}}\left(\{\mathbf{x}_i\}_{i=1}^n\right)-\mathbf{\Sigma}\right\|_F^2\right\}. \tag{2.1}$$

It is difficult to compute the MSE of $\widehat{\mathbf{\Sigma}}\left(\{\mathbf{x}_i\}_{i=1}^n\right)$ without additional constraints and therefore we restrict ourselves to a specific class of estimators that employ shrinkage [13, 17]. The unstructured classical estimator of $\mathbf{\Sigma}$ is the sample covariance $\widehat{\mathbf{S}}$ defined as

$$\widehat{\mathbf{S}}=\frac{1}{n}\sum_{i=1}^n \mathbf{x}_i\mathbf{x}_i^T. \tag{2.2}$$

This estimator is unbiased $E\{\widehat{\mathbf{S}}\}=\mathbf{\Sigma}$, and is also the maximum likelihood solution if $n \geq p$. However, it does not necessarily achieve low MSE due to its high variance and is usually ill-conditioned for large p problems. On the other hand, we may consider a naive but most well-conditioned estimate for $\mathbf{\Sigma}$:

$$\widehat{\mathbf{F}}=\frac{\mathrm{Tr}\left(\widehat{\mathbf{S}}\right)}{p}\mathbf{I}. \tag{2.3}$$

This "structured" estimate will result in reduced variance with the expense of increasing the bias. A reasonable tradeoff between low bias and low variance is achieved by shrinkage of $\widehat{\mathbf{S}}$ towards $\widehat{\mathbf{F}}$, resulting in the following class of estimators:

$$\widehat{\mathbf{\Sigma}}=(1-\hat{\rho})\widehat{\mathbf{S}}+\hat{\rho}\widehat{\mathbf{F}}. \tag{2.4}$$

The estimator $\widehat{\mathbf{\Sigma}}$ is characterized by the shrinkage coefficient $\hat{\rho}$, which is a parameter between 0 and 1 and can be a function of the observations $\{\mathbf{x}_i\}_{i=1}^n$. The matrix $\widehat{\mathbf{F}}$ is referred to as the shrinkage target.[1]

[1] The convex combination in (2.4) can be generalized to the linear combination of $\widehat{\mathbf{S}}$ and $\widehat{\mathbf{F}}$. The reader is referred to [51] for further discussion.

Throughout this chapter, we restrict our attention to shrinkage estimates of the form (2.4). Our goal is to find a shrinkage coefficient $\hat{\rho}$ that minimizes the MSE (2.1). As we show in the next section, the optimal $\hat{\rho}$ minimizing the MSE depends in general on the unknown $\mathbf{\Sigma}$ and therefore in general cannot be implemented. Instead, we propose two different approaches to approximate the optimal shrinkage coefficient.

2.3 Shrinkage algorithms

2.3.1 The Oracle estimator

We begin by deriving a clairvoyant oracle estimator that uses an optimal nonrandom coefficient to minimize the mean-squared error. In the following subsections we will show how to approximate the oracle using implementable data-driven methods.

The oracle estimate $\widehat{\mathbf{\Sigma}}_O$ is the solution to

$$\begin{aligned} \min_{\rho} \quad & E\left\{\left\|\widehat{\mathbf{\Sigma}}_O - \mathbf{\Sigma}\right\|_F^2\right\} \\ \text{s.t.} \quad & \widehat{\mathbf{\Sigma}}_O = (1-\rho)\,\widehat{\mathbf{S}} + \rho\widehat{\mathbf{F}} \end{aligned} . \tag{2.5}$$

The optimal parameter ρ_O is provided in the following theorem.

Theorem II.1. *Let $\widehat{\mathbf{S}}$ be the sample covariance of a set of p-dimensional vectors $\{\mathbf{x}_i\}_{i=1}^n$. If $\{\mathbf{x}_i\}_{i=1}^n$ are i.i.d. Gaussian vectors with covariance $\mathbf{\Sigma}$, then the solution to (2.5) is*

$$\rho_O = \frac{E\left\{\mathrm{Tr}\left(\left(\mathbf{\Sigma} - \widehat{\mathbf{S}}\right)\left(\widehat{\mathbf{F}} - \widehat{\mathbf{S}}\right)\right)\right\}}{E\left\{\left\|\widehat{\mathbf{S}} - \widehat{\mathbf{F}}\right\|_F^2\right\}} \tag{2.6}$$

$$= \frac{(1-2/p)\,\mathrm{Tr}\left(\mathbf{\Sigma}^2\right) + \mathrm{Tr}^2\left(\mathbf{\Sigma}\right)}{(n+1-2/p)\mathrm{Tr}(\mathbf{\Sigma}^2) + (1-n/p)\ \mathrm{Tr}^2(\mathbf{\Sigma})}. \tag{2.7}$$

Proof. Equation (2.6) was established in [17] for any distribution of $\{\mathbf{x}_i\}_{i=1}^n$. Under the additional Gaussian assumption, (2.7) can be obtained from straightforward

evaluation of the expectations:

$$
\begin{aligned}
E\left\{\mathrm{Tr}\left(\left(\boldsymbol{\Sigma}-\widehat{\mathbf{S}}\right)\left(\widehat{\mathbf{F}}-\widehat{\mathbf{S}}\right)\right)\right\} &= \frac{\mathrm{Tr}\left(\boldsymbol{\Sigma}\right)}{p}E\left\{\mathrm{Tr}\left(\widehat{\mathbf{S}}\right)\right\} \\
&- \frac{E\left\{\mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)\right\}}{p} - E\left\{\mathrm{Tr}\left(\boldsymbol{\Sigma}\widehat{\mathbf{S}}\right)\right\} + E\left\{\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)\right\},
\end{aligned} \tag{2.8}
$$

and

$$
\begin{aligned}
&E\left\{\left\|\widehat{\mathbf{S}}-\widehat{\mathbf{F}}\right\|_F^2\right\} \\
&= E\left\{\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)\right\} - 2E\left\{\mathrm{Tr}\left(\widehat{\mathbf{S}}\widehat{\mathbf{F}}\right)\right\} + E\left\{\mathrm{Tr}\left(\widehat{\mathbf{F}}^2\right)\right\} \\
&= E\left\{\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)\right\} - \frac{E\left\{\mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)\right\}}{p}.
\end{aligned} \tag{2.9}
$$

Equation (2.7) is a result of using the following identities [52]:

$$
E\left\{\mathrm{Tr}\left(\widehat{\mathbf{S}}\right)\right\} = \mathrm{Tr}\left(\boldsymbol{\Sigma}\right), \tag{2.10}
$$

$$
E\left\{\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)\right\} = \frac{n+1}{n}\mathrm{Tr}\left(\boldsymbol{\Sigma}^2\right) + \frac{1}{n}\mathrm{Tr}^2\left(\boldsymbol{\Sigma}\right), \tag{2.11}
$$

and

$$
E\left\{\mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)\right\} = \mathrm{Tr}^2\left(\boldsymbol{\Sigma}\right) + \frac{2}{n}\mathrm{Tr}\left(\boldsymbol{\Sigma}^2\right). \tag{2.12}
$$

□

Note that (2.6) specifies the optimal shrinkage coefficient for any sample distribution while (2.7) only holds for the Gaussian distribution. It can be shown that expression (2.7) will be a close approximation to (6) when the the sample distribution is close to the Gaussian distribution in variational norm.

2.3.2 The Rao-Blackwell Ledoit-Wolf (RBLW) estimator

The oracle estimator defined by (2.5) is optimal but cannot be implemented, since the solution specified by both (2.6) and (2.7) depends on the unknown $\boldsymbol{\Sigma}$.

Without any knowledge of the sample distribution, Ledoit and Wolf [17] proposed to approximate the oracle using the following consistent estimate of (2.6):

$$\hat{\rho}_{LW} = \frac{\sum_{i=1}^{n} \left\| \mathbf{x}_i \mathbf{x}_i^T - \widehat{\mathbf{S}} \right\|_F^2}{n^2 \left[\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right) - \mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)/p \right]}. \tag{2.13}$$

They then proved that when both $n, p \to \infty$ and $p/n \to c$, $0 < c < \infty$, (2.13) converges to (2.6) in probability regardless of the sample distribution. The LW estimator $\widehat{\mathbf{\Sigma}}_{LW}$ is then defined by plugging $\hat{\rho}_{LW}$ into (2.4). In [17] Ledoit and Wolf also showed that the optimal ρ_O (2.6) is always between 0 and 1. To further improve the performance, they suggested using a modified shrinkage parameter

$$\hat{\rho}^*_{LW} = \min\left(\hat{\rho}_{LW}, 1\right) \tag{2.14}$$

instead of $\hat{\rho}_{LW}$.

The Rao-Blackwell LW (RBLW) estimator described below provably improves on the LW method under the Gaussian model. The motivation for the RBLW originates from the fact that under the Gaussian assumption on $\{\mathbf{x}_i\}_{i=1}^n$, a sufficient statistic for estimating $\mathbf{\Sigma}$ is the sample covariance $\widehat{\mathbf{S}}$. Intuitively, the LW estimator is a function of not only $\widehat{\mathbf{S}}$ but other statistics and therefore, by the Rao-Blackwell theorem, can be improved. Specifically, the Rao-Blackwell theorem [47] states that if $g(X)$ is an estimator of a parameter θ, then the conditional expectation of $g(X)$ given $T(X)$, where T is a sufficient statistic, is never worse than the original estimator $g(X)$ under any convex loss criterion. Applying the Rao-Blackwell theorem to the LW estimator yields the following result.

Theorem II.2. *Let $\{\mathbf{x}_i\}_{i=1}^n$ be independent p-dimensional Gaussian vectors with covariance $\mathbf{\Sigma}$, and let $\widehat{\mathbf{S}}$ be the sample covariance of $\{\mathbf{x}_i\}_{i=1}^n$. The conditioned expec-*

tation of the LW covariance estimator is

$$\widehat{\mathbf{\Sigma}}_{RBLW} = E\left[\widehat{\mathbf{\Sigma}}_{LW} \middle| \widehat{\mathbf{S}}\right] \tag{2.15}$$

$$= (1-\hat{\rho}_{RBLW})\widehat{\mathbf{S}} + \hat{\rho}_{RBLW}\widehat{\mathbf{F}} \tag{2.16}$$

where

$$\hat{\rho}_{RBLW} = \frac{(n-2)/n \cdot \mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right) + \mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)}{(n+2)\left[\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right) - \mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)/p\right]}. \tag{2.17}$$

This estimator satisfies

$$E\left\{\left\|\widehat{\mathbf{\Sigma}}_{RBLW} - \mathbf{\Sigma}\right\|_F^2\right\} \leq E\left\{\left\|\widehat{\mathbf{\Sigma}}_{LW} - \mathbf{\Sigma}\right\|_F^2\right\}, \tag{2.18}$$

for every $\mathbf{\Sigma}$.

The proof of Theorem II.2 is given in the Appendix.

Similarly to the LW estimator, we propose the modification

$$\hat{\rho}^*_{RBLW} = \min\left(\hat{\rho}_{RBLW}, 1\right) \tag{2.19}$$

instead of $\hat{\rho}_{RBLW}$.

2.3.3 The Oracle-Approximating Shrinkage (OAS) estimator

The basic idea of the LW estimator is to asymptotically approximate the oracle, which is designed for large sample size. For a large number of samples the LW asymptotically achieves the minimum MSE with respect to shrinkage estimators. Clearly, the RBLW also inherits this property. However, for very small n, which is often the case of interest, there is no guarantee that such optimality still holds. To illustrate this point, consider the extreme example when only one sample is available. For $n = 1$ we have both $\hat{\rho}^*_{LW} = 1$ and $\hat{\rho}^*_{RBLW} = 1$, which indicates that $\widehat{\mathbf{\Sigma}}_{LW} = \widehat{\mathbf{\Sigma}}_{RBLW} = \widehat{\mathbf{S}}$. This however contradicts our expectations since if a single sample is

available, it is more reasonable to expect more confidence to be put on the more parsimonious $\widehat{\mathbf{F}}$ rather than $\widehat{\mathbf{S}}$.

In this section, we aim at developing a new estimation method which can approximate the oracle for finite n. Rather than employing asymptotic solutions we consider an alternative approach based on [46]. In (2.7), we obtained a closed-form formula of the oracle estimator under Gaussian assumptions. The idea behind the OAS is to approximate this oracle via an iterative procedure. We initialize the iterations with an initial guess of $\mathbf{\Sigma}$ and iteratively refine it. The initial guess $\widehat{\mathbf{\Sigma}}_0$ might be the sample covariance, the RBLW estimate or any other symmetric non-negative definite estimator. We replace $\mathbf{\Sigma}$ in the oracle solution by $\widehat{\mathbf{\Sigma}}_0$ yielding $\widehat{\mathbf{\Sigma}}_1$, which in turn generates $\widehat{\mathbf{\Sigma}}_2$ through our proposed iteration. The iteration process is continued until convergence. The limit, denoted as $\widehat{\mathbf{\Sigma}}_{OAS}$, is the OAS solution. Specifically, the proposed iteration is,

$$\hat{\rho}_{j+1} = \frac{(1-2/p)\mathrm{Tr}\left(\widehat{\mathbf{\Sigma}}_j\widehat{\mathbf{S}}\right) + \mathrm{Tr}^2\left(\widehat{\mathbf{\Sigma}}_j\right)}{(n+1-2/p)\mathrm{Tr}\left(\widehat{\mathbf{\Sigma}}_j\widehat{\mathbf{S}}\right) + (1-n/p)\ \mathrm{Tr}^2\left(\widehat{\mathbf{\Sigma}}_j\right)}, \tag{2.20}$$

$$\widehat{\mathbf{\Sigma}}_{j+1} = (1-\hat{\rho}_{j+1})\widehat{\mathbf{S}} + \hat{\rho}_{j+1}\widehat{\mathbf{F}}. \tag{2.21}$$

Comparing with (2.7), notice that in (2.20) $\mathrm{Tr}(\mathbf{\Sigma})$ and $\mathrm{Tr}(\mathbf{\Sigma}^2)$ are replaced by $\mathrm{Tr}(\widehat{\mathbf{\Sigma}}_j)$ and $\mathrm{Tr}(\widehat{\mathbf{\Sigma}}_j\widehat{\mathbf{S}})$, respectively. Here $\mathrm{Tr}(\widehat{\mathbf{\Sigma}}_j\widehat{\mathbf{S}})$ is used instead of $\mathrm{Tr}(\widehat{\mathbf{\Sigma}}_j^2)$ since the latter would always force $\hat{\rho}_j$ to converge to 1 while the former leads to a more meaningful limiting value.

Theorem II.3. *For any initial guess $\hat{\rho}_0$ that is between 0 and 1, the iterations specified by (2.20), (2.21) converge as $j \to \infty$ to the following estimate:*

$$\widehat{\mathbf{\Sigma}}_{OAS} = (1-\hat{\rho}^*_{OAS})\widehat{\mathbf{S}} + \hat{\rho}^*_{OAS}\widehat{\mathbf{F}}, \tag{2.22}$$

where

$$\hat{\rho}^*_{OAS} = \min\left(\frac{(1-2/p)\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)+\mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)}{(n+1-2/p)\left[\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)-\mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)/p\right]},1\right). \tag{2.23}$$

In addition, $0 < \hat{\rho}^*_{OAS} \leq 1$.

Proof. Plugging in $\widehat{\mathbf{\Sigma}}_j$ from (2.21) into (2.20) and simplifying yields

$$\hat{\rho}_{j+1} = \frac{1-(1-2/p)\hat{\phi}\hat{\rho}_j}{1+n\hat{\phi}-(n+1-2/p)\hat{\phi}\hat{\rho}_j}, \tag{2.24}$$

where

$$\hat{\phi} = \frac{\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)-\mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)/p}{\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)+(1-2/p)\ \mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)}. \tag{2.25}$$

Since $\mathrm{Tr}(\widehat{\mathbf{S}}^2) \geq \mathrm{Tr}^2(\widehat{\mathbf{S}})/p$, $0 \leq \hat{\phi} < 1$.

Next we introduce a new variable

$$\hat{b}_j = \frac{1}{1-(n+1-2/p)\hat{\phi}\hat{\rho}_j}. \tag{2.26}$$

Then,

$$\hat{\rho}_j = \frac{1-\hat{b}_j^{-1}}{(n+1-2/p)\hat{\phi}}. \tag{2.27}$$

Substituting (2.27) to (2.24), after simplification we obtain that

$$1-\hat{b}_{j+1}^{-1} = \frac{n\hat{\phi}+(1-2/p)\hat{\phi}\hat{b}_j^{-1}}{n\hat{\phi}+\hat{b}_j^{-1}}, \tag{2.28}$$

which leads to the following geometric series

$$\hat{b}_{j+1} = \frac{n\hat{\phi}}{1-(1-2/p)\hat{\phi}}\hat{b}_j + \frac{1}{1-(1-2/p)\hat{\phi}}. \tag{2.29}$$

It is easy to see that

$$\lim_{j\to\infty}\hat{b}_j = \begin{cases} \infty, & \text{if } \dfrac{n\hat{\phi}}{1-(1-2/p)\hat{\phi}} \geq 1 \\ \dfrac{1}{1-(n+1-2/p)\hat{\phi}}, & \text{if } \dfrac{n\hat{\phi}}{1-(1-2/p)\hat{\phi}} < 1 \end{cases}. \tag{2.30}$$

Therefore $\hat{\rho}_j$ also converges as $j \to \infty$ and $\hat{\rho}^*_{OAS}$ is given by

$$\hat{\rho}^*_{OAS} = \lim_{j\to\infty} \hat{\rho}_j = \begin{cases} \dfrac{1}{(n+1-2/p)\hat{\phi}}, & \text{if } (n+1-2/p)\hat{\phi} > 1 \\ 1, & \text{if } (n+1-2/p)\hat{\phi} \leq 1 \end{cases}. \tag{2.31}$$

We can write (2.31) equivalently as

$$\hat{\rho}^*_{OAS} = \min\left(\frac{1}{(n+1-2/p)\hat{\phi}}, 1\right). \tag{2.32}$$

Equation (2.23) is obtained by substituting (2.25) into (2.31). □

Note that (2.31) $\hat{\rho}^*_{OAS}$ is naturally bounded within $[0, 1]$. This is different from $\hat{\rho}^*_{LW}$ and $\hat{\rho}^*_{RBLW}$, where the $[0, 1]$ constraint is imposed in an *ad hoc* fashion. We also note that the iterative procedure described in (2.20) and (2.21) is only used to establish the limiting closed form OAS solution. In practice, the OAS estimate is calculated using (2.22) and (2.23), so that iterations are unnecessary.

2.3.4 Shrinkage and sphericity statistics

We now turn to theoretical comparisons between RBLW and OAS. The only difference is in their shrinkage coefficients. Although derived from distinct approaches, it is easy to see that $\hat{\rho}^*_{OAS}$ shares the same structure as $\hat{\rho}^*_{RBLW}$. In fact, they can both be expressed as the parameterized function

$$\hat{\rho}^*_E = \min\left(\alpha + \frac{\beta}{\hat{U}}, 1\right) \tag{2.33}$$

with $\hat{U}$ defined as

$$\hat{U} = \frac{1}{p-1}\left(\frac{p \cdot \mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)}{\mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)} - 1\right). \tag{2.34}$$

For $\hat{\rho}^*_E = \hat{\rho}^*_{OAS}$, α and β of (2.33) are given by

$$\begin{aligned} \alpha = \alpha_{OAS} &= \frac{1}{n+1-2/p} \\ \beta = \beta_{OAS} &= \frac{p+1}{(n+1-2/p)(p-1)} \end{aligned}, \tag{2.35}$$

while for $\hat{\rho}_E^* = \hat{\rho}_{RBLW}^*$:

$$\begin{aligned} \alpha &= \alpha_{RBLW} = \frac{n-2}{n(n+2)} \\ \beta &= \beta_{RBLW} = \frac{(p+1)n-2}{n(n+2)(p-1)} \end{aligned} . \tag{2.36}$$

Thus the only difference between $\hat{\rho}_{OAS}^*$ and $\hat{\rho}_{RBLW}^*$ is the choice of α and β. The statistic $\hat{U}$ arises in tests of sphericity of $\mathbf{\Sigma}$ [49, 50], *i.e.*, testing whether or not $\mathbf{\Sigma}$ is a scaled identity matrix. In particular, $\hat{U}$ is the locally most powerful invariant test statistic for sphericity under orthogonal transformations [48]. The smaller the value of $\hat{U}$, the more likely that $\mathbf{\Sigma}$ is proportional to an identity matrix $\mathbf{I}$. Similarly, in our shrinkage algorithms, the smaller the value of $\hat{U}$, the more shrinkage occurs in $\widehat{\mathbf{\Sigma}}_{OAS}$ and $\widehat{\mathbf{\Sigma}}_{RBLW}$.

Under the following asymptotic conditions [17]:

$$n \to \infty,\ p \to \infty,\ \text{and}\ \frac{p}{n} \to \text{constant}, \tag{2.37}$$

OAS and RBLW are equivalent, since $\hat{\rho}_{OAS}^*$ and $\hat{\rho}_{RBLW}^*$ converge to each other. In [17] the authors have proved that the LW estimator is asymptotically optimal under condition (2.37). Note that RBLW dominates LW for Gaussian samples. Therefore, RBLW, OAS and LW are asymptotically equivalent to each other when the samples are Gaussian.

It is also important to note that while RBLW and OAS share the same structure and are asymptotically equivalent, their finite sample performance may be entirely different, especially when n is small. For example, in the extreme case that only one sample is available, $\hat{\rho}_{OAS}^* = 1$ while $\hat{\rho}_{RBLW}^* = 0$, regardless of the value of the sample. This observation will be further explored in the numerical simulations below.

2.4 Numerical Simulations

In this section we implement and test the proposed covariance estimators. We first compare the estimated MSE of the RBLW and OAS techniques with the LW method. We then consider their application to the problem of adaptive beamforming, and show that they lead to improved performance of Capon beamformers.

2.4.1 MSE Comparison

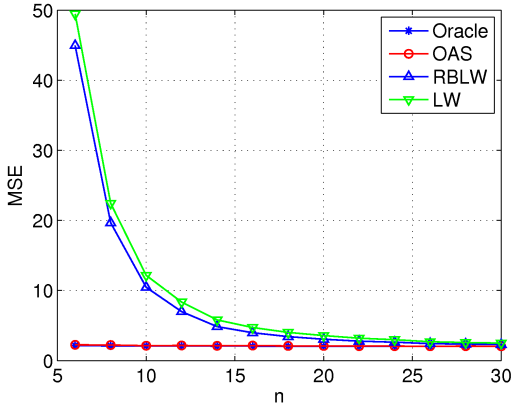
To test the MSE of the covariance estimators we designed two sets of experiments with different shapes of $\mathbf{\Sigma}$. Such covariance matrices have been used to study covariance estimators in [25]. We use (2.14), (2.19) and (2.23) to calculate the shrinkage coefficients for the LW, the RBLW and the OAS estimators. For comparison, the oracle estimator (2.5) uses the true $\mathbf{\Sigma}$ and is included as a benchmark lower bound on MSE for comparison. For all simulations, we set $p = 100$ and let n range from 6 to 30. Each simulation is repeated 5000 times and the MSE and shrinkage coefficients are plotted as a function of n. The 95% confidence intervals of the MSE and shrinkage coefficients were found to be smaller than the marker size and are omitted in the figures.

In the first experiment, an autoregressive covariance structured $\mathbf{\Sigma}$ is used. We let $\mathbf{\Sigma}$ be the covariance matrix of a Gaussian AR(1) process,

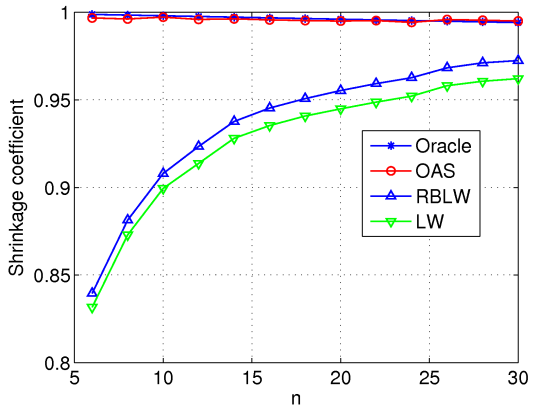
$$\mathbf{\Sigma}_{ij} = r^{|i-j|}, \tag{2.38}$$

where $\mathbf{\Sigma}_{ij}$ denotes the entry of $\mathbf{\Sigma}$ in row i and column j. We take $r = 0.1, 0.5$ and 0.9 for the different simulations reported below. Figs. 2.1(a)-2.3(a) show the MSE of the estimators for different values of r. Figs. 2.1(b)-2.3(b) show the corresponding shrinkage coefficients.

In Fig. 2.4 we plot the MSE of the first three iterations obtained by the iterative procedure in (2.21) and (2.20). For comparison we also plot the results of the OAS and the oracle estimator. We set $r = 0.5$ in this example and start the iterations with the initial guess $\widehat{\mathbf{\Sigma}}_0 = \widehat{\mathbf{S}}$. From Fig. 2.4 it can be seen that as the iterations proceed, the MSE gradually decreases towards that of the OAS estimator, which is very close to that of the oracle.



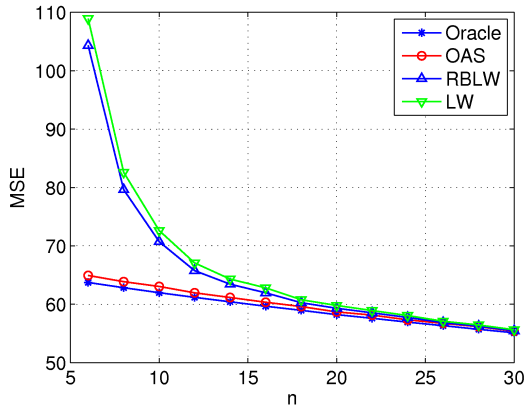
(a) MSE



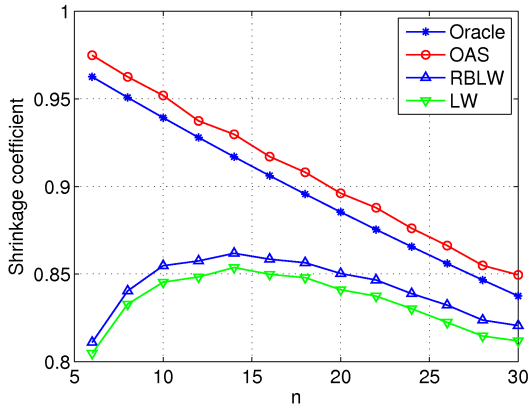
(b) Shrinkage coefficients

Figure 2.1: AR(1) process: Comparison of covariance estimators when $p = 100$, $r = 0.1$.

In the second experiment, we set $\mathbf{\Sigma}$ as the covariance matrix associated with the increment process of fractional Brownian motion (FBM) exhibiting long-range dependence. Such processes are often used to model internet traffic [53] and other

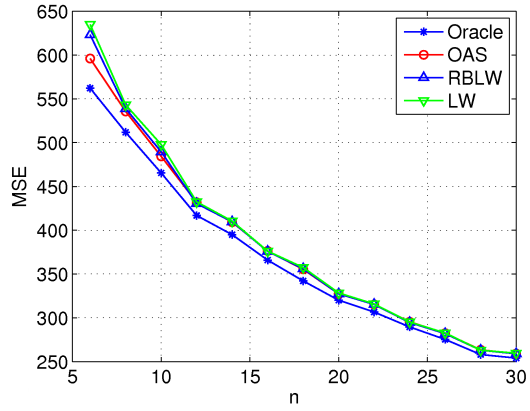


(a) MSE

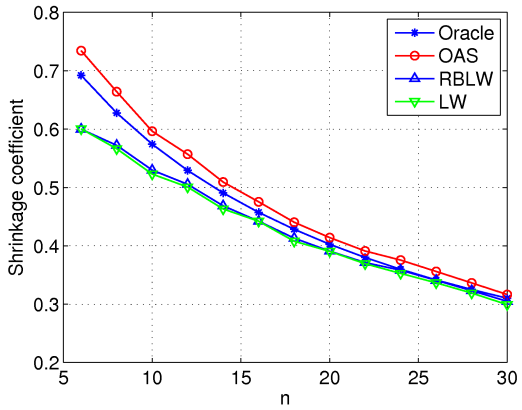


(b) Shrinkage coefficients

Figure 2.2: AR(1) process: Comparison of covariance estimators when $p = 100$, $r = 0.5$.



(a) MSE



(b) Shrinkage coefficients

Figure 2.3: AR(1) process: Comparison of covariance estimators when $p = 100$, $r = 0.9$.

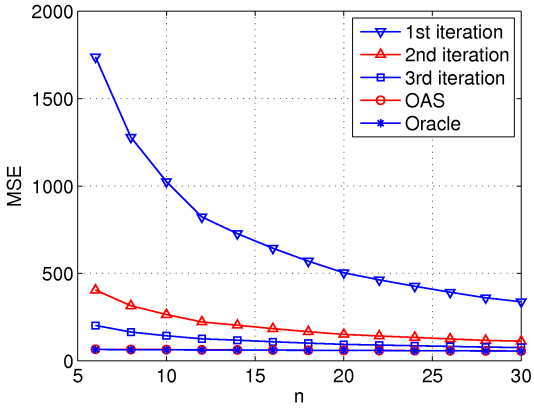
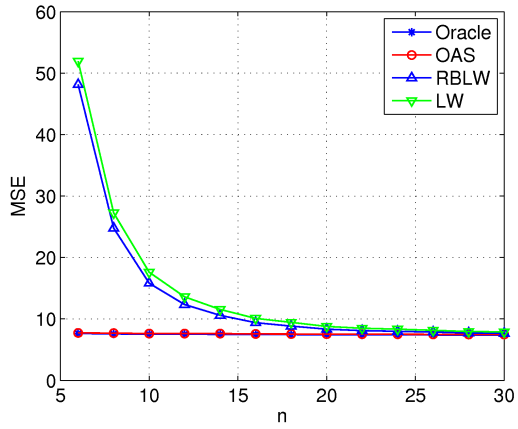
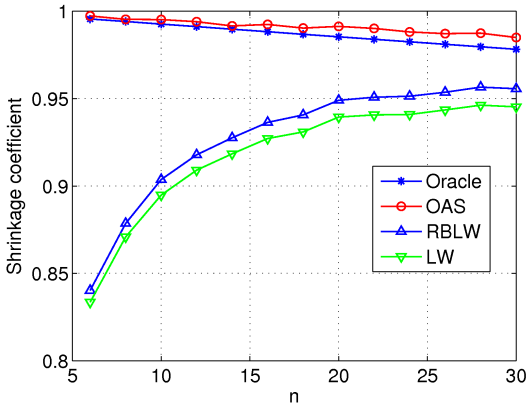


Figure 2.4: AR(1) process: Comparison of MSE in different iterations, when $p = 100$, $r = 0.5$

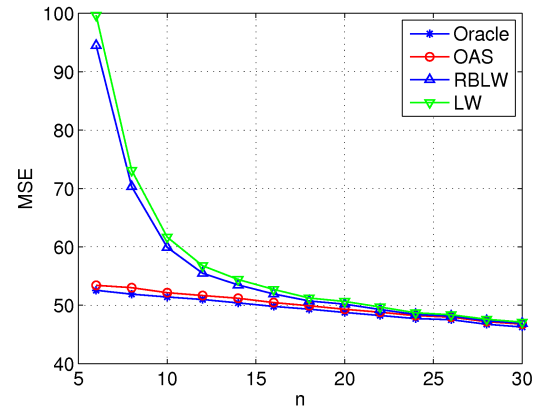


(a) MSE

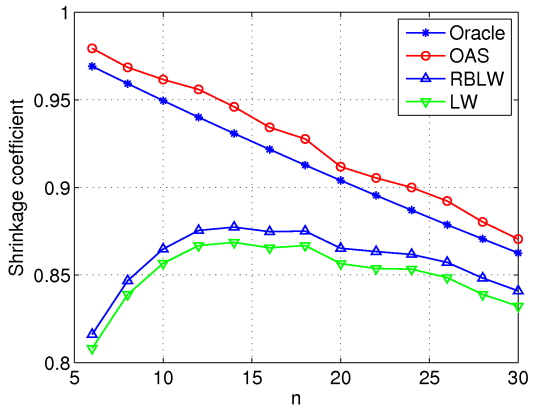


(b) Shrinkage coefficients

Figure 2.5: Incremental FBM process: Comparison of covariance estimators when $p = 100$, $h = 0.6$.

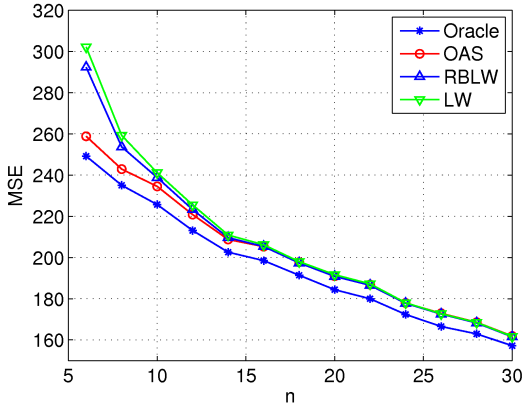


(a) MSE

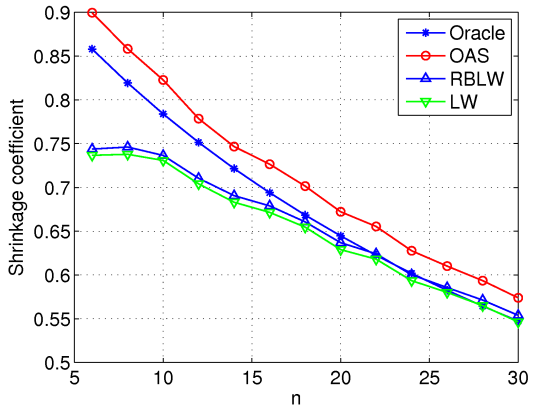


(b) Shrinkage coefficients

Figure 2.6: Incremental FBM process: Comparison of covariance estimators when $p = 100$, $h = 0.7$.



(a) MSE



(b) Shrinkage coefficients

Figure 2.7: Incremental FBM process: Comparison of covariance estimators when $p = 100$, $h = 0.8$.

complex phenomena. The form of the covariance matrix is given by

$$\mathbf{\Sigma}_{ij} = \frac{1}{2}\left[(|i-j|+1)^{2h} - 2|i-j|^{2h} + (|i-j|-1)^{2h}\right], \tag{2.39}$$

where $h \in [0.5, 1]$ is the so-called Hurst parameter. The typical value of h is below 0.9 in practical applications. We choose h equal to 0.6, 0.7 and 0.8. The MSE and shrinkage coefficients are plotted in Figs. 2.5(a)-2.7(a) and Figs. 2.5(b)-2.7(b), respectively.

From the simulation results in the above two experiments, it is evident that the OAS estimator performs very closely to the ideal oracle estimator. When n is small, the OAS significantly outperforms the LW and the RBLW. The RBLW improves slightly upon the LW, but this is not obvious at the scale of the plots shown in the figures. As expected, all the estimators converge to a common value when n increases.

As indicated in (2.5) and shown from simulation results, the oracle shrinkage coefficient ρ_O decreases in the sample number n. This makes sense since $(1 - \rho_O)$ can be regarded as a measure of "confidence" assigned to $\widehat{\mathbf{S}}$. Intuitively, as more observations are available, one acquires higher confidence in the sample covariance $\widehat{\mathbf{S}}$ and therefore ρ_O decreases. This characteristic is exhibited by $\hat{\rho}^*_{OAS}$ but not by $\hat{\rho}^*_{RBLW}$ and $\hat{\rho}^*_{LW}$. This may partially explain why OAS outperforms RBLW and LW for small samples.

It can be shown that all the estimators perform better when the sphericity of $\mathbf{\Sigma}$ increases, corresponding to small values of r and h. Indeed, the eigenvalues of $\mathbf{\Sigma}$ are more dispersed as r and h increases. As the dispersion of the eigenvalues is inversely related to the sphericity, larger sphericity in $\mathbf{\Sigma}$ indicates that better performance can be obtained by shrinking the sample covariance towards identity.

Our experience through numerous simulations with arbitrary parameters suggests

Table 2.1: Incremental FRM process: comparison of MSE and shrinkage coefficients when $h = 0.9, n = 20, p = 100$.

	MSE	Shrinkage coefficient
Oracle	428.9972	0.2675
OAS	475.2691	0.3043
RBLW	472.8206	0.2856
LW	475.5840	0.2867

that in practice the OAS is preferable to the RBLW. However, as the RBLW is provably better than the LW there exists counter examples. For the incremental FBM covariance $\mathbf{\Sigma}$ in (2.39), we set $h = 0.9, n = 20, p = 100$. The simulation is repeated for 5000 times and the result is shown in Table 1, where $\text{MSE}(\widehat{\mathbf{\Sigma}}_{RBLW}) < \text{MSE}(\widehat{\mathbf{\Sigma}}_{OAS}) < \text{MSE}(\widehat{\mathbf{\Sigma}}_{LW})$. The differences are very small but establish that the OAS estimator does not always dominate the RBLW. However, we suspect that this will only occur when $\mathbf{\Sigma}$ has a very small sphericity, a case of less interest in practice as small sphericity of $\mathbf{\Sigma}$ would suggest a different shrinkage target than $\widehat{\mathbf{F}}$.

Finally, we would like to point out that there is no free lunch. The superiority of the RBLW and OAS methods over the LW is only guaranteed when the samples are truly Gaussian. When the Gaussian assumption does not hold, RBLW and OAS may perform worse than LW. We performed an experiment (data not shown) with a heavy tailed (multivariate Student-T) sample distribution that tends to produce outliers. We observed that the RBLW no longer dominates the LW, due to outliers in the data, and the OAS outperforms the LW only when $n < 10$. Thus outlier resistant modifications of OAS and RBLW are a worthy topic of investigation.

2.4.2 Application to the Capon beamformer

Next we applied the proposed shrinkage estimators to the signal processing problem of adaptive beamforming. Assume that a narrow-band signal of interest $s(t)$ impinges on an unperturbed uniform linear array (ULA) [54] comprised of p sensors.

The complex valued vector of n snapshots of the array output is

$$\mathbf{x}(t) = \mathbf{a}(\theta_s)s(t) + \mathbf{n}(t), \quad \text{for} \quad t = 1, \ldots, n, \tag{2.40}$$

where θ_s is parameter vector determining the location of the signal source and $\mathbf{a}(\theta)$ is the array response for a generic source location θ. Specifically,

$$\mathbf{a}(\theta) = [1, e^{-j\omega}, e^{-j2\omega}, \ldots, e^{-j(p-1)\omega}]^T, \tag{2.41}$$

where ω is the spatial frequency. The noise/interference vector $\mathbf{n}(t)$ is assumed to be zero mean i.i.d. Gaussian distributed. We model the unknown $s(t)$ as a zero mean i.i.d. Gaussian process.

In order to recover the unknown $s(t)$ the Capon beamformer [54] linearly combines the array output $\mathbf{x}(t)$ using a vector of weights $\mathbf{w}$, calculated by

$$\mathbf{w} = \frac{\mathbf{\Sigma}^{-1}\mathbf{a}(\theta_s)}{\mathbf{a}(\theta_s)^H\mathbf{\Sigma}^{-1}\mathbf{a}(\theta_s)}, \tag{2.42}$$

where $\mathbf{\Sigma}$ is the covariance of $\mathbf{x}(t)$. The covariance $\mathbf{\Sigma}$ is unknown while the array response $\mathbf{a}(\theta)$ and the source direction-of-arrival (DOA) θ_s are known. After obtaining the weight vector $\mathbf{w}$, the signal of interest $s(t)$ is estimated by $\mathbf{w}^H\mathbf{x}(t)$.

To implement (2.42) the matrix $\mathbf{\Sigma}$ needs to be estimated. In [42] it was shown that using the LW estimator could substantially improve Capon beamformer performance over conventional methods. As we will see below, the OAS and the RBLW shrinkage estimators can yield even better results.

Note that the signal and the noise processes are complex valued and $\mathbf{\Sigma}$ is thus a complex (Hermitian symmetric) covariance matrix. To apply the OAS and RBLW estimators we use the same approach as used in [42] to extend the real LW covariance estimator to the complex case. Given a $p \times 1$ complex random vector $\mathbf{x}$, we represent

it as a $2p \times 1$ vector of its real and imaginary parts

$$\mathbf{x}_s = \begin{pmatrix} \mathrm{Re}\,(\mathbf{x}) \\ \\ \mathrm{Im}\,(\mathbf{x}) \end{pmatrix}. \tag{2.43}$$

Then the estimate of the complex covariance can be represented as

$$\widehat{\mathbf{\Sigma}}_s = \begin{pmatrix} \widehat{\mathbf{\Sigma}}_{rr} & \widehat{\mathbf{\Sigma}}_{ri} \\ \\ \widehat{\mathbf{\Sigma}}_{ir} & \widehat{\mathbf{\Sigma}}_{ii} \end{pmatrix} \tag{2.44}$$

where $\widehat{\mathbf{\Sigma}}_{rr}$, $\widehat{\mathbf{\Sigma}}_{ri}$, $\widehat{\mathbf{\Sigma}}_{ir}$ and $\widehat{\mathbf{\Sigma}}_{ii}$ are $p \times p$ sub-matrices. The real representation (2.44) can be mapped to the full complex covariance matrix $\mathbf{\Sigma}$ as

$$\widehat{\mathbf{\Sigma}} = \left(\widehat{\mathbf{\Sigma}}_{rr} + \widehat{\mathbf{\Sigma}}_{ii}\right) + j\left(\widehat{\mathbf{\Sigma}}_{ir} - \widehat{\mathbf{\Sigma}}_{ri}\right). \tag{2.45}$$

Using this representation we can easily extend the real valued LW, RBLW and OAS estimators to complex scenarios. As pointed in [42], this approach does not preserve the circular property of $\widehat{\mathbf{\Sigma}}_s$, *i.e.,* ,

$$\widehat{\mathbf{\Sigma}}_{rr} = \widehat{\mathbf{\Sigma}}_{ii},\ \widehat{\mathbf{\Sigma}}_{ri} = -\widehat{\mathbf{\Sigma}}_{ir}. \tag{2.46}$$

We note that exploiting the circular property in $\widehat{\mathbf{\Sigma}}_s$ would yield better estimators for complex covariance matrices. For purposes of this simple illustration we implement the simple methodology described above for a beamforming problem.

We conduct the beamforming simulation as follows. A ULA of $p = 10$ sensor elements with half wavelength spacing is assumed and three signals were simulated as impinging on the array. The signal of interest has a DOA $\theta_s = 20^\circ$ and a power $\sigma_s^2 = 10$ dB above the complex Gaussian sensor noise. The other two signals are mutually independent interferences. One is at DOA angle of $\theta_{i1} = -30^\circ$ and the other one is close to the source of interest with its angular location corresponding to a spatial frequency of

$$\omega_{i2} = \pi \sin(\theta_s) + 2\pi\frac{\gamma}{p}$$

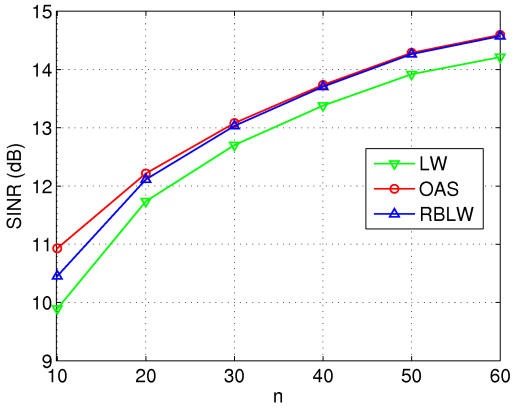


Figure 2.8: Comparison between different covariance shrinkage estimators in the Capon beamformer. SINR is plotted versus number of snapshots n. OAS achieves as much as 1 dB improvement over the LW.

where γ is set to 0.9. Each signal has power 15 dB above the sensor noise.

We implemented the complex versions of the LW, the RBLW and the OAS covariance estimators, described above, and used them in place of $\mathbf{\Sigma}$ in the Capon beamformer expression (2.42). The beamforming performance gain is measured by the SINR defined as [42]

$$\text{mean SINR} = \frac{1}{K}\sum_{k=1}^{K}\frac{\sigma_s^2\left|\hat{\mathbf{w}}_k^H\mathbf{a}\left(\theta_s\right)\right|^2}{\hat{\mathbf{w}}_k^H[\mathbf{\Sigma}-\sigma_s^2\mathbf{a}(\theta_s)\mathbf{a}(\theta_s)^H]\hat{\mathbf{w}}_k}, \tag{2.47}$$

where K is the number of Monte-Carlo simulations and $\hat{\mathbf{w}}_k$ is the weight vector obtained by (2.42) in the kth simulation. Here $K = 5000$ and n varies from 10 to 60 in step of 5 snap shots. The gain is shown in Fig. 2.8. In [42] it was reported that the LW estimator achieves the best SINR performances among several contemporary Capon-type beamformers. It can be seen in Fig. 2.8 that the RBLW and the OAS do even better, improving upon the LW estimator. Note also that the greatest improvement for OAS in the small n regime is observed.

2.5 Conclusion

In this work we introduced two new shrinkage algorithms to estimate covariance matrices when the samples are Gaussian distributed. The RBLW estimator was shown to improve upon the state-of-the-art LW method by virtue of the Rao-Blackwell theorem. The OAS estimator was developed by iterating on the optimal oracle estimate, where the limiting form was determined analytically. The RBLW provably dominates the LW, and the OAS empirically outperforms both the RBLW and the LW in most experiments we have conducted. The proposed OAS and RBLW estimators have simple explicit expressions and are easy to implement. Furthermore, they share similar structure differing only in the form of the shrinkage coefficients. We applied these estimators to the Capon beamformer and obtained significant gains in performance as compared to the LW Capon beamformer implementation.

Another important direction is to explore the outlier resistant modifications to the OAS and RBLW estimators presented here. We will investigate more robust shrinkage estimators in the next chapter.

2.6 Appendix

In this appendix we prove Theorem II.2. Theorem II.2 is non-trivial and requires careful treatment using results from the theory of Haar measure and singular Wishart distributions. The proof will require several intermediate results stated as lemmas. We begin with a definition.

Definition II.4. Let $\{\mathbf{x}_i\}_{i=1}^n$ be a sample of p-dimensional i.i.d. Gaussian vectors with mean zero and covariance $\mathbf{\Sigma}$. Define a $p \times n$ matrix $\mathbf{X}$ as

$$\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n)\,. \tag{2.48}$$

Denote $r = \min(p, n)$ and define the singular value decomposition on $\mathbf{X}$ as

$$\mathbf{X} = \mathbf{H}\mathbf{\Lambda}\mathbf{Q}, \tag{2.49}$$

where $\mathbf{H}$ is a $p \times r$ matrix such that $\mathbf{H}^T\mathbf{H} = \mathbf{I}$, $\mathbf{\Lambda}$ is a $r \times r$ diagonal matrix in probability 1, comprised of the singular values of $\mathbf{X}$, and $\mathbf{Q}$ is a $r \times n$ matrix such that $\mathbf{Q}\mathbf{Q}^T = \mathbf{I}$.

Next we state and prove three lemmas.

Lemma II.5. *Let* $(\mathbf{H}, \mathbf{\Lambda}, \mathbf{Q})$ *be matrices defined in Definition II.4. Then* $\mathbf{Q}$ *is independent of* $\mathbf{H}$ *and* $\mathbf{\Lambda}$.

Proof. For the case $n \leq p$, $\mathbf{H}$ is a $p \times n$ matrix, $\mathbf{\Lambda}$ is a $n \times n$ square diagonal matrix and $\mathbf{Q}$ is a $n \times n$ orthogonal matrix. The pdf of $\mathbf{X}$ is

$$p(\mathbf{X}) = \frac{1}{(2\pi)^{pn/2}\det(\Sigma)^{n/2}} e^{-\frac{1}{2}\mathrm{Tr}\left(\mathbf{X}\mathbf{X}^T\Sigma^{-1}\right)}. \tag{2.50}$$

Since $\mathbf{X}\mathbf{X}^T = \mathbf{H}\mathbf{\Lambda}\mathbf{\Lambda}^T\mathbf{H}^T$, the joint pdf of $(\mathbf{H}, \mathbf{\Lambda}, \mathbf{Q})$ is

$$\begin{aligned} &p(\mathbf{H}, \mathbf{\Lambda}, \mathbf{Q}) = \\ &\frac{1}{(2\pi)^{pn/2}\det(\Sigma)^{n/2}} e^{-\frac{1}{2}\mathrm{Tr}\left(\mathbf{H}\mathbf{\Lambda}\mathbf{\Lambda}^T\mathbf{H}^T\Sigma^{-1}\right)} J(\mathbf{X} \to \mathbf{H}, \mathbf{\Lambda}, \mathbf{Q}), \end{aligned} \tag{2.51}$$

where $J(\mathbf{X} \to \mathbf{H}, \mathbf{\Lambda}, \mathbf{Q})$ is the Jacobian converting from $\mathbf{X}$ to $(\mathbf{H}, \mathbf{\Lambda}, \mathbf{Q})$. According to Lemma 2.4 of [55],

$$\begin{aligned} &J(\mathbf{X} \to \mathbf{H}, \mathbf{\Lambda}, \mathbf{Q}) = \\ &2^{-n}\det(\mathbf{\Lambda})^{p-n} \prod_{j<k}^{n} \left(\lambda_j^2 - \lambda_k^2\right) g_{n,p}(\mathbf{H})\, g_{n,n}(\mathbf{Q}), \end{aligned} \tag{2.52}$$

where λ_j denotes the j-th diagonal element of $\mathbf{\Lambda}$ and $g_{n,p}(\mathbf{H})$ and $g_{n,n}(\mathbf{Q})$ are functions of $\mathbf{H}$ and $\mathbf{Q}$ defined in [55].

Substituting (2.52) into (2.51), $p(\mathbf{H}, \mathbf{\Lambda}, \mathbf{Q})$ can be factorized into functions of $(\mathbf{H}, \mathbf{\Lambda})$ and $\mathbf{Q}$. Therefore, $\mathbf{Q}$ is independent of $\mathbf{H}$ and $\mathbf{\Lambda}$.

Similarly, one can show that $\mathbf{Q}$ is independent of $\mathbf{H}$ and $\mathbf{\Lambda}$ when $n > p$. □

Lemma II.6. *Let* $\mathbf{Q}$ *be a matrix defined in Definition II.4. Denote* $\mathbf{q}$ *as an arbitrary column vector of* $\mathbf{Q}$ *and* q_j *as the* j*-th element of* $\mathbf{q}$*. Then*

$$E\left\{q_j^4\right\} = \frac{3}{n(n+2)} \tag{2.53}$$

and

$$E\left\{q_k^2 q_j^2\right\} = \frac{1}{n(n+2)}, \quad k \neq j. \tag{2.54}$$

Proof. The proof is different for the cases that $n \leq p$ and $n > p$, which are treated separately.

(1) Case $n \leq p$*:*

In this case, $\mathbf{Q}$ is a real Haar matrix and is isotropically distributed [56, 57, 58], *i.e.*, for any orthogonal matrices $\mathbf{\Phi}$ and $\mathbf{\Psi}$ which are independent with $\mathbf{Q}$, $\mathbf{\Phi Q}$ and $\mathbf{Q\Psi}$ have the same pdf of $\mathbf{Q}$:

$$p(\mathbf{\Phi Q}) = p(\mathbf{Q\Psi}) = p(\mathbf{Q}). \tag{2.55}$$

Following [59] in the complex case, we now use (2.55) to calculate the fourth order moments of elements of $\mathbf{Q}$. Since $\mathbf{Q}$ and

$$\begin{bmatrix} \cos\theta & \sin\theta & & & \\ -\sin\theta & \cos\theta & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} \mathbf{Q}$$

are also identically distributed, we have

$$
\begin{aligned}
& E\left\{\mathbf{Q}_{11}^4\right\} \\
& = E\left\{\left(\mathbf{Q}_{11}\cos\theta + \mathbf{Q}_{21}\sin\theta\right)^4\right\} \\
& = \cos^4\theta E\left\{\mathbf{Q}_{11}^4\right\} + \sin^4\theta E\left\{{\mathbf{Q}_{21}}^4\right\} \\
& \quad + 6\cos^2\theta\sin^2\theta E\left\{\mathbf{Q}_{11}^2\mathbf{Q}_{21}^2\right\} \\
& \quad + 2\cos^3\theta\sin\theta E\left\{\mathbf{Q}_{11}^3\mathbf{Q}_{21}\right\} + 2\cos\theta\sin^3\theta E\left\{\mathbf{Q}_{11}\mathbf{Q}_{21}^3\right\}
\end{aligned} \tag{2.56}
$$

By taking $\theta = -\theta$ in (2.56), it is easy to see that

$$2\cos^3\theta\sin\theta E\left\{\mathbf{Q}_{11}^3\mathbf{Q}_{21}\right\} + 2\cos\theta\sin^3\theta E\left\{\mathbf{Q}_{11}\mathbf{Q}_{21}^3\right\} = 0.$$

The elements of $[\mathbf{Q}_{ii}]$ are identically distributed. We thus have $E\left\{\mathbf{Q}_{11}^4\right\} = E\left\{{\mathbf{Q}_{21}}^4\right\}$, and hence

$$
\begin{aligned}
& E\left\{\mathbf{Q}_{11}^4\right\} \\
& = \left(\cos^4\theta + \sin^4\theta\right) E\left\{\mathbf{Q}_{11}^4\right\} + 6\cos^2\theta\sin^2\theta E\left\{\mathbf{Q}_{11}^2\mathbf{Q}_{21}^2\right\}.
\end{aligned} \tag{2.57}
$$

By taking $\theta = \pi/3$,

$$E\left\{\mathbf{Q}_{11}^4\right\} = 3E\left\{\mathbf{Q}_{11}^2\mathbf{Q}_{21}^2\right\}. \tag{2.58}$$

Now we consider $E\left\{\left(\sum_{j=1}^n \mathbf{Q}_{j1}^2\right)^2\right\}$. Since $\mathbf{Q}^T\mathbf{Q} = \mathbf{Q}\mathbf{Q}^T = \mathbf{I}$, $\sum\limits_{j=1}^n \mathbf{Q}_{j1}^2 = 1$. This implies

$$
\begin{aligned}
1 &= \sum_{j=1}^n E\left\{\mathbf{Q}_{j1}^4\right\} + \sum_{j\neq k} E\left\{\mathbf{Q}_{j1}^2\mathbf{Q}_{k1}^2\right\} \\
&= nE\left\{\mathbf{Q}_{11}^4\right\} + n(n-1)E\left\{\mathbf{Q}_{11}^2\mathbf{Q}_{21}^2\right\}.
\end{aligned} \tag{2.59}
$$

Substituting (2.58) into (2.59), we obtain that

$$E\left\{\mathbf{Q}_{11}^4\right\} = \frac{3}{n(n+2)}, \tag{2.60}$$

and

$$E\left\{\mathbf{Q}_{11}^2\mathbf{Q}_{21}^2\right\} = \frac{1}{n(n+2)}. \tag{2.61}$$

It is easy to see that $E\left\{q_j^4\right\} = E\left\{\mathbf{Q}_{11}^4\right\}$ and $E\left\{q_j^2q_k^2\right\} = E\left\{\mathbf{Q}_{11}^2\mathbf{Q}_{21}^2\right\}$. Therefore (2.53) and (2.54) are proved for the case of $n \leq p$.

(2) Case $n > p$:

The pdf of $\mathbf{q}$ can be obtained by Lemma 2.2 of [55]

$$p(\mathbf{q}) = C_1 \det\left(\mathbf{I} - \mathbf{q}\mathbf{q}^T\right)^{(n-p-2)/2} I(\mathbf{q}\mathbf{q}^T \prec \mathbf{I}), \tag{2.62}$$

where

$$C_1 = \frac{\pi^{-p/2}\Gamma\{n/2\}}{\Gamma\{(n-p)/2\}}, \tag{2.63}$$

and $I(\cdot)$ is the indicator function specifying the support of $\mathbf{q}$. Eq. (2.62) indicates that the elements of $\mathbf{q}$ are identically distributed. Therefore, $E\left\{q_j^4\right\} = E\left[q_1^4\right\}$ and $E\left\{q_j^2q_k^2\right\} = E\left\{q_1^2q_2^2\right\}$. By the definition of expectation,

$$E\left\{q_1^4\right\} = C_1 \int_{\mathbf{q}\mathbf{q}^T \prec \mathbf{I}} q_1^4 \det\left(\mathbf{I} - \mathbf{q}\mathbf{q}^T\right)^{(n-p-2)/2} d\mathbf{q}, \tag{2.64}$$

and

$$E\left\{q_1^2q_2^2\right\} = C_1 \int_{\mathbf{q}\mathbf{q}^T \prec \mathbf{I}} q_1^2q_2^2 \det\left(\mathbf{I} - \mathbf{q}\mathbf{q}^T\right)^{(n-p-2)/2} d\mathbf{q}. \tag{2.65}$$

Noting that

$$\mathbf{q}\mathbf{q}^T \prec \mathbf{I} \Leftrightarrow \mathbf{q}^T\mathbf{q} < 1 \tag{2.66}$$

and

$$\det\left(\mathbf{I} - \mathbf{q}\mathbf{q}^T\right) = 1 - \mathbf{q}^T\mathbf{q}, \tag{2.67}$$

we have

$$
\begin{aligned}
E\left\{q_1^4\right\} &= C_1 \int_{\mathbf{q}^T\mathbf{q}<1} q_1^4 (1-\mathbf{q}^T\mathbf{q})^{\frac{1}{2}(n-p-2)} d\mathbf{q} \\
&= C_1 \int_{\sum_{j=1}^{p} q_j^2<1} q_1^4 \left(1-\sum_{j=1}^{p} q_j^2\right)^{\frac{1}{2}(n-p-2)} dq_1 \ldots dq_p.
\end{aligned} \tag{2.68}
$$

By changing variable of integration $(q_1, q_2, \cdots, q_p)$ to $(r, \theta_1, \theta_2, \cdots, \theta_{p-1})$ such that

$$
\begin{cases}
q_1 &= r\cos\theta_1 \\
q_2 &= r\sin\theta_1\cos\theta_2 \\
q_3 &= r\sin\theta_1\sin\theta_2\cos\theta_3 \\
\vdots & \quad \vdots \\
q_{p-1} &= r\sin\theta_1\sin\theta_2\cdots\sin\theta_{p-2}\cos\theta_{p-1} \\
q_p &= r\sin\theta_1\sin\theta_2\cdots\sin\theta_{p-2}\sin\theta_{p-1}
\end{cases}, \tag{2.69}
$$

we obtain

$$
\begin{aligned}
E\left\{q_1^4\right\} &= C_1 \int_0^{\pi} d\theta_1 \int_0^{\pi} d\theta_2 \cdots \int_0^{\pi} d\theta_{p-2} \int_0^{2\pi} d\theta_{p-1} \\
&\cdot \int_0^1 r^4 \cos^4\theta_1 \left(1-r^2\right)^{\frac{1}{2}(n-p-2)} \left|\frac{\partial\left(q_1, \cdots, q_p\right)}{\partial\left(r, \theta_1, \cdots, \theta_{p-1}\right)}\right| dr,
\end{aligned} \tag{2.70}
$$

where

$$
\left|\frac{\partial\left(q_1, \cdots, q_p\right)}{\partial\left(r, \theta_1, \cdots, \theta_{p-1}\right)}\right| = r^{p-1}\sin^{p-2}\theta_1 \sin^{p-3}\theta_2 \cdots \sin\theta_{p-2}
$$

is the Jacobian associated with the change of variable.

Therefore,

$$
\begin{aligned}
E\left\{q_1^4\right\} &= C_1 \cdot \int_0^{\pi} \cos^4\theta_1 \sin^{p-2}\theta_1 d\theta_1 \cdot \int_0^{\pi} \sin^{p-3}\theta_2 d\theta_2 \\
&\cdot \int_0^{\pi} \sin^{p-4}\theta_3 d\theta_3 \cdots \int_0^{\pi} \sin\theta_{p-2} d\theta_{p-2} \int_0^{2\pi} d\theta_{p-1} \\
&\cdot \int_0^1 r^{p+3}\left(1-r^2\right)^{\frac{1}{2}(n-p-2)} dr \\
&= \frac{\pi^{-p/2}\Gamma\{n/2\}}{\Gamma\{(n-p)/2\}} \cdot \frac{3\pi^{\frac{1}{2}}}{4} \frac{\Gamma\{(p-1)/2\}}{\Gamma\{(p+4)/2\}} \cdot \pi^{\frac{1}{2}} \frac{\Gamma\{(p-2)/2\}}{\Gamma\{(p-1)/2\}} \\
&\cdot \pi^{\frac{1}{2}} \frac{\Gamma\{(p-3)/2\}}{\Gamma\{(p-2)/2\}} \cdots \pi^{\frac{1}{2}} \frac{\Gamma\{3/2\}}{\Gamma\{5/2\}} \cdot \pi^{\frac{1}{2}} \frac{\Gamma\{1\}}{\Gamma\{3/2\}} \cdot 2\pi \\
&\cdot \int_0^1 r^{p+3}\left(1-r^2\right)^{\frac{1}{2}(n-p-2)} dr \\
&= \frac{3}{2} \frac{\Gamma\{n/2\}}{\Gamma\{(n-p)/2\}\Gamma\{p/2+2\}} \int_0^1 r^{p+3}\left(1-r^2\right)^{\frac{1}{2}(n-p-2)} dr \\
&= \frac{3}{2} \frac{\Gamma\{n/2\}}{\Gamma\{(n-p)/2\}\Gamma\{p/2+2\}} \cdot \frac{1}{2} \frac{\Gamma\{(n-p)/2\}\Gamma\{p/2+2\}}{\Gamma\{n/2+2\}} \\
&= \frac{3\Gamma\{n/2\}}{4\Gamma\{n/2+2\}} \\
&= \frac{3}{n(n+2)}.
\end{aligned}
\tag{2.71}
$$

Similarly,

$$
\begin{aligned}
E\left\{q_1^2q_2^2\right\} &= C_1\int_{\sum_{k=1}^{p}q_k^2<1} q_1^2q_2^2\left(1-\sum_{k=1}^{p}q_k^2\right)^{\frac{1}{2}(n-p-2)} dq_1\dots dq_p \\
&= C_1\int_0^{\pi}d\theta_1\int_0^{\pi}d\theta_2\cdots\int_0^{\pi}d\theta_{p-2}\int_0^{2\pi}d\theta_{p-1} \\
&\cdot\int_0^1 r^2\cos^2\theta_1 r^2\sin^2\theta_1\cos^2\theta_2\big(1-r^2\big)^{\frac{1}{2}(n-p-2)} \\
&\cdot\left|\frac{\partial\left(q_1,\cdots,q_p\right)}{\partial\left(r,\theta_1,\cdots,\theta_{p-1}\right)}\right| dr \\
&= C_1\cdot\int_0^{\pi}\cos^2\theta_1\sin^p\theta_1 d\theta_1\cdot\int_0^{\pi}\cos^2\theta_2\sin^{p-3}\theta_2 d\theta_2 \\
&\cdot\int_0^{\pi}\sin^{p-4}\theta_3 d\theta_3\cdot\int_0^{\pi}\sin^{p-5}\theta_4 d\theta_4\cdots\int_0^{\pi}\sin\theta_{p-2}d\theta_{p-2} \\
&\cdot\int_0^{2\pi}d\theta_{p-1}\cdot\int_0^1 r^{p+3}\left(1-r^2\right)^{\frac{1}{2}(n-p-2)}dr \\
&= \frac{\pi^{-p/2}\Gamma\{n/2\}}{\Gamma\{(n-p)/2\}}\cdot\frac{\pi^{\frac{1}{2}}}{2}\frac{\Gamma\{(p+1)/2\}}{\Gamma\{p/2+2\}}\cdot\frac{\pi^{\frac{1}{2}}}{2}\frac{\Gamma\{(p-2)/2\}}{\Gamma\{(p+1)/2\}} \\
&\cdot\pi^{\frac{1}{2}}\frac{\Gamma\{(p-3)/2\}}{\Gamma\{(p-2)/2\}}\cdot\pi^{\frac{1}{2}}\frac{\Gamma\{(p-4)/2\}}{\Gamma\{(p-3)/2\}}\cdots\pi^{\frac{1}{2}}\frac{\Gamma\{1\}}{\Gamma\{3/2\}} \\
&\cdot 2\pi\cdot\frac{1}{2}\frac{\Gamma\{(n-p)/2\}\Gamma\{p/2+2\}}{\Gamma\{n/2+2\}} \\
&= \frac{1}{n(n+2)}.
\end{aligned}
\tag{2.72}
$$

Therefore, (2.53) and (2.54) are proved for the case when $n > p$. This completes the proof of Lemma II.6. □

Lemma II.7. *Let $\widehat{\mathbf{S}}$ be the sample covariance of a set of p-dimensional vectors $\{\mathbf{x}_i\}_{i=1}^n$. If $\{\mathbf{x}_i\}_{i=1}^n$ are i.i.d. Gaussian vectors with covariance $\mathbf{\Sigma}$,*

$$
E\left\{\left\|\mathbf{x}_i\right\|_2^4\middle|\,\widehat{\mathbf{S}}\right\} = \frac{n}{n+2}\left[2\mathrm{Tr}(\widehat{\mathbf{S}}^2)+\mathrm{Tr}^2(\widehat{\mathbf{S}})\right]. \tag{2.73}
$$

Proof. For simplicity, we work with the scaled covariance matrix $\mathbf{M}$ defined as

$$
\mathbf{M} = \sum_{i=1}^{n}\mathbf{x}_i\mathbf{x}_i^T = n\widehat{\mathbf{S}}, \tag{2.74}
$$

and calculate $E\left\{\left\|\mathbf{x}_i\right\|_2^4 \middle| \mathbf{M}\right\}$ instead of $E\left\{\left\|\mathbf{x}_i\right\|_2^4 \middle| \widehat{\mathbf{S}}\right\}$. We are then going to prove that

$$E\left\{\left\|\mathbf{x}_i\right\|_2^4 \middle| \mathbf{M}\right\} = \frac{1}{n\left(n+2\right)}\left(2\mathrm{Tr}\left(\mathbf{M}^2\right) + \mathrm{Tr}^2\left(\mathbf{M}\right)\right). \tag{2.75}$$

We use Lemma II.5 and Lemma II.6 to establish (2.75).

Let $\mathbf{X}$ and $(\mathbf{H}, \mathbf{\Lambda}, \mathbf{Q})$ be matrices defined in Definition II.4. Let $\mathbf{q}$ be the i-th column of $\mathbf{Q}$ defined in Definition II.4. Then

$$\mathbf{x}_i = \mathbf{H}\mathbf{\Lambda}\mathbf{q}. \tag{2.76}$$

Let

$$\mathbf{D} = \mathbf{\Lambda}^2. \tag{2.77}$$

Then

$$\mathbf{M} = \mathbf{X}\mathbf{X}^T = \mathbf{H}\mathbf{\Lambda}^2\mathbf{H}^T = \mathbf{H}\mathbf{D}\mathbf{H}^T, \tag{2.78}$$

and

$$\mathbf{x}_i^T\mathbf{x}_i = \mathbf{q}^T\mathbf{\Lambda}^T\mathbf{H}^T\mathbf{H}\mathbf{\Lambda}\mathbf{q} = \mathbf{q}^T\mathbf{D}\mathbf{q}. \tag{2.79}$$

Therefore we have

$$E\left\{\left\|\mathbf{x}_i\right\|_2^4 \middle| \mathbf{M}\right\} = E\left\{\left(\mathbf{q}^T\mathbf{D}\mathbf{q}\right)^2 \middle| \mathbf{M}\right\}. \tag{2.80}$$

According to Lemma II.5, $\mathbf{Q}$ is independent of $\mathbf{H}$ and $\mathbf{\Lambda}$. Since $\mathbf{q}$ is a function of $\mathbf{Q}$, $\mathbf{M}$ and $\mathbf{D}$ are functions of $\mathbf{H}$ and $\mathbf{\Lambda}$, $\mathbf{q}$ is independent of $\mathbf{M}$ and $\mathbf{D}$.

From the law of total expectation,

$$E\left\{\left(\mathbf{q}^T\mathbf{D}\mathbf{q}\right)^2 \middle| \mathbf{M}\right\} = E\left\{E\left\{\left(\mathbf{q}^T\mathbf{D}\mathbf{q}\right)^2 \middle| \mathbf{M}, \mathbf{D}\right\} \middle| \mathbf{M}\right\}. \tag{2.81}$$

Expand $\mathbf{q}^T\mathbf{D}\mathbf{q}$ as

$$\mathbf{q}^T\mathbf{D}\mathbf{q} = \sum_{j=1}^{n} d_j q_j^2, \tag{2.82}$$

where d_j is the j-th diagonal element of $\mathbf{D}$. Since $\mathbf{q}$ is independent of $\mathbf{M}$ and $\mathbf{D}$, according to Lemma II.6,

$$\begin{aligned}
&E\left\{ \left(\mathbf{q}^T\mathbf{D}\mathbf{q}\right)^2 \middle| \mathbf{M}, \mathbf{D} \right\} \\
&= E\left\{ \sum_{j=1}^{n} d_j^2 q_j^4 + \sum_{j\neq k} d_j d_k q_j^2 q_k^2 \middle| \mathbf{M}, \mathbf{D} \right\} \\
&= \sum_{j=1}^{n} d_j^2 E\left\{q_j^4\right\} + \sum_{j\neq k} d_j d_k E\left\{q_j^2 q_k^2\right\} \\
&= \frac{1}{n\left(n+2\right)} \left(3\sum_{j=1}^{n} d_j^2 + \sum_{j\neq k} d_j d_k \right) \\
&= \frac{1}{n\left(n+2\right)} \left(2\mathrm{Tr}\left(\mathbf{D}^2\right) + \mathrm{Tr}^2\left(\mathbf{D}\right) \right).
\end{aligned} \tag{2.83}$$

Since $\mathrm{Tr}\left(\mathbf{D}\right) = \mathrm{Tr}\left(\mathbf{M}\right)$ and $\mathrm{Tr}\left(\mathbf{D}^2\right) = \mathrm{Tr}\left(\mathbf{M}^2\right)$, substituting (2.83) into (2.81), we have

$$\begin{aligned}
&E\left\{ \left(\mathbf{q}^T\mathbf{D}\mathbf{q}\right)^2 \middle| \mathbf{M} \right\} \\
&= E\left\{ \frac{1}{n\left(n+2\right)} \left(2\mathrm{Tr}\left(\mathbf{D}^2\right) + \mathrm{Tr}^2\left(\mathbf{D}\right) \right) \middle| \mathbf{M} \right\} \\
&= E\left\{ \frac{1}{n\left(n+2\right)} \left(2\mathrm{Tr}\left(\mathbf{M}^2\right) + \mathrm{Tr}^2\left(\mathbf{M}\right) \right) \middle| \mathbf{M} \right\} \\
&= \frac{1}{n\left(n+2\right)} \left(2\mathrm{Tr}\left(\mathbf{M}^2\right) + \mathrm{Tr}^2\left(\mathbf{M}\right) \right).
\end{aligned} \tag{2.84}$$

□

Lemma 3 now allows us to prove Theorem II.2.

2.6.1 Proof of Theorem II.2

Proof.

$$\begin{aligned}\widehat{\mathbf{\Sigma}}_{RBLW} =& E\left\{\widehat{\mathbf{\Sigma}}_{LW}\middle|\widehat{\mathbf{S}}\right\}\\ =& E\left\{(1-\hat{\rho}_{LW})\,\widehat{\mathbf{S}}+\hat{\rho}_{LW}\widehat{\mathbf{F}}\middle|\widehat{\mathbf{S}}\right\}\\ =& \left(1-E\left\{\hat{\rho}_{LW}|\,\widehat{\mathbf{S}}\right\}\right)\widehat{\mathbf{S}}+E\left\{\hat{\rho}_{LW}\widehat{\mathbf{F}}\middle|\widehat{\mathbf{S}}\right\}.\end{aligned} \tag{2.85}$$

Therefore we obtain the shrinkage coefficient of $\widehat{\mathbf{\Sigma}}_{RBLW}$:

$$\begin{aligned}\hat{\rho}_{RBLW} =& E\left\{\hat{\rho}_{LW}|\,\widehat{\mathbf{S}}\right\}\\ =& \frac{\sum_{i=1}^{n} E\left\{\left\|\mathbf{x}_i\mathbf{x}_i^T-\widehat{\mathbf{S}}\right\|_F^2\middle|\widehat{\mathbf{S}}\right\}}{n^2\left[\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)-\mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right)/p\right]}.\end{aligned} \tag{2.86}$$

Note that

$$\begin{aligned}&\sum_{i=1}^{n} E\left\{\left\|\mathbf{x}_i\mathbf{x}_i^T-\widehat{\mathbf{S}}\right\|_F^2\middle|\widehat{\mathbf{S}}\right\}\\ &\quad=\sum_{i=1}^{n} E\left\{\|\mathbf{x}_i\|_2^4|\,\widehat{\mathbf{S}}\right\}-n\mathrm{Tr}(\widehat{\mathbf{S}}^2).\end{aligned} \tag{2.87}$$

From Lemma II.7, we have

$$\begin{aligned}&\sum_{i=1}^{n} E\left\{\left\|\mathbf{x}_i\mathbf{x}_i^T-\widehat{\mathbf{S}}\right\|_F^2\middle|\widehat{\mathbf{S}}\right\}\\ &\quad=\frac{n(n-2)}{n+2}\mathrm{Tr}\left(\widehat{\mathbf{S}}^2\right)+\frac{n^2}{n+2}\mathrm{Tr}^2\left(\widehat{\mathbf{S}}\right).\end{aligned} \tag{2.88}$$

Equation (2.17) is then obtained by substituting (2.88) into (2.86). □

CHAPTER III

Shrinkage Algorithms for Covariance Estimation: Elliptical Samples

This chapter addresses high dimensional covariance estimation for elliptical distributed samples, which are also known as spherically invariant random vectors (SIRV) or compound-Gaussian processes. Specifically we consider shrinkage methods that are suitable for high dimensional problems with a small number of samples (large p small n). We start from a classical robust covariance estimator (Tyler 1987, [31]), which is distribution-free within the family of elliptical distribution but inapplicable when $n < p$. Using a shrinkage coefficient, we regularize Tyler's fixed point iterations. We prove that, for all n and p, the proposed fixed point iterations converge to a unique limit regardless of the initial condition. Next, we propose a simple, closed-form and data dependent choice for the shrinkage coefficient, which is based on a minimum mean squared error framework. Simulations demonstrate that the proposed method achieves low estimation error and is robust to heavy-tailed samples. Finally, as a real world application we demonstrate the performance of the proposed technique in the context of activity/intrusion detection using a wireless sensor network.

3.1 Introduction

Estimating a covariance matrix (or a dispersion matrix) is a fundamental problem in statistical signal processing and much effort has been devoted to high-dimensional covariance estimation, which uses Steinian shrinkage [13, 17, 60] or other types of regularized methods such as [25, 18]. However, most of the high-dimensional estimators assume Gaussian distributed samples. This limits their usage in many important applications involving non-Gaussian and heavy-tailed samples. One exception is the Ledoit-Wolf estimator [17], where the authors shrink the sample covariance towards a scaled identity matrix and proposed a shrinkage coefficient which is asymptotically optimal for any distribution. On the other hand, traditional robust covariance estimators [30, 31, 32] designed for non-Gaussian samples generally require $n \gg p$ and are not suitable for "large p small n" problems. Therefore, the goal of our work is to develop a covariance estimator for problems that are both high dimensional and non-Gaussian. In this chapter, we model the samples using the elliptical distribution [61], which is also referred to as the spherically invariant random vector model (SIRV) [62, 63] or the compound-Gaussian process model [64]. As a flexible and popular alternative, the elliptical family encompasses a large number of important distributions such as Gaussian distribution, the multivariate Cauchy distribution, the multivariate exponential distribution, the multivariate Student-T distribution, the K-distribution and the Weibull distribution. The capability of modelling heavy-tails makes the elliptical distribution appealing in signal processing and related fields. Typical applications include radar detection [65, 64, 66, 67], speech signal processing [68], remote sensing [69], wireless fading channels modelling [63], financial engineering [70] and so forth.

A well-studied covariance estimator for elliptical distributions is the ML estimator based on normalized samples [31, 71, 72]. The estimator is derived as the solution to a fixed point equation by using fixed point iterations. It is distribution-free within the class of elliptical distributions and its performance advantages are well known in the $n \gg p$ regime. However, it is not suitable for the "large p small n" setting. Indeed, when $n < p$, the ML estimator as defined does not even exist. To avoid this problem the authors of [73] propose an iterative regularized ML estimator that employs diagonal loading and uses a heuristic procedure for selecting the regularization parameter. While they did not establish convergence and uniqueness [73] they empirically demonstrated that their algorithm has superior performance in the context of a radar application. Our approach is similar to [73] but is conceived in a Steinian shrinkage framework, where we establish convergence and uniqueness of the resultant iterative estimator. We also propose a general procedure of selecting the shrinkage coefficient for heavy-tailed homogeneous samples. For a fixed shrinkage coefficient, we prove that the regularized fixed iterations converge to a unique solution for all n and p, regardless of the initial condition. Next, following Ledoit-Wolf [17], we provide a simple closed-form expression for the shrinkage coefficient, based on minimizing mean-squared-error. The resultant coefficient is a function of the unknown true covariance and cannot be implemented in practice. Instead, we develop a data-dependent "plug-in" estimator approximation. Simulation results demonstrate that our estimator achieves superior performance for samples distributed within the elliptical family. Furthermore, for the case that the samples are truly Gaussian, we report very little performance degradation with respect to the shrinkage estimators designed specifically for Gaussian samples [60].

As a real world application we demonstrate the proposed estimator for activ-

ity/intrusion detection using an active wireless sensor network. We show that the measured data exhibit strong non-Gaussian behavior and demonstrate significant performance advantages of the proposed robust covariance estimator when used in a covariance-based anomaly detection algorithm.

The rest of the chapter is organized as follows. Section 3.2 provides a brief review of elliptical distributions and of Tyler's covariance estimation method. The regularized covariance estimator is introduced and derived in Section 3.3. We provide simulations and experimental results in Section 3.4 and Section 3.5, respectively. Section 3.6 summarizes our principal conclusions. The proof of theorems and lemmas are provided in the Appendix.

Notations: In the following, we depict vectors in lowercase boldface letters and matrices in uppercase boldface letters. $(\cdot)^T$ and $(\cdot)^H$ denote the transpose and conjugate transpose operator, respectively. $\mathrm{Tr}(\cdot)$ and $\det(\cdot)$ are the trace and the determinant of a matrix, respectively.

3.2 ML covariance estimation for elliptical distributions

3.2.1 Elliptical distribution

Let $\mathbf{x}$ be a $p \times 1$ real random vector generated by the following model

$$\mathbf{x} = \nu \mathbf{u}, \tag{3.1}$$

where ν is a real, positive random variable and $\mathbf{u}$ is a $p \times 1$ zero-mean, real Gaussian random vector with positive definite covariance $\mathbf{\Sigma}$. We assume that ν and $\mathbf{u}$ are statistically independent. The resulting random vector $\mathbf{x}$ is elliptically distributed and its probability density function (pdf) can be expressed by

$$p(\mathbf{x}) = \phi\left(\mathbf{x}^T \mathbf{\Sigma}^{-1} \mathbf{x}\right), \tag{3.2}$$

where $\phi(\cdot)$ is the characteristic function (Definition 2, pp. 5, [74]) related to the pdf of ν. The elliptical family encompasses many useful distributions in signal processing and related fields and includes: the Gaussian distribution itself, the K distribution, the Weibull distribution and many others. As stated above, elliptically distributed samples are also referred to as Spherically Invariant Random Vectors (SIRV) or compound Gaussian processes in signal processing.

3.2.2 ML estimation

Let $\{\mathbf{x}_i\}_{i=1}^n$ be a set of n independent and identically distributed (i.i.d.) samples drawn according to (3.1). As the covariance of $\mathbf{x}$ may not exist, our problem is formulated to estimate the covariance (dispersion) matrix $\mathbf{\Sigma}$ of $\mathbf{u}$ from $\{\mathbf{x}_i\}_{i=1}^n$. The model (3.1) is invariant to scaling of the covariance matrix $\mathbf{\Sigma}$. Therefore, without loss of generality, we assume that the covariance matrix is trace-normalized in the sense that $\mathrm{Tr}(\mathbf{\Sigma}) = p$.

The commonly used sample covariance, defined as

$$\widehat{\mathbf{S}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{x}_i \mathbf{x}_i^T, \tag{3.3}$$

is known to be a poor estimator of $\mathbf{\Sigma}$, especially when the samples are high-dimensional (large p) and/or heavy-tailed. Tyler's method [31] addresses this problem by working with the normalized samples:

$$\mathbf{s}_i = \frac{\mathbf{x}_i}{\|\mathbf{x}_i\|_2} = \frac{\mathbf{u}_i}{\|\mathbf{u}_i\|_2}, \tag{3.4}$$

for which the term ν in (3.1) drops out. The pdf of $\mathbf{s}_i$ is given by [74]

$$p(\mathbf{s}_i; \mathbf{\Sigma}) = \frac{\Gamma(p/2)}{2\pi^{p/2}} \cdot \sqrt{\det(\mathbf{\Sigma}^{-1})} \cdot \left(\mathbf{s}_i^T \mathbf{\Sigma}^{-1} \mathbf{s}_i\right)^{-p/2}. \tag{3.5}$$

Taking the derivative and equating to zero, the maximum likelihood estimator based

on $\{\mathbf{s}_i\}_{i=1}^n$ is the solution to

$$\mathbf{\Sigma} = \frac{p}{n} \cdot \sum_{i=1}^{n} \frac{\mathbf{s}_i \mathbf{s}_i^T}{\mathbf{s}_i^T \mathbf{\Sigma}^{-1} \mathbf{s}_i}. \tag{3.6}$$

When $n > p$, the ML estimator can be found using the following fixed point iterations:

$$\widehat{\mathbf{\Sigma}}_{j+1} = \frac{p}{n} \cdot \sum_{i=1}^{n} \frac{\mathbf{s}_i \mathbf{s}_i^T}{\mathbf{s}_i^T \widehat{\mathbf{\Sigma}}_j^{-1} \mathbf{s}_i}, \tag{3.7}$$

where the initial $\widehat{\mathbf{\Sigma}}_0$ is usually set to the identity matrix. Assuming that $n > p$ and that any p samples out of $\{\mathbf{s}_i\}_{i=1}^n$ are linearly independent with probability one, it can be shown that the iteration process in (3.7) converges and that the limiting value is unique up to constant scale, which does not depend on the initial value of $\widehat{\mathbf{\Sigma}}_0$. In practice, a final normalization step is needed, which ensures that the iteration limit $\widehat{\mathbf{\Sigma}}_\infty$ satisfies $\mathrm{Tr}(\widehat{\mathbf{\Sigma}}_\infty) = p$.

The ML estimate corresponds to the Huber-type M-estimator and has many good properties when $n \gg p$, such as asymptotic normality and strong consistency. Furthermore, it has been pointed out [31] that the ML estimate (3.7) is the "most robust" covariance estimator in the class of elliptical distributions in the sense of minimizing the maximum asymptotic variance. We note that (3.7) can be also motivated from other approaches as proposed in [71, 72].

3.3 Robust shrinkage covariance estimation

Here we extend Tyler's method to the high dimensional setting using shrinkage regularization. It is easy to see that there is no solution to (3.6) when $n < p$ (its left-hand-side is full rank whereas its right-hand-side of is rank deficient). This motivates us to develop a regularized covariance estimator for elliptical samples.

Following [17, 60], we propose to regularize the fixed point iterations as

$$\widetilde{\mathbf{\Sigma}}_{j+1} = (1-\rho)\frac{p}{n}\sum_{i=1}^{n}\frac{\mathbf{s}_i\mathbf{s}_i^T}{\mathbf{s}_i^T\widehat{\mathbf{\Sigma}}_j^{-1}\mathbf{s}_i} + \rho\mathbf{I}, \tag{3.8}$$

$$\widehat{\mathbf{\Sigma}}_{j+1} = \frac{\widetilde{\mathbf{\Sigma}}_{j+1}}{\mathrm{Tr}(\widetilde{\mathbf{\Sigma}}_{j+1})/p}, \tag{3.9}$$

where ρ is the so-called shrinkage coefficient, which is a constant between 0 and 1. When $\rho = 0$ and $n > p$ the proposed shrinkage estimator reduces to Tyler's unbiased method in (3.6) and when $\rho = 1$ the estimator reduces to the trivial uncorrelated case yielding a scaled identity matrix. The term $\rho\mathbf{I}$ ensures that $\widehat{\mathbf{\Sigma}}_{j+1}$ is always well-conditioned and thus allows continuation of the iterative process without the need for restarts. Therefore, the proposed iteration can be applied to high dimensional estimation problems. We emphasize that the normalization (3.9) is important and necessary for convergence. We establish provable convergence and uniqueness of the limit in the following theorem.

Theorem III.1. *Let $0 < \rho < 1$ be a shrinkage coefficient. Then, the fixed point iterations in (3.8) and (3.9) converge to a unique limit for any positive definite initial matrix $\widehat{\mathbf{\Sigma}}_0$.*

The proof of Theorem III.1 follows directly from the concave Perron-Frobenius theory [75] and is provided in the Appendix. We note that the regularization presented in (3.8) and (3.9) is similar to diagonal loading [73]. However, unlike the diagonal loading approach of [73], the proposed shrinkage approach provides a systematic way to choose the regularization parameter ρ, discussed in the next section.

3.3.1 Choosing the shrinkage coefficient

We now turn to the problem of choosing a good, data-dependent, shrinkage coefficient ρ, as as an alternative to cross-validation schemes which incur intensive

computational costs. As in Ledoit-Wolf [17], we begin by assuming we "know" the true covariance $\mathbf{\Sigma}$. Then we define the following clairvoyant "estimator":

$$\widetilde{\mathbf{\Sigma}}(\rho) = (1-\rho)\frac{p}{n}\sum_{i=1}^{n}\frac{\mathbf{s}_i\mathbf{s}_i^T}{\mathbf{s}_i^T\mathbf{\Sigma}^{-1}\mathbf{s}_i} + \rho\mathbf{I}, \tag{3.10}$$

where the coefficient ρ is chosen to minimize the minimum mean-squared error:

$$\rho_O = \arg\min_{\rho} E\left\{\left\|\widetilde{\mathbf{\Sigma}}(\rho) - \mathbf{\Sigma}\right\|_F^2\right\}. \tag{3.11}$$

The following theorem shows that there is a closed-form solution to the problem (3.11), which we refer to as the "oracle" coefficient.

Theorem III.2. *For i.i.d. elliptical distributed samples the solution to (3.11) is*

$$\rho_O = \frac{p^2 + (1-2/p)\mathrm{Tr}(\mathbf{\Sigma}^2)}{(p^2 - np - 2n) + (n+1+2(n-1)/p)\mathrm{Tr}(\mathbf{\Sigma}^2)}, \tag{3.12}$$

under the condition $\mathrm{Tr}(\mathbf{\Sigma}) = p$.

The proof of Theorem III.2 requires the calculation of the fourth moments of an isotropically distributed random vector [56, 57, 58] and is provided in the Appendix.

The oracle coefficient cannot be implemented since ρ_O is a function of the unknown true covariance $\mathbf{\Sigma}$. Therefore, we propose a plug-in estimate for ρ_O:

$$\hat{\rho} = \frac{p^2 + (1-2/p)\mathrm{Tr}(\widehat{\mathbf{M}}^2)}{(p^2 - np - 2n) + (n+1+2(n-1)/p)\mathrm{Tr}(\widehat{\mathbf{M}}^2)}, \tag{3.13}$$

where $\widehat{\mathbf{M}}$ can be any consistent estimator of $\mathbf{\Sigma}$, *e.g.,* the trace-normalized Ledoit-Wolf estimator. Another appealing candidate for plug-in is the (trace-normalized) normalized sample covariance $\widehat{\mathbf{R}}$ [76] defined by:

$$\widehat{\mathbf{R}} = \frac{p}{n}\sum_{i=1}^{n}\mathbf{s}_i\mathbf{s}_i^T. \tag{3.14}$$

We note that the only requirement on the covariance estimator $\widehat{\mathbf{M}}$ is that it provide a good approximation to $\mathrm{Tr}(\mathbf{\Sigma}^2)$. It does not have to be well-conditioned nor does it have to be an accurate estimator of the true covariance matrix $\mathbf{\Sigma}$.

By using the plug-in estimate $\hat{\rho}$ in place of ρ, the robust shrinkage estimator is computed via the fixed point iteration in (3.8) and (3.9).

We note that our proposed minimum MSE based approach to estimate the shrinkage coefficient is completely general and makes a minimum of assumptions on the provenance of the data. In specific applications such as adaptive beamforming, a specifically tailored coefficient estimator may be advantageous [73].

3.3.2 Extension to the complex case

Here we consider the scenario where the random vector $\mathbf{x}$ in (3.1) is complex elliptical distributed. In this case, ν is still a real, positive random variable but $\mathbf{u}$ is a complex Gaussian random vector with covariance matrix $\mathbf{\Sigma}$. Note that $\mathbf{\Sigma}$ is Hermitian and is assumed to be trace-normalized, *i.e.,* $\mathrm{Tr}(\mathbf{\Sigma}) = p$. The complex version of our fixed point iterations is

$$\begin{aligned} \widetilde{\mathbf{\Sigma}}_{j+1} &= (1-\rho)\frac{p}{n}\sum_{i=1}^{n}\frac{\mathbf{s}_i\mathbf{s}_i^H}{\mathbf{s}_i^H\widehat{\mathbf{\Sigma}}_j^{-1}\mathbf{s}_i} + \rho\mathbf{I}, \\ \widehat{\mathbf{\Sigma}}_{j+1} &= \frac{\widetilde{\mathbf{\Sigma}}_{j+1}}{\mathrm{Tr}(\widetilde{\mathbf{\Sigma}}_{j+1})/p}, \end{aligned} \tag{3.15}$$

where $\mathbf{s}_i$ is defined by (3.4). As in the real case, the shrinkage coefficient ρ is chosen to minimize (3.11), where the complex clairvoyant estimator $\widetilde{\mathbf{\Sigma}}(\rho)$ is re-defined as

$$\widetilde{\mathbf{\Sigma}}(\rho) = (1-\rho)\frac{p}{n}\sum_{i=1}^{n}\frac{\mathbf{s}_i\mathbf{s}_i^H}{\mathbf{s}_i^H\mathbf{\Sigma}^{-1}\mathbf{s}_i} + \rho\mathbf{I}. \tag{3.16}$$

The following theorem extends Theorems 1 and 2 to the complex case.

Theorem III.3. *For any $0 < \rho < 1$ the complex valued iterations in (3.15) converge to a unique limit for any positive definitive Hermitian matrix $\widehat{\mathbf{\Sigma}}_0$. For $\widetilde{\mathbf{\Sigma}}(\rho)$ defined in (3.16), the solution to (3.11) is*

$$\rho_O = \frac{p^2 - 1/p\mathrm{Tr}(\mathbf{\Sigma}\mathbf{\Sigma}^H)}{(p^2 - pn - n) + (n + (n-1)/p)\mathrm{Tr}(\mathbf{\Sigma}\mathbf{\Sigma}^H)} \tag{3.17}$$

under the condition $\mathrm{Tr}(\mathbf{\Sigma}) = p$.

The proof of Theorem III.3 is similar to that of Theorem III.1 and Theorem III.2 and is provided in the Appendix. In practice, ρ_O can be approximated by plugging any consistent estimator of $\mathrm{Tr}(\mathbf{\Sigma}\mathbf{\Sigma}^H)$ in (3.17).

3.4 Numerical simulation

In this section we use simulations to demonstrate the superior performance of the proposed shrinkage approach. First we show that our estimator outperforms other estimators for the case of heavy-tailed samples generated by a multivariate Student-T distribution, where ν in (3.1) is a function of a Chi-square random variable:

$$\nu = \sqrt{\frac{d}{\chi_d^2}}; \tag{3.18}$$

The degree-of-freedom d of this multivariate Student-T statistic is set to 3. The dimensionality p is chosen to be 100 and we let $\mathbf{\Sigma}$ be the covariance matrix of an AR(1) process,

$$\mathbf{\Sigma}(i,j) = r^{|i-j|}, \tag{3.19}$$

where $\mathbf{\Sigma}(i,j)$ denotes the entry of $\mathbf{\Sigma}$ in row i and column j. The parameter r is set to 0.7 in this simulation. The sample size n varies from 5 to 225 with step size 10. All the simulations are repeated for 100 trials and the average empirical performance is reported.

We use (3.13) with $\widehat{\mathbf{M}} = \widehat{\mathbf{R}}$ and employ iterations defined by (3.8) and (3.9) with $\rho = \hat{\rho}$. For comparison, we also plot the results of the trace-normalized oracle in (3.12), the trace-normalized Ledoit-Wolf estimator [17], and the non-regularized solution in (3.7) (when $n > p$). As the Ledoit-Wolf estimator operates on the sample covariance which is sensitive to outliers, we also compare to a trace-normalized version of a clairvoyant Ledoit-Wolf estimator implemented according to the procedure

in [17] with known ν. More specifically, the samples $\mathbf{x}_i$ are firstly normalized by the known realizations ν_i, yielding truly Gaussian samples; then the sample covariance of the normalized $\mathbf{x}_i$'s is computed, which is used to estimate the Ledoit-Wolf shrinkage parameters and estimate the covariance via equation (14) in [17]. The MSE is plotted in Fig. 3.1 for the case that $r = 0.7$. It can be observed that the proposed method performs significantly better than the Ledoit-Wolf estimator in Fig. 3.1, and that the performance is very close to the ideal oracle estimator using the optimal shrinkage parameter (3.12). Even the clairvoyant Ledoit-Wolf implemented with known ν_i does not outperform the proposed estimator in the small sample size regime. These results demonstrate the robustness of the proposed approach. Although the Ledoit-Wolf estimator performs better when $r = 0$, the case where $\mathbf{\Sigma} = \mathbf{I}$, the proposed approach still significantly outperforms it, especially for small sample size n (results not shown).

As a graphical illustration, in Fig. 3.2 we provide covariance visualizations for a realization of the estimated covariances using the Ledoit-Wolf method and the proposed approach. The covariance matrix estimates are rendered as a heatmap in Fig. 3.2(a). The sample size in this example is set to 50, which is smaller than the dimension 100. Compared to the true covariance, it is clear that the proposed covariance estimator preserves the structure of the true covariance, while in the Ledoit-Wolf covariance procudure produces "block pattern" artifacts caused by heavy-tails of the multivariate Student-T.

When $n > p$, we also observe a substantial improvement of the proposed method over the ML covariance estimate, which provides further evidence of the power of Steinian shrinkage for reducing MSE.

In order to assess the tradeoff between accuracy and robustness we investigate the

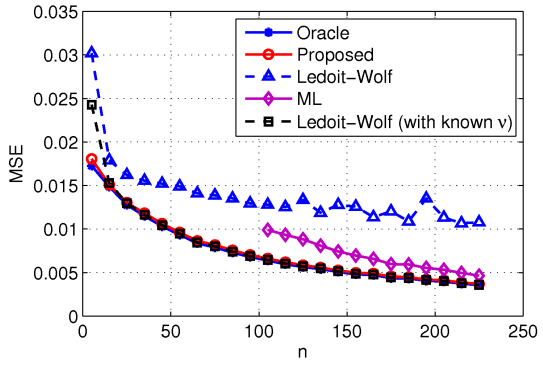
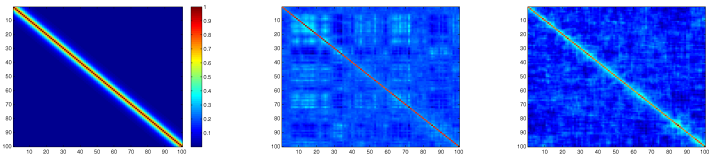


Figure 3.1: Multivariate Student-T samples: Comparison of different trace-normalized covariance estimators when $p = 100$, where r is set to 0.7.



(a) True covariance (b) Ledoit-Wolf (c) Proposed

Figure 3.2: Multivariate Student-T samples: Heatmap visualizations of the covariance matrix estimates using the Ledoit-Wolf and the proposed approaches. $p = 100, n = 50$. Note that $n < p$ in this case.

case when the samples are truly Gaussian distributed. We use the same simulation parameters as in the previous example, the only difference being that the samples are now generated from a Gaussian distribution. The performance comparison is shown in Fig. 3.3, where four different (trace-normalized) methods are included: the oracle estimator derived from Gaussian assumptions (Gaussian oracle) [60], the iterative approximation of the Gaussian oracle (Gaussian OAS) proposed in [60], the Ledoit-Wolf estimator and the proposed method. It can be seen that for truly Gaussian samples the proposed method performs very closely to the Gaussian OAS, which is specifically designed for Gaussian distributions. Indeed, for small sample size ($n < 20$), the proposed method performs even better than the Ledoit-Wolf estimator. This indicates that, although the proposed robust method is developed for the entire elliptical family, it actually sacrifices very little performance for the

case that the distribution is Gaussian.

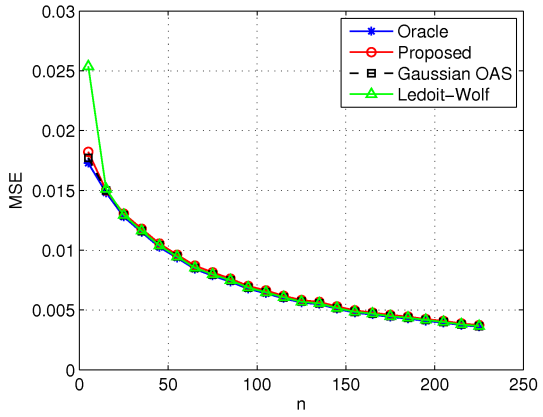


Figure 3.3: Gaussian samples: Comparison of trace-normalized different covariance estimators when $p = 100$.

3.5 Application to anomaly detection in wireless sensor networks

In this section we demonstrate the proposed robust covariance estimator in a real application: activity detection using a wireless sensor network.

The experiment was set up on an Mica2 sensor network platform, as shown in Fig. 3.4, which consists of 14 sensor nodes randomly deployed inside and outside a laboratory at the University of Michigan. Wireless sensors communicated with each other asynchronously by broadcasting an RF signal every 0.5 seconds. The received signal strength (RSS), defined as the voltage measured by a receiver's received signal strength indicator circuit (RSSI), was recorded for each pair of transmitting and receiving nodes. There were $14 \times 13 = 182$ pairs of RSSI measurements over a 30 minute period, and samples were acquired every 0.5 sec. During the experiment period, persons walked into and out of the lab at random times, causing anomaly patterns in the RSSI measurements. Finally, for ground truth, a web camera was employed to record the actual activity.

Fig. 3.5 shows all the received signals and the ground truth indicator extracted

Figure 3.4: Experimental platform: wireless Mica2 sensor nodes.

from the video. The objective of this experiment was intrusion (anomaly) detection. We emphasize that, with the exception of the results shown in Fig. 3.10, the ground truth indicator is only used for performance evaluation and the detection algorithms presented here were completely *unsupervised*.

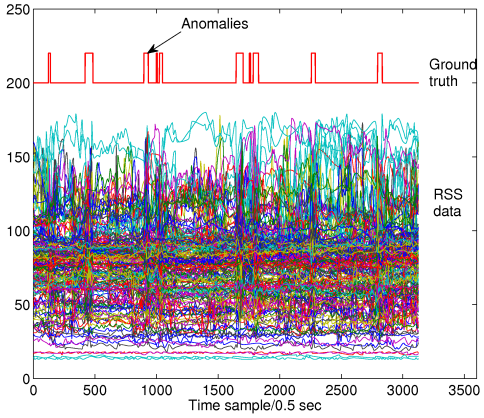


Figure 3.5: At bottom 182 RSS sequences sampled from each pair of transmitting and receiving nodes in intrusion detection experiment. Ground truth indicators at top are extracted from video captured from a web camera that recorded the scene.

To remove temperature drifts [77] of receivers we detrended the data as follows. Let $x_i[k]$ be the k-th sample of the i-th RSS signal and denote

$$\mathbf{x}[k] = (x_1[k], x_2[k], \ldots, x_{182}[k])^T . \tag{3.20}$$

The local mean value of $\mathbf{x}[k]$ is defined by

$$\bar{\mathbf{x}}[k] = \frac{1}{2m+1} \sum_{i=k-m}^{k+m} \mathbf{x}[k], \tag{3.21}$$

where the integer m determines local window size and is set to 50 in this study. We

detrend the data by subtracting this local mean

$$\mathbf{y}[k] = \mathbf{x}[k] - \bar{\mathbf{x}}[k], \tag{3.22}$$

yielding a detrended sample $\mathbf{y}[k]$ used in our anomaly detection.

We established that the detrended measurements were heavy-tailed non-Gaussian by performing several statistical tests. First the Lilliefors test [78] of Gaussianity was performed on the detrended RSS measurements. This resulted in rejection of the Gaussian hypothesis at a level of significance of 10^{-6}. As visual evidence, we show the quantile-quantile plot (QQ plot) for one of the detrended RSS sequences in Fig. 3.6 which illustrates that the samples are non-Gaussian. In Fig. 3.7, we plot the histograms and scatter plots of two of the detrended RSS sequences, which shows the heavy-tail nature of the sample distribution. This strongly suggests that the RSS samples can be better described by a heavy-tailed elliptical distribution than by a Gaussian distribution. As additional evidence, we fitted a Student-T distribution to the first detrended RSS sequence, and used maximum likelihood to estimate the degree-of-freedom as $d = 2$ with a 95% confidence interval (CI) $[1.8460, 2.2879]$.

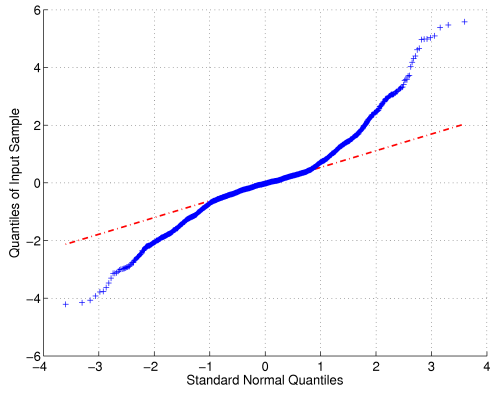


Figure 3.6: QQ plot of data versus the standard Gaussian distribution.

Consider the following discriminant based on the detrended data:

$$t_k = \mathbf{s}^T[k]\boldsymbol{\Sigma}^{-1}\mathbf{s}[k], \tag{3.23}$$

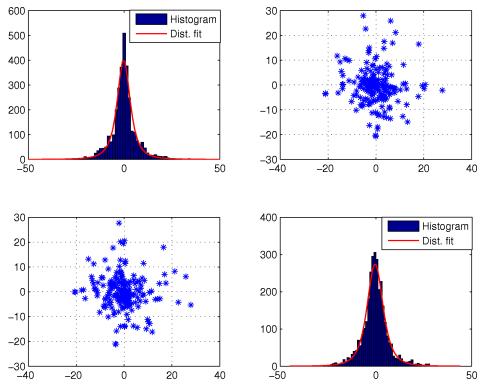


Figure 3.7: Histograms and scatter plots of the first two de-trended RSS sequences, which are fit by a multivariate Student-T distribution with degree-of-freedom $d = 2$.

where, similarly to (3.4), $\mathbf{s}[k] = \mathbf{y}[k]/\|\mathbf{y}[k]\|_2$ and $\boldsymbol{\Sigma}$ is given by the solution to (3.6). A time sample is declared to be anomalous if the test statistic t_k exceeds a specified threshold. The statistic (3.23) is equivalent to a robustified version of the Mahalanobis distance anomaly detector [79]. Note that direct application of shrinkage to the sample covariance of $\mathbf{y}[k]$ would be problematic since a multivariate Student-T vector with 2 degrees of freedom has no second order moments. The test statistic (3.23) can be interpreted as a shrinkage robustified Mahalanobis distance test applied to the better behaved variable $\mathbf{s}[k]$ that has finite moments of all orders. Specifically, we constructed the 182×182 sample covariance by randomly subsampling 200 time slices from the RSS data shown in Fig. 3.5. Note, that these 200 samples correspond to a training set that is contaminated by anomalies at the same anomaly rate (approximately 10%) as the entire sample set. The detection performance was evaluated using the receiver operating characteristic (ROC) curve, where the averaged curves from 200 independent Monte-Carlo trials are shown in Fig. 3.8. For comparison, we also implemented the activity detector (3.23) with other covariance estimates including: the sample covariance, the Ledoit-Wolf estimator and Tyler's ML estimator.

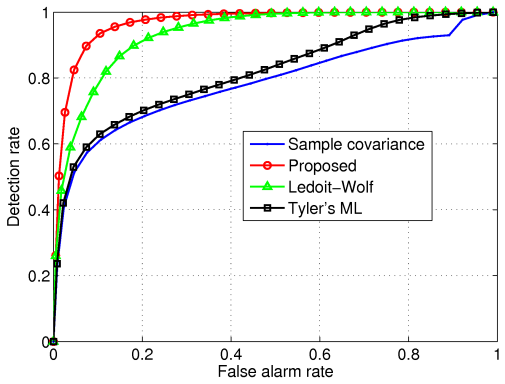


Figure 3.8: Performance comparison for different covariance estimators, $p = 182, n = 200$.

From the mean ROCs we can see that the detection performances are rank ordered as follows: Proposed > Ledoit-Wolf > Tyler's ML > Sample covariance. The sample covariance performs poorly in this setting due to the small sample size ($n = 200, p = 182$) and its sensitivity to the heavy-tailed distribution shown in Fig. 3.6 and 3.7. The Tyler ML method and the Ledoit-Wolf estimator improve upon the sample covariance since they compensate for heavy tails and for small sample size, respectively. Our proposed method compensates for both effects simultaneously and achieves the best detection performance.

We also plot the 90% confidence envelopes, determined by cross-validation, on the ROCs in Fig. 3.9. The width of the confidence interval reflects the sensitivity of the anomaly detector to variations in the training set. Indeed, the upper and lower endpoints of the confidence interval are the optimistic and the pessimistic predictions of detection performance. The proposed method achieves the smallest width among the four computed 90% confidence envelopes.

Finally, for completeness we provide performance comparison of covariance-based *supervised* activity detection algorithms in Fig. 3.10. The training period is selected to be $[251, 450]$ based on ground truth where no anomalies appear. It can be observed

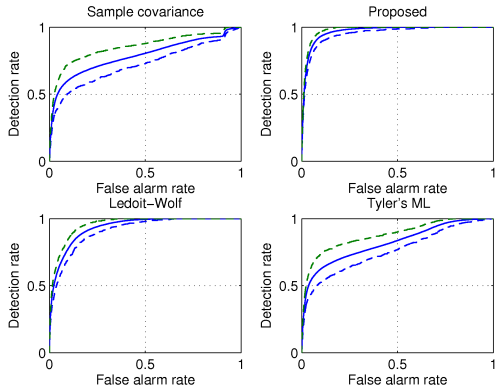


Figure 3.9: Performance comparison for different covariance estimators, including the mean value and 90% confidence intervals. (a): Sample covariance. (b): Proposed. (c): Ledoit-Wolf. (d): Tyler's ML. The 200 training samples are randomly selected from the entire data set.

that, by excluding the outliers caused by anomalies, the performance of the Ledoit-Wolf based intrusion detection algorithm is close to that of the proposed method. We conclude that the activity detection performance of the proposed covariance estimator is more robust than the other three estimators with respect to outlier contamination in the training samples.

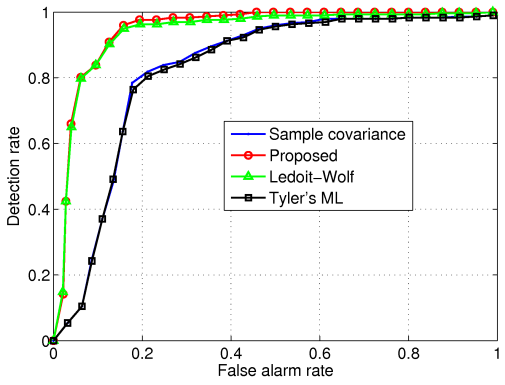


Figure 3.10: Performance comparison for different covariance estimators, $p = 182, n = 200$. The covariance matrix is estimated in a *supervised* manner.

3.6 Conclusion

In this work, we proposed a shrinkage covariance estimator which is robust over the class of elliptically distributed samples. The proposed estimator is obtained

by fixed point iterations, and we established theoretical guarantees for existence, convergence and uniqueness. The optimal shrinkage coefficient was derived using a minimum mean-squared-error framework and has a closed-form expression in terms of the unknown true covariance. This expression can be well approximated by a simple plug-in estimator. Simulations suggest that the iterative approach converges to a limit which is robust to heavy-tailed multivariate Student-T samples. Furthermore, we show that for the Gaussian case, the proposed estimator performs very closely to previous estimators designed expressly for Gaussian samples.

As a real world application we demonstrated the performance of the proposed estimator in intrusion detection using a wireless sensor network. Implementation of a standard covariance-based detection algorithm using our robust covariance estimator achieved superior performances as compared to conventional covariance estimators.

The basis of the proposed method is the ML estimator originally proposed by Tyler in [31]. However, the approach presented in this chapter can be extended to other M-estimators.

One of the main contributions of our work is the proof of uniqueness and convergence of the estimator. This proof extends the results of [31, 80] to the regularized case. Recently, an alternative proof to the non-regularized case using convexity on manifolds was presented in [81]. This latter proof highlights the geometrical structure of the problem and gives additional insight.

3.7 Appendix

3.7.1 Proof of Theorem III.1

In this appendix we prove Theorem III.1. The original convergence proof for the non-regularized case in [31, 80] is based on careful exploitation of the specific form of (3.6). In the contrast, our proof for the regularized case is based on a direct

connection from concave Perron-Frobenius theory [75, 82]. We begin by summarizing the required concave Perron-Frobenius result in the following lemma.

Lemma III.4 ([75]). *Let $(E, \|\cdot\|)$ be a Banach space with $K \subset E$ being a closed, convex cone on which $\|\cdot\|$ is increasing, i.e., for which $x \leq y$ implies $\|x\| \leq \|y\|$, where the operator $\leq$ on the convex cone K means that if $x \leq y$ then $y - x \in K$. Define $U = \{x | x \in K, ||x|| = 1\}$. Let $T : K \to K$ be a concave operator such that*

$$T(\mu x + (1-\mu)y) \geq \mu T(x) + (1-\mu)T(y), \quad \text{for all } \mu \in [0,1], \text{ all } x, y \in K. \tag{3.24}$$

If for some $e \in K - \{0\}$ and constants $a > 0$, $b > 0$ there is

$$ae \leq T(x) \leq be, \quad \text{for all } x \in U, \tag{3.25}$$

then there exists a unique $x^* \in U$ to which the iteration of the normalized operator $\tilde{T}(x) = T(x)/\|T(x)\|, x \in K - \{0\}$ converges:

$$\lim_{k\to\infty} \tilde{T}^k(x) = x^*, \text{ for all } x \in K - \{0\}. \tag{3.26}$$

Lemma III.4 can be obtained by combining results from Lemma 2 and Theorem in Section 4 of [75]. Here we show that the proof of Theorem III.1 is a direct result of applying Lemma III.4 with proper definitions of E, K, U and T:

- E: the set of all symmetric matrices;
- K: the set of all positive semi-definite matrices on E;
- $\|\mathbf{\Sigma}\|$: the normalized nuclear norm of $\mathbf{\Sigma}$, *i.e.*,

  $$\|\mathbf{\Sigma}\| = \frac{1}{p}\sum_{j=1}^{p} |\lambda_j|, \tag{3.27}$$

  where λ_j is the j-th eigenvalue of $\mathbf{\Sigma}$ and $|\cdot|$ is the absolute value operator. Note that for any $\mathbf{\Sigma} \in K$, the nuclear norm $\|\cdot\|$ is identical to $\mathrm{Tr}(\cdot)/p$ and is increasing;

- U: the set $U = \{\mathbf{\Sigma} | \mathbf{\Sigma} \in K, \|\mathbf{\Sigma}\| = 1\}$;
- T: the mapping from K to K defined by

$$T(\mathbf{\Sigma}) = (1-\rho)\frac{p}{n}\sum_{i=1}^{n} w(\mathbf{s}_i, \mathbf{\Sigma})\mathbf{s}_i\mathbf{s}_i^T + \rho\mathbf{I}, \tag{3.28}$$

where the weight function $w(\mathbf{s}_i, \mathbf{\Sigma})$ is defined as

$$w(\mathbf{s}_i, \mathbf{\Sigma}) = \inf_{\mathbf{z}^T\mathbf{s}_i \neq 0} \frac{\mathbf{z}^T\mathbf{\Sigma}\mathbf{z}}{(\mathbf{s}_i^T\mathbf{z})^2}, \tag{3.29}$$

for any $\mathbf{\Sigma} \in K$.

Proof. With the above definitions we show that Theorem III.1 is a direct result of Lemma III.4. We begin by showing that the mapping operator T is concave. Indeed, it is sufficient to show that $w(\mathbf{s}_i, \mathbf{\Sigma})$ in concave in $\mathbf{\Sigma}$, which is true because it is the infinimum of affine functions of $\mathbf{\Sigma}$.

Next, we show that T satisfies condition (3.25) with $e = \mathbf{I}$. It is easy to see that

$$\rho\mathbf{I} \leq T(\mathbf{\Sigma}), \tag{3.30}$$

for any $\mathbf{\Sigma} \in U$. Then we show that

$$w(\mathbf{s}_i, \mathbf{\Sigma}) \leq p, \tag{3.31}$$

for any $\mathbf{\Sigma} \in U$. Indeed,

$$w(\mathbf{s}_i, \mathbf{\Sigma}) = \inf_{\mathbf{z}^T\mathbf{s}_i \neq 0} \frac{\mathbf{z}^T\mathbf{\Sigma}\mathbf{z}}{(\mathbf{s}_i^T\mathbf{z})^2} \leq \frac{\mathbf{s}_i^T\mathbf{\Sigma}\mathbf{s}_i}{(\mathbf{s}_i^T\mathbf{s}_i)^2} \leq \frac{\lambda_{\max}}{\mathbf{s}_i^T\mathbf{s}_i} = \lambda_{\max}, \tag{3.32}$$

where $\lambda_{\max}$ is the maximum eigenvalue of $\mathbf{\Sigma}$. The last equality in the right-hand-side of (3.32) comes from the fact that $\mathbf{s}_i$ is of unit norm by definition (3.4). (3.31) is thus obtained by noticing that $\mathbf{\Sigma} \in U$ and $\lambda_{\max} \leq p$. Substituting (3.31) into (3.28) we have

$$T(\mathbf{\Sigma}) \leq (1-\rho)p^2\widehat{\mathbf{R}} + \rho\mathbf{I} \leq \left((1-\rho)p^2\alpha_{\max} + \rho\right)\mathbf{I}, \tag{3.33}$$

where

$$\widehat{\mathbf{R}} = \frac{1}{n}\sum_{i=1}^{n} \mathbf{s}_i \mathbf{s}_i^T,$$

and $\alpha_{\max}$ is the maximum eigenvalue of $\widehat{\mathbf{R}}$. Again, as $\mathbf{s}_i$ is of unit norm, $\alpha_{\max} \leq \mathrm{Tr}(\widehat{\mathbf{R}}) = 1$ and

$$T(\mathbf{\Sigma}) \leq \left((1-\rho)p^2 + \rho\right)\mathbf{I}. \tag{3.34}$$

Therefore, we have shown that T satisfies condition (3.25), where $e = \mathbf{I}$, $a = \rho$ and $b = (1-\rho)p^2 + \rho$. In addition, (3.25) establishes that the mapping T from U always yields a positive definite matrix. Therefore, the convergent limit of the fixed-point iteration is positive definite.

Finally, we note that, for any $\mathbf{\Sigma} \succ \mathbf{0}$, we have

$$\|\mathbf{\Sigma}\| = \frac{\mathrm{Tr}(\mathbf{\Sigma})}{p}, \tag{3.35}$$

and

$$w(\mathbf{s}_i, \mathbf{\Sigma}) = \inf_{\mathbf{z}^T\mathbf{s}_i \neq 0} \frac{\mathbf{z}^T\mathbf{\Sigma}\mathbf{z}}{(\mathbf{s}_i^T\mathbf{z})^2} = \frac{1}{\mathbf{s}_i^T\mathbf{\Sigma}^{-1}\mathbf{s}_i}. \tag{3.36}$$

The limit (3.26) is then identical to the limit of proposed iterations (3.8) and (3.9) for any $\mathbf{\Sigma} \succ \mathbf{0}$. Therefore, Theorem III.1 has been proved. □

3.7.2 Proof of Theorem III.2

Proof. To ease the notation we define $\widetilde{\mathbf{C}}$ as

$$\widetilde{\mathbf{C}} = \frac{p}{n}\sum_{i=1}^{n} \frac{\mathbf{s}_i \mathbf{s}_i^T}{\mathbf{s}_i^T\mathbf{\Sigma}^{-1}\mathbf{s}_i}. \tag{3.37}$$

The shrinkage estimator in (3.10) is then

$$\widetilde{\mathbf{\Sigma}}(\rho) = (1-\rho)\widetilde{\mathbf{C}} + \rho\mathbf{I}. \tag{3.38}$$

By substituting (3.38) into (3.10) and taking derivatives of ρ, we obtain that

$$\begin{aligned}\rho_O &= \frac{E\left\{\mathrm{Tr}\left((\mathbf{I}-\widetilde{\mathbf{C}})(\mathbf{\Sigma}-\widetilde{\mathbf{C}})\right)\right\}}{E\left\{\left\|\mathbf{I}-\widetilde{\mathbf{C}}\right\|_F^2\right\}} \\ &= \frac{m_2 - m_{11} - m_{12} + \mathrm{Tr}(\mathbf{\Sigma})}{m_2 - 2m_{11} + p},\end{aligned} \tag{3.39}$$

where

$$m_2 = E\left\{\mathrm{Tr}(\widetilde{\mathbf{C}}^2)\right\}, \tag{3.40}$$

$$m_{11} = E\left\{\mathrm{Tr}(\widetilde{\mathbf{C}})\right\}, \tag{3.41}$$

and

$$m_{12} = E\left\{\mathrm{Tr}(\widetilde{\mathbf{C}}\mathbf{\Sigma})\right\}. \tag{3.42}$$

Next, we calculate the moments. We begin by eigen-decomposing $\mathbf{\Sigma}$ as

$$\mathbf{\Sigma} = \mathbf{U}\mathbf{D}\mathbf{U}^T, \tag{3.43}$$

and denote

$$\mathbf{\Lambda} = \mathbf{U}\mathbf{D}^{1/2}. \tag{3.44}$$

Then, we define

$$\mathbf{z}_i = \frac{\mathbf{\Lambda}^{-1}\mathbf{s}_i}{\|\mathbf{\Lambda}^{-1}\mathbf{s}_i\|_2} = \frac{\mathbf{\Lambda}^{-1}\mathbf{u}_i}{\|\mathbf{\Lambda}^{-1}\mathbf{u}_i\|_2}. \tag{3.45}$$

Noting that $\mathbf{u}_i$ is a Gaussian distributed random vector with covariance $\mathbf{\Sigma}$, it is easy to see that $\|\mathbf{z}_i\|_2 = 1$ and $\mathbf{z}_i$ and $\mathbf{z}_j$ are independent with each other for $i \neq j$. Furthermore, $\mathbf{z}_i$ is isotropically distributed [56, 57, 58] and satisfies [60]

$$E\left\{\mathbf{z}_i\mathbf{z}_i^T\right\} = \frac{1}{p}\mathbf{I}, \tag{3.46}$$

$$
\begin{aligned}
E\left\{\left(\mathbf{z}_i^T\mathbf{D}\mathbf{z}_i\right)^2\right\} &= \frac{1}{p(p+2)}\left(2\mathrm{Tr}(\mathbf{D}^2)+\mathrm{Tr}^2(\mathbf{D})\right) \\
&= \frac{1}{p(p+2)}\left(2\mathrm{Tr}(\mathbf{\Sigma}^2)+\mathrm{Tr}^2(\mathbf{\Sigma})\right),
\end{aligned} \tag{3.47}
$$

and

$$
E\left\{\left(\mathbf{z}_i^T\mathbf{D}\mathbf{z}_j\right)^2\right\} = \frac{1}{p^2}\mathrm{Tr}(\mathbf{D}^2) = \frac{1}{p^2}\mathrm{Tr}(\mathbf{\Sigma}^2),\ i \neq j. \tag{3.48}
$$

Expressing $\widetilde{\mathbf{C}}$ in terms of $\mathbf{z}_i$, there is

$$
\widetilde{\mathbf{C}} = \frac{p}{n}\mathbf{\Lambda}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^T\mathbf{\Lambda}^T. \tag{3.49}
$$

Then,

$$
E\left\{\widetilde{\mathbf{C}}\right\} = \frac{p}{n}\mathbf{\Lambda}\sum_{i=1}^{n}E\left\{\mathbf{z}_i\mathbf{z}_i^T\right\}\mathbf{\Lambda}^T = \mathbf{\Sigma}, \tag{3.50}
$$

and accordingly we have

$$
m_{11} = E\left\{\mathrm{Tr}(\widetilde{\mathbf{C}})\right\} = \mathrm{Tr}(\mathbf{\Sigma}), \tag{3.51}
$$

and

$$
m_{12} = E\left\{\mathrm{Tr}(\widetilde{\mathbf{C}}\mathbf{\Sigma})\right\} = \mathrm{Tr}(\mathbf{\Sigma}^2). \tag{3.52}
$$

For m_2 there is

$$
\begin{aligned}
m_2 &= \frac{p^2}{n^2}E\left\{\mathrm{Tr}\left(\mathbf{\Lambda}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^T\mathbf{\Lambda}^T\mathbf{\Lambda}\sum_{j=1}^{n}\mathbf{z}_j\mathbf{z}_j^T\mathbf{\Lambda}^T\right)\right\} \\
&= \frac{p^2}{n^2}E\left\{\mathrm{Tr}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{z}_i\mathbf{z}_i^T\mathbf{\Lambda}^T\mathbf{\Lambda}\mathbf{z}_j\mathbf{z}_j^T\mathbf{\Lambda}^T\mathbf{\Lambda}\right)\right\} \\
&= \frac{p^2}{n^2}E\left\{\mathrm{Tr}\left(\sum_{i=1}^{n}\sum_{j=1}^{n}\mathbf{z}_i\mathbf{z}_i^T\mathbf{D}\mathbf{z}_j\mathbf{z}_j^T\mathbf{D}\right)\right\} \\
&= \frac{p^2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left\{\left(\mathbf{z}_i^T\mathbf{D}\mathbf{z}_j\right)^2\right\}.
\end{aligned} \tag{3.53}
$$

Now substitute (3.47) and (3.48) to (3.53):

$$\begin{aligned} m_2 &= \frac{p^2}{n^2}\left(\frac{n}{p(p+2)}\left(2\mathrm{Tr}(\mathbf{\Sigma}^2)+\mathrm{Tr}^2(\mathbf{\Sigma})\right)+\frac{n(n-1)}{p^2}\mathrm{Tr}(\mathbf{\Sigma}^2)\right) \\ &= \frac{1}{n(1+2/p)}\left(2\mathrm{Tr}(\mathbf{\Sigma}^2)+\mathrm{Tr}^2(\mathbf{\Sigma})\right)+(1-\frac{1}{n})\mathrm{Tr}(\mathbf{\Sigma}^2) \\ &= \left(1-\frac{1}{n}+\frac{2}{n(1+2/p)}\right)\mathrm{Tr}(\mathbf{\Sigma}^2)+\frac{\mathrm{Tr}^2(\mathbf{\Sigma})}{n(1+2/p)}. \end{aligned} \tag{3.54}$$

Recalling $\mathrm{Tr}(\mathbf{\Sigma}) = p$, (3.12) is finally obtained by substituting (3.51), (3.52) and (3.54) into (3.39). □

3.7.3 Proof of Theorem III.3

The proof of convergence and uniqueness is a simple extension of the proof in Appendix A by re-defining E as the set of all Hermitian matrices and the mapping function $T(\mathbf{\Sigma})$ as

$$T(\mathbf{\Sigma}) = (1-\rho)\frac{p}{n}\sum_{i=1}^{n} w(\mathbf{s}_i, \mathbf{\Sigma})\mathbf{s}_i\mathbf{s}_i^H + \rho\mathbf{I}, \tag{3.55}$$

where

$$w(\mathbf{s}_i, \mathbf{\Sigma}) = \inf_{\mathbf{z}^H\mathbf{s}_i \neq 0} \frac{\mathbf{z}^H\mathbf{\Sigma}\mathbf{z}}{|\mathbf{s}_i^H\mathbf{z}|^2}. \tag{3.56}$$

For the oracle coefficient, define

$$\widetilde{\mathbf{C}} = \frac{n}{p}\sum_{i=1}^{n}\frac{\mathbf{s}_i\mathbf{s}_i^H}{\mathbf{s}_i^H\mathbf{\Sigma}^{-1}\mathbf{s}_i}. \tag{3.57}$$

Then $\widetilde{\mathbf{\Sigma}}(\rho) = (1-\rho)\widetilde{\mathbf{C}} + \rho\mathbf{I}$. It can be shown that

$$\begin{aligned} \rho_O &= \frac{E\left\{\mathrm{Re}\left(\mathrm{Tr}\left((\mathbf{I}-\widetilde{\mathbf{C}})(\mathbf{\Sigma}-\widetilde{\mathbf{C}})^H\right)\right)\right\}}{E\left\{\|\mathbf{I}-\widetilde{\mathbf{C}}\|_F^2\right\}} \\ &= \frac{m_2 - m_{11} - m_{12} + \mathrm{Tr}(\mathbf{\Sigma}^H)}{m_2 - 2m_{11} + p}, \end{aligned} \tag{3.58}$$

where m_2, m_{11} and m_{12} are re-defined as

$$m_2 = E\left\{\mathrm{Tr}\left(\widetilde{\mathbf{C}}\widetilde{\mathbf{C}}^H\right)\right\}, \tag{3.59}$$

$$m_{11} = E\left\{\mathrm{Re}\left(\mathrm{Tr}\left(\widetilde{\mathbf{C}}^H\right)\right)\right\} \tag{3.60}$$

and

$$m_{12} = E\left\{\mathrm{Re}\left(\mathrm{Tr}\left(\widetilde{\mathbf{C}}\boldsymbol{\Sigma}^H\right)\right)\right\}, \tag{3.61}$$

respectively.

Next, we eigen-decompose the Hermitian positive definite matrix $\boldsymbol{\Sigma}$ as

$$\boldsymbol{\Sigma} = \mathbf{U}\mathbf{D}\mathbf{U}^H, \tag{3.62}$$

where $\mathbf{D}$ is a real diagonal matrix and $\mathbf{U}$ is a complex unitary matrix. Define

$$\boldsymbol{\Lambda} = \mathbf{U}\mathbf{D}^{1/2} \tag{3.63}$$

and

$$\mathbf{z}_i = \frac{\boldsymbol{\Lambda}^{-1}\mathbf{s}_i}{\|\boldsymbol{\Lambda}^{-1}\mathbf{s}_i\|_2}. \tag{3.64}$$

$\{\mathbf{z}_i\}_{i=1}^n$ are then complex valued isotropically distributed random vectors and are independent to each other [56, 57, 58]. Using results from [83], it can be shown that

$$E\left\{\mathbf{z}_i\mathbf{z}_i^H\right\} = \frac{1}{p}\mathbf{I}, \tag{3.65}$$

$$\begin{aligned} E\left\{\left|\mathbf{z}_i^H\mathbf{D}\mathbf{z}_i\right|^2\right\} &= \frac{1}{p(p+1)}\left(\mathrm{Tr}(\mathbf{D}^2) + \mathrm{Tr}^2(\mathbf{D})\right) \\ &= \frac{1}{p(p+1)}\left(\mathrm{Tr}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H) + \mathrm{Tr}^2(\boldsymbol{\Sigma})\right) \end{aligned} \tag{3.66}$$

and

$$E\left\{\left|\mathbf{z}_i^H\mathbf{D}\mathbf{z}_j\right|^2\right\} = \frac{1}{p^2}\mathrm{Tr}(\mathbf{D}^2) = \frac{1}{p^2}\mathrm{Tr}(\boldsymbol{\Sigma}\boldsymbol{\Sigma}^H),\ i \neq j. \tag{3.67}$$

Eq. (3.65) $\sim$ (3.67) are complex versions of (3.46) $\sim$ (3.48). Expressing $\widetilde{\mathbf{C}}$ in terms of $\mathbf{z}_i$, there is

$$\widetilde{\mathbf{C}} = \frac{p}{n}\mathbf{\Lambda}\sum_{i=1}^{n}\mathbf{z}_i\mathbf{z}_i^H\mathbf{\Lambda}^H, \tag{3.68}$$

and accordingly

$$E\left\{\widetilde{\mathbf{C}}\right\} = \mathbf{\Sigma}. \tag{3.69}$$

As $\mathrm{Re}(\cdot)$, $\mathrm{Tr}(\cdot)$ and $E\{\cdot\}$ are exchangeable to each other, we have

$$m_{11} = \mathrm{Re}\left(\mathrm{Tr}(\mathbf{\Sigma}^H)\right) = \mathrm{Tr}(\mathbf{\Sigma}) \tag{3.70}$$

and

$$m_{12} = \mathrm{Tr}(\mathbf{\Sigma}\mathbf{\Sigma}^H). \tag{3.71}$$

For m_2, using a similar derivation as in Appendix B, it can be shown that

$$\begin{aligned} m_2 &= \frac{p^2}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}E\left\{\left|\mathbf{z}_i^H\mathbf{D}\mathbf{z}_j\right|^2\right\} \\ &= \left(1-\frac{1}{n(p+1)}\right)\mathrm{Tr}(\mathbf{\Sigma}\mathbf{\Sigma}^H) + \frac{p}{n(p+1)}\mathrm{Tr}^2(\mathbf{\Sigma}). \end{aligned} \tag{3.72}$$

As $\mathrm{Tr}(\mathbf{\Sigma}^H) = \mathrm{Tr}(\mathbf{\Sigma}) = p$, (3.17) can be finally obtained by substituting (3.70), (3.71) and (3.72) to (3.58).

CHAPTER IV

Recursive $\ell_{1,\infty}$ Group lasso

In many applications of signal processing and machine learning, the unknown system may be time-varying and the data are acquired in a streaming fashion. In those scenarios, online estimation and efficient update can be very important. In this chapter we address the time-varying covariance estimation problem where the samples are non-stationary. Specifically we consider covariance matrices of variables associated by underlying graphical models. This model has attracted growing interests in problems of sensor/agent networks [7, 8], bioinformatics [39] and many others. Instead of estimating the entire matrix directly, we split the matrix estimation into multiple Lasso-type problems, each of which performs a column-wise estimation of the inverse covariance. We then propose a recursive adaptive $\ell_{1,\infty}$ group lasso algorithm, a generalization to standard ℓ_1 lasso method, for real-time penalized least squares prediction that produces a time sequence of optimal sparse predictor coefficient vectors. At each time index the proposed algorithm computes an exact update of the optimal $\ell_{1,\infty}$-penalized recursive least squares (RLS) predictor. Each update minimizes a convex but non-differentiable function optimization problem. We develop an online homotopy method to reduce the computational complexity. Numerical simulations demonstrate that the proposed algorithm outperforms the ℓ_1 regularized RLS

algorithm for a group sparse system identification problem and has lower implementation complexity than direct group lasso solvers. Finally, we demonstrate online covariance estimation using multiple recursive lasso estimators in anomaly detection for the Abilene Networks.

4.1 Introduction

Many applications in signal processing and related fields deal with streaming samples where the data are acquired online and the system of interest may be time-varying. In those scenarios online processing plays a major role and covers the main theme of adaptive system identification [84] and has attracted growing interest in various data mining problems [85, 86, 87]. In those settings, most of available covariance estimators cannot be directly applied either because of the non-stationary samples or the intensive computational complexity for online calculation.

As the system is time-varying, the size of locally stationary samples can be far less than the number of parameters to be estimated. As we have demonstrated in Chapter I, this is one major cause of "large p small n" problems. Therefore, regularization is generally required for online covariance estimation so that a prior model can be imposed to mitigate the high dimension curse. In this work we are interested in estimating covariance of variables associating with an underlying graphical model, which has received considerable interests [3, 20, 18]. Among them, the ℓ_1-type regularization is popular and can achieve consistent estimation of sparse graphs even when $n \ll p$ in certain conditions. For some kinds of data, it is reasonable to assume that the variables can be grouped into clusters, which share similar connectivity or correlation patterns. Examples can be found in analysis of gene expression data [86] as well as sensor/agent networks [7, 8] and recent works have extended the ℓ_1

penalized framework to the case of group sparsity by regularizing the $\ell_{1,2}$ norm [88] or the $\ell_{1,\infty}$ norm [89]. The analogous result in linear regression is usually referred to as the group lasso that promotes block sparsity in the underlying graphs. Indeed, the standard ℓ_1 lasso can be treated as the special case of the group lasso when each group only contains a single variable.

In this chapter we address the online covariance estimation problem where an underlying graphical model is assumed to associate the variables. Instead of estimating the entire matrix directly, we split the matrix estimation into a set of recursive lasso problems which perform column-wise estimation of the inverse covariance. For each recursive lasso problem we employ the $\ell_{1,\infty}$ norm rather than the ℓ_1 norm for the flexibility to promote sparsity as well as group sparsity. Our recursive group lasso algorithm is suitable for online applications where data are acquired sequentially. The proposed algorithm is based on the homotopy approach which updates from previous solutions and has lower implementation complexity than direct group lasso solvers.

The chapter is organized as follows. In Section 4.2 we illustrate the connection between covariance estimation and linear regression and formulates the online covariance estimation into multiple recursive regression problems in Section 4.3. Section 4.4 develops the homotopy based algorithm to solve the recursive $\ell_{1,\infty}$ group lasso in an online recursive manner. Section 4.5 provides numerical simulation results. In Section 4.6 we perform an online covariance-based anomaly detection using the Abilene Networks data. Section 4.7 summarizes our principal conclusions. The proofs of theorems and some details of the proposed algorithm are provided in Appendix.

Notations: In the following, matrices and vectors are denoted by boldface upper case letters and boldface lower case letters, respectively; $(\cdot)^T$ denotes the transpose

operator, and $\|\cdot\|_1$ and $\|\cdot\|_\infty$ denote the ℓ_1 and ℓ_∞ norm of a vector, respectively; for a set $\mathcal{A}$, $|\mathcal{A}|$ denotes its cardinality and ϕ denotes the empty set; $\mathbf{x}_\mathcal{A}$ denotes the sub-vector of $\mathbf{x}$ from the index set $\mathcal{A}$ and $\mathbf{R}_{\mathcal{A}\mathcal{B}}$ denotes the sub-matrix of $\mathbf{R}$ formed from the row index set $\mathcal{A}$ and column index set $\mathcal{B}$.

4.2 Connecting covariance estimation to regression

Consider a $(p+1)$-dimensional random vector $\mathbf{x} = (x_1, x_2, ..., x_{p+1})^T$ that has a jointly distribution with mean zero and covariance $\mathbf{\Sigma}$, where $\mathbf{\Sigma}$ is a $(p+1)\times(p+1)$ positive definite matrix. Let $\mathbf{C}$ be the inverse covariance such that $\mathbf{C} = \mathbf{\Sigma}^{-1}$. Then, for $1 \le i \le p+1$, x_i can be expressed by

$$x_i = \sum_{j\neq i} \beta_{ij} x_j + \epsilon_i \tag{4.1}$$

where ϵ_i is un-correlated with any x_j, $j \neq i$ and

$$\beta_{ij} = -\frac{c_{ij}}{c_{ii}}. \tag{4.2}$$

In addition, there is

$$\mathrm{var}(\epsilon_i) = \frac{1}{c_{ii}}. \tag{4.3}$$

Eq. (4.1) - (4.3) are the keys to connect covariance estimation to regression problems. Let $\hat{\beta}_{ij}$ be the estimated coefficients from the regression problem (4.1) and define $\hat{\mathrm{var}}(\epsilon_i)$ be the sample variance of the residuals. We then use the following relationship for (inverse) covariance estimation:

$$\hat{c}_{ii} = \frac{1}{\hat{\mathrm{var}}(\epsilon_i)},\ \hat{c}_{ij} = -\frac{1}{2}\left(\hat{c}_{ii}\hat{\beta}_{ij} + \hat{c}_{jj}\hat{\beta}_{ji}\right), \tag{4.4}$$

where the second equation in (4.4) guarantees the estimate $\hat{\mathbf{C}}$ is symmetric. Indeed, if $\hat{\beta}_{ij}$ is calculated from the least squares method, the resultant estimate $\hat{\mathbf{C}}$ coincides with the inverse of the sample covariance if it is well-conditioned.

For high-dimensional problems with limited sample size, the regression problem may be ill-posed which requires regularization. The regularization of $\mathbf{C}$ can be performed by regularization of the matrix of β_{ij}. It is well known that when $\hat{\beta}_{ij}$ is obtained from the least squares method, the corresponding covariance estimate then coincides with the sample covariance. A similar example is to use the ridge regression with ℓ_2 penalties to estimate $\hat{\beta}_{ij}$ and the resultant covariance estimate becomes the diagonally loaded sample covariance, which is related to the Steinian shrinkage discussed in previous chapters. Recently, Meinshausen [3] proposed to employ the ℓ_1 lasso to estimate β_{ij} in (4.1) and to detect zeros of $\mathbf{C}$, based on the assumption that a sparse graph model associates all the variables of interest. We note that there are alternative approaches to covariance estimation with sparse Gaussian graphs, where the lasso-type penalties are imposed directly on $\mathbf{C}$ rather than on the β_{ij} [18, 20]. As [18] has pointed out, the two problem are not equivalent to each other. For example, decoupled regression methods to estimate the covariance may not guarantee the positive definiteness of the covariance estimate. On the other hand, direct covariance estimation generally requires more computational cost and memory storage. In this work, we are driven by computational considerations and adopt Meinshausen's approach to split the matrix estimation into a set of decoupled regression problems. Positive definiteness can be imposed in a variety of ways, *e.g.,* by adding a diagonal matrix scaled by the maximum eigenvalue or the trace of the estimated matrix [90], but this is not investigated here.

4.3 Problem formulation

Transferring covariance estimation to multiple vector regression problems is a way to achieve recursive estimation of a time-varying covariance matrix. Indeed, estimat-

ing a time-varying vector using non-stationary samples has been long investigated in theories of adaptive system identification, where Recursive Least Squares (RLS) is a widely used method because of its fast convergence and low steady-state error. RLS recursively fitted a measurement stream to a linear model in such a way to minimize a weighted average of squared residual prediction errors. In this work we propose a $\ell_{1,\infty}$ regularized RLS which can be solved efficiently using an online homotopy approach. For ease of notations and the main focus of our contribution, we formulate our problems in the framework of regularized RLS algorithms.

4.3.1 Recursive Least Squares

Let $\mathbf{w}$ be a p-dimensional coefficient vector.[1] Let $\mathbf{y}$ be an n-dimensional vector comprised of observations $\{y_j\}_{j=1}^n$. Let $\{\mathbf{x}_j\}_{j=1}^n$ be a sequence of p-dimensional predictor variables. In standard adaptive filtering terminology, y_j, $\mathbf{x}_j$ and $\mathbf{w}$ are the primary signal, the reference signal, and the adaptive filter weights. The RLS algorithm solves the following quadratic minimization problem recursively over time $n = p, p+1, \ldots$:

$$\hat{\mathbf{w}}_n = \arg\min_{\mathbf{w}} \sum_{j=1}^{n} \gamma^{n-j}(y_j - \mathbf{w}^T\mathbf{x}_j)^2, \tag{4.5}$$

where $\gamma \in (0, 1]$ is the forgetting factor controlling the trade-off between transient and steady-state behaviors.

To serve as a template for the sparse RLS extensions described below we briefly review the RLS update algorithm. Define $\mathbf{R}_n$ and $\mathbf{r}_n$ as

$$\mathbf{R}_n = \sum_{j=1}^{n} \gamma^{n-j}\mathbf{x}_j\mathbf{x}_j^T \tag{4.6}$$

[1] In the rest of the chapters we focus on regression problems and no longer use the notations of β_{ij} unless explicit declaration. The vector $\mathbf{w}$ here can represent any regression coefficients $\{\beta_{ij}\}_{j\neq i}$ for a responsive variable.

and

$$\mathbf{r}_n = \sum_{j=1}^{n} \gamma^{n-j} \mathbf{x}_j y_j. \tag{4.7}$$

The solution $\hat{\mathbf{w}}_n$ to (4.5) can be then expressed as

$$\hat{\mathbf{w}}_n = \mathbf{R}_n^{-1} \mathbf{r}_n. \tag{4.8}$$

The matrix $\mathbf{R}_n$ and the vector $\mathbf{r}_n$ are updated as

$$\mathbf{R}_n = \gamma \mathbf{R}_{n-1} + \mathbf{x}_n \mathbf{x}_n^T,$$

and

$$\mathbf{r}_n = \gamma \mathbf{r}_{n-1} + \mathbf{x}_n y_n^T.$$

Applying the Sherman-Morrison-Woodbury formula [91],

$$\mathbf{R}_n^{-1} = \gamma^{-1} \mathbf{R}_{n-1}^{-1} - \gamma^{-1} \alpha_n \mathbf{g}_n \mathbf{g}_n^T, \tag{4.9}$$

where

$$\mathbf{g}_n = \mathbf{R}_{n-1}^{-1} \mathbf{x}_n \tag{4.10}$$

and

$$\alpha_n = \frac{1}{\gamma + \mathbf{x}_n^T \mathbf{g}_n}. \tag{4.11}$$

Substituting (4.9) into (4.8), we obtain the weight update [92]

$$\hat{\mathbf{w}}_n = \hat{\mathbf{w}}_{n-1} + \alpha_n \mathbf{g}_n e_n, \tag{4.12}$$

where

$$e_n = y_n - \hat{\mathbf{w}}_{n-1}^T \mathbf{x}_n. \tag{4.13}$$

Equations (4.9)-(4.13) define the RLS algorithm which has computational complexity of order $\mathcal{O}(p^2)$.

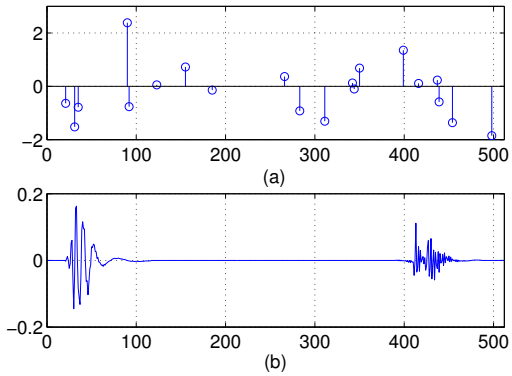


Figure 4.1: Examples of (a) a general sparse system and (b) a group-sparse system.

4.3.2 Non-recursive $\ell_{1,\infty}$ group lasso

The $\ell_{1,\infty}$ group lasso is a regularized least squares approach which uses the $\ell_{1,\infty}$ mixed norm to promote group-wise sparse pattern on the predictor coefficient vector. The $\ell_{1,\infty}$ norm of a vector $\mathbf{w}$ is defined as

$$\|\mathbf{w}\|_{1,\infty} = \sum_{m=1}^{M} \|\mathbf{w}_{\mathcal{G}_m}\|_\infty,$$

where $\{\mathcal{G}_m\}_{m=1}^M$ is a group partition of the index set $\mathcal{G} = \{1, \ldots, p\}$, *i.e.*,

$$\bigcup_{m=1}^{M} \mathcal{G}_m = \mathcal{G}, \quad \mathcal{G}_m \cap \mathcal{G}_{m'} = \phi \text{ if } m \neq m',$$

and $\mathbf{w}_{\mathcal{G}_m}$ is a sub-vector of $\mathbf{w}$ indexed by $\mathcal{G}_m$. The $\ell_{1,\infty}$ norm is a mixed norm: it encourages correlation among coefficients inside each group via the ℓ_∞ norm within each group and promotes sparsity across each group using the ℓ_1 norm. The mixed norm $\|\mathbf{w}\|_{1,\infty}$ is convex in $\mathbf{w}$ and reduces to $\|\mathbf{w}\|_1$ when each group contains only one coefficient, *i.e.*, $|\mathcal{G}_1| = |\mathcal{G}_2| = \cdots = |\mathcal{G}_M| = 1$.

The $\ell_{1,\infty}$ group lasso solves the following penalized least squares problem:

$$\hat{\mathbf{w}}_n = \arg\min_{\mathbf{w}} \frac{1}{2} \sum_{j=1}^{n} \gamma^{n-j} (y_j - \mathbf{w}^T \mathbf{x}_j)^2 + \lambda \|\mathbf{w}\|_{1,\infty}, \tag{4.14}$$

where λ is a regularization parameter. Eq. (4.14) is a convex problem and can be

solved by standard convex optimizers or path tracing algorithms [1]. Direct solution of (4.14) has computational complexity of $\mathcal{O}(p^3)$.

4.3.3 Recursive $\ell_{1,\infty}$ group lasso

In this subsection we obtain a recursive solution for (4.14) that gives an update $\hat{\mathbf{w}}_n$ from $\hat{\mathbf{w}}_{n-1}$. The approach taken is a group-wise generalization of recent works [93, 94] that uses the homotopy approach to sequentially solve the lasso problem. Using the definitions (4.6) and (4.7), the problem (4.14) is equivalent to

$$\begin{aligned}\hat{\mathbf{w}}_n &= \arg\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}^T\mathbf{R}_n\mathbf{w} - \mathbf{w}^T\mathbf{r}_n + \lambda\|\mathbf{w}\|_{1,\infty} \\ &= \arg\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}^T\left(\gamma\mathbf{R}_{n-1} + \mathbf{x}_n^T\mathbf{x}_n\right)\mathbf{w} \\ &\qquad - \mathbf{w}^T(\gamma\mathbf{r}_{n-1} + \mathbf{x}_n y_n) + \lambda\|\mathbf{w}\|_{1,\infty}.\end{aligned} \tag{4.15}$$

Let $f(\theta, \lambda)$ be the solution to the following parameterized problem

$$\begin{aligned}f(\theta,\lambda) = \arg\min_{\mathbf{w}} &\frac{1}{2}\mathbf{w}^T\left(\gamma\mathbf{R}_{n-1} + \theta\mathbf{x}_n\mathbf{x}_n^T\right)\mathbf{w} \\ &- \mathbf{w}^T(\gamma\mathbf{r}_{n-1} + \theta\mathbf{x}_n y_n) + \lambda\|\mathbf{w}\|_{1,\infty}\end{aligned} \tag{4.16}$$

where θ is a constant between 0 and 1. $\hat{\mathbf{w}}_n$ and $\hat{\mathbf{w}}_{n-1}$ of problem (4.15) can be expressed as

$$\hat{\mathbf{w}}_{n-1} = f(0, \gamma\lambda),$$

and

$$\hat{\mathbf{w}}_n = f(1, \lambda).$$

Our proposed method computes $\hat{\mathbf{w}}_n$ from $\hat{\mathbf{w}}_{n-1}$ in the following two steps:

Step 1. Fix $\theta = 0$ and calculate $f(0, \lambda)$ from $f(0, \gamma\lambda)$. This is accomplished by computing the regularization path between $\gamma\lambda$ and λ using homotopy methods introduced for the non-recursive $\ell_{1,\infty}$ group lasso. The solution path is piecewise linear and the algorithm is described in [1].

Step 2. Fix λ and calculate the solution path between $f(0,\lambda)$ and $f(1,\lambda)$. This is the key problem addressed in this work.

To ease the notations we denote $\mathbf{x}_n$ and y_n by $\mathbf{x}$ and y, respectively, and define the following variables:

$$\mathbf{R}(\theta) = \gamma \mathbf{R}_{n-1} + \theta \mathbf{x}\mathbf{x}^T \tag{4.17}$$

$$\mathbf{r}(\theta) = \gamma \mathbf{r}_{n-1} + \theta \mathbf{x} y. \tag{4.18}$$

Problem (4.16) is then

$$f(\theta,\lambda) = \arg\min_{\mathbf{w}} \frac{1}{2}\mathbf{w}^T\mathbf{R}(\theta)\mathbf{w} - \mathbf{w}^T\mathbf{r}(\theta) + \lambda\|\mathbf{w}\|_{1,\infty}. \tag{4.19}$$

In Section 4.4 we will show how to propagate $f(0,\lambda)$ to $f(1,\lambda)$ using the homotopy approach applied to (4.19).

4.4 Online homotopy update

4.4.1 Set notation

We begin by introducing a series of set definitions. Figure 4.2 provides an example. We divide the entire group index set into $\mathcal{P}$ and Q, respectively, where $\mathcal{P}$ contains active groups and $\mathcal{Q}$ is its complement. For each active group $m \in \mathcal{P}$, we partition the group into two parts: the maximal values, with indices $\mathcal{A}_m$, and the rest of the values, with indices $\mathcal{B}_m$:

$$\mathcal{A}_m = \arg\max_{i \in \mathcal{G}_m} |w_i|, m \in \mathcal{P},$$

and

$$\mathcal{B}_m = \mathcal{G}_m - \mathcal{A}_m.$$

The set $\mathcal{A}$ and $\mathcal{B}$ are defined as the union of the $\mathcal{A}_m$ and $\mathcal{B}_m$ sets, respectively:

$$\mathcal{A} = \bigcup_{m \in \mathcal{P}} \mathcal{A}_m, \quad \mathcal{B} = \bigcup_{m \in \mathcal{P}} \mathcal{B}_m.$$

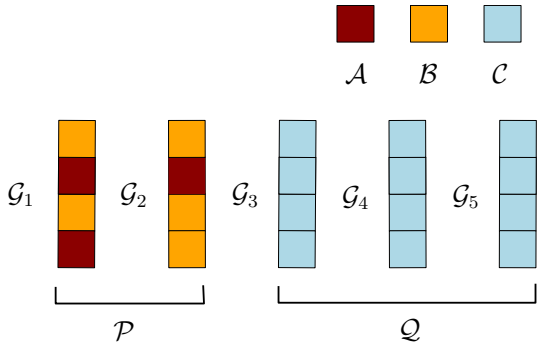


Figure 4.2: Illustration of the partitioning of a 20 element coefficient vector $\mathbf{w}$ into 5 groups of 4 indices. The sets $\mathcal{P}$ and $\mathcal{Q}$ contain the active groups and the inactive groups, respectively. Within each of the two active groups the maximal coefficients are denoted by the dark red color.

Finally, we define

$$\mathcal{C} = \bigcup_{m \in \mathcal{Q}} \mathcal{G}_m.$$

and

$$\mathcal{C}_m = \mathcal{G}_m \cap \mathcal{C}.$$

4.4.2 Optimality condition

The objective function in (4.19) is convex but non-smooth as the $\ell_{1,\infty}$ norm is non-differentiable. Therefore, problem (4.19) reaches its global minimum at $\mathbf{w}$ if and only if the sub-differential of the objective function contains the zero vector. Let $\partial\|\mathbf{w}\|_{1,\infty}$ denote the sub-differential of the $\ell_{1,\infty}$ norm at $\mathbf{w}$. A vector $\mathbf{z} \in \partial\|\mathbf{w}\|_{1,\infty}$ only if $\mathbf{z}$ satisfies the following conditions [89, 1]:

$$\|\mathbf{z}_{\mathcal{A}_m}\|_1 = 1, m \in \mathcal{P}, \tag{4.20}$$

$$\operatorname{sgn}\left(\mathbf{z}_{\mathcal{A}_m}\right) = \operatorname{sgn}\left(\mathbf{w}_{\mathcal{A}_m}\right), m \in \mathcal{P}, \tag{4.21}$$

$$\mathbf{z}_{\mathcal{B}} = \mathbf{0}, \tag{4.22}$$

$$\|\mathbf{z}_{\mathcal{C}_m}\|_1 \leq 1, m \in \mathcal{Q}, \tag{4.23}$$

where $\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{P}$ and $\mathcal{Q}$ are θ-dependent sets defined on $\mathbf{w}$ as defined in Section 4.4.1.

For notational convenience we drop θ in $\mathbf{R}(\theta)$ and $\mathbf{r}(\theta)$ leaving the θ-dependency implicit. The optimality condition is then written as

$$\mathbf{R}\mathbf{w} - \mathbf{r} + \lambda \mathbf{z} = \mathbf{0}, \quad \mathbf{z} \in \partial \|\mathbf{w}\|_{\mathbf{1},\infty}. \tag{4.24}$$

As $\mathbf{w}_{\mathcal{C}} = \mathbf{0}$ and $\mathbf{z}_{\mathcal{B}} = \mathbf{0}$, (4.24) implies the three conditions

$$\mathbf{R}_{\mathcal{A}\mathcal{A}}\mathbf{w}_{\mathcal{A}} + \mathbf{R}_{\mathcal{A}\mathcal{B}}\mathbf{w}_{\mathcal{B}} - \mathbf{r}_{\mathcal{A}} + \lambda \mathbf{z}_{\mathcal{A}} = \mathbf{0}, \tag{4.25}$$

$$\mathbf{R}_{\mathcal{B}\mathcal{A}}\mathbf{w}_{\mathcal{A}} + \mathbf{R}_{\mathcal{B}\mathcal{B}}\mathbf{w}_{\mathcal{B}} - \mathbf{r}_{\mathcal{B}} = \mathbf{0}, \tag{4.26}$$

$$\mathbf{R}_{\mathcal{C}\mathcal{A}}\mathbf{w}_{\mathcal{A}} + \mathbf{R}_{\mathcal{C}\mathcal{B}}\mathbf{w}_{\mathcal{B}} - \mathbf{r}_{\mathcal{C}} + \lambda \mathbf{z}_{\mathcal{C}} = \mathbf{0}. \tag{4.27}$$

The vector $\mathbf{w}_{\mathcal{A}}$ lies in a low dimensional subspace. Indeed, by definition of $\mathcal{A}_m$, if $|\mathcal{A}_m| > 1$

$$|w_i| = |w_{i'}|, \ i, i' \in \mathcal{A}_m.$$

Therefore, for any active group $m \in \mathcal{P}$,

$$\mathbf{w}_{\mathcal{A}_m} = \mathbf{s}_{\mathcal{A}_m} \alpha_m \tag{4.28}$$

where

$$\alpha_m = \|\mathbf{w}_{\mathcal{G}_m}\|_\infty,$$

and

$$\mathbf{s}_{\mathcal{A}} = \operatorname{sgn}(\mathbf{w}_{\mathcal{A}}).$$

Using matrix notation, we represent (4.28) as

$$\mathbf{w}_{\mathcal{A}} = \mathbf{S}\mathbf{a}. \tag{4.29}$$

where

$$\mathbf{S} = \begin{pmatrix} \mathbf{s}_{A_1} & & \\ & \ddots & \\ & & \mathbf{s}_{A_{|P|}} \end{pmatrix} \tag{4.30}$$

is a $|\mathcal{A}| \times |\mathcal{P}|$ sign matrix and the vector $\mathbf{a}$ is comprised of $\alpha_m, m \in \mathcal{P}$.

The solution to (4.19) can be determined in closed form if the sign matrix $\mathbf{S}$ and sets $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{P}, \mathcal{Q})$ are available. Indeed, from (4.20) and (4.21)

$$\mathbf{S}^T \mathbf{z}_{\mathcal{A}} = \mathbf{1}, \tag{4.31}$$

where $\mathbf{1}$ is a $|\mathcal{P}| \times 1$ vector comprised of 1's. With (4.29) and (4.31), (4.25) and (4.26) are equivalent to

$$\begin{aligned} &\mathbf{S}^T \mathbf{R}_{\mathcal{A}\mathcal{A}} \mathbf{S}\mathbf{a} + \mathbf{S}^T \mathbf{R}_{\mathcal{A}\mathcal{B}} \mathbf{w}_{\mathcal{B}} - \mathbf{S}^T \mathbf{r}_{\mathcal{A}} + \lambda \mathbf{1} = 0, \\ &\mathbf{R}_{\mathcal{B}\mathcal{A}} \mathbf{S}\mathbf{a} + \mathbf{R}_{\mathcal{B}\mathcal{B}} \mathbf{w}_{\mathcal{B}} - \mathbf{r}_{\mathcal{B}} = \mathbf{0}. \end{aligned} \tag{4.32}$$

Therefore, by defining the (a.s. invertible) matrix

$$\mathbf{H} = \begin{pmatrix} \mathbf{S}^T \mathbf{R}_{\mathcal{A}\mathcal{A}} \mathbf{S} & \mathbf{S}^T \mathbf{R}_{\mathcal{A}\mathcal{B}} \\ \mathbf{R}_{\mathcal{B}\mathcal{A}} \mathbf{S} & \mathbf{R}_{\mathcal{B}\mathcal{B}} \end{pmatrix}, \tag{4.33}$$

and

$$\mathbf{b} = \begin{pmatrix} \mathbf{S}^T \mathbf{r}_{\mathcal{A}} \\ \mathbf{r}_{\mathcal{B}} \end{pmatrix}, \mathbf{v} = \begin{pmatrix} \mathbf{a} \\ \mathbf{w}_B \end{pmatrix}, \tag{4.34}$$

(4.32) is equivalent to $\mathbf{H}\mathbf{v} = \mathbf{b} - \lambda \mathbf{e}$, where $\mathbf{e} = (\mathbf{1}^T, \mathbf{0}^T)^T$, so that

$$\mathbf{v} = \mathbf{H}^{-1}(\mathbf{b} - \lambda \mathbf{e}). \tag{4.35}$$

As $\mathbf{w}_{\mathcal{A}} = \mathbf{S}\mathbf{a}$, the solution vector $\mathbf{w}$ can be directly obtained from $\mathbf{v}$ via (4.34). For the sub-gradient vector, it can be shown that

$$\lambda \mathbf{z}_{\mathcal{A}} = \mathbf{r}_{\mathcal{A}} - (\mathbf{R}_{\mathcal{A}\mathcal{A}} \mathbf{S} \; \mathbf{R}_{\mathcal{A}\mathcal{B}}) \, \mathbf{v}, \tag{4.36}$$

$$\mathbf{z}_{\mathcal{B}} = \mathbf{0} \tag{4.37}$$

and

$$\lambda \mathbf{z}_{\mathcal{C}} = \mathbf{r}_{\mathcal{C}} - (\mathbf{R}_{\mathcal{C}\mathcal{A}} \mathbf{S} \; \mathbf{R}_{\mathcal{C}\mathcal{B}}) \, \mathbf{v}. \tag{4.38}$$

4.4.3 Online update

Now we consider (4.19) using the results in 4.4.2. Let θ_0 and θ_1 be two constants such that $\theta_1 > \theta_0$. For a given value of $\theta \in [\theta_0, \theta_1]$ define the class of sets $\mathcal{S} = (\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{P}, \mathcal{Q})$ and make θ explicit by writing $\mathcal{S}(\theta)$. Recall that $\mathcal{S}(\theta)$ is specified by the solution $f(\theta, \lambda)$ defined in (19). Assume that $\mathcal{S}(\theta)$ does not change for $\theta \in [\theta_0, \theta_1]$. The following theorem propagates $f(\theta_0, \lambda)$ to $f(\theta_1, \lambda)$ via a simple algebraic relation.

Theorem IV.1. *Let θ_0 and θ_1 be two constants such that $\theta_1 > \theta_0$ and for any $\theta \in [\theta_0, \theta_1]$ the solutions to (4.19) share the same sets $\mathcal{S} = (\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{P}, \mathcal{Q})$. Let $\mathbf{v}'$ and $\mathbf{v}$ be vectors defined as $f(\theta_1, \lambda)$ and $f(\theta_0, \lambda)$, respectively. Then*

$$\mathbf{v}' = \mathbf{v} + \frac{\theta_1 - \theta_0}{1 + \sigma_H^2 \theta_1}(y - \hat{y})\mathbf{g}, \tag{4.39}$$

and the corresponding sub-gradient vector has the explicit update

$$\lambda \mathbf{z}'_{\mathcal{A}} = \lambda \mathbf{z}_{\mathcal{A}} + \frac{\theta_1 - \theta_0}{1 + \sigma_H^2 \theta_1} (y - \hat{y}) \{\mathbf{x}_{\mathcal{A}} - (\mathbf{R}_{\mathcal{A}\mathcal{A}}\mathbf{S}\ \mathbf{R}_{\mathcal{A}\mathcal{B}})\mathbf{g}\} \tag{4.40}$$

and

$$\lambda \mathbf{z}'_{\mathcal{C}} = \lambda \mathbf{z}_{\mathcal{C}} + \frac{\theta_1 - \theta_0}{1 + \sigma_H^2 \theta_1} (y - \hat{y}) \{\mathbf{x}_{\mathcal{C}} - (\mathbf{R}_{\mathcal{C}\mathcal{A}}\mathbf{S}\ \mathbf{R}_{\mathcal{C}\mathcal{B}})\mathbf{g}\}, \tag{4.41}$$

where $\mathbf{R} = \mathbf{R}(0)$ as defined in (4.17), $(\mathbf{x}, y)$ is the new sample as defined in (4.17) and (4.18), the sign matrix $\mathbf{S}$ is obtained from the solution at $\theta = \theta_0$, $\mathbf{H}_0$ is calculated from (4.33) using $\mathbf{S}$ and $\mathbf{R}(0)$, and $\mathbf{d}$, $\mathbf{u}$, $\hat{y}$ and σ_H^2 are defined by

$$\mathbf{d} = \begin{pmatrix} \mathbf{S}^T \mathbf{x}_{\mathcal{A}} \\ \\ \mathbf{x}_{\mathcal{B}} \end{pmatrix}, \tag{4.42}$$

$$\mathbf{g} = \mathbf{H}_0^{-1}\mathbf{d}, \tag{4.43}$$

$$\hat{y} = \mathbf{d}^T \mathbf{v}, \tag{4.44}$$

$$\sigma_H^2 = \mathbf{d}^T \mathbf{g}. \tag{4.45}$$

The proof of Theorem IV.1 is provided in Appendix A. Theorem IV.1 provides the closed form update for the solution path $f(\theta_0, \lambda) \to f(\theta_1, \lambda)$, under the assumption that the associated sets $\mathcal{S}(\theta)$ remain unaltered over the path.

Next, we partition the range $\theta \in [0, 1]$ into contiguous segments over which $\mathbf{S}(\theta)$ is piecewise constant. Within each segment we can use Theorem 1 to propagate the solution from left endpoint to right endpoint. Below we specify an algorithm for finding the endpoints of each of these segments.

Fix an endpoint θ_0 of one of these segments. We seek a *critical point* θ_1 that is defined as the maximum θ_1 ensuring $\mathcal{S}(\theta)$ remains unchanged within $[\theta_0, \theta_1]$. By increasing θ_1 from θ_0, the sets $\mathcal{S}(\theta)$ will not change until at least one of the following conditions are met:

Condition 1. There exists $i \in \mathcal{A}$ such that $z_i' = 0$;

Condition 2. There exists $i \in \mathcal{B}_m$ such that $|w_i'| = \alpha_m'$;

Condition 3. There exists $m \in \mathcal{P}$ such that $\alpha_m' = 0$;

Condition 4. There exists $m \in \mathcal{Q}$ such that $\|\mathbf{z}_{\mathcal{C}_m}'\|_1 = 1$.

Condition 1 is from (4.21) and (4.22), Condition 2 and 3 are based on definitions of $\mathcal{A}$ and $\mathcal{P}$, respectively, and Condition 4 comes from (4.20) and (4.23). Following [95, 1], the four conditions can be assumed to be mutually exclusive. The actions with respect to Conditions 1-4 are given by

Action 1. Move the entry i from $\mathcal{A}$ to $\mathcal{B}$:

$$\mathcal{A} \leftarrow \mathcal{A} - \{i\}, \mathcal{B} \leftarrow \mathcal{B} \cup \{i\};$$

Action 2. Move the entry i from $\mathcal{B}$ to $\mathcal{A}$:

$$\mathcal{A} \leftarrow \mathcal{A} \cup \{i\}, \mathcal{B} \leftarrow \mathcal{B} - \{i\};$$

Action 3. Remove group m from the active group list

$$\mathcal{P} \leftarrow \mathcal{P} - \{m\}, \mathcal{Q} \leftarrow \mathcal{Q} \cup \{m\},$$

and update the related sets

$$\mathcal{A} \leftarrow \mathcal{A} - \mathcal{A}_m, \mathcal{C} \leftarrow \mathcal{C} \cup \mathcal{A}_m;$$

Action 4. Select group m

$$\mathcal{P} \leftarrow \mathcal{P} \cup \{m\}, \mathcal{Q} \leftarrow \mathcal{Q} - \{m\},$$

and update the related sets

$$\mathcal{A} \leftarrow \mathcal{A} \cup \mathcal{C}_m, \mathcal{C} \leftarrow \mathcal{C} - \mathcal{C}_m.$$

By Theorem IV.1, the solution update from θ_0 to θ_1 is in closed form. The critical point of θ_1 can be determined in a straightforward manner (details are provided in Appendix B). Let $\theta_1^{(k)}, k = 1, ..., 4$ be the minimum value that is greater than θ_0 and meets Condition 1-4, respectively. The critical point θ_1 is then

$$\theta_1 = \min_{k=1,...,4} \theta_1^{(k)}.$$

4.4.4 Homotopy algorithm implementation

We now have all the ingredients for the homotopy update algorithm and the pseudo code is given in Algorithm 1.

Next we analyze the computational cost of Algorithm 1. The complexity to compute each critical point is summarized in Table 4.1, where N is the dimension of $\mathbf{H}_0$. As $N = |\mathcal{P}| + |\mathcal{B}| \leq |\mathcal{A}| + |\mathcal{B}|$, N is upper bounded by the number of non-zeros in the solution vector. The vector $\mathbf{g}$ can be computed in $\mathcal{O}(N^2)$ time using

Algorithm 1: Homotopy update from $f(0,\lambda)$ to $f(1,\lambda)$.

Input : $f(0,\lambda), \mathbf{R}(0), \mathbf{x}, \mathbf{y}$
output: $f(1,\lambda)$

Initialize $\theta_0 = 0$, $\theta_1 = 0$, $\mathbf{R} = \mathbf{R}(0)$;
Calculate $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{P}, \mathcal{Q})$ and $(\mathbf{v}, \lambda\mathbf{z}_{\mathcal{A}}, \lambda\mathbf{z}_{\mathcal{C}})$ from $f(0,\lambda)$;
while $\theta_0 < 1$ **do**
  Calculate the environmental variables $(\mathbf{S}, \mathbf{H}_0, \mathbf{d}, \mathbf{g}, \hat{y}, \sigma_H^2)$ from $f(\theta_0,\lambda)$ and $\mathbf{R}$;
  Calculate $\{\theta_1^{(k)}\}_{k=1}^4$ that meets Condition 1-4, respectively;
  Calculate the critical point θ_1 that meets Condition k_*: $k_* = \arg\min_k \theta_1^{(k)}$ and $\theta_1 = \theta_1^{(k_*)}$;
  if $\theta_1 \leq 1$ **then**
    Update $(\mathbf{v}, \lambda\mathbf{z}_{\mathcal{A}}, \lambda\mathbf{z}_{\mathcal{C}})$ using (4.39), (4.40) and (4.41);
    Update $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{P}, \mathcal{Q})$ by Action k_*;
    $\theta_0 = \theta_1$;
  else
    break;
  end
end
$\theta_1 = 1$;
Update $(\mathbf{v}, \lambda\mathbf{z}_{\mathcal{A}}, \lambda\mathbf{z}_{\mathcal{C}})$ using (4.39);
Calculate $f(1,\lambda)$ from $\mathbf{v}$.

the matrix-inverse lemma [91] and the fact that, for each action, $\mathbf{H}_0$ is at most perturbed by a rank-two matrix. This implies that the computation complexity per critical point is $\mathcal{O}(p \max\{N, \log p\})$ and the total complexity of the online update is $\mathcal{O}(k_2 \cdot p \max\{N, \log p\})$, where k_2 is the number of critical points of θ in the solution path $f(0,\lambda) \to f(1,\lambda)$. This is the computational cost required for Step 2 in Section 4.3.3.

A similar analysis can be performed for the complexity of Step 1, which requires $\mathcal{O}(k_1 \cdot p \max\{N, \log p\})$ where k_1 is the number of critical points in the solution path $f(0,\gamma\lambda) \to f(0,\lambda)$. Therefore, the overall computation complexity of the recursive $\ell_{1,\infty}$ group lasso is $\mathcal{O}(k \cdot p \max\{N, \log p\})$, where $k = k_1 + k_2$, *i.e.,* the total number of critical points in the solution path $f(0,\gamma\lambda) \to f(0,\lambda) \to f(1,\lambda)$.

An instructive benchmark is to directly solve the n-samples problem (4.16) from the solution path $f(1,\infty)$ (*i.e.,* a zero vector) $\to f(1,\lambda)$ [1], without using the previous solution $\hat{\mathbf{w}}_{n-1}$. This algorithm, called iCap in [1], requires $\mathcal{O}(k' \cdot p \max\{N, \log p\})$,

$\mathbf{g} = \mathbf{H}_0^{-1}\mathbf{d}$	$\mathcal{O}(N^2)$
$\mathbf{x}_{\mathcal{A}} - (\mathbf{R}_{\mathcal{A}\mathcal{A}}\mathbf{S}\ \mathbf{R}_{\mathcal{A}\mathcal{B}})\mathbf{g}$	$\mathcal{O}(\lvert\mathcal{A}\rvert N)$
$\mathbf{x}_{\mathcal{C}} - (\mathbf{R}_{\mathcal{C}\mathcal{A}}\mathbf{S}\ \mathbf{R}_{\mathcal{C}\mathcal{B}})\mathbf{g}$	$\mathcal{O}(\lvert\mathcal{C}\rvert N)$
$\theta_1^{(1)}$	$\mathcal{O}(\lvert\mathcal{A}\rvert)$
$\theta_1^{(2)}$	$\mathcal{O}(\lvert\mathcal{B}\rvert)$
$\theta_1^{(3)}$	$\mathcal{O}(\lvert\mathcal{P}\rvert)$
$\theta_1^{(4)}$	$\mathcal{O}(\lvert\mathcal{C}\rvert \log \lvert\mathcal{C}\rvert)$

Table 4.1: Computation costs of online homotopy update for each critical point.

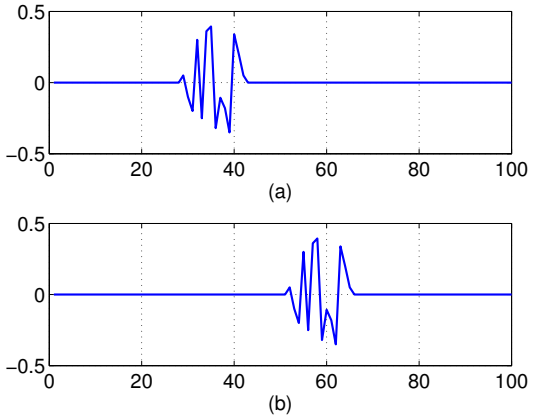


Figure 4.3: Responses of the time varying system. (a): Initial response. (b): Response after the 200th iteration. The groups for Algorithm 1 were chosen as 20 equal size contiguous groups of coefficients partitioning the range $1, \ldots, 100$.

where k' is the number of critical points in $f(1, \infty) \to f(1, \lambda)$. Empirical comparisons between k and k', provided in the following section, indicate that iCap requires significantly more computation than our proposed Algorithm 1.

4.5 Numerical simulations

In this section we demonstrate our proposed recursive $\ell_{1,\infty}$ group lasso algorithm by numerical simulation. We simulated the model $y_j = \mathbf{w}_*^T \mathbf{x}_j + v_j$, $j = 1, \ldots, 400$, where v_j is a zero mean Gaussian noise and $\mathbf{w}_*$ is a sparse $p = 100$ element vector containing only 14 non-zero coefficients clustered between indices 29 and 42. See Fig. 4.3 (a). After 200 time units, the locations of the non-zero coefficients of $\mathbf{w}_*$ is shifted to the right, as indicated in Fig. 4.3 (b).

The input vectors were generated as independent identically distributed Gaussian

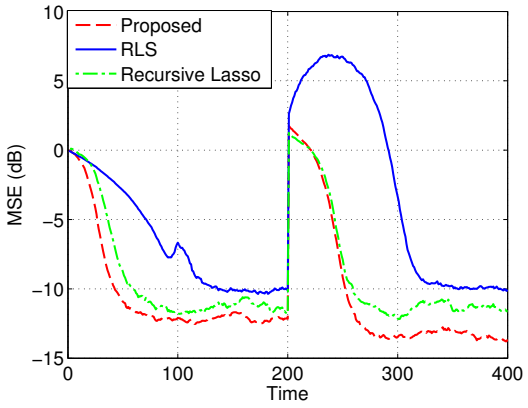


Figure 4.4: Averaged MSE of the proposed algorithm, RLS and recursive lasso.

random vectors with zero mean and identity covariance matrix, and the variance of observation noise v_j is 0.01. We created the groups in the recursive $\ell_{1,\infty}$ group lasso as follows. We divide the 100 RLS filter coefficients $\mathbf{w}$ into 20 groups with group boundaries $1, 5, 10, \ldots$, where each group contains 5 coefficients. The forgetting factor γ and the regularization parameter λ were set to 0.9 and 0.1, respectively. We repeated the simulation 100 times and the averaged mean squared errors of the RLS, sparse RLS and proposed RLS shown in Fig. 4.4. We implemented the standard RLS and sparse RLS using the ℓ_1 regularization, where the forgetting factors are also set to 0.9. We implemented sparse RLS [93] by choosing the regularization parameter λ which achieves the lowest steady-state error, resulting in $\lambda = 0.05$.

It can be seen from Fig. 4.4 that our proposed sparse RLS method outperforms standard RLS and sparse RLS in both convergence rate and steady-state MSE. This demonstrates the power of our group sparsity penalty. At the change point of 200 iterations, both the proposed method and sparse RLS of [93] show superior tracking performances as compared to the standard RLS. We also observe that the proposed method achieves even smaller MSE after the change point occurs. This is due to the fact that the active cluster spans across group boundaries in the initial system (Fig.

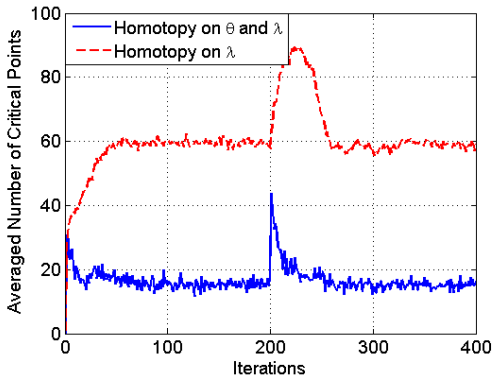


Figure 4.5: Averaged number of critical points for the proposed recursive method of implementing $\ell_{1,\infty}$ lasso and the iCap[1] non-recursive method of implementation.

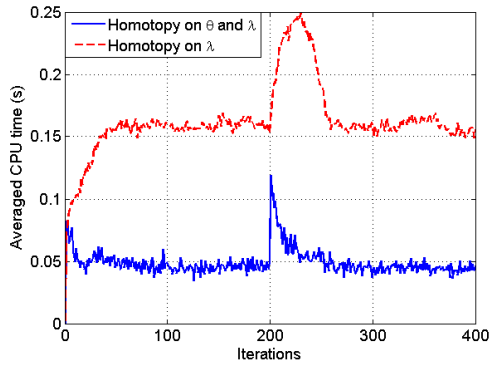


Figure 4.6: Averaged CPU time for the proposed recursive method of implementing $\ell_{1,\infty}$ lasso and the iCap non-recursive method of implementation. Algorithms are implemented using Intel (R) Core (TM) i5 CPU at 2.53GHz.

4.3 (a)), while the active clusters in the shifted system overlap with fewer groups.

Fig. 4.5 shows the average number of critical points (accounting for both trajectories in θ and λ) of the proposed algorithm, *i.e.,* the number k as defined in Section 4.4.4. As a comparison, we implement the iCap method of [1], a homotopy based algorithm that traces the solution path only over λ. The average number of critical points for iCap is plotted in Fig. 4.5, which is the number k' in Section 4.4.4. Both the proposed algorithm and iCap yield the same solution but have different computational complexities proportional to k and k', respectively. It can be seen that the proposed algorithm saves as much as 75% of the computation costs for equivalent

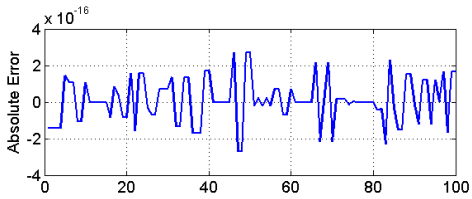


Figure 4.7: The absolute value of the difference vector between solutions from the proposed algorithm and iCap. The differences are within the range of quantization errors.

performance. We also compute the averaged CPU time of the proposed algorithm and iCap that shows similar advantages in Fig. 4.6. Finally, to validate the accuracy of the proposed algorithm we randomly choose an iteration index and plot the absolute value of the difference between solution vectors obtained by the proposed method and iCap for one trial of simulation as shown in Fig. 4.7. Considering the quantization error, it is observed that our algorithm and iCap yield the same result with machine precision.

4.6 Application to covariance-based anomaly detection in the Abilene Networks

In this section we perform an online covariance-based anomaly detection to demonstrate the proposed method. We consider a real world data set of Abilene, the Internet 2 backbone network. This networks carries traffic from universities in the United States. Fig. 4.8 shows its connectivity map consisting of 11 routers and 41 links. Each line corresponds to two links and there are additional links from each of the nodes to itself. In our experiments we use the 1008 samples of the traffic on each of the 41 Abilene links during April 7 - 13, 2003.

To remove the drifts and seasonal effects we detrended the data as follows. Let $x_{i,k}$ be the k-th sample of the i-th traffic link and denote

$$\mathbf{x}_k = (x_{1,k}, x_{2,k}, ..., x_{182,k})^T. \tag{4.46}$$

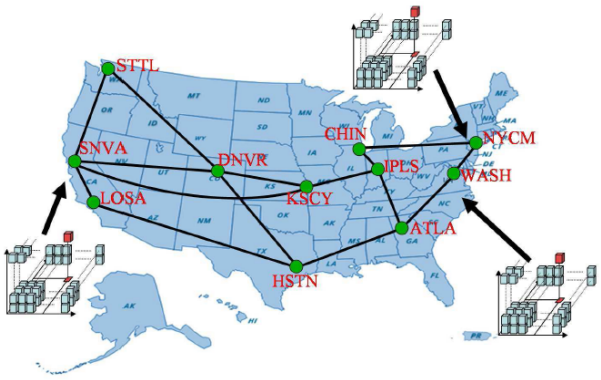


Figure 4.8: Map of Abilene router network.

The local mean value of $x_{i,k}$ is defined by

$$\bar{x}_{i,k} = \frac{1}{m+1} \sum_{j=0}^{m} x_{i,k-j}, \tag{4.47}$$

where m determines the local window size which is set to 40 in this study. Using the same local window we also calculated the empirical variance of each $x_{i,k}$, denoting as $\hat{s}^2_{i,k}$. We then detrend and normalize the raw samples $\mathbf{x}_k$ by

$$y_{i,k} = \frac{x_{i,k} - \bar{x}_{i,k}}{\hat{s}_{i,k}} \tag{4.48}$$

yielding a detrended sample $\mathbf{y}_k$ that is used for online anomaly detection.

We use a similar detect statistic t_k defined in Chapter III, *i.e.*,

$$t_k = \mathbf{y}_k^T \mathbf{C} \mathbf{y}_k, \tag{4.49}$$

where $\mathbf{C}$ is the inverse covariance such that $\mathbf{C}^{-1} = E\left\{\mathbf{y}_k \mathbf{y}_k^T\right\}$. Note that $\mathbf{C}$ may be time-varying and dependent on k. Now the anomaly detection problem turns to online estimation of the inverse covariance $\mathbf{C}$ and we use (4.1) to transfer the matrix estimation of $\mathbf{C}$ to p regression problems to estimate β_{ij}, $i, j = 1, ..., p$. As demonstrated in [96], the Abilene Network data can be well fit by an undirected Gaussian graph model where $\mathbf{C}$ is sparse, which produces sparse patterns in β_{ij} as

well. We thus apply the proposed recursive lasso algorithm for online estimation of β_{ij}, which yielding $\hat{\beta}_{ij}$. The inverse covariance estimate $\hat{\mathbf{C}}$ is estimated via $\hat{\beta}_{ij}$ using (4.4) and then plugged into (4.49) for anomaly detection.

As the sample dimension in this study is only 41, we treat each variable as a single group and the recursive $\ell_{1,\infty}$ lasso naturally reduces to recursive lasso. The forgetting factor γ is set to 0.95, where the effective sample size for each update is roughly estimated as $1/(1-\gamma) = 20$, and the regularization parameter λ is set to 1. The test statistic t_k for detection is plotted in Fig. 4.9 (a). As we do not have ground truth for $\mathbf{\Sigma}$, we use the sample covariance of the entire 1008 samples to approximate the true $\mathbf{\Sigma}$, and the corresponding test statistic is plotted in Fig. 4.9 (b). For comparison we also implement the OAS covariance estimator using local samples within a sliding windows of size $1/(1-\gamma) = 20$, and its test statistics is plotted in Fig. 4.9 (c). It can be observed that Fig. 4.9 (a) and (b) share similar global patterns but differ in local details. Considering that the sample covariance is obtained using 1008 samples, Fig. 4.9 (b) has significantly more smoothing than Fig. 4.9 (a), which is smoothed over only 21 samples. Thus Fig. 4.9 (a) has extracted local covariance patterns, indicated by arrows, that are not visible in Fig. 4.9 (b) and only weakly visible in Fig. 4.9 (c). On the other hand, the global patterns of Fig. 4.9 (b) and (c) are quite different to each other, which implies the shrinkage method (equivalent to ridge regressions on β_{ij}) is not appropriate in estimating the covariance of the Abilene Network data. To further illustrate this point, we also visualize heatmaps of the inverse sample covariance (using the entire 1008 samples), the estimates of $\mathbf{C}$ at the 900th sample using the proposed approach and OAS, respectively. It is clear that the global inverse sample covariance contains a sparse pattern which justifies that ℓ_1 type penalized regression (as used in the proposed approach) is more appropriate

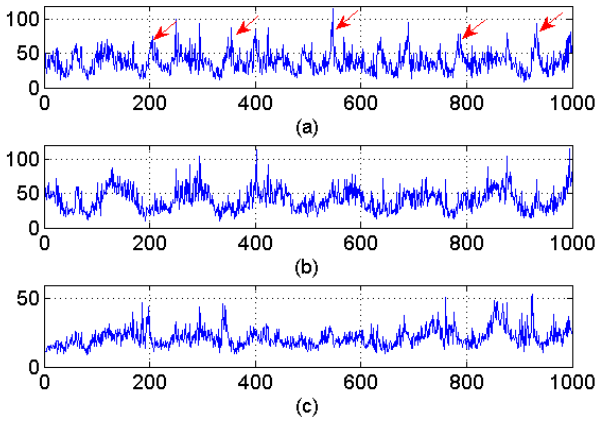
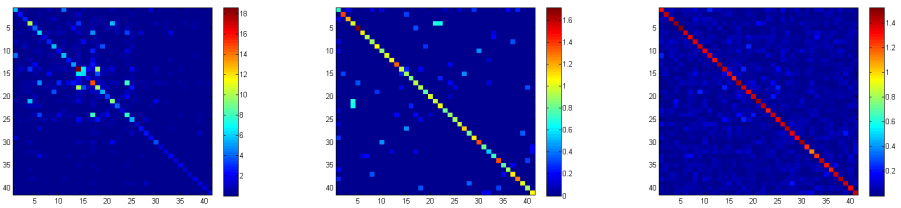


Figure 4.9: Covariance-based anomaly detection using the real world dataset of Abilene, where the traffic data on the 41 Abilene links were recorded during April 7-13, 2007: (a) online covariance estimation using the proposed algorithm; (b) sample covariance using the entire data; (c) online covariance estimation using the OAS estimator proposed in Chapter 2. (a) and (b) share similar global patterns while local details (possible anomalies) are more evident in (a), which are marked by arrows. (c) also captures similar details of (a) but loses the global pattern of (b).

than ℓ_2 norm based ridge regressions (equivalent to Steinian shrinkage in OAS) in this study.



(a) Inverse of the global sample covariance

(b) Proposed

(c) Inverse of the local OAS estimator

Figure 4.10: Heatmaps of inverse covariance estimates (shown in absolute values): (a) the inverse sample covariance using the entire data; (b) the estimated local inverse covariance at the 900th sample using the proposed approach; (c) the estimated local inverse covariance at the 900th sample using the OAS estimator.

4.7 Conclusion

In this chapter we considered the online (inverse) covariance estimation for time-varying systems, where the variables are associated with a sparse Gaussian graphical model. We proposed to use results of multiple decoupled regression problems to ap-

proximate the inverse covariance estimate. Each regression problem was solved by a regularized recursive least squares problem and the solution vector was updated online efficiently. Specifically we proposed a $\ell_{1,\infty}$ regularized RLS algorithm for online sparse linear prediction. We developed a homotopy based method to sequentially update the solution vector as new measurements are acquired. Our proposed algorithm uses the previous estimate as a "warm-start", from which we compute the homotopy update to the current solution. The proposed algorithm can process streaming measurements with time varying predictors and is computationally efficient compared to non-recursive group lasso solvers. Numerical simulations demonstrated that the proposed method outperformed the standard and ℓ_1 regularized RLS for identifying an unknown group sparse system, in terms of both tracking and steady-state mean squared error. Finally, we incorporated the proposed recursive lasso into online covariance estimation and demonstrated its performance using the Abilene Networks data.

4.8 Appendix

4.8.1 Proof of Theorem IV.1

We begin by deriving (4.39). According to (4.35),

$$\mathbf{v}' = \mathbf{H}'^{-1}(\mathbf{b}' - \lambda \mathbf{e}'). \tag{4.50}$$

As $\mathbf{S}$ and $(\mathcal{A}, \mathcal{B}, \mathcal{C}, \mathcal{P}, \mathcal{Q})$ remain constant within $[\theta_0, \theta_1]$,

$$\mathbf{e}' = \mathbf{e}, \tag{4.51}$$

$$\mathbf{b}' = \mathbf{b} + \delta \mathbf{d} y, \tag{4.52}$$

and

$$\mathbf{H}' = \mathbf{H} + \delta \mathbf{d}\mathbf{d}^T,$$

where

$$\delta = \theta_1 - \theta_0,$$

$\mathbf{H}$ and $\mathbf{b}$ are calculated using $\mathbf{S}$ within $[\theta_0, \theta_1]$ and $\mathbf{R}(\theta_0)$ and $\mathbf{r}(\theta_0)$, respectively. We emphasize that $\mathbf{H}$ is based on $\mathbf{R}(\theta)$ and is different from $\mathbf{H}_0$ defined in Theorem IV.1. According to the Sherman-Morrison-Woodbury formula,

$$\mathbf{H}'^{-1} = \mathbf{H}^{-1} - \frac{\delta}{1+\sigma^2\delta}(\mathbf{H}^{-1}\mathbf{d})(\mathbf{H}^{-1}\mathbf{d})^T, \tag{4.53}$$

where $\sigma^2 = \mathbf{d}^T\mathbf{H}^{-1}\mathbf{d}$. Substituting (4.51), (4.52) and (4.53) into (4.50), after simplification we obtain

$$\begin{aligned}
\mathbf{v}' &= \left(\mathbf{H}^{-1} - \frac{\delta}{1+\sigma^2\delta}(\mathbf{H}^{-1}\mathbf{d})(\mathbf{H}^{-1}\mathbf{d})^T\right)(\mathbf{b} + \delta\mathbf{d}y - \lambda\mathbf{e}) \\
&= \mathbf{H}^{-1}(\mathbf{b} - \lambda\mathbf{e}) + \mathbf{H}^{-1}\delta\mathbf{d}y \\
&\quad - \frac{\delta}{1+\sigma^2\delta}\mathbf{H}^{-1}\mathbf{d}\mathbf{d}^T\mathbf{H}^{-1}(\mathbf{b} - \lambda\mathbf{e}) - \frac{\sigma^2\delta^2}{1+\sigma^2\delta}\mathbf{H}^{-1}\mathbf{d}y \\
&= \mathbf{v} + \frac{\delta}{1+\sigma^2\delta}(y - \mathbf{d}^T\mathbf{v})\mathbf{H}^{-1}\mathbf{d} \\
&= \mathbf{v} + \frac{\delta}{1+\sigma^2\delta}(y - \hat{y})\mathbf{H}^{-1}\mathbf{d},
\end{aligned} \tag{4.54}$$

where $\hat{y} = \mathbf{d}^T\mathbf{v}$ as defined in (4.44).

Note that $\mathbf{H}$ is defined in terms of $\mathbf{R}(\theta_0)$ rather than $\mathbf{R}(0)$ and

$$\mathbf{H} = \mathbf{H}_0 + \theta_0\mathbf{d}\mathbf{d}^T,$$

so that

$$\mathbf{H}^{-1} = \mathbf{H}_0^{-1} - \frac{\theta_0}{1+\sigma_H^2\theta_0}\mathbf{g}\mathbf{g}^T, \tag{4.55}$$

where $\mathbf{g}$ and σ_H^2 are defined by (4.43) and (4.45), respectively. As $\sigma_H^2 = \mathbf{d}^T\mathbf{g}$,

$$\mathbf{H}^{-1}\mathbf{d} = \mathbf{H}_0^{-1}\mathbf{d} - \frac{\sigma_H^2\theta_0}{1+\sigma_H^2\theta_0}\mathbf{g}. \tag{4.56}$$

Accordingly,

$$\sigma^2 = \mathbf{d}^T\mathbf{H}^{-1}\mathbf{d} = \sigma_H^2 - \frac{\sigma_H^2\theta_0}{1+\sigma_H^2\theta_0}\sigma_H^2 = \frac{\sigma_H^2}{1+\sigma_H^2\theta_0}. \tag{4.57}$$

Substituting (4.56) and (4.57) to (4.54), we finally obtain

$$\mathbf{v}' = \mathbf{v} + \frac{\delta}{1+\sigma_H^2\theta_1}(y-\hat{y})\mathbf{g} = \mathbf{v} + \frac{\theta_1-\theta_0}{1+\sigma_H^2\theta_1}(y-\hat{y})\mathbf{g}.$$

Equations (4.40) and (4.41) can be established by direct substitutions of (4.39) into their definitions (4.36) and (4.38) and thus the proof of Theorem IV.1 is complete.

4.8.2 Computation of critical points

For ease of notation we work with ρ, defined by

$$\rho = \frac{\theta_1-\theta_0}{1+\sigma_H^2\theta_1}. \tag{4.58}$$

It is easy to see that over the range $\theta_1 > \theta_0$, ρ is monotonically increasing in $(0, 1/\sigma_H^2)$. Therefore, (4.58) can be inverted by

$$\theta_1 = \frac{\rho+\theta_0}{1-\sigma_H^2\rho}, \tag{4.59}$$

where $\rho \in (0, 1/\sigma_H^2)$ to ensure $\theta_1 > \theta_0$.

Suppose we have obtained $\rho^{(k)}, k = 1, ..., 4$, $\theta_1^{(k)}$ can be calculated using (4.59) and the critical point θ_1 is then

$$\theta_1 = \min_{k=1,...,4} \theta_1^{(k)}.$$

We now calculate the critical value of ρ for each condition one by one.

Critical point for Condition 1

Define the temporary vector

$$\mathbf{t}_{\mathcal{A}} = (y-\hat{y})\left\{\mathbf{x}_{\mathcal{A}} - (\mathbf{R}_{\mathcal{A}\mathcal{A}}\mathbf{S}\ \mathbf{R}_{\mathcal{A}\mathcal{B}})\,\mathbf{g}\right\}.$$

According to (4.40),

$$\lambda \mathbf{z}'_{\mathcal{A}} = \lambda \mathbf{z}_{\mathcal{A}} + \rho \mathbf{t}_{\mathcal{A}}.$$

Condition 1 is met for any $\rho = \rho_i^{(1)}$ such that

$$\rho_i^{(1)} = -\frac{\lambda z_i}{t_i}, i \in \mathcal{A}.$$

Therefore, the critical value of ρ that satisfies Condition 1 is

$$\rho^{(1)} = \min \left\{ \rho_i^{(1)} \,\middle|\, i \in \mathcal{A}, \rho_i^{(1)} \in (0, 1/\sigma_H^2) \right\}.$$

Critical point for Condition 2

By the definition (4.34), $\mathbf{v}$ is a concatenation of α_m and $\mathbf{w}_{\mathcal{B}_m}, m \in \mathcal{P}$:

$$\mathbf{v}^T = \left((\alpha_m)_{m \in \mathcal{P}}, \mathbf{w}_{\mathcal{B}_1}^T, ..., \mathbf{w}_{\mathcal{B}_{|\mathcal{P}|}}^T \right), \tag{4.60}$$

where $(\alpha_m)_{m \in \mathcal{P}}$ denotes the vector comprised of $\alpha_m, m \in \mathcal{P}$. Now we partition the vector $\mathbf{g}$ in the same manner as (4.60) and denote τ_m and $\mathbf{u}_m$ as the counter part of α_m and $\mathbf{w}_{\mathcal{B}_m}$ in $\mathbf{g}$, *i.e.*,

$$\mathbf{g}^T = \left((\tau_m)_{m \in \mathcal{P}}, \mathbf{u}_1^T, ..., \mathbf{u}_{|\mathcal{P}|}^T \right).$$

Eq. (4.39) is then equivalent to

$$\alpha'_m = \alpha_m + \rho \tau_m, \tag{4.61}$$

and

$$w'_{\mathcal{B}_m,i} = w_{\mathcal{B}_m,i} + \rho u_{m,i},$$

where $u_{m.i}$ is the i-th element of the vector $\mathbf{u}_m$. Condition 2 indicates that

$$\alpha'_m = \pm w'_{\mathcal{B}_m,i},$$

and is satisfied if $\rho = \rho_{m,i}^{(2+)}$ or $\rho = \rho_{m,i}^{(2-)}$, where

$$\rho_{m,i}^{(2+)} = \frac{\alpha_m - w_{\mathcal{B}_m,i}}{u_{m,i} - \tau_m}, \ \rho_{m,i}^{(2-)} = -\frac{\alpha_m + w_{\mathcal{B}_m,i}}{u_{m,i} + \tau_m}.$$

Therefore, the critical value of ρ for Condition 2 is

$$\rho^{(2)} = \min\left\{\rho_{m,i}^{(2\pm)} \,\middle|\, m \in \mathcal{P}, i = 1, ..., |\mathcal{B}_m|, \rho_{m,i}^{(2\pm)} \in (0, 1/\sigma_H^2)\right\}.$$

Critical point for Condition 3

According to (4.61), $\alpha'_m = 0$ yields $\rho = \rho_i^{(3)}$ determined by

$$\rho_m^{(3)} = -\frac{\alpha_m}{\tau_m}, m \in \mathcal{P},$$

and the critical value for $\rho^{(3)}$ is

$$\rho^{(3)} = \min\left\{\rho_m^{(3)} \,\middle|, m \in \mathcal{P}, \rho_m^{(3)} \in (0, 1/\sigma_H^2)\right\}.$$

Critical point for Condition 4

Define

$$\mathbf{t}_{\mathcal{C}} = (y - \hat{y})\left\{\mathbf{x}_{\mathcal{C}} - (\mathbf{R}_{\mathcal{C}\mathcal{A}}\mathbf{S}\ \mathbf{R}_{\mathcal{C}\mathcal{B}})\,\mathbf{g}\right\}.$$

Eq. (4.41) is then

$$\lambda \mathbf{z}'_{\mathcal{C}_m} = \lambda \mathbf{z}_{\mathcal{C}_m} + \rho \mathbf{t}_{\mathcal{C}_m},$$

and Condition 4 is equivalent to

$$\sum_{i \in \mathcal{C}_m} |\rho t_i + \lambda z_i| = \lambda. \tag{4.62}$$

To solve (4.62) we develop a fast method that requires complexity of $\mathcal{O}(N \log N)$, where $N = |\mathcal{C}_m|$. The algorithm is given in Appendix C. For each $m \in \mathcal{Q}$, let $\rho_m^{(4)}$ be the minimum positive solution to (4.62). The critical value of ρ for Condition 4 is then

$$\rho^{(4)} = \min\left\{\rho_m^{(4)} \,\middle|\, m \in \mathcal{Q}, \rho_m^{(4)} \in (0, 1/\sigma_H^2)\right\}.$$

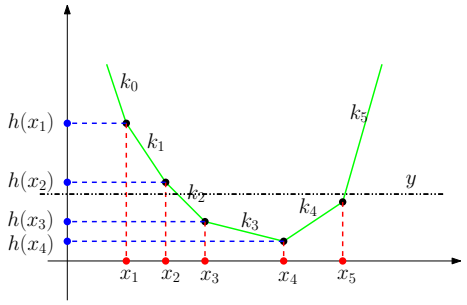


Figure 4.11: An illustration of the fast algorithm for critical condition 4.

4.8.3 Fast algorithm for critical condition 4

Here we develop an algorithm to solve problem (4.62). Consider solving the more general problem:

$$\sum_{i=1}^{N} a_i|x - x_i| = y, \tag{4.63}$$

where a_i and x_i are constants and $a_i > 0$. Please note that the notations here have no connections to those in previous sections. Define the following function

$$h(x) = \sum_{i=1}^{N} a_i|x - x_i|.$$

The problem is then equivalent to finding $h^{-1}(y)$, if it exists.

An illustration of the function $h(x)$ is shown in Fig. 4.11, where k_i denotes the slope of the ith segment. It can be shown that $h(x)$ is piecewise linear and convex in x. Therefore, the equation (4.63) generally has two solutions if they exist, denoted as $x_{\min}$ and $x_{\max}$. Based on piecewise linearity we propose a search algorithm to solve (4.63). The pseudo code is shown in Algorithm 2 and its computation complexity is dominated by the sorting operation which requires $\mathcal{O}(N \log N)$.

Algorithm 2: Solve x from $\sum_{i=1}^{N} a_i|x - x_i| = y$.

Input : $\{a_i, x_i\}_{i=1}^{N}$, y
output: $x_{\min}, x_{\max}$

Sort $\{x_i\}_{i=1}^{N}$ in the ascending order: $x_1 \leq x_2 \leq ... \leq x_N$;
Re-order $\{a_i\}_{i=1}^{N}$ such that a_i corresponds to x_i;
Set $k_0 = -\sum_{i=1}^{N} a_i$;
for $i = 1, ..., N$ **do**
  $k_i = k_{i-1} + 2a_i$;
end
Calculate $h_1 = \sum_{i=2}^{N} a_i|x_1 - x_i|$;
for $i = 2, ..., N$ **do**
  $h_i = h_{i-1} + k_{i-1}(x_i - x_{i-1})$
end
if $\min_i k_i > y$ **then**
  No solution;
  Exit;
else
  if $y > h_1$ **then**
    $x_{\min} = x_1 + (y - h_1)/k_0$;
  else
    Seek j such that $y \in [h_j, h_{j-1}]$;
    $x_{\min} = x_j + (y - h_j)/k_{j-1}$;
  end
  if $y > h_N$ **then**
    $x_{\max} = x_N + (y - h_N)/k_N$;
  else
    Seek j such that $y \in [h_{j-1}, h_j]$;
    $x_{\max} = x_{j-1} + (y - h_{j-1})/k_{j-1}$;
  end
end

CHAPTER V

Regularized Least-Mean-Square Algorithms

In this chapter we continue to study time-varying regression problems where we consider adaptive system identification with convex constraints and propose a family of regularized Least-Mean-Square (LMS) algorithms. We show that with a properly selected regularization parameter the regularized LMS provably dominates its conventional counterpart in terms of mean square deviations. We establish simple and closed-form expressions for choosing this regularization parameter. For identifying an unknown sparse system we propose sparse and group-sparse LMS algorithms, which are special examples of the regularized LMS family. Simulation results demonstrate the advantages of the proposed filters in both convergence rate and steady-state error under sparsity assumptions on the true coefficient vector.

5.1 Introduction

The Least Mean Square (LMS) algorithm, introduced by Widrow and Hoff [92], is a popular method for adaptive system identification. Its applications include echo cancelation, channel equalization, interference cancelation and so forth. Although there exist algorithms with faster convergence rates such as the Recursive Least Square (RLS) methods, LMS-type methods are popular because of its ease of implementation, low computational costs and robustness.

In many scenarios often prior information about the unknown system is available. One important example is when the impulse response of the unknown system is known to be sparse, containing only a few large coefficients interspersed among many small ones. Exploiting such prior information can improve the filtering performance and has been investigated for several years. Early work includes heuristic online selection of active taps [97, 98, 99] and sequential partial updating [100, 101]; other algorithms assign proportional step sizes of different taps according to their magnitudes, such as the Proportionate Normalized LMS (PNLMS) and its variations [102, 103].

Motivated by LASSO [104] and recent progress in compressive sensing [40, 105], the authors in [106] introduced an ℓ_1-type regularization to the LMS framework resulting in two sparse LMS methods called ZA-LMS and RZA-LMS. This methodology was also applied to other adaptive filtering frameworks such as RLS [107, 108] and projection-based adaptive algorithms [109]. Inheriting the advantages of conventional LMS methods such as robustness and low computational costs, the sparse LMS filters were empirically demonstrated to achieve superior performances in both convergence rate and steady-state behavior, compared to the standard LMS when the system is sparse. However, while the regularization parameter needs to be tuned there is no systematical way to choose the parameter. Furthermore, the analysis of [106] is only based on the ℓ_1 penalty and not applicable to other regularization schemes.

In this work, we extend the methods presented in [106] to a broad family of regularization penalties and consider LMS and Normalized LMS algorithms (NLMS) [92] under general convex constraints. In addition, we allow the convex constraints to

be time-varying. This results in a regularized LMS/NLMS[1] update equation with an additional sub-gradient term. We show that the regularized LMS provably dominates its conventional counterpart if a proper regularization parameter is selected. We also establish a simple and closed-form formula to choose this parameter. For white input signals, the proposed parameter selection guarantees dominance of the regularized LMS over the conventional LMS. Next, we show that the sparse LMS filters in [106], *i.e.,* ZA-LMS and RZA-LMS, can be obtained as special cases of the regularized LMS family introduced here. Furthermore, we consider a group-sparse adaptive FIR filter response that is useful for practical applications [103, 110]. To enforce group sparsity we use $\ell_{1,2}$ type regularization functions [88] in the regularized LMS framework. For sparse and group-sparse LMS methods, we propose alternative closed-form expressions for selecting the regularization parameters. This guarantees provable dominance for both white and correlated input signals. Finally, we demonstrate performance advantages of our proposed sparse and group-sparse LMS filters using numerical simulation. In particular, we show that the regularized LMS method is robust to model mis-specification and outperforms the contemporary projection based methods [109] for equivalent computational cost.

This chapter is organized as follows. Section 5.2 formulates the problem and introduces the regularized LMS algorithm. In Section 5.3 we develop LMS filters for sparse and group-sparse system identification. Section 5.4 provides numerical simulation results and Section 5.5 summarizes our principal conclusion. The proofs of theorems are provided in the Appendix.

Notations: In the following parts of the chapter, matrices and vectors are denoted by boldface upper case letters and boldface lower case letters, respectively; $(\cdot)^T$

[1]We treat NLMS as a special case of the general LMS algorithm and will not distinguish the two unless required for clarity.

denotes the transpose operator, and $\|\cdot\|_1$ and $\|\cdot\|_2$ denote the ℓ_1 and ℓ_2 norm of a vector, respectively.

5.2 Regularized LMS

5.2.1 LMS framework

We begin by briefly reviewing the framework of the LMS filter, which forms the basis of our derivations to follow. Denote the coefficient vector and the input signal vector of the adaptive filter as

$$\hat{\mathbf{w}}_n = [\hat{w}_{n,0}, \hat{w}_{n,1}, \cdots, \hat{w}_{n,N-1}]^T \tag{5.1}$$

and

$$\mathbf{x}_n = [x_n, x_{n-1}, \cdots, x_{n-N+1}]^T, \tag{5.2}$$

respectively, where n is the time index, x_n is the input signal, $\hat{w}_{n,i}$ is the i-th coefficient at time n and N is the length of the filter. The goal of the LMS algorithm is to identify the true system impulse response $\mathbf{w}$ from the input signal x_n and the desired output signal y_n, where

$$y_n = \mathbf{w}^T\mathbf{x}_n + v_n. \tag{5.3}$$

v_n is the observation noise which is assumed to be independent with x_n.

Let e_n denote the instantaneous error between the filter output $\hat{\mathbf{w}}_n^T\mathbf{x}_n$ and the desired output y_n:

$$e_n = y_n - \hat{\mathbf{w}}_n^T\mathbf{x}_n. \tag{5.4}$$

In the standard LMS framework, the cost function L_n is defined as the instantaneous square error

$$L_n(\hat{\mathbf{w}}_n) = \frac{1}{2}e_n^2$$

and the filter coefficient vector is updated in a stochastic gradient descent manner:

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n - \mu_n \nabla L_n(\mathbf{w}_n) = \hat{\mathbf{w}}_n + \mu_n e_n \mathbf{x}_n, \tag{5.5}$$

where μ_n is the step size controlling the convergence and the steady-state behavior of the LMS algorithm. We refer to (5.5) as the conventional LMS algorithm and emphasize that μ_n can be both time-varying and functions of $\mathbf{x}_n$. For example,

$$\mu_n = \frac{\alpha_n}{\|\mathbf{x}_n\|_2^2} \tag{5.6}$$

yields the normalized LMS (NLMS) algorithm with variable step size α_n.

5.2.2 Regularized LMS

Conventional LMS algorithms do not impose any model on the true system response $\mathbf{w}$. However, in practical scenarios often prior knowledge of $\mathbf{w}$ is available. For example, if the system is known to be sparse, the ℓ_1 norm of $\mathbf{w}$ can be upper bounded by some constant [104]. In this work, we study the adaptive system identification problem where the true system is constrained by

$$f_n(\mathbf{w}) \leq \eta_n, \tag{5.7}$$

where $f_n(\cdot)$ is a convex function and η_n is a constant. We note that the subscript n in $f_n(\cdot)$ allows adaptive constraints that can vary in time. Based on (5.7) we propose a regularized instantaneous cost function

$$L_n^{\text{reg}}(\hat{\mathbf{w}}_n) = \frac{1}{2} e_n^2 + \gamma_n f_n(\hat{\mathbf{w}}_n) \tag{5.8}$$

and update the coefficient vector by

$$\begin{aligned} \hat{\mathbf{w}}_{n+1} &= \hat{\mathbf{w}}_n - \mu_n \nabla L_n^{\text{reg}}(\hat{\mathbf{w}}_n) \\ &= \hat{\mathbf{w}}_n + \mu_n e_n \mathbf{x}_n - \rho_n \partial f_n(\hat{\mathbf{w}}_n), \end{aligned} \tag{5.9}$$

where $\partial f_n(\cdot)$ is the sub-gradient of the convex function $f_n(\cdot)$, γ_n is the regularization parameter and $\rho_n = \gamma_n \mu_n$.

Eq. (5.9) is the proposed regularized LMS. Compared to its conventional counterpart, the regularization term, $-\rho_n \partial f_n(\hat{\mathbf{w}}_n)$, always promotes the coefficient vector to satisfy the constraint (5.7). The parameter ρ_n is referred to as the regularization step size. Instead of tuning ρ_n in an *ad hoc* manner, we establish a systematic approach to choosing ρ_n.

Theorem V.1. *Assume both $\{x_n\}$ and $\{v_n\}$ are Gaussian independent and identically distributed (i.i.d.) processes that are mutually independent. For any $n > 1$*

$$E\left\|\hat{\mathbf{w}}_n - \mathbf{w}\right\|_2^2 \leq E\left\|\hat{\mathbf{w}}'_n - \mathbf{w}\right\|_2^2 \tag{5.10}$$

if $\hat{\mathbf{w}}_0 = \hat{\mathbf{w}}'_0$ and $\rho_n \in [0, 2\rho_n^*]$, where $\mathbf{w}$ is the true coefficient vector and $\hat{\mathbf{w}}'_n$ and $\hat{\mathbf{w}}_n$ are filter coefficients updated by (5.5) and (5.9) with the same step size μ_n, respectively. ρ_n^* is calculated by

$$\rho_n^* = \max\left\{(1 - \mu_n \sigma_x^2)\frac{f_n(\hat{\mathbf{w}}_n) - \eta_n}{\|\partial f_n(\hat{\mathbf{w}}_n)\|_2^2}, 0\right\} \tag{5.11}$$

if μ_n are constant values (LMS), or

$$\rho_n^* = \max\left\{(1 - \alpha_n/N)\frac{f_n(\hat{\mathbf{w}}_n) - \eta_n}{\|\partial f_n(\hat{\mathbf{w}}_n)\|_2^2}, 0\right\} \tag{5.12}$$

if μ_n is chosen using (5.6) (NLMS), where N is the filter length, σ_x^2 is the variance of $\{x_n\}$ and η_n is an upper bound of $f_n(\mathbf{w})$ defined in (5.7).

The proof of Theorem V.1 is provided in the Appendix.

Remark 1. Theorem V.1 shows that with the same initial condition and step size μ_n, the regularized LMS algorithm provably dominates conventional LMS when the input signal is white. The parameter ρ_n^* in (5.11) or (5.12) can be used as the

value for ρ_n in (5.9) to guarantee that regularized LMS will have lower MSD than conventional LMS. The value ρ_n^* only requires specification of the noise variance and η_n which upper bounds the true value $f_n(\mathbf{w})$. Simulations in latter sections show that the performance of the regularized LMS is robust to misspecified values of η_n.

Remark 2. Eq. (5.11) and (5.12) indicate that to ensure superiority the regularization is only "triggered" if $f_n(\hat{\mathbf{w}}_n) > \eta_n$. When $f_n(\hat{\mathbf{w}}_n) \leq \eta_n$, $\rho_n^* = 0$ and the regularized LMS reduces to the conventional LMS.

Remark 3. The closed form expression for ρ_n^* is derived based on the white input assumption. Simulation results in latter sections show that the (5.11) and (5.12) are also empirically good choices even for correlated input signals. Indeed, in the next section we will show that provable dominance can be guaranteed for correlated inputs when the regularization function is suitably selected.

5.3 Sparse system identification

A sparse system contains only a few large coefficients interspersed among many negligible ones. Such sparse systems are arise in many applications such as digital TV transmission channels [110] and acoustic echo channels [103]. Sparse systems can be further divided into general sparse systems and group-sparse systems, as shown in Fig. 5.1 (a) and Fig. 5.1 (b), respectively. Here we apply our regularized LMS to both general and group sparse system identification. We show that ZA-LMS and RZA-LMS in [106] are special examples of regularized LMS. We then propose group-sparse LMS algorithms for identifying group-sparse systems.

5.3.1 Sparse LMS

For a general sparse system, the locations of active non-zero coefficients are unknown but one may know an upper bound on their number. Specifically, we will

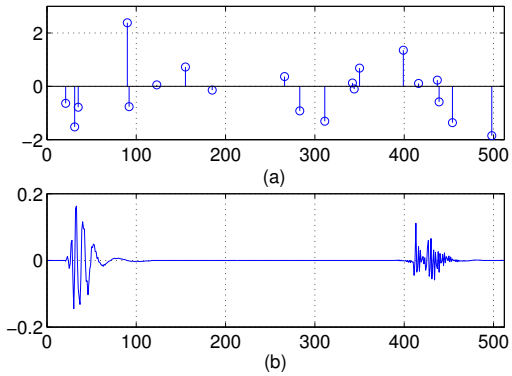


Figure 5.1: Examples of (a) a general sparse system and (b) a group-sparse system.

assume that the impulse response $\mathbf{w}$ satisfies

$$\|\mathbf{w}\|_0 \leq k, \tag{5.13}$$

where $\|\cdot\|_0$ is the ℓ_0 norm denoting the number of non-zero entries of a vector, and k is a known upper bound. As the ℓ_0 norm is non-convex it is not suited to the proposed framework. Following [104] and [40], we instead adopt the ℓ_1 norm as a surrogate approximation to the ℓ_0 norm:

$$\|\mathbf{w}\|_1 = \sum_{i=0}^{N-1} |w_i|. \tag{5.14}$$

Using the regularization penalty $f_n(\mathbf{w}) = \|\mathbf{w}\|_1$ in regularized LMS (5.9), we obtain

$$\hat{\mathbf{w}}_{n+1} = \hat{\mathbf{w}}_n + \mu_n e_n \mathbf{x}_n - \rho_n \operatorname{sgn} \hat{\mathbf{w}}_n, \tag{5.15}$$

where the component-wise $\operatorname{sgn}(\cdot)$ function is defined as

$$\operatorname{sgn}(x) = \begin{cases} x/|x| & x \neq 0 \\ 0 & x = 0 \end{cases}. \tag{5.16}$$

Equation (5.15) yields the ZA-LMS introduced in [106]. The regularization parameter ρ_n can be calculated by (5.11) for LMS and by (5.12) for NLMS, where $f_n(\hat{\mathbf{w}}_n) = \|\hat{\mathbf{w}}_n\|_1$ and η_n is an estimate of the true $\|\mathbf{w}\|_1$.

An alternative approach to approximating the ℓ_0 norm is to consider the following function [111, 106, 109]:

$$\|\mathbf{w}\|_0 \simeq \sum_{i=0}^{N-1} \frac{1}{|w_i| + \delta} \cdot |w_i|, \tag{5.17}$$

where δ is a sufficiently small positive real number. Interpreting (5.17) as a weighted ℓ_1 approximation, we propose the regularization function $f_n(\mathbf{w})$

$$f_n(\mathbf{w}) = \sum_{i=0}^{N-1} \beta_{n,i} \cdot |w_i|, \tag{5.18}$$

and

$$\beta_{n,i} = \frac{1}{|\hat{w}_{n,i}| + \delta}, \tag{5.19}$$

where $\hat{w}_{n,i}$ is the i-th coefficient of $\hat{\mathbf{w}}_n$ defined in (5.1). Using (5.18) in (5.9) yields

$$\hat{w}_{n+1,i} = \hat{w}_{n,i} + \mu_n e_n x_{n-i} - \rho_n \beta_{n,i} \operatorname{sgn} \hat{w}_{n,i}, \tag{5.20}$$

which is a component-wise update of the RZA-LMS proposed in [106]. Again, ρ_n can be computed using (5.11) for LMS or (5.12) for NLMS, where η_n is an estimate of the true $\|\mathbf{w}\|_0$, *i.e.,* the number of the non-zero coefficients.

5.3.2 Group-sparse LMS

In many practical applications, a sparse system often exhibits a grouping structure, *i.e.,* coefficients in the same group are highly correlated and take on the values zero or non-zero as a group, as shown in Fig. 5.1 (b). The motivation for developing group-sparse LMS is to take advantage of such a structure.

We begin by employing the mixed $\ell_{1,2}$ norm for promoting group-sparsity, which was originally proposed in [88] and has been widely adopted for various structured sparse regression problems [112, 113]. The $\ell_{1,2}$ norm of a vector $\mathbf{w}$ is defined as

$$\|\mathbf{w}\|_{1,2} = \sum_{j=1}^{J} \|\mathbf{w}_{I_j}\|_2, \tag{5.21}$$

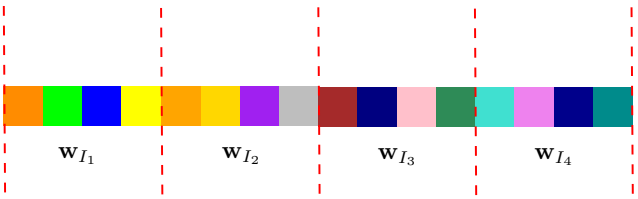


Figure 5.2: A toy example illustrating the $\ell_{1,2}$ norm of a 16×1 coefficient vector $\mathbf{w}$: $\|\mathbf{w}\|_{1,2} = \sum_{j=1}^{4} \|\mathbf{w}_{I_j}\|_2$.

where $\{I_j\}_{j=1}^{J}$ is a group partition of the whole index set $I = \{0, 1, \ldots, N-1\}$:

$$\bigcup_{j=1}^{J} I_j = I, \quad I_j \cap I_{j'} = \phi \text{ when } j \neq j', \tag{5.22}$$

and $\mathbf{w}_{I_j}$ is a sub-vector of $\mathbf{w}$ indexed by I_j. The $\ell_{1,2}$ norm is a mixed norm: it encourages correlation among coefficients inside each group via the ℓ_2 norm and promotes sparsity across those groups using the ℓ_1 norm. $\|\mathbf{w}\|_{1,2}$ is convex in $\mathbf{w}$ and reduces to $\|\mathbf{w}\|_1$ when each group contains only one coefficient, *i.e.,*

$$|I_1| = |I_2| = \cdots = |I_J| = 1, \tag{5.23}$$

where $|\cdot|$ denotes the cardinality of a set. Employing $f_n(\mathbf{w}) = \|\mathbf{w}\|_{1,2}$, the $\ell_{1,2}$ regularized LMS, which we refer to as GZA-LMS, is

$$\hat{\mathbf{w}}_{n+1,I_j} = \hat{\mathbf{w}}_{n,I_j} + \mu_n e_n \mathbf{x}_{I_j} - \rho_n \frac{\hat{\mathbf{w}}_{n,I_j}}{\|\hat{\mathbf{w}}_{n,I_j}\|_2 + \delta}, \; j = 1, ..., J, \tag{5.24}$$

and δ is a sufficiently small number ensuring a non-zero denominator. To the best of our knowledge this is the first time that the $\ell_{1,2}$ norm has been proposed for the LMS adaptive filters.

To further promote group selection we consider the following weighted $\ell_{1,2}$ regularization as a group-wise generalization of (5.18):

$$f_n(\mathbf{w}) = \sum_{j=1}^{J} \beta_{n,j} \|\mathbf{w}_{I_j}\|_2, \tag{5.25}$$

where $\beta_{n,j}$ is a re-weighting parameter defined by

$$\beta_{n,j} = \frac{1}{\|\hat{\mathbf{w}}_{n,I_j}\|_2 + \delta}, \tag{5.26}$$

and the corresponding regularized LMS update is then

$$\hat{\mathbf{w}}_{n+1,I_j} = \hat{\mathbf{w}}_{n,I_j} + \mu_n e_n \mathbf{x}_{I_j} - \rho_n \beta_{n,j} \frac{\hat{\mathbf{w}}_{n,I_j}}{\|\hat{\mathbf{w}}_{n,I_j\|_2} + \delta}, \ j = 1, ..., J, \tag{5.27}$$

which is referred to as GRZA-LMS.

As both the $\ell_{1,2}$ norm and the weighted $\ell_{1,2}$ norm are convex, Theorem V.1 applies under the assumption of white input signals and ρ_n can be calculated by (5.11) or (5.12). The parameter η_n can be chosen as an estimate of the true $\|\mathbf{w}\|_{1,2}$ for GZA-LMS (5.24), or the number of non-zero groups of $\mathbf{w}$ for GRZA-LMS (5.27).

Finally, we note that GZA-LMS and GRZA-LMS reduce to ZA-LMS and RZA-LMS, respectively, if each group contains only one element.

5.3.3 Choosing regularization parameter for correlated input

Theorem V.1 gives a closed form expression for ρ_n and (5.11) or (5.12) is applicable for any convex $f_n(\mathbf{w})$. However, the dominance over conventional LMS is only guaranteed when the input signal is white. Here we develop an alternative formula to determine ρ_n that applies to correlated input signals for sparse and group-sparse LMS, *i.e.,* (5.15), (5.20), (5.24) and (5.27).

We begin by considering the weighted $\ell_{1,2}$ regularization (5.25) and the corresponding GRZA-LMS update (5.27). Indeed, the other three algorithms, *i.e.,* (5.24), (5.20) and (5.15), can be treated as special cases of (5.27). For general wide-sense stationary (WSS) input signals, the regularization parameter ρ_n of (5.27) can be selected according the following theorem.

Theorem V.2. *Assume* $\{x_n\}$ *and* $\{v_n\}$ *are WSS stochastic processes which are mutually independent. Let* $\hat{\mathbf{w}}_n$ *and* $\hat{\mathbf{w}}'_n$ *be filter coefficients updated by (5.27) and (5.5) with the same* μ_n*, respectively. Then,*

$$E\left\|\hat{\mathbf{w}}_{n+1}-\mathbf{w}\right\|_2^2 \leq E\left\|\hat{\mathbf{w}}'_{n+1}-\mathbf{w}\right\|_2^2 \tag{5.28}$$

if $\hat{\mathbf{w}}_n=\hat{\mathbf{w}}'_n$ *and* $\rho_n \in [0, 2\rho_n^*]$*,* $\mathbf{w}$ *is the true coefficient vector and* ρ_n^* *is*

$$\rho_n^* = \max\left\{\frac{f_n(\hat{\mathbf{w}}_n)-\eta_n-\mu_n r_n}{\|\partial f_n(\hat{\mathbf{w}}_n)\|_2^2}, 0\right\}, \tag{5.29}$$

where $f_n(\hat{\mathbf{w}}_n)$ *is determined by (5.25),* η_n *is an upper bound of* $f_n(\mathbf{w})$ *and*

$$r_n = \hat{\mathbf{w}}_n^T\mathbf{x}_n \cdot \mathbf{x}_n^T\partial f_n(\hat{\mathbf{w}}_n) + \eta_n \cdot \max_j\left\{\frac{\|\mathbf{x}_{I_j}\|_2}{\beta_{n,j}}\right\} \cdot |\mathbf{x}_n^T\partial f_n(\hat{\mathbf{w}}_n)|. \tag{5.30}$$

The proof of Theorem V.2 can be found in the Appendix. We make the following remarks.

Remark 4. Theorem V.2 is derived from the general form (5.27) and can be directly specialized to (5.24), (5.20) and (5.15). Specifically,

- GZA-LMS (5.24) can be obtained by assigning $\beta_{n,j}=1$;
- RZA-LMS (5.20) can be obtained when $|I_j|=1, j=1,...,J$;
- ZA-LMS (5.15) can be obtained when both $|I_j|=1, j=1,...,J$ and $\beta_{n,j}=1$.

Remark 5. Theorem V.2 is valid for any WSS input signals. However, the dominance result in (5.28) is weaker than that in Theorem V.1, as it requires $\hat{\mathbf{w}}_n=\hat{\mathbf{w}}'_n$ at each iteration.

Remark 6. Eq. (5.29) can be applied to both LMS and NLMS, depending on if μ_n are deterministic functions of $\mathbf{x}_n$ as specified in (5.6). This is different from Theorem V.1 where we have separate expressions for LMS and NLMS.

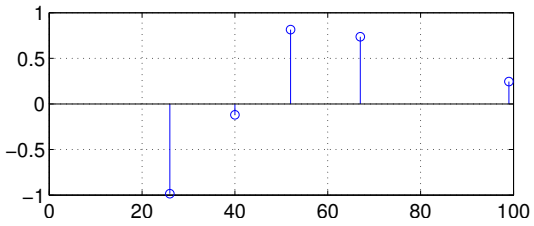


Figure 5.3: The general sparse system used for simulations.

Remark 7. ρ_n^* in (5.29) is non-zero only if $f_n(\hat{\mathbf{w}}_n)$ is greater than $\eta_n + \mu_n r_n$ (rather than η_n as presented in Theorem V.1). This may yield a more conservative performance.

5.4 Numerical simulations

In this section we demonstrate our proposed sparse LMS algorithms by numerical simulations. Multiple experiments are designed to evaluate their performances over a wide range of conditions.

5.4.1 Identifying a general sparse system

Here we perform evaluation of the proposed filters for general sparse system identification, as illustrated in Fig. 5.1 (a). There are 100 coefficients in the time varying system and only five of them are non-zero. The five non-zero coefficients are assigned to random locations and their values are also randomly drawn from a standard Gaussian distribution. The resultant true coefficient vector is plotted in Fig. 5.3.

White input signals

Initially we simulate white Gaussian input signal $\{x_n\}$ with zero mean and unit variance. The measurement noise $\{v_n\}$ is an independent Gaussian random process of zero mean and variance $\sigma_v^2 = 0.1$. For ease of parameter selection, we implement NLMS-type filters in our simulation. Three filters (NLMS, ZA-NLMS and RZA-

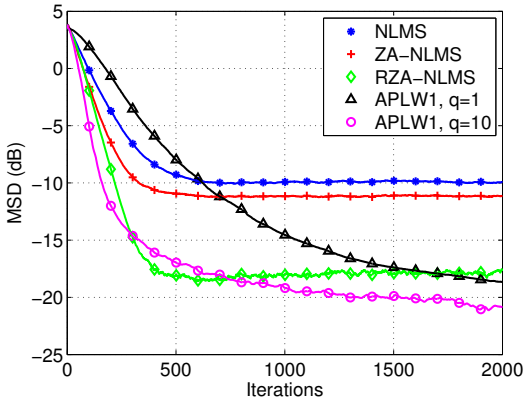


Figure 5.4: White input signals: performance comparison for different filters.

NLMS) are implemented and their common step-size μ_n is set via (5.6) with $\alpha_n = 1$. The regularization parameter ρ_n is computed using (5.12), where η_n is set to $\eta_n = \|\mathbf{w}\|_1$ (*i.e.,* the true value) for ZA-NLMS and $\eta_n = 5$ for RZA-NLMS. For comparison we also implement a recently proposed sparse adaptive filter, referred to as APWL1 [109], which sequentially projects the coefficient vector onto weighted ℓ_1 balls. We note that our simulation setting is identical to that used in [109] and thus we adopt the same tuning parameters for APWL1. In addition, the weights $\beta_{n,i}$ for RZA-NLMS is scheduled in the same manner as that in [109] for a fair comparison. The simulations are run 100 times and the average estimates of mean square deviation (MSD) are shown in Fig. 5.4.

It can be observed that ZA-NLMS improves upon NLMS in both convergence rate and steady-state behavior and RZA-NLMS does even better. The parameter q of APLW1 is the number of samples used in each iteration. One can see that RZA-NLMS outperforms APLW1 when $q = 1$, *i.e.,* the case that APLW1 operates with the same memory storage as RZA-NLMS. With larger p APLW1 begins to perform better and exceeds RZA-NLMS when $q \geq 10$. However, there is a trade-off between the system complexity and filtering performance, as APWL1 requires $\mathcal{O}(qN)$ for

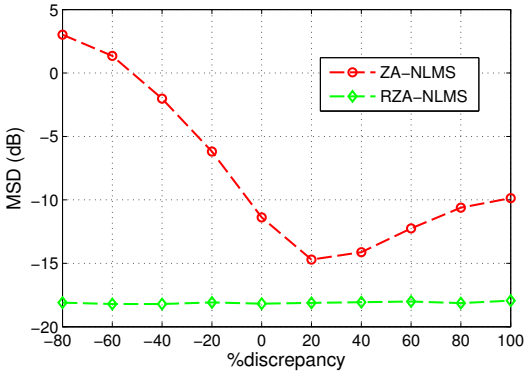


Figure 5.5: Sensitivity of ZA-NLMS and RZA-NLMS to η_n: MSD for ZA-NLMS and RZA-NLMS at the 750th iteration for white input signals.

memory storage and $\mathcal{O}(N \log_2 N + qN)$ for computation, in contrast to LMS-type methods which require only $\mathcal{O}(N)$ for both memory and computation.

Next, we investigate the sensitivity to η_n for ZA-NLMS and RZA-NLMS. The result shown in Fig. 5.5 indicates that ZA-NLMS is more sensitive to η_n than RZA-NLMS, which is highly robust to misspecified η_n.

Further analysis reveals that the projection based methods such APWL1 may exhibit unstable converging behaviors. Fig. 5.6 shows two independent trials of the simulation implemented in Fig. 5.4. It can be seen that there exist several local minima in APWL1. For example, Fig. 5.6 (b) seems to indicate that APWL1 ($q = 10$) converges at the 400th iteration with MSD $\simeq -12$ dB, yet its MSD actually reaches values as low as -25 dB at the 900th iteration. This slow convergence phenomenon is due to the fact that the weighted ℓ_1 ball is determined in an online fashion and the projection operator is sensitive to mis-specifications of the convex set. In the contrast, our regularized LMS uses sub-gradient rather than projection to pursue sparsity, translating into improved convergence.

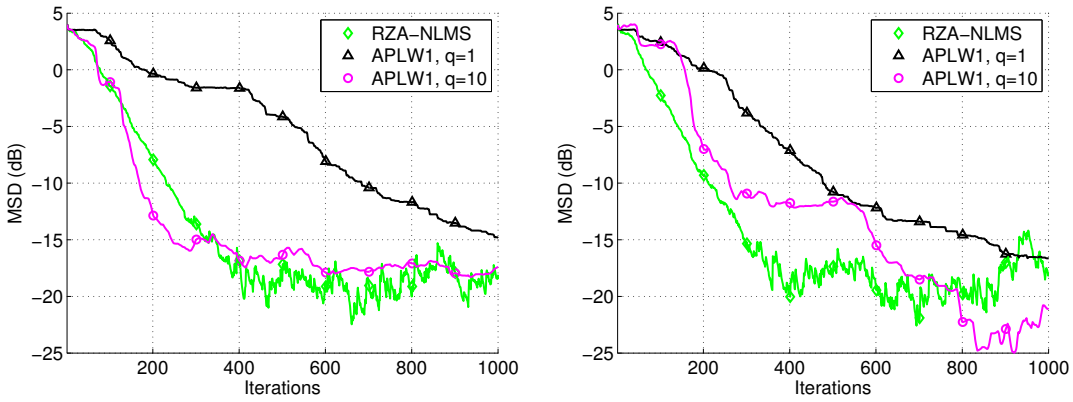


Figure 5.6: Two different trials of RZA-NLMS and APWL1 for white input signals. APWL1 exhibits unstable convergence.

Correlated input signals

Next, we evaluate the filtering performance using correlated input signals. We generate the sequence $\{x_n\}$ as an AR(1) process

$$x_n = 0.8x_{n-1} + u_n, \tag{5.31}$$

which is then normalized to unit variance, where $\{u_n\}$ is a Gaussian i.i.d. process. The measurement system is the same as before and the variance of the noise is also $\sigma_v^2 = 0.1$.

We compare our RZA-NLMS with APWL1 ($q = 10$) and standard NLMS is also included as a benchmark. All the filter parameters are set to the same values as that in the previous simulation, except we employ both (5.12) and (5.29) to calculate ρ_n in RZA-NLMS. The simulations are run 100 times and the average MSD curves are plotted in Fig. 5.7. While Theorem V.1 is derived based on white input assumptions, using (5.12) to determine ρ_n achieves an empirically better performance compared to using (5.29) – whose use guarantees dominance but yields a conservative result. This confirms our conjecture in Remark 7. We also observe a severe performance degradation of APWL1 for correlated input signals. Fig. 5.8 draws two independent

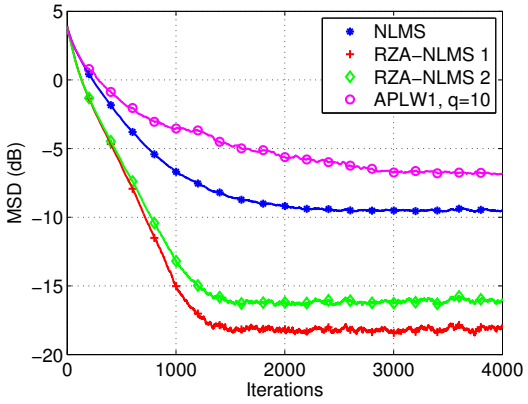


Figure 5.7: Correlated input signals: performance comparison for different filters, where RZA-NLMS 1 and RZA-NLMS 2 use (5.12) and (5.29) to determine ρ_n, respectively.

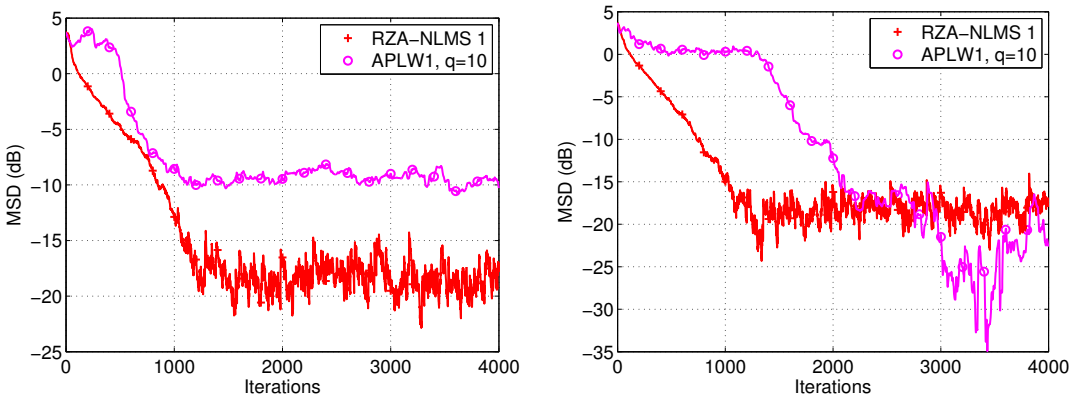


Figure 5.8: Two different trials of RZA-NLMS and APWL1 for correlated input signals.

trials in this simulation. The phenomenon described in Fig. 5.6 becomes more frequent when the input signal is correlated, which drags down the average performance of APWL1 significantly. Finally, we note that the filtering performance of a group sparse system (*e.g.*, Fig. 5.1 (b)) may be very different from that of a general sparse system. This will investigated in Section 5.4.2.

Tracking performance

Finally, we study the tracking performance of the proposed filters. The time-varying system is initialized using the same parameters as used to generate Fig.

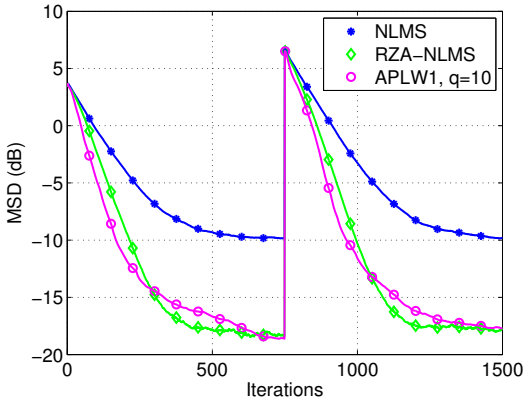


Figure 5.9: Comparison of tracking performances when the input signal is white.

5.3. At the 750th iteration the system encounters a sudden change, where all the active coefficients are left-shifted for 10 taps. We use white input signals to excite the unknown system and all the filter parameters are set in an identical manner to Section 5.4.1. The simulation is repeated 100 times and the averaged result is shown in Fig. 5.9. It can be observed that both RZA-NLMS and APWL1 ($q = 10$) achieve better tracking performance than the conventional NLMS.

5.4.2 Identifying a group-sparse system

Here we test performance of the group-sparse LMS filters developed in Section 5.3.2. The unknown system contains 200 coefficients that are distributed into two groups. The locations of the two groups are randomly selected, which start from the 36th tap and the 107th tap, respectively. Both of the two groups contain 15 coefficients and their values are randomly drawn from a standard Gaussian distribution. Fig. 5.10 shows the response of the true system.

The input signal $\{x_n\}$ is initially set to an i.i.d. Gaussian process and the variance of observation noise is $\sigma_v^2 = 0.1$. Three filters, GRZA-NLMS, RZA-NLMS and NLMS, are implemented, where the performance of NLMS is treated as a benchmark. In GRZA-NLMS, we divide the 200 coefficients equally into 20 groups, where each

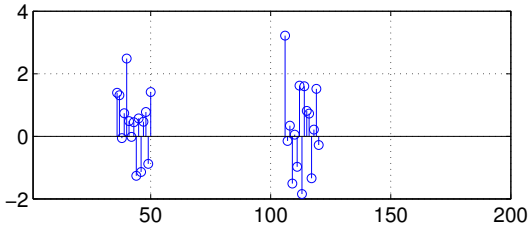


Figure 5.10: The group-sparse system used for simulations. There are two active blocks; each of them contains 15 non-zero coefficients.

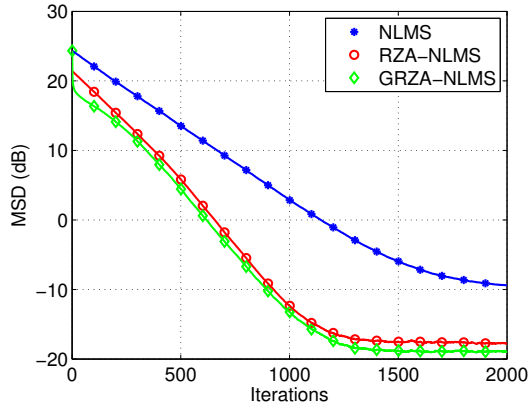


Figure 5.11: MSD comparison for the group-sparse system for white input signals.

of them contains 10 coefficients. The step size μ_n of the three filters are all set according to (5.6) with $\alpha_n = 1$. We use (5.12) to calculate ρ_n, where η_n is set to 30 (the number of non-zero coefficients) for RZA-NLMS and 2 (the number of non-zero blocks) for GRZA-NLMS, respectively. We repeat the simulation 200 times and the averaged MSD is shown in Fig. 5.11. It can be seen that GRZA-NLMS and RZA-NLMS outperform the standard NLMS for 10 dB in the steady-state MSD, while GRZA-NLMS only improves upon RZA-NLMS, but only marginally. This is partially due to the fact that in the white input scenario each coefficient is updated in an independent manner.

We next consider the case of correlated input signals, where $\{x_n\}$ is generated by (5.31) and then normalized to have unit variance. The parameters for all the filters are set to the same values as in the white input example and the averaged MSD

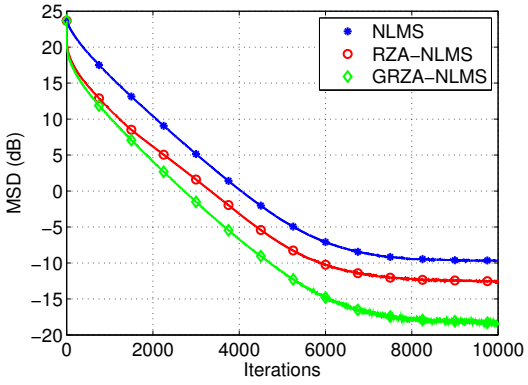


Figure 5.12: MSD comparison for the group-sparse system for correlated input signals.

curves are plotted in Fig. 5.12. In the contrast to the white input example, here RZA-NLMS slightly outperforms NLMS but there is a significant improvement of GRZA-NLMS over RZA-NLMS. This demonstrates the power of promoting group-sparsity especially when the input signal is correlated.

Finally, we evaluate the tracking performance of the adaptive filters. We use white signals as the system input and initialize the time-varying system using that in Fig. 5.10. At the 2000th iteration, the system response is right-shifted for 50 taps, while the values of coefficients inside each block are unaltered. We then keep the block locations and reset the values of non-zero coefficients randomly at the 4000th iteration. From Fig. 5.13 we observe that the tracking rate of RZA-NLMS and GRZA-NLSM are comparable to each other when the system changes across blocks, and GRZA-NLMS shows a better tracking performance than RZA-NLMS when the system response changes only inside its active groups.

5.5 Conclusion

In this work we proposed a general class of LMS-type filters regularized by convex sparsifying penalties. We derived closed-form expressions for choosing the regularization parameter that guarantees provable dominance over conventional LMS filters.

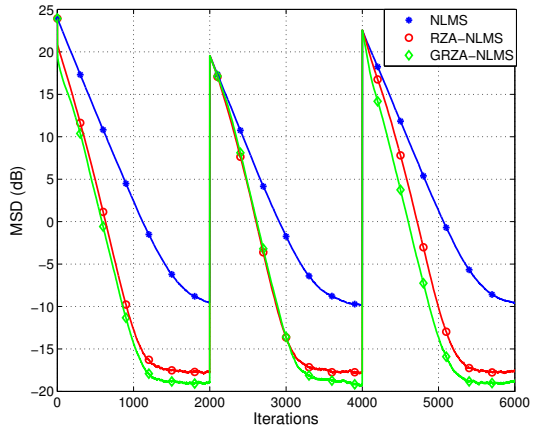


Figure 5.13: Tracking performance comparison for the group-sparse system for white input signals.

We applied the proposed regularized LMS filters to sparse and group-sparse system identification and demonstrated their performances using numerical simulations.

Our regularized LMS filter is derived from the LMS framework and inherits its simplicity, low computational cost and low memory requirements, and robustness to parameter mismatch. It is likely that the convergence rate and steady-state performance can be improved by extension to second-order methods, such as RLS and Kalman filters. Efficient extensions of our results for sparse/group-sparse RLS filters are a worthy topic of future study.

5.6 Appendix

5.6.1 Proof of Theorem V.1

We prove Theorem V.1 for LMS, *i.e.,* the case that μ_n are constants. NLMS, where μ_n is determined by (5.6), can be derived in a similar manner.

According to (5.9),

$$\begin{aligned} &\hat{\mathbf{w}}_{n+1} - \mathbf{w} \\ &= (\mathbf{I} - \mu_n \mathbf{x}_n \mathbf{x}_n^T)(\hat{\mathbf{w}}_n - \mathbf{w}) - \rho_n \partial f_n(\hat{\mathbf{w}}_n) + \mu_n v_n \mathbf{x}_n. \end{aligned} \tag{5.32}$$

Noting that $\hat{\mathbf{w}}_n$, $\mathbf{x}_n$ and v_n are mutually independent, we have

$$\begin{aligned} &E\left\{\|\hat{\mathbf{w}}_{n+1}-\mathbf{w}\|^2|\hat{\mathbf{w}}_n\right\}= \\ &(\hat{\mathbf{w}}_n-\mathbf{w})^T E\left\{\left(\mathbf{I}-\mu_n\mathbf{x}_n\mathbf{x}_n^T\right)^2\right\}(\hat{\mathbf{w}}_n-\mathbf{w})+\mu_n^2\sigma_v^2 E\left\{\|\mathbf{x}_n\|^2\right\} \\ &+2\rho_n(\mathbf{w}-\hat{\mathbf{w}}_n)^T E\left\{\mathbf{I}-\mu_n\mathbf{x}_n\mathbf{x}_n^T\right\}\partial f_n(\hat{\mathbf{w}}_n)+\rho_n^2\|\partial f_n(\hat{\mathbf{w}}_n)\|^2. \end{aligned} \tag{5.33}$$

As $\{x_n\}$ is a Gaussian i.i.d. process, $\mathbf{x}_n$ is a Gaussian random vector with mean zero and covariance $\sigma_x^2\mathbf{I}$. Thus,

$$E\left\{\left(\mathbf{I}-\mu_n\mathbf{x}_n\mathbf{x}_n^T\right)^2\right\}=(1-2\sigma_x^2\mu_n+N\sigma_x^4\mu_n^2)\mathbf{I}, \tag{5.34}$$

$$E\left\{\mathbf{I}-\mu_n\mathbf{x}_n\mathbf{x}_n^T\right\}=(1-\sigma_x^2\mu_n)\mathbf{I}, \tag{5.35}$$

and

$$E\left\{\|\mathbf{x}_n\|^2\right\}=N\sigma_x^2. \tag{5.36}$$

Substituting (5.34), (5.35) and (5.36) into (5.33), we have

$$\begin{aligned} &E\left\{\|\hat{\mathbf{w}}_{n+1}-\mathbf{w}\|^2|\hat{\mathbf{w}}_n\right\}= \\ &(1-2\sigma_x^2\mu_n+N\sigma_x^4\mu_n^2)\ \|\hat{\mathbf{w}}_n-\mathbf{w}\|^2+N\mu_n^2\sigma_x^2\sigma_v^2 \\ &+2\rho_n(1-\sigma_x^2\mu_n)(\mathbf{w}-\hat{\mathbf{w}}_n)^T\partial f_n(\hat{\mathbf{w}}_n)+\rho_n^2\|\partial f_n(\hat{\mathbf{w}}_n)\|^2. \end{aligned} \tag{5.37}$$

As $f_n(\cdot)$ is a convex function, by the definition of sub-gradient, we have

$$(\mathbf{w}-\hat{\mathbf{w}}_n)^T\partial f_n(\hat{\mathbf{w}}_n)\le f_n(\mathbf{w})-f_n(\hat{\mathbf{w}}_n)\le\eta_n-f_n(\hat{\mathbf{w}}_n). \tag{5.38}$$

Therefore,

$$\begin{aligned} &E\left\{\|\hat{\mathbf{w}}_{n+1}-\mathbf{w}\|^2|\hat{\mathbf{w}}_n\right\}\le \\ &(1-2\sigma_x^2\mu_n+N\sigma_x^4\mu_n^2)\ \|\hat{\mathbf{w}}_n-\mathbf{w}\|^2+N\mu_n^2\sigma_x^2\sigma_v^2 \\ &-2\rho_n(1-\sigma_x^2\mu_n)(f_n(\hat{\mathbf{w}}_n)-\eta_n)+\rho_n^2\|\partial f_n(\hat{\mathbf{w}}_n)\|^2. \end{aligned} \tag{5.39}$$

Define

$$C(\rho_n) = -2\rho_n(1-\sigma_x^2\mu_n)(f_n(\hat{\mathbf{w}}_n)-\eta_n) + \rho_n^2\|\partial f_n(\hat{\mathbf{w}}_n)\|^2, \tag{5.40}$$

and take expectation on both sides of (5.39) with respect to $\hat{\mathbf{w}}_n$ to obtain

$$\begin{aligned} &E\left\{\|\hat{\mathbf{w}}_{n+1}-\mathbf{w}\|^2\right\} \\ &\le (1-2\sigma_x^2\mu_n+N\sigma_x^4\mu_n^2)E\left\{\|\hat{\mathbf{w}}_n-\mathbf{w}\|^2\right\} + N\mu_n^2\sigma_x^2\sigma_v^2 \\ &\quad + E\left\{C(\rho_n)\right\}. \end{aligned} \tag{5.41}$$

It is easy to check that $C(\rho_n) \le 0$ if $\rho_n \in [0, 2\rho_n^*]$, where ρ_n^* is defined in (5.11). Therefore,

$$\begin{aligned} &E\left\{\|\hat{\mathbf{w}}_{n+1}-\mathbf{w}\|^2\right\} \\ &\le (1-2\sigma_x^2\mu_n+N\sigma_x^4\mu_n^2)E\left\{\|\hat{\mathbf{w}}_n-\mathbf{w}\|^2\right\} + N\mu_n^2\sigma_x^2\sigma_v^2 \end{aligned} \tag{5.42}$$

if $\rho_n \in [0, 2\rho_n^*]$. For the standard LMS, there is

$$\begin{aligned} &E\left\{\|\hat{\mathbf{w}}'_{n+1}-\mathbf{w}\|^2\right\} \\ &= (1-2\sigma_x^2\mu_n+N\sigma_x^4\mu_n^2)E\left\{\|\hat{\mathbf{w}}'_n-\mathbf{w}\|^2\right\} + N\mu_n^2\sigma_x^2\sigma_v^2. \end{aligned} \tag{5.43}$$

Therefore, under the condition that $E\{\|\hat{\mathbf{w}}_0-\mathbf{w}\|^2\} = E\{\|\hat{\mathbf{w}}'_0-\mathbf{w}\|^2\}$, (5.10) can be obtained from (5.42) and (5.43) using a simple induction argument.

5.6.2 Proof of Theorem V.2

We start our proof from (5.32) and calculate the following conditional MSD:

$$\begin{aligned} &E\left\{\|\hat{\mathbf{w}}_{n+1}-\mathbf{w}\|^2|\hat{\mathbf{w}}_n, \mathbf{x}_n\right\} = \\ &(\hat{\mathbf{w}}_n-\mathbf{w})^T(\mathbf{I}-\mu_n\mathbf{x}_n\mathbf{x}_n^T)^2(\hat{\mathbf{w}}_n-\mathbf{w}) + \mu_n^2\sigma_v^2\|\mathbf{x}_n\|^2 + D(\rho_n), \end{aligned} \tag{5.44}$$

where

$$D(\rho_n) = 2\rho_n(\mathbf{w}-\hat{\mathbf{w}}_n)^T(\mathbf{I}-\mu_n\mathbf{x}_n\mathbf{x}_n^T)\partial f_n(\hat{\mathbf{w}}_n) + \rho_n^2\|\partial f_n(\hat{\mathbf{w}}_n)\|^2. \tag{5.45}$$

For the cross term $2\rho_n(\mathbf{w}-\hat{\mathbf{w}}_n)^T(\mathbf{I}-\mu_n\mathbf{x}_n\mathbf{x}_n^T)\partial f_n(\hat{\mathbf{w}}_n)$ we have

$$
\begin{aligned}
&2\rho_n(\mathbf{w}-\hat{\mathbf{w}}_n)^T(\mathbf{I}-\mu_n\mathbf{x}_n\mathbf{x}_n^T)\partial f_n(\hat{\mathbf{w}}_n)\\
&=2\rho_n(\mathbf{w}-\hat{\mathbf{w}}_n)^T\partial f_n(\hat{\mathbf{w}}_n)+2\rho_n\mu_n\hat{\mathbf{w}}_n^T\mathbf{x}_n\cdot\mathbf{x}_n^T\partial f_n(\hat{\mathbf{w}}_n)\\
&\quad-2\rho_n\mu_n\mathbf{w}^T\mathbf{x}_n\cdot\mathbf{x}_n^T\partial f_n(\hat{\mathbf{w}}_n) \qquad (5.46)\\
&\le 2\rho_n(\eta_n-f_n(\hat{\mathbf{w}}_n))+2\rho_n\mu_n\hat{\mathbf{w}}_n^T\mathbf{x}_n\cdot\mathbf{x}_n^T\partial f_n(\hat{\mathbf{w}}_n)\\
&\quad+2\rho_n\mu_n\left|\mathbf{w}^T\mathbf{x}_n\right|\cdot\left|\mathbf{x}_n^T\partial f_n(\hat{\mathbf{w}}_n)\right|.
\end{aligned}
$$

We now establish upper-bounds for $|\mathbf{w}^T\mathbf{x}_n|$. Indeed,

$$
\begin{aligned}
\left|\mathbf{w}^T\mathbf{x}_n\right| &= \left|\sum_{j=1}^{J}\mathbf{w}_{I_j}^T\mathbf{x}_{n,I_j}\right|\\
&\le \sum_{j=1}^{J}\left|\beta_{n,j}\mathbf{w}_{I_j}^T\frac{1}{\beta_{n,j}}\mathbf{x}_{n,I_j}\right|\\
&\le \sum_{j=1}^{J}\beta_{n,j}\|\mathbf{w}_{I_j}\|_2\frac{\|\mathbf{x}_{n,I_j}\|_2}{\beta_{n,j}} \qquad (5.47)\\
&\le \left\{\sum_{j=1}^{J}\beta_{n,j}\|\mathbf{w}_{I_j}\|_2\right\}\max_j\frac{\|\mathbf{x}_{n,I_j}\|_2}{\beta_{n,j}}\\
&= f_n(\mathbf{w}_n)\max_j\frac{\|\mathbf{x}_{n,I_j}\|_2}{\beta_{n,j}}\le\eta_n\max_j\frac{\|\mathbf{x}_{n,I_j}\|_2}{\beta_{n,j}}.
\end{aligned}
$$

Substituting (5.46) and (5.47) into (5.45) we obtain that

$$
D(\rho_n)\le -2\rho_n(f_n(\hat{\mathbf{w}}_n)-\eta_n-\mu_n r_n)+\rho_n^2\|\partial f_n(\hat{\mathbf{w}}_n)\|_2^2, \qquad (5.48)
$$

where r_n is defined in (5.30). Note that $D(\rho_n)\le 0$ if $\rho_n\in[0,2\rho_n^*]$ (ρ_n^* is defined in (5.29)). There is

$$
\begin{aligned}
&E\left\{\|\hat{\mathbf{w}}_{n+1}-\mathbf{w}\|^2|\hat{\mathbf{w}}_n,\mathbf{x}_n\right\}\\
&\le(\hat{\mathbf{w}}_n-\mathbf{w})^T(\mathbf{I}-\mu_n\mathbf{x}_n\mathbf{x}_n^T)^2(\hat{\mathbf{w}}_n-\mathbf{w})+\mu_n^2\sigma_v^2\|\mathbf{x}_n\|^2, \qquad (5.49)
\end{aligned}
$$

if $\rho_n \in [0, 2\rho_n^*]$. Therefore,

$$
\begin{aligned}
& E\left\{\|\hat{\mathbf{w}}_{n+1} - \mathbf{w}\|^2|\right\} \\
& \leq E\left\{(\hat{\mathbf{w}}_n - \mathbf{w})^T(\mathbf{I} - \mu_n \mathbf{x}_n \mathbf{x}_n^T)^2(\hat{\mathbf{w}}_n - \mathbf{w})\right\} \\
& \quad + \mu_n^2 \sigma_v^2 E\left\{\|\mathbf{x}_n\|^2\right\} \\
& = E\left\{(\hat{\mathbf{w}}'_n - \mathbf{w})^T(\mathbf{I} - \mu_n \mathbf{x}_n \mathbf{x}_n^T)^2(\hat{\mathbf{w}}'_n - \mathbf{w})\right\} \\
& \quad + \mu_n^2 \sigma_v^2 E\left\{\|\mathbf{x}_n\|^2\right\} \\
& = E\left\{\|\hat{\mathbf{w}}'_{n+1} - \mathbf{w}\|^2|\right\},
\end{aligned}
\tag{5.50}
$$

which proves Theorem V.2.

CHAPTER VI

Sampling Ultra-fast Sparse Spectrum Signals: Calibration and Signal Recovery

In this chapter we introduce a sub-Nyquist sampling system based on inferring high-dimensional covariance structure from its low-dimensional random projections. The state-of-the-art system, referred to as the modulated wideband converter (MWC), origins from recent progresses of compressive sensing theories. In [2], a hardware prototype was developed to implement the MWC system in the board level. In this work, we focuses on the calibration and signal reconstruction problem of the MWC prototype, which is one of the main challenges in successful transfering the conceptual design to the real world. Our work witnessed the birth, to the best of our knowledge, the first reported sub-Nyquist sampling hardware that is able to compete with cutting-edge commercial analog-to-digital converters (ADCs).

6.1 Introduction

Efficient sampling of wideband analog signals is a challenging problem, since their Nyquist rates may exceed the specifications of the best analog-to-digital converters (ADCs) nowadays by orders of magnitude. The modulated wideband converter (MWC) [2, 114] is a recent sub-Nyquist system for sampling multiband signals of wide spectral ranges. The MWC, depicted in Fig. 6.1 and further described in

Section 6.2, consists of simple mixers and lowpass filters. By exploiting frequency sparsity of multiband signals, the MWC is able to significantly reduce the conversion rate.

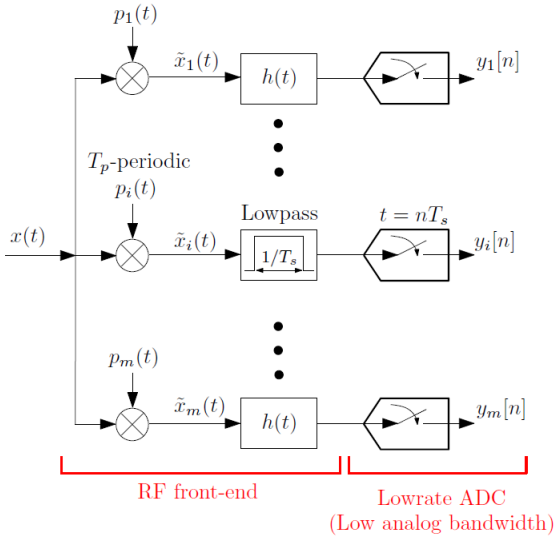


Figure 6.1: A block diagram of the modulated wideband converter. The MWC consists of m parallel channels, which mix the input against m periodic waveforms. The mixed signal is then lowpass filtered and sampled at a low rate.

The MWC system is designed for sampling sparse multiband signals which consist of a relatively small number of narrowband transmissions spread across a wide spectrum range. The underlying principle is based on recently developed theories of compressive sensing [40, 41]. The system exploits spread-spectrum techniques from communication theory [115]. An analog mixing radio-frequency front-end aliases the spectrum, such that a spectrum portion from each band appears in baseband. The systems consist of multiple channels where each of them delivers low-rate digital samples. It is demonstrated that the covariance matrix of the multi-channel low-rate digital samples is a low-dimensional random projection from a high-dimensional but sparse in both columns and rows covariance matrix. In principle, a sufficiently large number of channels allows to identity the high-dimensional covariance structure and

to recover the input sparse multi-band signal. In [2], a board-level prototype was developed to demonstrate MWC in real scenarios.

While theories of analog compressive sensing [41, 116] appear mature and a board-level prototype has been developed, there is still a long way to go for the final success of the practical system. The original theorems and algorithms of MWC assume the system works in the ideal setting. On the other hand, imperfections exist everywhere in practical circuits and the unconventional architecture design makes system calibration and signal reconstruction, especially in the digital domain, a challenging problem. This motivates the work presented in this chapter. Specifically we focus on the following major issues which manifested themselves as tread stones of bridging the gap between theories and practical systems.

- *Sensing matrix estimation.* Accurate estimation of the sensing matrix is crucial to the success of compressive sensing applications. Ideally, the sensing matrix of MWC is comprised of Fourier coefficients of periodic random mixing waveforms and can be obtained in the design stage. However, in practice we observe significant distortions between the measured mixing waveforms and the designed ones. Furthermore, the non-linearity of the mixtures and the existence of equivalent filters of board-level circuits also cause deviations of the sensing matrix from the Fourier coefficients of the mixing waveforms. To obtain accurate estimation we propose an end-to-end calibration scheme via multiple experiments. The resultant estimate of the sensing matrix guarantees performances of support recovery and accurate signal reconstruction using the MWC prototype.

- *Filter compensation.* The original MWC requires ideal analog lowpass filters to accomplish the reconstruction process. In practice, implementing ideal filters is generally difficult and the usual option is to employ high order Butterworth

or Chebyshev filters. Direct use of such off-the-shelf filters does not guarantee perfect reconstruction (PR) in the recovered signal. Indeed, this problem is encountered in the practical implementation of the MWC. Therefore, it is important to develop methods to compensate for imperfect lowpass filters. In this work we extend the MWC to enable the use of practical filters. Under the assumption of near perfect stopband response, we show that, with only a moderate amount of oversampling, the imperfections caused by non-ideal filters can be effectively corrected in the digital domain. We derive a perfect reconstruction condition that must be satisfied by lowpass filters in the MWC and propose a compensation method operating in the digital domain that uses a simple bank of FIR filters. Numerical simulations and real measured data demonstrate that the proposed compensation method can significantly reduce the reconstruction error using low-order FIR filters.

- *Synchronization and signal reconstruction.* The MWC is a linear but time-variant system. When a signal enters the RF front-end, there is an unknown time delay between the input signal and mixing waveforms. Such a delay will cause unexpected phase distortions in the system transfer functions and collapse the support estimation and signal recovery. In this work we analyze the phase-distorted system in closed-form and propose a jointly estimation method of time delay and active supports. The effectiveness of the proposed method is validated using real experimental data.

The rest of the chapter is organized as follows. Section 6.2 provides a brief introduction of the theoretical system as well as the hardware prototype. In Section 6.3, we study the aforementioned calibration problems for the practical MWC system. Section 6.4 summarizes our principal conclusions in this chapter.

6.2 Modulated Wideband Converter

The MWC is a sub-Nyquist sampling system for sampling sparse wideband analog signals. It consists of two stages: sampling and reconstruction. In this section, we briefly introduce the principle and system of the MWC. A complete description can be found in [114, 117].

6.2.1 Sampling

In the sampling stage, the signal $x(t)$ enters m channels simultaneously. In the ith channel, $x(t)$ is multiplied by a T_p- periodic mixing function $p_i(t)$. After mixing, the output is lowpass filtered with cutoff frequency $1/(2T_s)$ and then uniformly sampled at rate $1/T_s$. The overall sampling frequency of the MWC is then m/T_s.

The input $x(t)$ is assumed to be a sparse wideband analog signal bandlimited to $[-f_{\text{NYQ}}/2, f_{\text{NYQ}}/2]$, where f_{NYQ} can be very large, much larger than the sampling frequency m/T_s. The support of $x(t)$ resides within N frequency intervals, or bands, such that the width of each band does not exceed B Hz. The band positions are arbitrary and in particular unknown in advance. For example, in communications N represents the number of concurrent transmissions and B is specified by the specific modulation techniques in use.

The sub-Nyquist sampling of the MWC relies on the following key observation. The mixing operation scrambles the spectrum of $x(t)$ such that the baseband frequencies that reside below the filter cutoff contain a mixture of the spectral contents from the entire Nyquist range.

To further illustrate this point, let us consider a single channel, and let $P_i(f)$ be the spectrum of the mixing function $p_i(t)$. Since $p_i(t)$ is T_p- periodic, $P_i(f)$ can be

expressed as

$$P_i(f) = \sum_{l=-\infty}^{+\infty} c_{i,l}\delta(f - lf_p), \tag{6.1}$$

where $f_p = 1/T_p$, $c_{i,l}$ are arbitrary coefficients and $\delta(\cdot)$ is the Dirac delta function. The spectrum of the mixed signal $\tilde{x}_i(t) = x(t)p_i(t)$ is then

$$\tilde{X}_i(f) = P_i(f) * X(f) = \sum_{l=-\infty}^{+\infty} c_{i,l}X(f - lf_p), \tag{6.2}$$

where $X(f)$ is the spectrum of $x(t)$. Lowpass filtering with a filter transfer function $H(f)$ results in the following relation:

$$Y_i\left(e^{j2\pi fT_s}\right) = \sum_{l=-\infty}^{+\infty} c_{i,l}X(f - lf_p)H(f), \tag{6.3}$$

where $Y_i\left(e^{j2\pi fT_s}\right)$ is the DTFT transform of $y_i[n]$. We note that (6.5) is based on the assumption that no aliasing occurs at the sampling rate f_s, *i.e.,* f_s exceeds the stopband width of $H(f)$. Denote $y_i(f) = Y_i\left(e^{j2\pi fT_s}\right)$ and suppose $H(f)$ is an ideal lowpass filter which is a unit rectangular function over $\mathcal{F}_s = [-f_s/2, f_s/2]$. There is

$$y_i(f) = \sum_{l=-L_0}^{L_0} c_{i,l}X(f - lf_p), f \in \mathcal{F}_s, i = 1, ..., m, \tag{6.4}$$

where L_0 is the smallest integer satisfying $2L_0 + 1 > f_{\text{NYQ}}/f_p$. In the basic configuration we choose $f_s = f_p$. This gives

$$\mathbf{y}(f) = \mathbf{A}\mathbf{z}(f), f \in \mathcal{F}_p, \tag{6.5}$$

where $\mathbf{A}$ is a matrix comprised of $c_{i,l}$ and the vector $\mathbf{y}(f)$ and $\mathbf{z}(f)$ consist of $y_i(f), i = 1, ..., m$, and $X(f - lf_p), l = -L_0, ..., L_0$, respectively. The relation (6.5) ties the output spectrum to the unknown $X(f)$, which is the key to recover $x(t)$. The elements in $\mathbf{z}(f)$ covers all the spectral information of $x(t)$.

The design of MWC also supports the advanced configuration where $f_s > f_p$. This would be beneficial to reduce the number of channels at the expense of a higher

sampling rate in each channel and additional digital processing. Without losing generality, we assume that $f_s = Qf_p$, with odd $Q = 2Q_0 + 1$. For the ith physical channel, it can be shown that

$$y_i(f + qf_p) = \sum_{l=-L_0}^{L_0} c_{i,l+q} X(f - lf_p), q = -Q_0, ..., Q_0. \tag{6.6}$$

This expands each physical channel into Q virtual ones:

$$\begin{pmatrix} y_i(f - Q_0 f_p) \\ \vdots \\ y_i(f) \\ \vdots \\ y_i(f + Q_0 f_p) \end{pmatrix} = \begin{bmatrix} c_{i,L_0-Q_0} & & \cdots & & c_{i,-L_0-Q_0} \\ \vdots & \ddots & & & \vdots \\ c_{i,L_0} & \cdots & c_{i,-1} \quad c_{i,0} \quad c_{i,1} & \cdots & c_{i,-L_0} \\ \vdots & & & \ddots & \vdots \\ c_{i,L_0+Q_0} & & \cdots & & c_{i,-L_0+Q_0} \end{bmatrix} \begin{pmatrix} | \\ \mathbf{z}(f) \\ | \end{pmatrix}, f \in \mathcal{F}_p,$$

and the form of the key relation in (6.5) remains unchanged.

6.2.2 Reconstruction

The reconstruction stage consists of two steps and is implemented completely in the time domain. First, the spectral support is determined, and then the signal is recovered from the samples by a closed-form expression.

Spectral support recovery relies on recent ideas developed in the context of analog compressed sensing [41] and are implemented by a series of digital computations, which are grouped together under the Continuous-to-Finite (CTF) block [114]. Let the support of $\mathbf{z}(f)$ be $S = \bigcup_{f \in \mathcal{F}_p} \text{supp}(\mathbf{z}(f))$, where $\text{supp}(\cdot)$ is the set of indices of the nonzero entries of a vector. In other words, if $i \notin S$ then $z_i(f) = 0$ for all $f \in \mathcal{F}_p$. By exploiting the sparsity of $\mathbf{z}(f)$, the CTF efficiently infers the support S from a low-complexity finite program.

Consider the covariance of $\mathbf{y}[n]$, defined by

$$\mathbf{\Sigma}_y = \sum_n \mathbf{y}[n]\mathbf{y}^H[n]. \tag{6.7}$$

According to Parseval's theorem,

$$\mathbf{\Sigma}_y = \int_{f\in\mathcal{F}_p} \mathbf{y}(f)\mathbf{y}(f)^H df = \mathbf{A}\mathbf{\Sigma}_z\mathbf{A}^H, \tag{6.8}$$

where

$$\mathbf{\Sigma}_z = \int_{f\in\mathcal{F}_p} \mathbf{z}(f)\mathbf{z}(f)^H df. \tag{6.9}$$

Eq. (6.8) is the key equation of MWC. As $\mathbf{z}(f)$ is a sparse vector, $\mathbf{\Sigma}_z(i,j) \neq 0$ if and only if $i,j \in S$, where $\mathbf{\Sigma}_z(i,j)$ is the (i,j)-th entry of $\mathbf{\Sigma}_z$. This indicates that $\mathbf{\Sigma}_z$ is a high-dimensional matrix but only a portion of rows and columns are non-zero, and $\mathbf{\Sigma}_y$ is a low-dimensional mapping of the high-dimensional but sparse matrix $\mathbf{\Sigma}_z$. The primary objective, referred to as the support recovery, is to infer S from $\mathbf{\Sigma}_y$, given the known sensing matrix $\mathbf{A}$. Indeed, by eigen-decomposing $\mathbf{\Sigma}_y$ as $\mathbf{\Sigma}_y = \mathbf{V}\mathbf{V}^H$, it is proved in [114] that the following equation

$$\mathbf{V} = \mathbf{A}\mathbf{U} \tag{6.10}$$

has a unique solution matrix $\mathbf{U}$ with minimal number of non-identically zero rows, and that the locations of these rows coincide with the support S. The sparse solution of the under determined problem (6.10) can be efficiently solved by algorithms in [118, 119, 120, 121].

Once the support S is determined, it follows from (6.5) that

$$\begin{aligned} \mathbf{z}_S[n] &= \mathbf{A}_S^{\dagger}\mathbf{y}[n] \\ z_i[n] &= 0,\ i \notin S, \end{aligned} \tag{6.11}$$

where $\mathbf{z}[n] = (z_1[n], \ldots, z_L[n])^T$ and $z_i[n]$ is the inverse DTFT of $z_i(f)$. $\mathbf{z}_S[n]$ and $\mathbf{A}_S$ mean the subvector and submatrix comprised of the rows of $\mathbf{z}[n]$ and $\mathbf{A}$ indexed by S, respectively. The notation $(\cdot)^{\dagger}$ denotes the pseudo inverse. Equation (6.11)

Table 6.1: Prototype parameters

Parameter	Choice
Signal model	$N = 6$, $B = 19$ MHz, $f_{\mathrm{NYQ}} = 2$ GHz
Number of channels m (basic)	4
Number of channels mQ (advanced)	12
Waveform type	periodic sign alternation
Alternation rate	2.075 GHz
Sign pattern length	108
Period f_p	$2.075/108 = 19.212$ MHz
Filter cutoff	33 MHz
Sampling rate/channel f_s	70 MHz

allows $z_i[n]$ to be generated at the low rate f_s. Every $z_i[n]$ is then interpolated to a continuous baseband signal at rate f_s (*e.g.*, using digital-to-analog devices) yielding (complex valued) $z_i(t)$:

$$z_i(t) = \sum_{n=-\infty}^{\infty} z_i[n] h(t - nT_s), \tag{6.12}$$

where $h(t) = \mathrm{sinc}(\pi t/T_s)$. Finally, $x(t)$ is reconstructed by modulating $z_i(t)$ to their corresponding bands:

$$\hat{x}(t) = \sum_{i \in S, i > L_0} \mathrm{Re}\left\{z_i(t)\right\} \cos(2\pi i f_p t) + \mathrm{Im}\left\{z_i(t)\right\} \sin(2\pi i f_p t), \tag{6.13}$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ denote the real and imaginary part of their argument, respectively.

6.2.3 Prototype

To validate the MWC concept a board-level prototype of the system was developed in [2]. The hardware consists of a pair of printed boards: the analog board and digital board. It aims at supporting input signals with 2GHz sampling rate and 120MHz spectrum occupancy, with arbitrary transmission frequencies. The sampling rate is as low as 280MHz and the system uses the advanced configuration where $Q = 3$, yielding 12 equivalent channels. The specifications that were used in the circuit realization are summarized in Table 6.1.

Fig. 6.2 shows a photo snapshot of the analog board, which is comprised of three consecutive stages: splitting the input into four channels, mixing with the sign patterns $p_i(t)$, and lowpass filtering. The input signal passes through a 100MHz highpass filter to reject the range of radio stations and airport transmissions, which are typically not of interest. A breakdown diode is used to project the circuit from instantaneous high input power. After passing a low-noise amplifier (LNA) and several digitally controlled attenuators, the signal is split to four channels using a passive splitter and two RF switches. For each channel, the signal is equalized and then mixed with the corresponding waveform $p_i(t)$ that is provided by the digital board. The mixing stage, which is the heart of the prototype, uses passive mixture devices (SYM25-DHW) whose working points are carefully adjusted. Two cascaded elliptic lowpass filters of order seven finally conclude the channel, yielding the output signal $y_i(t)$.

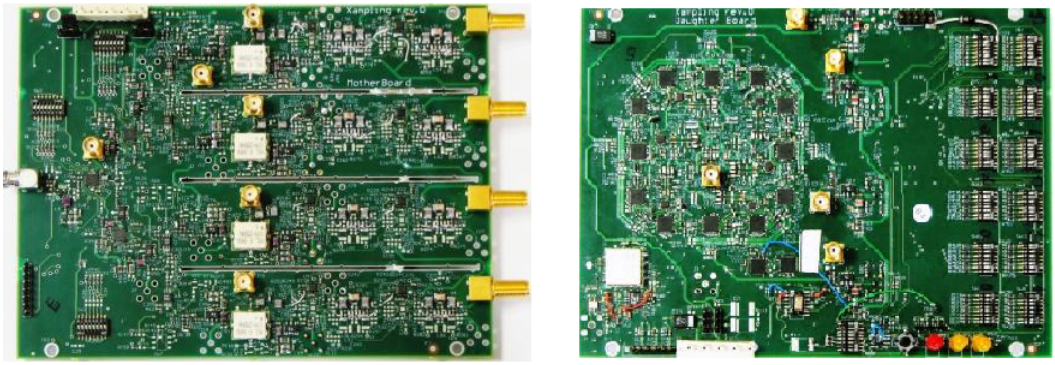


Figure 6.2: The analog (left) and digital (right) boards of the sampling prototype developed in [2].

The digital board, also shown in Fig. 6.2, is designed to general the mixing waveforms that feed into the mixtures of the analog board. It is comprised of a shift-register of 96 bits at emitter-coupled-logic technology, concatenating 12 packages of an 8-bit shift-register each. The initial value of the SR is `43 A7 A5 D7 96 AB 62 B7 2A B3 5C AC`. Each analog channel receives a different tap of the shift-registers

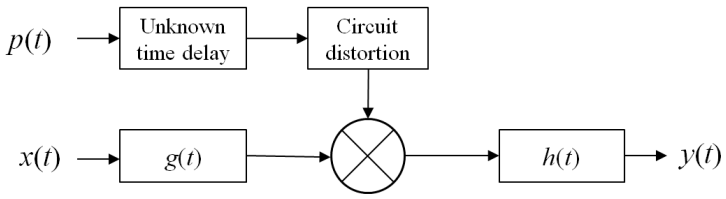


Figure 6.3: Modeling imperfections in the hardware calibration: equivalent front-end filter $g(t)$, delayed and distorted mixing waveform $p(t)$, the non-linear mixture and non-ideal lowpass filter $h(t)$.

and the waveforms $p_i(t)$ are shifted version of each other. The clock network for the shift-register packages is derived from a 2.075GHz sine waveform by locking a standard VCO to a 25 MHz temperature compensated crystal oscillator. The 12 shift-registers uses separate clocks routed in short straight lines to avoid unintended time skews.

6.3 MWC in practice: Calibration and reconstruction

The actual hardware realization of the MWC contains several imperfections, which are modeled in Fig. 6.3. For simplicity, we shall focus on a specific channel and omit the index i in the presentation below.

Several difficulties follows from the hardware model:

1. The circuit distortion makes the measured mixing waveform $p(t)$ far different from that in the conceptual design and this is the key reason for calibration.

2. The sampling grid $t = nT_s$ is chosen arbitrarily by the scope and is therefore not synchronized to the beginning of a period of $p(t)$. The unknown offset τ implies an uncertainty of the form $c_l \exp(j2\pi l f_p \tau)$.

3. The input signal has unknown phase with respect to the sampling grid. This is another source for errors due to missing synchronization.

4. The lowpass filter $h(t)$ has nonideal response. In addition, the chain of amplifi-

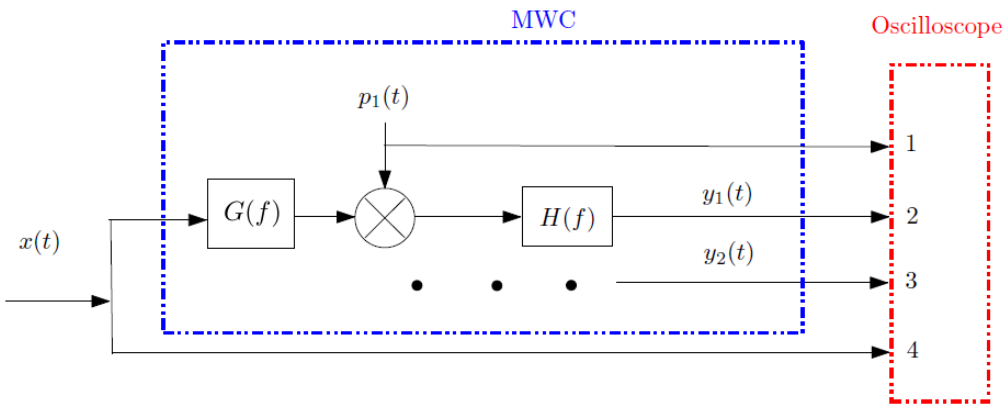


Figure 6.4: The proposed calibration system. The oscilloscope is triggered by $p_1(t)$ to synchronize the mixing waveform and the measured input and output signals.

cation stages from the input to the mixer is also nonideal, modeled by the filter $g(t)$.

5. The mixing procedure is nonlinear, so that instead of three output frequencies for a given sinusoid at the input, we get additional tones due to nonlinear effects. The power of these undesired tones is reduced from the energy of c_l.

6.3.1 Sensing matrix estimation

Due to the nonlinearity of the circuit, c_l cannot be obtained from direct measurement of $p(t)$. In order to estimate c_l as accurate as possible, we decided on the calibration system of Fig. 6.4, which is based on end-to-end measurements. In the sequel, we explain how this setup allows solving for c_l with respect to the above difficulties. In the proposed scheme, the output signals is measured by a four-terminal Agilent DSO80204 oscilloscope, which is triggered by the mixing waveform $p(t)$ of the first channel from the digital board. In this way, we synchronize $x(t)$ and $p(t)$ and eliminate the unknown offset τ in the calibration. The precision for the triggering time is as low as 1 ns, which is accurate enough in our setting. We then use a set of end-to-end measurements to estimate $\{c_0, c_1, c_2, ...\}$ and use $c_{-l} = c_l^*$ to obtain the

remaining coefficients with negative indices. The input signal is chosen as a sinusoid with a sweeping frequency f_k:

$$x^{(k)}(t) = 2A_k \cos\left(2\pi f_k t + \psi_k\right), \tag{6.14}$$

where

$$f_k = f_0 + k f_p, \tag{6.15}$$

where A_k and ψ are the magnitude and initial phase of the input sinusoid, f_0 is a pre-fixed frequency and the superscript k in (6.14) denotes the k-th measurement. For the kth measurement, the output signal is

$$y^{(k)}(t) = \left\{[x^{(k)}(t) * g(t)]p(t)\right\} * h(t). \tag{6.16}$$

The 3dB cutoff frequency of the lowpass filter is 33MHz. Noting that $f_p = 19.212$MHz, a sinusoid input signal generally result in three non-vanishing peaks within $[-33\text{MHz}, 33\text{MHz}]$ in the output. The spectrum of $y^{(k)}(t)$ can be then expressed as

$$Y^{(k)}(f) = \sum_{q=-1}^{1} F_q(k)\ \delta(f - f_q) + \sum_{q=-1}^{1} F_q(k)^*\ \delta(f + f_q). \tag{6.17}$$

Let l_0 be the band index of frequency f_0, where

$$l_0 = \left\lfloor \frac{f_0}{f_p} + \frac{1}{2} \right\rfloor. \tag{6.18}$$

Then, the parameters in (6.17) are expressed as

$$f_q = f_0 - l_0 f_p + q f_p, \tag{6.19}$$

and

$$F_q(k) = A_k e^{j\psi_k} \cdot G(f_k) \cdot H(f_q) \cdot c_{-l_0-k+q}, \tag{6.20}$$

	c_{-6}	c_{-5}	c_{-4}
Measurement #1	$0.3251 + j0.1454$	$0.7290 - j0.3957$	$0.6676 + j0.0308$
Measurement #2	$0.3246 + j0.1413$	$0.7219 - j0.4075$	$0.6663 + j0.0252$
Measurement #3	$0.3258 + j0.1407$	$0.7194 - j0.4081$	$0.6669 + j0.0219$

Table 6.2: Three independent calibration results for Channel 1.

where $q = -1, 0, 1$. $G(f)$ and $H(f)$ is the frequency response of the high-pass and the low-pass filter, respectively.

As f_0 is fixed, $\{f_q\}_{q=-1,0,1}$ are constant values. We then estimate $F_q(k)$ from samples of $y^{(k)}(t)$. The frequency responses of $G(f)$ and $H(f)$ are measured and stored by an Agilent HP8753E network analyzer. To calculate c_{-l_0-k+q} from $F_q(k)$, we set $k = l_0, l_0 + 1, l_0 + 2, ...$ and obtain $\{c_0, c_{-1}, c_{-2}, ...\}$ from each measurement. The procedure is summarized as follows:

1. Estimate A_k and ψ_k from the input $x(t)$.
2. Look up $H(f_q)$ and $G(f_k)$.
3. Substitute A_k, ψ_k, $G(f_k)$ and $H(f_q)$ into (6.20) to obtain $\{c_0, c_{-1}, c_{-2}, ...\}$.

The oscilloscope in use has only four input terminals with two of them reserved by $p(t)$ and $x(t)$. Thus only two terminals are available to measure the output signals. Three measurements are then conduct to calibrate channel 1 and 2, channel 1 and 3, and channel 1 and 4, respectively. The results are merged and show in Fig. 6.5, where coefficients in the central part are zero due to the 100MHz highpass filter $G(f)$. The size of the sensing matrix $\mathbf{A}$ is 12×121 and the calibrated system supports input signals from 100MHz to 1.16GHz. As a by product we get three copies of calibrated coefficients for channel 1, where a subset of those coefficients are provided in Table 6.2 as a validation of synchronization.

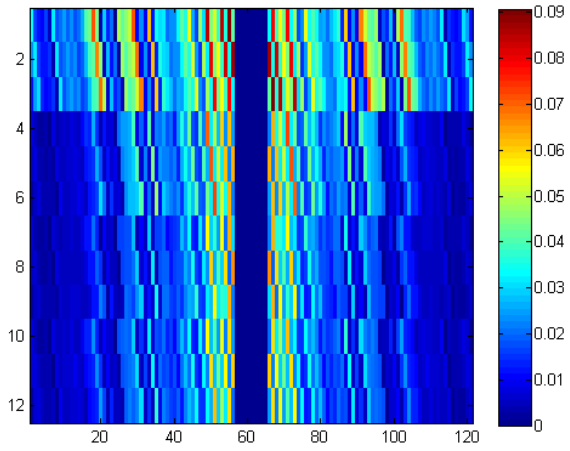


Figure 6.5: Heatmap of the calibrated 12×121 sensing matrix $\mathbf{A}$ (magnitude).

6.3.2 Filter compensation

The lowpass filters in the standard MWC are treated as ideal rect functions in the frequency domain. Imperfect lowpass filters not only affect the sensing matrix estimation but also impact perfect recovery in the reconstructed signal. Hereby we investigate how those imperfections results in signal recovery and propose compensation schemes to correct those imperfections. We start our discussion from the basic configuration where $Q = 1$ and then extend our results to the advanced configuration as used in the prototype.

The perfect recovery condition

The basic configuration applied here is slightly different than that in the conceptual design: while each physical channel still provide a single equation in (6.5), we no longer require $f_s = f_p$. Instead, we choose to oversample $y_i(t)$ at the rate f_s which is larger than both f_p and the stopband width of $H(f)$. Our analysis is based on (6.3), which applies to any analog filter $H(f)$. Indeed, (6.3) can be expressed as

$$Y\left(e^{j2\pi f T_s}\right) = \sum_{l=-L_0}^{L_0} c_l Q_l(f), f \in \mathcal{F}_s, \tag{6.21}$$

where $\mathcal{F}_s = [-f_s/2, f_s/2]$ and

$$Q_l(f) = X(f - lf_p)H(f). \tag{6.22}$$

Therefore, from (6.5) we actually solve $Q_l(f)$ rather than $z_l(f)$. After interpolation in (6.12) and modulation in (6.13), the resulted spectrum of the reconstructed signal is calculated as

$$\hat{X}(f) = \sum_{l=-L_0}^{L_0} Q_l\left(\frac{2\pi}{f_s}(f + lf_p)\right). \tag{6.23}$$

Substituting (6.22) in (6.23) we obtain

$$\hat{X}(f) = \left(\sum_{l=-L_0}^{L_0} H(f + lf_p)\right) X(f). \tag{6.24}$$

Since $X(f)$ is only non-zero within $[-f_{\mathrm{NYQ}}/2, f_{\mathrm{NYQ}}/2]$, the PR condition for $H(f)$ is then

$$\sum_{l=-L_0}^{L_0} H(f + lf_p) = 1, \ f \in \left[-\frac{f_{\mathrm{NYQ}}}{2}, \frac{f_{\mathrm{NYQ}}}{2}\right]. \tag{6.25}$$

We note that the PR condition in (6.25) coincides with the well-known Nyquist ISI criterion [122], and any lowpass filter that satisfies (6.25) is usually referred to as a *Nyquist filter*. Typical examples include raised cosine functions, Kaiser windows and others [122]. Any such filter will lead to PR without requiring any further processing.

Digital compensating FIR filters

In the above discussion we demonstrated that any Nyquist filter which satisfies (6.25) ensures PR. For lowpass filters that do not meet the PR condition, we now propose a simple compensation in the digital domain. The compensation scheme is illustrated in Fig. 6.6 for a single channel. Let $D(e^{j\omega})$ be the digital frequency response of the compensation filter, where we use the notation $e^{j\omega}$ to emphasis that

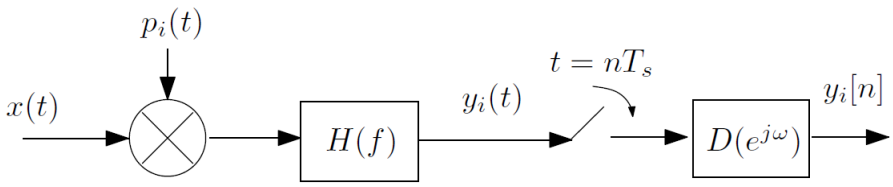


Figure 6.6: Illustration of one channel of the digital correction scheme.

the DTFT is 2π-periodic. The relationship in (6.21) still holds by replacing $Q_l(\omega)$ with

$$Q_l(f) = X\left(f - lf_p\right) H\left(f\right) D\left(e^{j2\pi T_s f}\right).$$

Therefore, to ensure perfect reconstruction we need to design a digital filter $D(e^{j\omega})$ such that the frequency response of the corrected analog filter

$$T(f) = H(f) D\left(e^{j2\pi T_s f}\right) \tag{6.26}$$

satisfies (6.25).

Here we show that we can implement $D(e^{jw})$ by an FIR filter. Let $\{d_n\}_{n=-N_0}^{N_0}$ be the coefficients of an FIR filter with order $2N_0 + 1$. The digital frequency response $D(e^{j\omega})$ is

$$D(e^{j\omega}) = \sum_{n=-N_0}^{N_0} d_n e^{-j\omega}. \tag{6.27}$$

Combining (6.26) and (6.27),

$$T(f) = \mathbf{h}(f)^H \mathbf{d}, \tag{6.28}$$

where $\mathbf{h}(f) = H(f)^*(e^{-j2\pi N_0 T_s f}, \ldots, e^{j2\pi N_0 T_s f})^T$, and $\mathbf{d}$ is the coefficient vector $\mathbf{d} = (d_{-N_0}, \ldots, d_{N_0})^T$. The design objective is to seek coefficients $\{d_n\}_{n=-N_0}^{N_0}$ such that $T(f)$ in (6.28) best meets the PR condition in terms of integrated squared error:

$$\min_{\mathbf{d}} \int_{-f_{\mathrm{NYQ}}/2}^{f_{\mathrm{NYQ}}/2} \left| \sum_{l=-L_0}^{L_0} \mathbf{h}(f - lf_p)^H \mathbf{d} - 1 \right|^2 \mathrm{d}f. \tag{6.29}$$

Since (6.29) is a least-squares problem, it has a closed-form solution. It can be shown that the optimal solution is:

$$\mathbf{d}_{\text{opt}} = \left[\int_{-f_{\text{NYQ}}/2}^{f_{\text{NYQ}}/2} \mathbf{g}(f)\mathbf{g}(f)^H \mathrm{d}f\right]^{-1} \int_{-f_{\text{NYQ}}/2}^{f_{\text{NYQ}}/2} \mathbf{g}(f)\mathrm{d}f \tag{6.30}$$

where $\mathbf{g}(f) = \sum_{l=-L_0}^{L_0} \mathbf{h}(f - lf_p)$. When $\mathbf{h}(f)$ contains $H(f)$ and is not specified analytically, computing the integrals in (6.30) can be performed using numerical methods.

The minimum of problem (6.29) also has a closed-form expression:

$$\epsilon = f_{\text{NYQ}} - \mathbf{d}_{\text{opt}}^H \int_{-f_{\text{NYQ}}/2}^{f_{\text{NYQ}}/2} \mathbf{g}(f)df. \tag{6.31}$$

This result can be used to bound the reconstruction error caused by $H(f)$ in advance. Indeed, by using Parseval's theorem it can be shown that

$$\int \left(\hat{x}(t) - x(t)\right)^2 dt \leq \epsilon \int x^2(t)dt, \tag{6.32}$$

which indicates that the reconstruction SNR $\geq 10\log_{10}(1/\epsilon)$.

Experimental validation

Hereby we demonstrate the proposed compensation method by experimental results, where two examples are studied.

In the first example we simulate the MWC system with non-ideal filters and evaluate the overall performance of the proposed compensation. The input $x(t)$ is a multiband signal consisting of 3 pairs of bands, each of width $B = 50$ MHz, defined as

$$x(t) = \sum_{i=1}^{3} \sqrt{E_i B}\text{sinc}(B(t - \tau_i))\cos(2\pi f_i(t - \tau_i)), \tag{6.33}$$

where the energy coefficients $E_i = \{1, 2, 3\}$, the time offsets $\tau_i = \{1.1, 0.3, 0.7\}$ μsecs, and the carriers are set to $f_i = \{1.8, 1.2, 2.8\}$ GHz. The Nyquist rate of $x(t)$ is

$f_{\text{NYQ}} = 10$ GHz. We choose $L_0 = 97$ and $f_p = f_{\text{NYQ}}/(2L_0 + 1) \simeq 51.3$ MHz. The number of channels is $m = 50$ and the same mixing functions $p_i(t)$ are used as in [114]. The main difference between the simulation in [114] and the one proposed here is that we use an 8-order Butterworth filter in each channel. The 3-dB bandwidth of the Butterworth filter is set to f_p. With a moderate oversampling, f_s is chosen as $f_s = 5/3f_p$. Finally, all the continuous signals are represented by a dense grid of 78975 samples observed within $[0, 1.6]$ μsecs, where the time resolution is $1/(5f_{\text{NYQ}})$. As

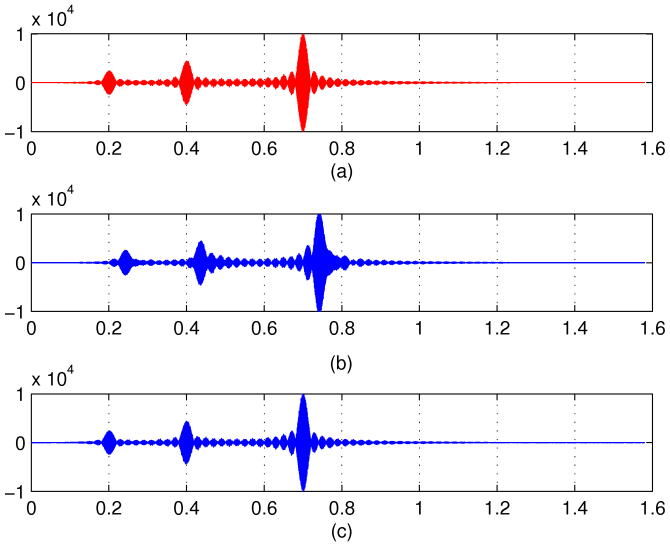


Figure 6.7: Reconstructions using Butterworth filters. (a) The multiband input signal $x(t)$. (b) Direct reconstruction signal. (c) Reconstructed signal after digital corrections.

predicted by our analysis, direct reconstruction using the standard approach yields distortions in the recovered signal, which can be found be comparing Fig. 6.7(b) with Fig. 6.7(c). We use a 21-order FIR filter to correct the non-ideal Butterworth filter in each channel. The coefficients are determined by (6.30) and the results are plotted in Fig. 6.8. We note that since the group delay of Butterworth filters is not constant, the coefficients of the correcting FIR filter is not symmetric. The reconstructed signal after applying digital corrections is plotted in Fig. 6.7(c). As

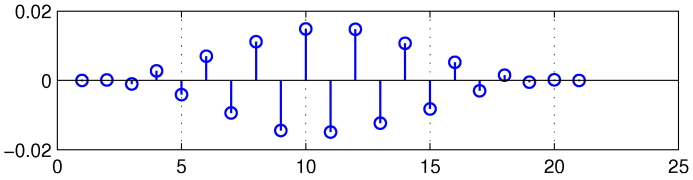


Figure 6.8: Coefficients of the correcting FIR filter.

expected, near perfect recovery is achieved. For further demonstration, we examine the PR condition of the employed Butterworth filter $H(f)$ and the corrected filter $T(f)$ (obtained by (6.26)) in Fig. 6.9, where $\sum_l H(f+lf_p)$ and $\sum_l T(f+lf_p)$ are plotted in dB. It can be seen that for $H(f)$ there exists significant distortions, which illustrates why direct reconstruction does not ensure PR.

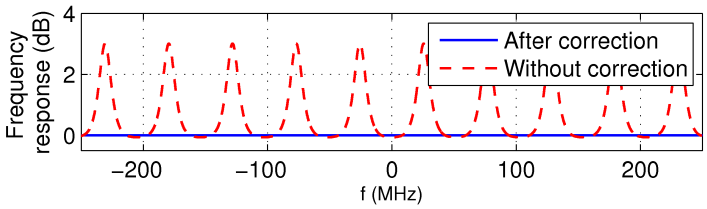


Figure 6.9: PR condition. Dot line represents $20\log_{10}|\sum_l H(f+lf_p)|$, and solid line represents $20\log_{10}|\sum_l T(f+lf_p)|$.

In the second example, we employ the proposed method to correct a real analog lowpass filter implemented in a recent hardware realization of the MWC system [2]. The frequency response is measured by an Agilent HP8753E network analyzer and the magnitude is shown in Fig. 6.10. Here we set $f_p = 60$ MHz and $f_s = 100$ MHz, The results of the correcting FIR filter and the PR condition test are shown in Fig. 6.11 and Fig. 6.12. These results indicate that our proposed compensator can be applied to practical applications in signal processing and communications.

Generalization to the advanced configuration

The MWC prototype uses the advanced configuration, where each channel requires three digital filters to split the output sequences. The following discussions extend the

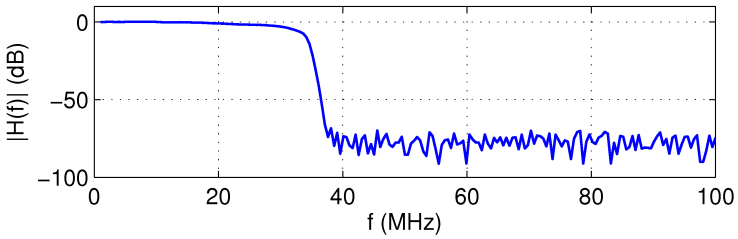


Figure 6.10: Frequency response of a real lowpass filter.

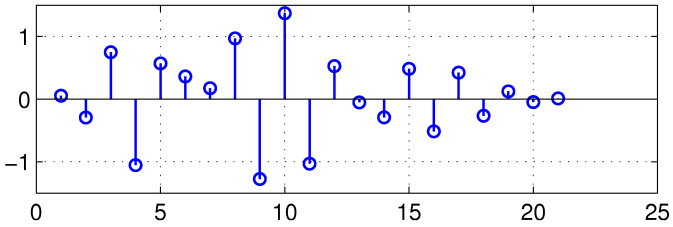


Figure 6.11: Coefficients of the correcting FIR filter for the real filter.

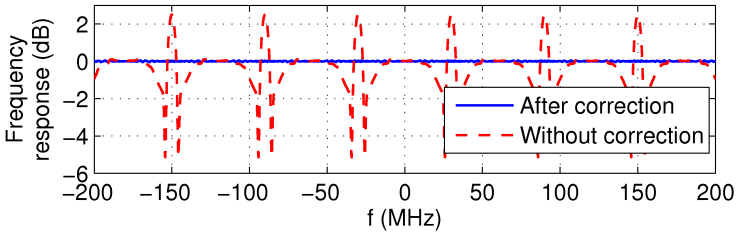


Figure 6.12: Tests of the PR condition for the real filter.

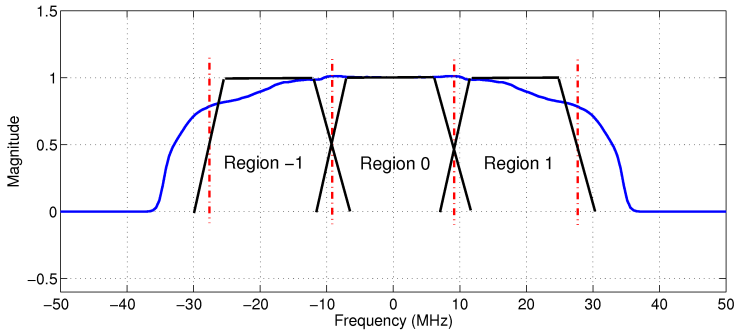


Figure 6.13: The magnitude of the frequency response of the low-pass filter $H(f)$. Illustrations for the corrected filters $T(f)$, $T(f+f_p)$ and $T(f-f_p)$.

results for the basic configuration into the advanced scenario that $Q = 3$. In contrast to the above discussion, we need to determine three digital filters to split and output samples and to compensate the non-ideal lowpass filter $H(f)$. Fig. 6.13 provides an illustration as well as the magnitude frequency response of $H(f)$ in the linear scale. Denote the DTFT of the digital filters in the three regions as $D_{-1}(e^{j\omega})$, $D_0(e^{j\omega})$ and $D_1(e^{j\omega})$, respectively. Using a similar derivation as for the basic configuration, the PR condition for the advanced configuration ($Q = 3$) is

$$H(f)D_0(e^{j2\pi T_s f}) = H(f-f_p)D_{-1}\left(e^{j2\pi T_s(f-f_p)}\right) = H(f+f_p)D_1\left(e^{j2\pi T_s(f+f_p)}\right), \tag{6.34}$$

and

$$\sum_{l=-L_0}^{L_0} T(f-lf_p) = 1, \tag{6.35}$$

where $T(f)$, referred to as the template filter, is defined as

$$T(f) = H(f)D_0(e^{j2\pi T_s f}). \tag{6.36}$$

The digital filters can be implemented using FIR filters:

$$D_q(e^{j2\pi\omega}) = \sum_{n=-N_0}^{N_0} d_n^{(q)} e^{-j2\pi\omega}, \tag{6.37}$$

where $q = 0, -1, 1$ and $2N_0 + 1$ is the order of the FIR filters. We propose to determine $d_n^{(q)}$ using least squares methods. We begin by calculating the coefficients

$\{d_n^{(1)}\}$ using the same approach as in Section 6.3.2. We then obtain the template filter

$$T(f) = H(f + f_p) \sum_{n=-N_0}^{N_0} d_n^{(1)}(e^{j2\pi T_s(f+f_p)}).$$

Finally, the filter coefficients $\{d_n^{(0)}\}$ and $\{d_n^{(-1)}\}$ are chosen in the manner that the equivalent corrected analog filters approximate $\hat{T}(f)$ in terms of least squares:

$$d_n^{(q)} = \arg\min_{d_n^{(q)}} \int_{-f_{\mathrm{NYQ}}}^{f_{\mathrm{NYQ}}} \left| H(f - kf_p) \sum_{n=-N_0}^{N_0} d_n^{(q)} e^{-j2\pi T_s f} - T(f) \right|^2 \mathrm{d}f, \ q = -1, 0. \tag{6.38}$$

Specifically, denote $\mathbf{d}^{(q)} = (d_{-N_0}^{(q)}, ..., d_{N_0}^{(q)})$. The solution to (6.38) has a closed form:

$$\mathbf{d}^{(q)} = \mathbf{R}_q^{-1}\mathbf{r}_q, \tag{6.39}$$

where

$$\mathbf{R}_q = \int_{-L_0 f_p}^{L_0 f_p} |H(f - qf_p)|^2 \mathbf{e}(f)\mathbf{e}(f)^H \mathrm{d}f, \tag{6.40}$$

and

$$\mathbf{r}_q = \int_{-L_0 f_p}^{L_0 f_p} T(f)H^*(f - qf_p)\mathbf{e}(f)\mathrm{d}f, \tag{6.41}$$

with $\mathbf{e}(f)$ defined as

$$\mathbf{e}(f) = (e^{-j2\pi N_0 T_s f}, ..., e^{j2\pi N_0 T_s f})^T. \tag{6.42}$$

In practice, we set $N_0 = 20$. Using the above approach, the corrected filter $\hat{T}(f)$ is shown in Fig. 6.14 and the FIR filter coefficients are shown in Fig. 6.15. Finally, Fig. 6.16 provides results of the PR test.

6.3.3 Synchronization and signal reconstruction

In the conceptual design of MWC, the input signal $x(t)$ is mixed by the periodic waveform $p(t)$ and then lowpass filtered by $h(t)$, yielding $y(t) = (x(t)p(t)) * h(t)$.

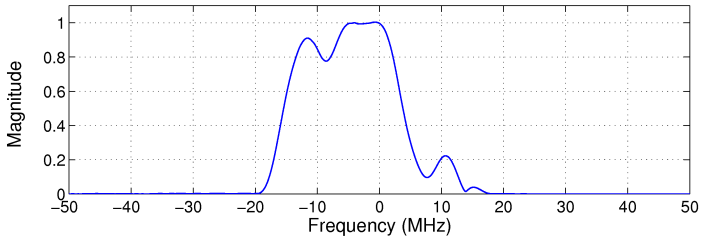


Figure 6.14: The magnitude of the corrected filter $T(f)$ by a least square design.

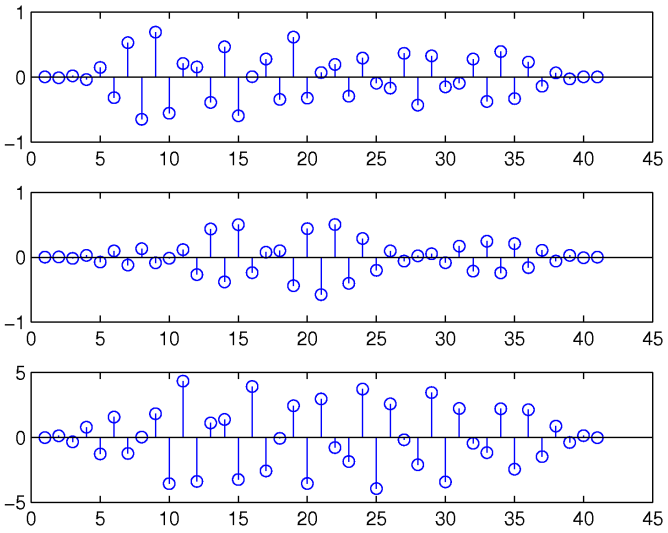


Figure 6.15: The FIR coefficients for the digital filters by the least square design. The first, second and third row corresponds to results for $k = 0$, $k = 1$, and $k = -1$, respectively.

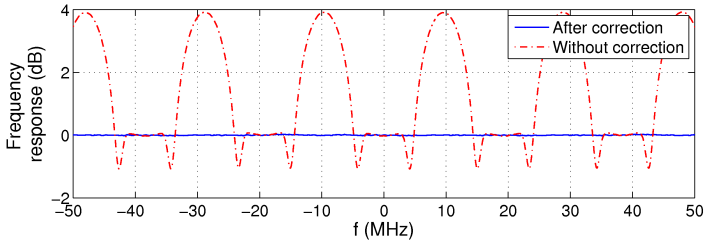


Figure 6.16: Tests of the PR condition in the advanced configuration ($Q = 3$): $20\log_{10}\left|\sum_{q=-1}^{1} T(f - qf_p)\right|$ v.s. frequency.

However, in practice, there is an unknown time offset between $x(t)$ and $p(t)$ when the input signal enters the system. Furthermore, such an offset varies for each independent experiment. In the calibration stage we avoid this problem by using $p(t)$ to trigger the oscilloscope. However, in signal recovery this trick is no long valid and the unknown time offset has to be estimated. In fact, this is the main problem we encountered in signal reconstruction using MWC.

We start by deriving the system transfer function that includes all the above imperfections.. For ease of derivation we work in the analog domain. Denote the unknown time offset as τ. According to our imperfect modeling, the output signal $y(t)$ becomes

$$y(t) = ((x(t) * g(t))p(t-\tau))) * h(t). \tag{6.43}$$

In the frequency domain,

$$y(f) = \sum_l c_l e^{-j2l\pi f_p \tau} X(f - lf_p)G(f - lf_p)H(f). \tag{6.44}$$

$y(t)$ is then processed by the q-th correcting filter, resulting in $y'_q(t)$ with its spectrum

$$y'_q(f) = y(f)D_q(e^{j2\pi T_s f}) = \sum_l c_l e^{-j2l\pi f_p \tau} X(f - lf_p)G(f - lf_p)H(f)D_q(e^{j2\pi T_s f}).$$

According to (6.34),

$$H(f)D_q(e^{j2\pi T_s f}) = T(f - qf_p). \tag{6.45}$$

There is

$$y'_q(f) = \sum_l c_l e^{-j2l\pi f_p \tau} X(f - lf_p)G(f - lf_p)T(f - qf_p).$$

Now consider $y'_q(f + qf_p)$:

$$\begin{aligned} y'_q(f + qf_p) &= \sum_l c_l e^{-j2l\pi f_p \tau} X(f - lf_p + qf_p)G(f - lf_p + qf_p)T(f), \\ &= \sum_{l=-L_0}^{L_0} c_{l+q} e^{-j2q\pi f_p \tau} X'(f - lf_p)e^{-j2l\pi f_p \tau}, \end{aligned} \tag{6.46}$$

where

$$X'(f - lf_p) = X(f - lf_p)G(f - lf_p)T(f). \tag{6.47}$$

Defining

$$\theta = 2\pi f_p \tau \tag{6.48}$$

and expressing (6.46) in the vector form, we have

$$\mathbf{y}'(f) = \mathbf{D}_1(\theta)\mathbf{A}\mathbf{D}_2(\theta)\mathbf{z}'(f), f \in \mathcal{F}_s, \tag{6.49}$$

or equivalently

$$\mathbf{y}'[n] = \mathbf{D}_1(\theta)\mathbf{A}\mathbf{D}_2(\theta)\mathbf{z}'[n], \tag{6.50}$$

in the time domain, where $\mathbf{y}'(f)$ consists of $y'_q(f + qf_p), q = -1, 0, 1$, for each physical channel, $\mathbf{z}'(f)$ is comprised of $X'(f - lf_p), l = -L_0, ..., L_0$, and $\mathbf{D}_1(\theta)$ and $\mathbf{D}_2(\theta)$ are two diagonal matrices defined by

$$\mathbf{D}_1(\theta) = \text{diag}\left(e^{j\theta}, 1, e^{j-\theta}, \cdots, e^{j\theta}, 1, e^{-j\theta}\right) \tag{6.51}$$

and

$$\mathbf{D}_2(\theta) = \text{diag}\left(e^{jL_0\theta}, \cdots, e^{j-L_0\theta}, \cdots, e^{jL_0\theta}, \cdots, e^{-jL_0\theta}\right). \tag{6.52}$$

Equation (6.50) is the new system transfer function connecting the input and output of the prototype. An important observation is that $\mathbf{D}_2(\theta)\mathbf{z}'(f)$ shares the same support of $\mathbf{z}(f)$ in (6.5). On the other hand, $\mathbf{D}_1(\theta)$ can lead to erroneous support recovery if the unknown phase θ is not properly estimated. Empirical results reveal that accurate estimation of θ is essential to the success of support recovery and the relationship between θ and recovery results may be complicated. Once

θ is obtained, the support recovery can be computed efficiently using fast greedy algorithms described in [120, 121]. Therefore, in this work we propose a brutal searching method to jointly estimate θ and the active spectrum slices. The procedure is depicted as follows.

1. Calculate the matrix $\mathbf{\Sigma}'_y = \sum_n \mathbf{y}'[n]\mathbf{y}'[n]^H$. Eigen-decompose $\mathbf{\Sigma}'_y$ as $\mathbf{\Sigma}'_y = \mathbf{V}\mathbf{V}^H$.
2. For $i = 1, 2, ..., N_0$
   (a) Set $\theta_i = 2\pi i/N_0$ and calculate $\mathbf{D}_1(\theta_i)$;
   (b) Solve the MMV problem: $\mathbf{D}_1^{-1}(\theta_i)\mathbf{V} = \mathbf{A}\mathbf{U}$, and calculate the errors of the solution, stored by $\texttt{score}(i)$;
3. Let $i_0 = \arg\min_i \texttt{score}(i)$. Run CTF to solve $\mathbf{D}_1(\theta_{i_0})^{-1}\mathbf{V} = \mathbf{A}\mathbf{U}$. The output support S is then the recovered support.
4. Use θ_{i_0} to calculate $\mathbf{D}_2(\theta_{i_0})$. Solve $\mathbf{z}[n]$ by

$$\mathbf{z}'[n] = \mathbf{D}_2^{-1}(\theta_{i_0})\mathbf{A}_S^{\dagger}\mathbf{D}_1^{-1}(\theta_{i_0})\mathbf{y}'[n], \tag{6.53}$$

   and reconstruct the highpass filtered input signal $x'(t) = x(t) * g(t)$ by submitting $\mathbf{z}'[n]$ into (6.12) and (6.13);
5. The input signal $x(t)$ is recovered by applying an inverse filter of $g(t)$ to $x'(t)$.

Finally, we verify our calibration and reconstruction scheme using real experimental data. Two signal generators were combined at the input terminal of the MWC prototype: an amplitude-modulated (AM) signal at 246.0MHz with 50kHz cosine envelope and a frequency-modulated (FM) source at 130.7MHz with 100kHz frequency deviation. The carrier positions were chosen so that their aliases overlay at baseband, as shown in Fig. 6.17. The digital recovery algorithm described above was implemented and detected the correct support set. The unknown carrier frequencies were estimated up to 10 kHz accuracy. In addition, correct reconstruction

of the AM and FM signal contents is validated and shown in Fig. 6.18 and Fig. 6.19, respectively. Please note that the carrier information is completely blind in the support estimation and recovery stage. The success of separating the AM and FM signals from their baseband mixtures demonstrates the effectiveness of our proposed calibration and reconstruction algorithms.

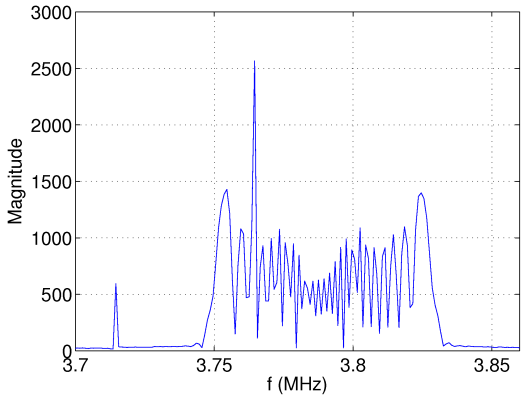


Figure 6.17: The spectrum of the output signal $y_1(t)$. The baseband spectrum of the AM and FM signals are mixed to each other.

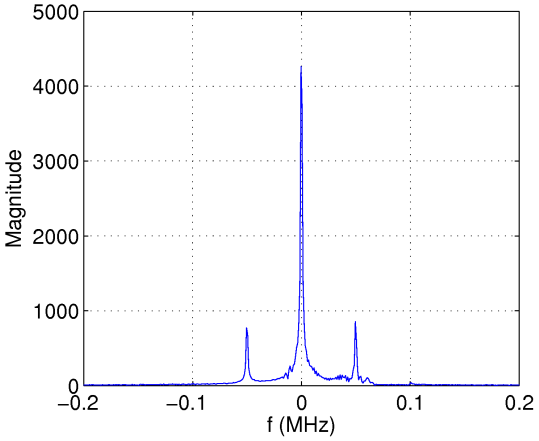


Figure 6.18: Reconstructed AM spectrum (baseband) from the MWC prototype.

6.4 Conclusion

In this chapter, we introduced the calibration and signal reconstruction methods for the MWC prototype, which is, to our best knowledge, the first reported

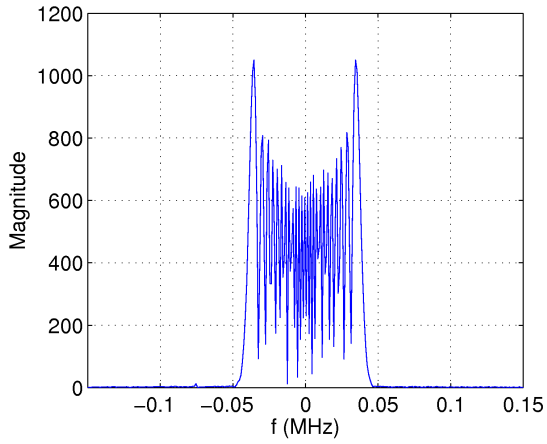


Figure 6.19: Reconstructed FM spectrum (baseband) from the MWC prototype.

sub-Nyquist hardware sampling system. The basic principle is to use compressive sensing algorithms to recover the sparse structure of a high-dimensional covariance matrix from its low-dimensional random projections. As an important step bridging theories and practical hardware prototype, we investigated the calibration and signal reconstruction problems for MWC. We identified and discussed three major problems: sensing matrix estimation, imperfect filter compensation and signal reconstruction. For each problem we proposed and derived effective solutions that conquered the difficulties caused by various imperfections of practical circuits. The proposed schemes and methods were demonstrated using real experimental data, which finally marked the success of MWC in practice.

CHAPTER VII

Conclusion and future works

Covariance matrix estimation has played an important role in many signal processing and machine learning applications. This thesis develops several important strategies and solutions for high-dimensional covariance estimation that accounts for small sample size, outlier sensitivity, and online implementation constraints.

In Chapter II, under a Gaussian assumption we improve upon the LW shrinkage covariance estimator using state-of-the-art Stein approaches. Chapter III then generalizes Chapter II for samples distributed in the elliptical family, yielding a robust covariance estimator in the high-dimensional setting. Unlike Chapter II and III which address covariance estimation for general purposes, in Chapter IV we consider covariance estimation with a time-varying graphical model. We divide online estimation of an unknown time-varying covariance into multiple adaptive system identification problem and propose the $\ell_{1,\infty}$ recursive group lasso that can be online updated efficiently using homotopy methods. Chapter V continues the discussion of regularized adaptive system identification and focuses on the stochastic gradient algorithm for low computational costs. In Chapter VI, we work on a real-time board-level sampling hardware that is based on covariance structure recovery of random projections. Our work bridges the gap between theories and practice and introduce

the state-of-the-art sub-Nyquist sampling equipment.

Both Chapter II and III are devoted to developing shrinkage covariance estimators to correct the systematical distortion of high-dimensional eigen-structures. The estimators presented in Chapter II can be used when the samples are Gaussian and no clear prior model can be imposed. When there are heavy-tails or outliers among the samples and the sample size is not enough to perform an outlier rejection, we recommend to use the estimator developed in Chapter III. However, all the methods proposed in Chapter II and III are based on Steinian type shrinkage estimator towards an identity matrix. We believe there are better options and in the future work we would like to develop data-dependent methods for shrinkage target selection.

In Chapter IV and V we developed regularized algorithms for identifying an unknown time-varying system. We split the (inverse) covariance estimation to a set of regression problems which can be then solved by the proposed system identification methods. The "divide-and-conquer" approach follows Meinshausen's strategy and achieves the capability of efficient online implementation. However, it may compromise the estimation accuracy. As pointed by [18], the multiple regression problems generally couple to each other if casting the covariance estimation into the regularized maximum-likelihood (ML) framework. As most regularized ML covariance estimators are computationally expensive for online implementation, developing an online covariance estimation method that efficiently updates the entire covariance matrix would be another future direction.

REFERENCES

[1] P. Zhao, G. Rocha, and B. Yu, "The composite absolute penalties family for grouped and hierarchical variable selection," *Annals of Statistics*, vol. 37, no. 6A, pp. 3468–3497, 2009.

[2] M. Mishali, Y. Eldar, O. Dounaevsky, and E. Shoshan, "Xampling: Analog to digital at sub-Nyquist rates," *Circuits, Devices and Systems, IET*, vol. 5, pp. 8–20, 2011.

[3] N. Meinshausen and P. Bühlmann, "High-dimensional graphs and variable selection with the lasso," *The Annals of Statistics*, vol. 34, no. 3, pp. 1436–1462, 2006.

[4] A. Dempster, "Covariance selection," *Biometrics*, vol. 28, no. 1, pp. 157–175, 1972.

[5] J. Friedman, "Regularized discriminant analysis," *Journal of the American statistical association*, vol. 84, no. 405, pp. 165–175, 1989.

[6] P. Stoica and R. Moses, *Introduction to spectral analysis*, vol. 57. Prentice Hall Upper Saddle River, New Jersey, 1997.

[7] K. Xu, M. Kliger, Y. Chen, P. Woolf, and A. Hero, "Revealing social networks of spammers through spectral clustering," in *Communications, 2009. ICC'09. IEEE International Conference on*, pp. 1–6, IEEE, 2009.

[8] K. Xu, M. Kliger, and A. Hero, "Evolutionary spectral clustering with adaptive forgetting factor," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 2174–2177, IEEE, 2010.

[9] G. Derado, F. Bowman, and C. Kilts, "Modeling the spatial and temporal dependence in fMRI data," *Biometrics*, vol. 66, no. 3, pp. 949–957, 2010.

[10] J. Schafer and K. Strimmer, "A shrinkage approach to large-scale covariance matrix estimation and implications for functional genomics," *Statistical applications in genetics and molecular biology*, vol. 4, no. 1, p. 1175, 2005.

[11] I. Johnstone and A. Lu, "Sparse principal components analysis," *J. Amer. Statist. Assoc*, 2008.

[12] N. Rao, J. Mingo, R. Speicher, and A. Edelman, "Statistical eigen-inference from large Wishart matrices," *Ann. Statist*, vol. 36, no. 6, pp. 2850–2885, 2008.

[13] W. James and C. Stein, "Estimation with quadratic loss," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, p. 361, University of California Press, 1956.

[14] C. Stein, "Estimation of a covariance matrix," *Rietz Lecture*, 1975.

[15] L. Haff, "Empirical Bayes estimation of the multivariate normal covariance matrix," *The Annals of Statistics*, vol. 8, no. 3, pp. 586–597, 1980.

[16] D. Dey and C. Srinivasan, "Estimation of a covariance matrix under Stein's loss," *The Annals of Statistics*, vol. 13, no. 4, pp. 1581–1591, 1985.

[17] O. Ledoit and M. Wolf, "A well-conditioned estimator for large-dimensional covariance matrices," *Journal of multivariate analysis*, vol. 88, no. 2, pp. 365–411, 2004.

[18] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, no. 3, p. 432, 2008.

[19] A. Rothman, P. Bickel, E. Levina, and J. Zhu, "Sparse permutation invariant covariance estimation," *Electronic Journal of Statistics*, vol. 2, pp. 494–515, 2008.

[20] O. Banerjee, L. El Ghaoui, and A. d'Aspremont, "Model selection through sparse maximum likelihood estimation for multivariate gaussian or binary data," *The Journal of Machine Learning Research*, vol. 9, pp. 485–516, 2008.

[21] M. Yuan and Y. Lin, "Model election and estimation in the Gaussian graphical model," *Biometrika*, 2007.

[22] W. Wu and M. Pourahmadi, "Nonparametric estimation of large covariance matrices of longitudinal data," *Biometrika*, vol. 90, no. 4, p. 831, 2003.

[23] J. Huang, N. Liu, M. Pourahmadi, and L. Liu, "Covariance matrix selection and estimation via penalised normal likelihood," *Biometrika*, vol. 93, no. 1, p. 85, 2006.

[24] P. Bickel and E. Levina, "Covariance regularization by thresholding," *Ann. Statist*, vol. 36, no. 6, pp. 2577–2604, 2008.

[25] P. Bickel and E. Levina, "Regularized estimation of large covariance matrices," *Annals of Statistics*, vol. 36, no. 1, pp. 199–227, 2008.

[26] A. Rothman, E. Levina, and J. Zhu, "Generalized thresholding of large covariance matrices," *Journal of the American Statistical Association*, vol. 104, no. 485, pp. 177–186, 2009.

[27] N. El Karoui, "Operator norm consistent estimation of large dimensional sparse covariance matrices," *Ann. Statist*, vol. 36, no. 6, pp. 2717–2756, 2008.

[28] R. Furrer and T. Bengtsson, "Estimation of high-dimensional prior and posterior covariance matrices in Kalman filter variants," *Journal of Multivariate Analysis*, vol. 98, no. 2, pp. 227–255, 2007.

[29] J. Fan, Y. Fan, and J. Lv, "High dimensional covariance matrix estimation using a factor model," *Journal of Econometrics*, vol. 147, no. 1, pp. 186–197, 2008.

[30] P. Huber and E. Ronchetti, *Robust statistics*. John Wiley & Sons Inc, 2009.

[31] D. Tyler, "A distribution-free M-estimator of multivariate scatter," *The Annals of Statistics*, vol. 15, no. 1, pp. 234–251, 1987.

[32] P. Rousseeuw, "Multivariate estimation with high breakdown point," *Mathematical statistics and applications*, vol. 8, pp. 283–297, 1985.

[33] P. Rousseeuw and K. Van Driessen, "A fast algorithm for the minimum covariance determinant estimator," *Technometrics*, vol. 41, no. 3, pp. 212–223, 1999.

[34] P. Rousseeuw and V. Yohai, "Robust regression by means of S-estimators," *Robust and nonlinear time series analysis*, vol. 26, pp. 256–272, 1984.

[35] M. Finegold and M. Drton, "Robust Graphical Modeling with Classical and Alternative T-Distributions," *Arxiv preprint arXiv:1009.3669*, 2010.

[36] Y. Abramovich, N. Spencer, and M. Turley, "Time-varying autoregressive (TVAR) models for multiple radar observations," *Signal Processing, IEEE Transactions on*, vol. 55, no. 4, pp. 1298–1311, 2007.

[37] Y. Abramovich, N. Spencer, and M. Turley, "Order estimation and discrimination between stationary and time-varying (TVAR) autoregressive models," *Signal Processing, IEEE Transactions on*, vol. 55, no. 6, pp. 2861–2876, 2007.

[38] Y. Abramovich, B. Johnson, and N. Spencer, "Two-dimensional multivariate parametric models for radar applicationsPart II: Maximum-entropy extensions for hermitian-block matrices," *Signal Processing, IEEE Transactions on*, vol. 56, no. 11, pp. 5527–5539, 2008.

[39] M. Kolar, A. Le Song, and E. Xing, "Estimating time-varying networks," *Annals*, vol. 4, no. 1, pp. 94–123, 2010.

[40] E. Candès, "Compressive sampling," *Int. Congress of Mathematics*, vol. 3, pp. 1433–1452, 2006.

[41] Y. Eldar, "Compressed sensing of analog signals in shift-invariant spaces," *Signal Processing, IEEE Transactions on*, vol. 57, no. 8, pp. 2986–2997, 2009.

[42] R. Abrahamsson, Y. Selen, and P. Stoica, "Enhanced covariance matrix estimators in adaptive beamforming," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 2, IEEE, 2007.

[43] O. Ledoit and M. Wolf, "Improved estimation of the covariance matrix of stock returns with an application to portfolio selection," *Journal of Empirical Finance*, vol. 10, no. 5, pp. 603–621, 2003.

[44] R. Tsay, *Analysis of financial time series*. Wiley-Interscience, 2005.

[45] F. Lin, S. Tsai, R. Otazo, A. Caprihan, L. Wald, J. Belliveau, and S. Posse, "Sensitivity-encoded (SENSE) proton echo-planar spectroscopic imaging (PEPSI) in the human brain," *Magnetic Resonance in Medicine*, vol. 57, no. 2, pp. 249–257, 2007.

[46] Y. Eldar and J. Chernoi, "A pre-test like estimator dominating the least-squares method," *Journal of Statistical Planning and Inference*, vol. 138, no. 10, pp. 3069–3085, 2008.

[47] H. Van Trees, *Detection, estimation, and modulation theory: Detection, estimation, and linear modulation theory*. Wiley-Interscience, 2001.

[48] S. John, "Some optimal multivariate tests," *Biometrika*, vol. 58, no. 1, p. 123, 1971.

[49] M. Srivastava and E. Carter, "Introduction to applied multivariate statistics.," *Elsevier Science Ltd*, 1983.

[50] O. Ledoit and M. Wolf, "Some hypothesis tests for the covariance matrix when the dimension is large compared to the sample size," *Annals of Statistics*, vol. 30, no. 4, pp. 1081–1102, 2002.

[51] P. Stoica, L. Jian, Z. Xumin, and J. Guerci, "On using a priori knowledge in space-time adaptive processing," *IEEE transactions on signal processing*, vol. 56, no. 6, pp. 2598–2602, 2008.

[52] G. Letac and H. Massam, "All invariant moments of the Wishart distribution," *Scandinavian Journal of Statistics*, vol. 31, no. 2, pp. 295–318, 2004.

[53] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, "On the self-similar nature of Ethernet traffic (extended version)," *Networking, IEEE/ACM Transactions on*, vol. 2, no. 1, pp. 1–15, 2002.

[54] P. Stoica and R. Moses, *Spectral analysis of signals*. Pearson/Prentice Hall, 2005.

[55] M. Srivastava, "Singular Wishart and multivariate beta distributions," *Annals of statistics*, pp. 1537–1560, 2003.

[56] T. Marzetta and B. Hochwald, "Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading," *Information Theory, IEEE Transactions on*, vol. 45, no. 1, pp. 139–157, 2002.

[57] B. Hassibi and T. Marzetta, "Multiple-antennas and isotropically random unitary inputs: The received signal density in closed form," *Information Theory, IEEE Transactions on*, vol. 48, no. 6, pp. 1473–1484, 2002.

[58] Y. Eldar and S. Shamai, "A covariance shaping framework for linear multiuser detection," *Information Theory, IEEE Transactions on*, vol. 51, no. 7, pp. 2426–2446, 2005.

[59] F. Hiai and D. Petz, *Asymptotic freeness almost everywhere for random matrices.* MaPhySto, Department of Mathematical Sciences, University of Aarhus, 1999.

[60] Y. Chen, A. Wiesel, Y. Eldar, and A. Hero, "Shrinkage Algorithms for MMSE Covariance Estimation," *Signal Processing, IEEE Transactions on*, vol. 58, no. 10, pp. 5016–5029, 2010.

[61] D. Kelker, "Distribution theory of spherical distributions and a location-scale parameter generalization," *Sankhyā: The Indian Journal of Statistics, Series A*, vol. 32, no. 4, pp. 419–430, 1970.

[62] K. Yao, "A representation theorem and its applications to spherically-invariant random processes," *Information Theory, IEEE Transactions on*, vol. 19, no. 5, pp. 600–608, 1973.

[63] K. Yao, M. Simon, and E. Bigiieri, "Unified theory on wireless communication fading statistics based on SIRP," in *Signal Processing Advances in Wireless Communications, 2004 IEEE 5th Workshop on*, pp. 135–139, IEEE, 2004.

[64] E. Conte, M. Lops, and G. Ricci, "Asymptotically optimum radar detection in compound-Gaussian clutter," *Aerospace and Electronic Systems, IEEE Transactions on*, vol. 31, no. 2, pp. 617–625, 2002.

[65] J. Billingsley, "Ground clutter measurements for surface-sited radar," *MIT Lincoln Lab Report.*, 1993.

[66] M. Rangaswamy, J. Michels, and B. Himed, "Statistical analysis of the non-homogeneity detector for STAP applications," *Digital Signal Processing*, vol. 14, no. 3, pp. 253–267, 2004.

[67] J. Wang, A. Dogandzic, and A. Nehorai, "Maximum likelihood estimation of compound-Gaussian clutter and target parameters," *Signal Processing, IEEE Transactions on*, vol. 54, no. 10, pp. 3884–3898, 2006.

[68] H. Brehm and W. Stammler, "Description and generation of spherically invariant speech-model signals," *Signal Processing*, vol. 12, no. 2, pp. 119–141, 1987.

[69] G. Vasile, J. Ovarlez, F. Pascal, C. Tison, L. Bombrun, M. Gay, and E. Trouvé, "Normalized coherency matrix estimation under the SIRV model. Alpine glacier POLSAR data analysis," in *Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International*, vol. 1, IEEE, 2009.

[70] G. Frahm, "Asymptotic distributions of robust shape matrices and scales," *Journal of Multivariate Analysis*, vol. 100, no. 7, pp. 1329–1337, 2009.

[71] F. Gini and M. Greco, "Covariance matrix estimation for CFAR detection in correlated heavy tailed clutter," *Signal Processing*, vol. 82, no. 12, pp. 1847–1859, 2002.

[72] E. Conte, A. De Maio, and G. Ricci, "Recursive estimation of the covariance matrix of a compound-Gaussian process and its application to adaptive CFAR detection," *Signal Processing, IEEE Transactions on*, vol. 50, no. 8, pp. 1908–1915, 2002.

[73] Y. Abramovich and N. Spencer, "Diagonally loaded normalised sample matrix inversion (LNSMI) for outlier-resistant adaptive filtering," in *Acoustics, Speech and Signal Processing, 2007. ICASSP 2007. IEEE International Conference on*, vol. 3, IEEE, 2007.

[74] G. Frahm, "Generalized elliptical distributions: theory and applications," *Unpublished Ph. D. thesis, University of Cologne*, 2004.

[75] U. Krause, "Relative stability for ascending and positively homogeneous operators on Banach spaces," *Journal of Mathematical Analysis and Applications*, vol. 188, no. 1, pp. 182–202, 1994.

[76] S. Bausson, F. Pascal, P. Forster, J. Ovarlez, and P. Larzabal, "First-and second-order moments of the normalized sample covariance matrix of spherically invariant random vectors," *Signal Processing Letters, IEEE*, vol. 14, no. 6, pp. 425–428, 2007.

[77] D. Johns and K. Martin, *Analog integrated circuit design*. Wiley-India, 2009.

[78] H. Lilliefors, "On the Kolmogorov-Smirnov test for normality with mean and variance unknown," *Journal of the American Statistical Association*, vol. 62, no. 318, pp. 399–402, 1967.

[79] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Computing Surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.

[80] F. Pascal, Y. Chitour, J. Ovarlez, P. Forster, and P. Larzabal, "Covariance structure maximum-likelihood estimates in compound Gaussian noise: existence and algorithm analysis," *Signal Processing, IEEE Transactions on*, vol. 56, no. 1, pp. 34–48, 2007.

[81] C. Auderset, C. Mazza, and E. Ruh, "Angular Gaussian and Cauchy estimation," *Journal of Multivariate Analysis*, vol. 93, no. 1, pp. 180–197, 2005.

[82] U. Krause, "Concave Perron-Frobenius theory and applications," *Nonlinear Analysis-Theory Methods and Applications*, vol. 47, no. 3, pp. 1457–1466, 2001.

[83] A. Tulino and S. Verdú, *Random matrix theory and wireless communications*. Now Publishers Inc, 2004.

[84] L. Ljung, "System identification," *Uncertainty and Control*, pp. 48–83, 1985.

[85] J. Duchi, S. Shalev-Shwartz, Y. Singer, and T. Chandra, "Efficient projections onto the l 1-ball for learning in high dimensions," in *Proceedings of the 25th international conference on Machine learning*, pp. 272–279, ACM, 2008.

[86] J. Duchi, S. Gould, and D. Koller, "Projected subgradient methods for learning sparse gaussians," in *Proc. of the Conf. on Uncertainty in AI*, Citeseer, 2008.

[87] J. Langford, L. Li, and T. Zhang, "Sparse online learning via truncated gradient," *The Journal of Machine Learning Research*, vol. 10, pp. 777–801, 2009.

[88] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 68, no. 1, pp. 49–67, 2006.

[89] S. Negahban and M. Wainwright, "Joint support recovery under high-dimensional scaling: Benefits and perils of $\ell_{1,\infty}$-regularization," *Advances in Neural Information Processing Systems*, pp. 1161–1168, 2008.

[90] A. Wiesel, Y. Eldar, and A. Hero, "Covariance estimation in decomposable Gaussian graphical models," *Signal Processing, IEEE Transactions on*, vol. 58, no. 3, pp. 1482–1492, 2010.

[91] W. Hager, "Updating the inverse of a matrix," *SIAM review*, vol. 31, no. 2, pp. 221–239, 1989.

[92] B. Widrow and S. Stearns, *Adaptive Signal Processing*. New Jersey: Prentice Hall, 1985.

[93] P. Garrigues and L. Ghaoui, "An homotopy algorithm for the Lasso with online observations," in *Neural Information Processing Systems (NIPS)*, vol. 21, 2008.

[94] S. Asif and J. Romberg, "Dynamic Updating for ℓ_1 Minimization," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 421–434, 2010.

[95] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Annals of statistics*, vol. 32, no. 2, pp. 407–451, 2004.

[96] A. Wiesel and A. Hero, "Decomposable principal component analysis," *Signal Processing, IEEE Transactions on*, vol. 57, no. 11, pp. 4369–4377, 2009.

[97] S. Kawamura and M. Hatori, "A TAP selection algorithm for adaptive filters," in *Proceedings of ICASSP*, vol. 11, pp. 2979–2982, 1986.

[98] J. Homer, I. Mareels, R. Bitmead, B. Wahlberg, and A. Gustafsson, "LMS estimation via structural detection," *IEEE Trans. on Signal Processing*, vol. 46, pp. 2651–2663, October 1998.

[99] Y. Li, Y. Gu, and K. Tang, "Parallel NLMS filters with stochastic active taps and step-sizes for sparse system identification," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 3, IEEE, 2006.

[100] D. Etter, "Identification of sparse impulse response systems using an adaptive delay filter," in *Proceedings of ICASSP*, pp. 1169–1172, 1985.

[101] M. Godavarti and A. O. Hero, "Partial update LMS algorithms," *IEEE Trans. on Signal Processing*, vol. 53, pp. 2382–2399, 2005.

[102] S. Gay, "An efficient, fast converging adaptive filter for network echocancellation," in *Proceedings of Asilomar*, vol. 1, pp. 394–398, 1998.

[103] D. Duttweiler, "Proportionate normalized least-mean-squares adaptation in echo cancelers," *IEEE Trans. on Speech and Audio Processing*, vol. 8, pp. 508–518, 2000.

[104] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Royal. Statist. Soc B.*, vol. 58, pp. 267–288, 1996.

[105] R. Baraniuk, "Compressive sensing," *IEEE Signal Processing Magazine*, vol. 25, pp. 21–30, March 2007.

[106] Y. Chen, Y. Gu, and A. Hero, "Sparse LMS for system identification," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pp. 3125–3128, IEEE, 2009.

[107] B. Babadi, N. Kalouptsidis, and V. Tarokh, "SPARLS: The sparse RLS algorithm," *Signal Processing, IEEE Transactions on*, vol. 58, no. 8, pp. 4013–4025, 2010.

[108] D. Angelosante, J. Bazerque, and G. Giannakis, "Online Adaptive Estimation of Sparse Signals: Where RLS Meets the ℓ_1 norm," *Signal Processing, IEEE Transactions on*, vol. 58, no. 7, pp. 3436–3447, 2010.

[109] Y. Kopsinis, K. Slavakis, and S. Theodoridis, "Online Sparse System Identification and Signal Reconstruction using Projections onto Weighted ℓ_1 Balls," *Arxiv preprint arXiv:1004.3040*, 2010.

[110] W. Schreiber, "Advanced television systems for terrestrial broadcasting: Some problems and some proposed solutions," *Proceedings of the IEEE*, vol. 83, no. 6, pp. 958–981, 1995.

[111] E. Candes, M. Wakin, and S. Boyd, "Enhancing sparsity by reweighted l 1 minimization," *Journal of Fourier Analysis and Applications*, vol. 14, no. 5, pp. 877–905, 2008.

[112] L. Meier, S. Van De Geer, and P. Buhlmann, "The group lasso for logistic regression," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 70, no. 1, pp. 53–71, 2008.

[113] F. Bach, "Consistency of the group Lasso and multiple kernel learning," *The Journal of Machine Learning Research*, vol. 9, pp. 1179–1225, 2008.

[114] M. Mishali and Y. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 375–391, 2010.

[115] A. Viterbi *et al.*, *CDMA: principles of spread spectrum communication*, vol. 129. Addison-Wesley Reading, MA, 1995.

[116] Y. Eldar and H. Bolcskei, "Block-sparsity: Coherence and efficient recovery," in *Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on*, pp. 2885–2888, IEEE, 2009.

[117] Y. Chen, M. Mishali, Y. Eldar, and A. Hero, "Modulated wideband converter with non-ideal lowpass filters," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, pp. 3630–3633, IEEE, 2010.

[118] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *Signal Processing, IEEE Transactions on*, vol. 54, no. 12, pp. 4634–4643, 2006.

[119] S. Cotter, B. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *Signal Processing, IEEE Transactions on*, vol. 53, no. 7, pp. 2477–2488, 2005.

[120] J. Tropp, A. Gilbert, and M. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.

[121] J. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Processing*, vol. 86, no. 3, pp. 589–602, 2006.

[122] J. Proakis and M. Salehi, *Digital communications*, vol. 4. McGraw-hill New York, 2001.