

EXPLOITATION OF GEOMETRY IN SIGNAL PROCESSING AND SENSING

YUEJIE CHI

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
ELECTRICAL ENGINEERING
ADVISER: ROBERT CALDERBANK

SEPTEMBER 2012

© Copyright by Yuejie Chi, 2012.
All rights reserved.

Abstract

Developments in hardware and massive sensors or transducers have quickly shifted many disciplines and applications from a scarcity of data to a deluge of data, presenting new challenges in signal processing of high-dimensional and large-scale datasets. One key insight is that the signals of interest often possess certain geometry, such that the information hidden in the signals usually exhibits a much lower dimension than that of the signals themselves. New hardware and sensor developments also offer new degrees of freedom in signal acquisition through design of pilot or probing sequences that are tuned to the underlying geometry of systems, which is defined by the choice of sensing sequences that optimizes the objective function of the system.

The first half of this thesis presents algorithms take advantage of geometry in the form of sparsity and low rank representations to minimize complexity and increase system capability. Performance is evaluated both in theory and in experiments. Chapter 2 discusses compressive sensing and sparse signal processing using Orthogonal Matching Pursuit, with applications to asynchronous multi-user detection in random access channels and diagnostic grade wireless ECG transmission and monitoring. Chapter 3 analyzes the sensitivity of compressive sensing to basis mismatch when the sparsity basis of the signal realized by the physics is differed from the one assumed in compressive sensing, and examines its implications for spectrum analysis and beamforming. Chapter 4 presents the Parallel Estimation and Tracking via Recursive Least Squares (PETRELS) algorithm for online estimation and tracking of a low-dimensional linear subspace from highly incomplete streaming data.

The second half of this thesis presents deterministic sensing sequences for active sensing and wireless communications. Chapter 5 presents coordination schemes of the transmission of a pair of Golay complementary waveforms to suppress the range sidelobes in a desired Doppler interval using the Prouhet-Thue-Morse sequence and its generalizations. Chapter 6 presents a family of minimum mean squared error (MMSE) optimal training sequences for channel state estimation in multi-user Multiple Input Multiple Output (MIMO)-Orthogonal Frequency Division Multiplexing (OFDM) systems.

Acknowledgements

I want to thank my advisor Prof. Robert Calderbank. I am very grateful to his guidance and the freedom he gave in research, which led me to discover the beauty of signal processing and information science. I am constantly amazed by his insights and vision into problems, and the way he finds the perfect balance between theory and practice. I remember his remark: “if theory is the x -axis and practice is the y -axis on a two-dimensional plane, what you want is to place yourself at a point with maximized area.” And he sets a great role model as a mathematician and an engineer.

I benefit tremendously from his generosity of providing opportunities to collaborate with many researchers outside Princeton, from which I developed independence and broadened my scope. I have been fortunate to explore not only Princeton, but also Stanford, Duke, and Colorado State University during my PhD studies, which is quite an experience for a graduate student.

He is a great educator who really cares about students and has their future in his mind. He consistently reminded me to improve my English as a non-native speaker. He encouraged my pursuit of academic career while allowing me to explore alternatives. He spent a lot of time help me practice my job talk, no matter how busy he already was. To these I am always indebted to him.

I want to thank Prof. Louis Scharf for his encouragements to my academic pursuits and for teaching me fly fishing. His enthusiasm and skepticism in research is a constant inspiration to me. I want to thank Prof. Ali Pezeshki for being a longtime friend and collaborator since my first project at Princeton till today. I want to thank Prof. Yonina Eldar for teaching me perfectionism in every technical detail. I want to thank Prof. Bede Liu and Prof. Ingrid Daubechies for being on my general exam committee. I want to thank Prof. Peter Ramadge for kindly serving as my dissertation reader, and thank Prof. Amit Singer and Prof. Paul Cuff for kindly serving on my dissertation committee. I want to thank my mentor Dr. Harinath Garudadri at Qualcomm Inc. for leading me to the area of wireless healthcare. I want to thank Prof. Andrea Goldsmith for hosting me at Stanford University. I also want to thank my mentor Dr. Fatih Porikli at Mitsubishi Electric Research Lab, Prof. Naofal Al-Dhahir at University of Texas at Dallas, Dr. Stephen Howard at DSTO Australia for teaching me various things during our collaborations.

I am grateful to my friends at and outside Princeton for their friendship and for making my life wonderful. Last but not least, I want to thank my parents for bringing me to this world and for their love ever since. To them I dedicate this thesis.

To my parents.

Contents

Abstract	iii
Acknowledgements	iv
1 Introduction	1
1.1 The Geometry of Signals	1
1.2 The Geometry of Systems	2
2 Sparse Signal Processing from Compressive Measurements	4
2.1 Introduction to Compressive Sensing	4
2.2 Performance of Orthogonal Matching Pursuit	7
2.3 Demodulation of Mutually Interfering Signals	11
2.3.1 System Model	12
2.3.2 Demodulation and Codeword Construction	13
2.3.3 Numerical Examples	14
2.4 Diagnostic Wireless ECG Monitoring	15
2.4.1 CS Approaches at Sensor and Receiver	18
2.4.2 Validation for Diagnostic Wireless ECG	20
3 Sensitivity of Compressive Sensing to Basis Mismatch	24
3.1 Motivation	24
3.2 Problem Formulation	27
3.3 Main Results	28
3.3.1 Degradation of the Best k -Term Approximation	28
3.3.2 Confidence Bounds for Image Inversion	30
3.3.3 Fourier Imaging and DFT Grid Mismatch	31
3.4 Numerical Examples	33
3.4.1 Tone Reconstruction with Basis Pursuit	34
3.4.2 Modal Analysis	37
3.4.3 A Synthetic Worst-Case Example	38
3.5 Conclusions	41
4 Low-rank Subspace Estimation from Partial Streaming Data	43
4.1 Introduction	43
4.2 Problem Statement and Related Work	44
4.2.1 Problem Statement	44
4.2.2 Conventional Subspace Identification and Tracking	46

4.2.3	Matrix Completion	47
4.3	The PETRELS Algorithm	47
4.3.1	PETRELS	48
4.3.2	Comparison with GROUSE	51
4.3.3	Simplified PETRELS	52
4.4	Convergence Analysis	53
4.4.1	PETRELS in the Full Observation Scenario	53
4.4.2	Simplified PETRELS in Rank-One Scenario	54
4.5	Numerical Results	56
4.5.1	Choice of Parameters	56
4.5.2	Direction-Of-Arrival Analysis	58
4.5.3	Matrix Completion	61
4.5.4	Simplified PETRELS	62
5	Waveform Coordination for Range Sidelobe Suppression in Radar	65
5.1	Introduction	65
5.2	Golay Complementary Waveforms	66
5.3	Doppler Resilience of Golay Waveforms	70
5.3.1	Range Sidelobe Suppression in Small Doppler	70
5.3.2	Range Sidelobes Suppression in Higher Doppler	74
5.4	Instantaneous Radar Polarimetry	76
5.4.1	Unitary Waveform Matrices	79
5.4.2	Doppler Resilient IRP	80
5.4.3	Four-Antenna IRP	81
6	Optimal Training Sequences in Multi-User MIMO-OFDM Systems	83
6.1	Introduction	83
6.2	System Model	85
6.3	Main Results	87
6.3.1	One OFDM Training Symbol	87
6.3.2	K OFDM Training Symbols with $K \geq 2$	88
6.3.3	Special case when $K = 2$	90
6.3.4	Discussions	90
6.4	Numerical Results	92
7	Appendix	94
7.1	Proofs in Chapter 2	94
7.1.1	Preparations for Proofs	94
7.1.2	Proof of Theorem 2.2.1	96
7.1.3	Proof of Theorem 2.2.2	98
7.1.4	Proof of Corollary 2.2.4	99
7.1.5	Proof of Corollary 2.2.5	99
7.2	Proofs in Chapter 3	99
7.2.1	Proof of Theorem 3.3.1	99
7.2.2	Proof of Theorem 3.3.2	101

7.3	Proofs in Chapter 4	101
7.3.1	Proof of Theorem 4.4.1	101
7.4	Proofs in Chapter 5	103
7.4.1	Proof of Theorem 5.3.3	103
7.4.2	Proof of Corollary 5.3.2	104
7.4.3	Proof of Theorem 5.3.4	104
7.5	Proofs in Chapter 6	104
7.5.1	Proof of Theorem 6.3.2	104
7.5.2	Proof of Theorem 6.3.3	105
Bibliography		106

Chapter 1

Introduction

1.1 The Geometry of Signals

The world of signal processing used to be limited by the complexity of algorithms but now it is limited by the volume of data to be processed. The developments in hardware and massive sensors or transducers have quickly shifted many scientific disciplines and engineering applications from a scarcity of data, where the amount of data to handle is below the capabilities of processing, communications and storage, to a deluge of data in recent years, posing new challenges for signal processing in understanding and extracting information from them.

One key insight that may be able to save the day is that the information hidden in the signals of interest usually exhibits a much lower dimension than that of the signals themselves. To put differently, it is possible to capture the high-dimensional signals by a small number of parameters. This defines the *geometry of signals*. One such geometry is *sparsity*, where the representation of the N -dimensional signal \mathbf{x} in a basis or an over-complete dictionary has no more than $k \ll N$ nonzero coefficients. In fact, the success of conventional compression schemes such as JPEG2000 is built on the premise that images are sparse or near sparse in the wavelet domain. Another geometry is *low rank*, where a series of N -dimensional signals $\mathbf{x}_i, i = 1, 2, \dots$ approximately spans a low-dimensional subspace, such as mutual distance measurements collected across time from a wireless sensor localization network. Consumer datasets from large-scale social networks also exhibit certain low rank structures. One example is movies ratings where sparsity derives from consensus in public opinion and the premise that a small number of factors have a disproportionate influence over personal preferences.

The geometry of signals can be harnessed to design a new framework for sensing and processing, called *Compressive Sensing*, and has shown its promise in Analog-to-Digital (A/D) conversion, magnetic resonance imaging (MRI), optical systems, etc. In stead of measuring the signal in full dimension, only a few compressive (sometimes random) linear measurements are taken and stored without compromising the inherited information in the signal. This simple sensing architecture performs compression at the same time, therefore bypassing the risk of running out of storage and com-

munication bandwidths, but requires more advanced signal processing methods that explicitly take the geometry of signals into account to recover the information.

However, sometimes the underlying geometry of signals assumed by the proposed algorithms is mismatched to the actual geometry determined by nature or the physical world. The Green's function that specifies the point spread function of a source is parameterized by continuous variables, but is sometimes modeled by a discrete grid. It is of great importance to understand when this mismatch would cause catastrophic error and when it would not, and to proceed with caution when this mismatch is inevitable. Moreover, in applications such as network monitoring and video streaming, the geometry of time-varying and dynamic signals may be changing slowly over time, and it is important to infer and track the changes in a timely fashion from a compressive number of streaming measurements.

The first half of the thesis from Chapter 2 to Chapter 4 presents algorithms that exploits sparsity and low rank in signal representations, and evaluates their performances both in theory and through numerical experiments. Chapter 2 discusses sparse signal processing and develops performance guarantees for signal reconstruction and model selection from compressive measurements using Orthogonal Matching Pursuit, with applications to asynchronous multi-user detection in random access channels and diagnostic grade wireless ECG transmission and monitoring. Chapter 3 analyzes the sensitivity of compressive sensing to basis mismatch when the sparsity basis of the signal proposed by the physics is differed from the one assumed in CS, and examined its implications for spectrum analysis and beamforming [1, 2, 3]. Chapter 4 presents the Parallel Estimation and Tracking via REcursive Least-Squares(PETRELS) algorithm for online estimation and tracking of a low-dimensional linear subspace from highly incomplete observations in streaming data [4, 5].

1.2 The Geometry of Systems

The task of signal processing is often centered around solving an optimization problem with an objective function $f(\mathbf{s})$, which depends on the sensing sequence \mathbf{s} used to probe the physical world. The sensing sequence \mathbf{s} should be chosen so that the objective $f(\mathbf{s})$ is optimized. The objective function defines *the geometry of systems*. In radar, the transmitted waveform should be selected such that its ambiguity function approximates a perfect thumbtack shape. In communications, the training sequence should be designed such that the channel state information is estimated with fidelity.

New hardware and sensor developments offer new degrees of freedom in signal acquisition, which enable new ways of designing sensing sequences that better match the geometry of systems. In active sensing, the Moyal's identity states that the volume under the ambiguity function of a single waveform is constant, implying it is impossible to obtain an ideal thumbtack shape from transmitting a single waveform. New generations of radar transmitters allow for transmission of different waveforms across time, frequency, space and polarization of the electromagnetic waves. While it is still impossible to create a perfect thumbtack shape for the whole delay-Doppler plane, it is actually possible to push the volume under the ambiguity function outside

of the delay-Doppler region of interest, by smartly coordinating the transmission of a family with only two waveforms.

In wireless communications, Multiple-Input Multiple-Output (MIMO) systems and Orthogonal Frequency Division Multiplexing (OFDM) support transmission of data symbols across multiple antennas and multiple carriers. Higher dimensionality makes it possible to introduce orthogonal space time codes and the algebraic structures of Fourier matrices into the design of optimal training sequences for channel estimation. The new designs match the geometric properties required by certain channel estimation methods, such as linear least-squares estimation methods.

The second half of the thesis from Chapter 5 to Chapter 6 presents deterministic sequences that are tailored to the geometry of systems in active sensing and wireless communications. Chapter 5 describes coordinated transmission of a pair of Golay complementary waveforms to suppress the range sidelobes in a desired Doppler interval in the radar ambiguity function. This involves Prouhet-Thue-Morse sequences and their generalization [6, 7]. Chapter 6 presents a family of minimum mean squared error (MMSE) optimal training sequences for multi-user MIMO-OFDM systems [8, 9].

Chapter 2

Sparse Signal Processing from Compressive Measurements

2.1 Introduction to Compressive Sensing

Sparse signal processing is a fundamental task in many applications involving high-dimensional data, such as image processing, gene analysis, and networks, where the signal of interests $\boldsymbol{\beta} \in \mathbb{C}^p$ can be represented as a sparse or compressible vector. If $\boldsymbol{\beta}$ is sparse, the cardinality of its support $T_{\boldsymbol{\beta}} = \{k : \beta_k \neq 0\}$ is assumed to be small. If $\boldsymbol{\beta}$ is compressible, then its entries obey a power law as $|\beta|_{(k)} \leq C_r k^{-r}$, where $|\beta|_{(k)}$ is the k th largest entry of absolute values of $\boldsymbol{\beta}$, $r > 1$ and C_r is a constant depending only on r . Let $\boldsymbol{\beta}_k$ be the best k -term approximation of $\boldsymbol{\beta}$, i.e. $\boldsymbol{\beta}_k = \operatorname{argmin}_{\boldsymbol{\beta}' \in \Sigma_k} \|\boldsymbol{\beta} - \boldsymbol{\beta}'\|_1$ where $\Sigma_k = \{\boldsymbol{\beta}' \in \mathbb{C}^p : |T_{\boldsymbol{\beta}'}| \leq k\}$. Then $\boldsymbol{\beta}_k$ is obtained by keeping the k -largest entries of $\boldsymbol{\beta}$ and padding with zeros elsewhere. A compressible signal $\boldsymbol{\beta}$ can be well approximated by $\boldsymbol{\beta}_k$, since

$$\|\boldsymbol{\beta} - \boldsymbol{\beta}_k\|_1 = \sum_{t=k+1}^p |\beta|_{(t)} \leq \sum_{t=k+1}^p C_r t^{-r} \leq \int_{x=k}^{\infty} C_r x^{-r} dx = \frac{C_r}{r-1} k^{-r+1}.$$

Much of the recent advances in Compressive Sensing (CS) [10, 11] have been centered around reconstructing a sparse signal $\boldsymbol{\beta} \in \mathbb{C}^p$ in the presence of noise $\boldsymbol{\eta}$ from a number of linear measurements $\boldsymbol{y} \in \mathbb{C}^n$ that is much smaller than the signal dimension p . The signal model is

$$\boldsymbol{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\eta}, \quad (2.1)$$

where $\mathbf{X} \in \mathbb{C}^{n \times p}$ is the measurement matrix.

There are commonly two goals when inverting (2.1) in different applications. The first one is support recovery, or model selection, where one is interested in estimating $T_{\boldsymbol{\beta}}$; one example is detecting the set of active users in multiaccess communication. The other is signal reconstruction, where one is interested in estimating the coefficients of $\boldsymbol{\beta}$; one example is image denoising.

In both cases, we seek the sparsest solution to $\boldsymbol{\beta}$ given the observation \mathbf{y} . We wish to solve

$$\boldsymbol{\beta}^* = \arg \min_{\boldsymbol{\beta}} \|\boldsymbol{\beta}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2 \leq \epsilon \quad (2.2)$$

where the noise $\|\boldsymbol{\eta}\|_2 \leq \epsilon$ is assumed bounded in the ℓ_2 norm. This is in general an NP-hard problem. However, if the measurement matrix \mathbf{X} satisfies certain properties, (2.2) can be solved exactly or near exactly with high probability using computationally feasible methods. There are two main classes of algorithms in this regard, one is convex optimization based algorithms such as Basis Pursuit (BP) [12], the other is greedy pursuit based algorithms such as Orthogonal Matching Pursuit (OMP) [13].

In Basis Pursuit, the solution to (2.2) is obtained using linear programming by solving the following ℓ_1 minimization problem:

$$\boldsymbol{\beta}^* = \arg \min_{\boldsymbol{\beta}} \|\boldsymbol{\beta}\|_1 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2 \leq \epsilon. \quad (2.3)$$

If \mathbf{X} satisfies the so-called *Restricted Isometry Property* (RIP) [11], [14], such that for any $2k$ -sparse vector \mathbf{v} ,

$$(1 - \delta_{2k})\|\mathbf{v}\|_2^2 \leq \|\mathbf{X}\mathbf{v}\|_2^2 \leq (1 + \delta_{2k})\|\mathbf{v}\|_2^2 \quad (2.4)$$

with $\delta_{2k} < \sqrt{2} - 1$, the solution $\boldsymbol{\beta}^*$ to (3.5) approximates $\boldsymbol{\beta}$ as [15], [14]

$$\|\boldsymbol{\beta} - \boldsymbol{\beta}^*\|_2 \leq C_0 k^{-1/2} \|\boldsymbol{\beta} - \boldsymbol{\beta}_k\|_1 + C_1 \epsilon. \quad (2.5)$$

The constants C_0 and C_1 are given by

$$C_0 = \frac{2(1 + \alpha_0)}{1 - \alpha_0} \quad \text{and} \quad C_1 = \frac{2\alpha_1}{1 - \alpha_0}. \quad (2.6)$$

where

$$\alpha_0 = \frac{\sqrt{2}\delta_{2k}}{1 - \delta_{2k}} \quad \text{and} \quad \alpha_1 = \frac{2\sqrt{1 + \delta_{2k}}}{1 - \delta_{2k}}. \quad (2.7)$$

It is shown in [11] that a random matrix drawn from suitable distributions (e.g. Gaussian or sub-Gaussian) will satisfy the RIP of order k with constant δ_k with high probability provided that $n \gtrsim \mathcal{O}(k \log(p/k)/\delta_k^2)$. Therefore $n \gtrsim \mathcal{O}(k \log(p/k))$ measurements are required to reconstruct all k -sparse vectors using BP from a random measurement matrix with high probability.

The representative of the second class of algorithms is OMP, formulated as Algorithm 1 and Algorithm 2 below. The two variants differ in the way they terminate the iterations; Algorithm 1 performs a fixed number of iterations, and Algorithm 2 iterates until the stopping rule is hit. It is shown in [16] for a fixed sparse vector $\boldsymbol{\beta}$, that a random matrix \mathbf{X} drawn from suitable distributions can recover $\boldsymbol{\beta}$ with high probability with $n \gtrsim \mathcal{O}(k \log p)$ measurements from the noiseless measurement $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$. Unfortunately, it does not provide a practical way to design or validate the properties required of a measurement matrix, and it does not guarantee a fixed measurement matrix can be used to recover all sparse vectors. The RIP can also be

used to characterize the performance of OMP. In [17], if the RIP of \mathbf{X} of order $k + 1$ is satisfied with $\delta_{k+1} < 1/(3\sqrt{k})$, then OMP will recover all k -sparse signal, again from noiseless measurements $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$. Then $n \gtrsim \mathcal{O}(k^2 \log(p/k))$ measurements are required to reconstruct all k -sparse vectors from OMP using a random measurement matrix with high probability. However, it is not computationally feasible to determine whether a matrix \mathbf{X} satisfies RIP, limiting the practical value of this criterion.

Algorithm 1 OMP with a fixed number of iterations

- 1: Input: an $n \times p$ matrix \mathbf{X} , a vector $\mathbf{y} \in \mathbb{C}^n$, and a sparsity level k
 - 2: Output: an estimate $\hat{\mathcal{S}}$ of the true model \mathcal{S}
 - 3: Initialization: $\hat{\mathcal{S}}_0 :=$ empty set, residual $\mathbf{r}_0 = \mathbf{y}$
 - 4: **for** $t := 1 : k$ **do**
 - 5: $\mathbf{f} := \mathbf{X}^H \mathbf{r}_{t-1}$
 - 6: $i := \arg \max_j |f_j|$
 - 7: $\hat{\mathcal{S}}_t := \hat{\mathcal{S}}_{t-1} \cup \{j\}$
 - 8: $\mathbf{r}_t := \mathbf{y} - \mathbf{X}_{\hat{\mathcal{S}}_t} (\mathbf{X}_{\hat{\mathcal{S}}_t}^H \mathbf{X}_{\hat{\mathcal{S}}_t})^{-1} \mathbf{X}_{\hat{\mathcal{S}}_t}^H \mathbf{y}$
 - 9: **end for**
 - 10: $\hat{\mathcal{S}} := \hat{\mathcal{S}}_k$
-

Algorithm 2 OMP with a stopping rule

- 1: Input: an $n \times p$ matrix \mathbf{X} , a vector $\mathbf{y} \in \mathbb{C}^n$, and a threshold δ
 - 2: Output: an estimate $\hat{\mathcal{S}}$ of the true model \mathcal{S}
 - 3: Initialization: $\hat{\mathcal{S}}_0 :=$ empty set, residual $\mathbf{r}_0 = \mathbf{y}$, set the iteration counter $t = 1$
 - 4: **while** $\|\mathbf{X}^H \mathbf{r}_{t-1}\|_\infty > \delta$ **do**
 - 5: $\mathbf{f} := \mathbf{X}^H \mathbf{r}_{t-1}$
 - 6: $i := \arg \max_j |f_j|$
 - 7: $\hat{\mathcal{S}}_t := \hat{\mathcal{S}}_{t-1} \cup \{j\}$
 - 8: $\mathbf{r}_t := \mathbf{y} - \mathbf{X}_{\hat{\mathcal{S}}_t} (\mathbf{X}_{\hat{\mathcal{S}}_t}^H \mathbf{X}_{\hat{\mathcal{S}}_t})^{-1} \mathbf{X}_{\hat{\mathcal{S}}_t}^H \mathbf{y}$
 - 9: $t := t + 1$
 - 10: **end while**
 - 11: $\hat{\mathcal{S}} := \hat{\mathcal{S}}_{t-1}$
-

In contrast, if \mathbf{X} is an $n \times p$ unit-norm measurement matrix, we define two coherence properties of $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_p]$ below that are easy to compute. The first one is worst-case coherence:

$$\mu = \max_{i \neq j} |\langle \mathbf{x}_i, \mathbf{x}_j \rangle| = \|\mathbf{X}^H \mathbf{X} - \mathbf{I}\|_\infty,$$

which captures the angles between different columns of \mathbf{X} . The second one is average coherence:

$$\nu = \frac{1}{p-1} \max_i \left| \sum_{j:j \neq i} \langle \mathbf{x}_i, \mathbf{x}_j \rangle \right| = \frac{1}{p-1} \|(\mathbf{X}^H \mathbf{X} - \mathbf{I})\mathbf{1}\|_\infty.$$

which captures the average correlation between one column of \mathbf{X} and the remaining columns.

The performance of OMP can also be characterized using the worst coherence of the measurement matrix \mathbf{X} . It is shown in [13] by that if $\mu < \frac{1}{2k-1}$, then OMP will recover any k -sparse vector $\boldsymbol{\beta}$ from the noiseless measurement $\mathbf{y} = \mathbf{X}\boldsymbol{\beta}$, and this result is confirmed to be sharp in [18]. Furthermore it is proven in [19] that provided the amplitudes of the nonzero entries of $\boldsymbol{\beta}$ are not too small, OMP will recover the support of the signal from noisy observations. From the Welch bound [20], $\mu \gtrsim \mathcal{O}(n^{-1/2})$, therefore in order to recover all k -sparse vectors, the number of measurements required is $n \gtrsim \mathcal{O}(k^2)$.

2.2 Performance of Orthogonal Matching Pursuit

OMP is appealing and competitive in many practical applications due to its simplicity and low computational cost [21]. However, compared with BP, the above performance guarantees for OMP suffer from the so-called “square-root bottleneck”, in that the number of measurements is quadratic in k , instead of linear in k as for BP. The goal of this section is to present performance guarantees of OMP for both support recovery and signal reconstruction that overcome the squared-root bottleneck. In particular, instead of finding a measurement matrix \mathbf{X} via random draws that is able to recover all k -sparse vectors, we aim to find a deterministic \mathbf{X} that is able to recover any k -sparse vectors with high probability from $n \gtrsim \mathcal{O}(k \log p)$ measurements, possibly corrupted by noise.

We are interested in \mathbf{X} that satisfies the Strong Coherence Property (SCP) if the equation below holds:

$$\mu \leq \frac{1}{240 \log p}, \quad \nu \leq \frac{\mu}{\sqrt{n}}. \quad (2.8)$$

Notice that the condition $\nu \leq \frac{\mu}{\sqrt{n}}$ can be achieved with essentially no cost via “wiggling” \mathbf{X} , i.e. flipping the signs of the columns of \mathbf{X} [22]. The “wiggling” procedure doesn’t change μ and $\|\mathbf{X}\|_2$. There are classes of deterministic matrices whose worst-case coherence achieves the Welch bound, and satisfy the SCP, examples including Gabor frames [23], Kerdock code sets [24], and Delsarte-Goethals code sets [24].

We define the minimum-to-average ratio MAR and the t th-largest-to-average ratio $\text{LAR}_{(t)}$ of a k -sparse vector $\boldsymbol{\beta}$ respectively as

$$\text{MAR} = \frac{|\beta|_{\min}^2}{\|\boldsymbol{\beta}\|_2^2/k}, \quad \text{LAR}_{(t)} = \frac{|\beta|_{(t)}^2}{\|\boldsymbol{\beta}\|_2^2/k}, \quad (2.9)$$

where $|\beta|_{(t)}$ and $|\beta|_{\min}$ are the t th largest entry and the minimum entry of $\boldsymbol{\beta}$ in absolute value. The signal-to-noise ratio SNR and minimum signal-to-noise ratio SNR_{\min} are

defined respectively as

$$\text{SNR}_{\min} = \frac{|\beta|_{\min}^2}{\mathbb{E}\|\boldsymbol{\eta}\|_2^2/k}, \quad \text{SNR} = \frac{\|\boldsymbol{\beta}\|_2^2}{\mathbb{E}\|\boldsymbol{\eta}\|_2^2}. \quad (2.10)$$

We have the following theorem for the OMP algorithm with a fixed number of iterations in Algorithm 1.

Theorem 2.2.1. *Suppose \mathbf{X} satisfies the strong coherence property for any $p \geq 128$, $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$. If the sparsity level of $\boldsymbol{\beta}$ satisfies*

$$k \leq \min \left\{ \frac{p}{c_2^2 \|\mathbf{X}\|_2^2 \log p}, \frac{1}{c_1^2 \mu^2 \log p} \right\} \quad (2.11)$$

for $c_1 = 50\sqrt{2}$ and $c_2 = 104\sqrt{2}$, and its nonzero entries satisfy

$$|\beta|_{(t+1)} > \frac{2\sigma\sqrt{(1+\alpha)\log p}}{1 - c_1\mu\sqrt{(k-t)\log p}}, \quad (2.12)$$

or write differently, as

$$\text{LAR}_{(t+1)} > \frac{4(1+\alpha)}{(1 - c_1\mu\sqrt{(k-t)\log p})^2} \cdot \left(\frac{k \log p}{n \text{SNR}} \right), \quad (2.13)$$

for $0 \leq t \leq k-1$ and $\alpha > 0$, then the OMP algorithm in Algorithm 1 successfully finds the support of $\boldsymbol{\beta}$ with probability at least $1 - k(p^{\alpha\pi})^{-1} - 2p^{-2\log 2} - 4p^{-1}$.

Proof. See Appendix 7.1.2. □

For the OMP algorithm with a stopping rule in Algorithm 2, we have the following theorem.

Theorem 2.2.2. *Suppose \mathbf{X} satisfies the strong coherence property for any $p \geq 128$, $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$. If the sparsity level of $\boldsymbol{\beta}$ satisfies (2.11) and its nonzero entries satisfy (2.13) for $\alpha > 0$, and choose $\delta = \sigma\sqrt{(1+\alpha)\log p}$, then the OMP algorithm in Algorithm 2 successfully finds the support of $\boldsymbol{\beta}$ with probability at least $1 - (k+1)(p^{\alpha\pi})^{-1} - 2p^{-2\log 2} - 4p^{-1}$ in k iterations.*

Proof. See Appendix 7.1.3. □

Since $\text{MAR} \leq \text{LAR}_{(t+1)}$ for all $0 \leq t \leq k-1$, we have the following corollary of Theorem 2.2.1 and Theorem 2.2.2.

Corollary 2.2.3. *Suppose \mathbf{X} satisfies the strong coherence property for any $p \geq 128$, $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$. If the sparsity level of $\boldsymbol{\beta}$ satisfies (2.11), and it satisfies*

$$\text{MAR} > \frac{4(1+\alpha)}{(1 - c_1\mu\sqrt{k \log p})^2} \cdot \left(\frac{k \log p}{n \text{SNR}} \right), \quad (2.14)$$

for $\alpha > 0$, then the support of $\boldsymbol{\beta}$ is successfully recovered by OMP with probability at least $1 - k(p^\alpha \pi)^{-1} - 2p^{-2 \log 2} - 4p^{-1}$ using Algorithm 1, and with probability at least $1 - (k + 1)(p^\alpha \pi)^{-1} - 2p^{-2 \log 2} - 4p^{-1}$ by Algorithm 2 with $\delta = \sigma \sqrt{(1 + \alpha) \log p}$ in k iterations.

Let $\theta = c_1 \mu \sqrt{k \log p} \in (0, 1)$, then (2.14) implies the sparsity level k satisfies

$$k < \frac{(1 - \theta)^2}{4(1 + \alpha)} \cdot \frac{\text{SNR}_{\min}}{\log p}. \quad (2.15)$$

Combining with (2.11), we have

$$k < \max_{0 < \theta < 1} \min \left\{ \frac{n(1 - \theta)^2 \text{SNR}_{\min}}{4(1 + \alpha) \log p}, \frac{\theta^2}{c_1^2 \mu^2 \log p}, \frac{p}{c_2^2 \|\mathbf{X}\|_2^2 \log p} \right\}, \quad (2.16)$$

where the first term is determined by SNR_{\min} , which is signal dependent; and the second term and the third term are determined by the worst-case coherence and the spectral norm of the measurement matrix \mathbf{X} . If \mathbf{X} is a tight frame, $\|\mathbf{X}\|_2^2 = p/n$, the third term becomes $k < \mathcal{O}(n/\log p)$. Now we write the worst-case coherence as $\mu = c_3 n^{-1/\gamma}$ for some $c_3 > 0$ and $\gamma \geq 2$. From the Welch bound [20] μ is lower bounded by $\mu \gtrsim \mathcal{O}(n^{-1/2})$, therefore the maximum sparsity level is determined by the second term in (2.16), given $k \lesssim \mathcal{O}((n/\log p)^{2/\gamma})$. Note that the sparsity level k doesn't depend strongly on the profile of signal strength of $\boldsymbol{\beta}$, i.e. MAR of the signal.

We have another corollary on partial recovery.

Corollary 2.2.4. *Suppose \mathbf{X} satisfies the strong coherence property for any $p \geq 128$, and $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$. If the sparsity level of $\boldsymbol{\beta}$ satisfies (2.11), and its nonzero entries satisfy (2.13) for $0 \leq t \leq k' - 1 \leq k - 1$ and $\alpha > 0$, then the OMP algorithm in both Algorithm 1 and Algorithm 2 successfully selects k' entries from the support of $\boldsymbol{\beta}$ with probability at least $1 - k'(p^\alpha \pi)^{-1} - 2p^{-2 \log 2} - 4p^{-1}$.*

Proof. See Appendix 7.1.4. □

It is worth noting that it is not necessarily the support of the k' -largest entries that is recovered from the first k' iterations. The next corollary provides the condition on detecting the k' -largest entries from the first k' iterations.

Corollary 2.2.5. *Suppose \mathbf{X} satisfies the SCP for any $p \geq 128$, and $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$. If the sparsity level of $\boldsymbol{\beta}$ satisfies (2.11), and its nonzero entries satisfy*

$$|\beta|_{(t+1)} > \frac{|\beta|_{(t+2)} + 2\sigma \sqrt{(1 + \alpha) \log p}}{1 - c_1 \mu \sqrt{(k - t) \log p}}, \quad (2.17)$$

for $0 \leq t \leq k' - 1 \leq k - 1$ and $\alpha > 0$, then the OMP algorithm in both Algorithm 1 and Algorithm 2 successfully selects k' largest entries from the support of $\boldsymbol{\beta}$ with probability at least $1 - k'(p^\alpha \pi)^{-1} - 2p^{-2 \log 2} - 4p^{-1}$.

Proof. See Appendix 7.1.5. □

Let Π be the support of $\boldsymbol{\beta}$, the signal $\boldsymbol{\beta}$ can be written as $\boldsymbol{\beta} = \mathbf{P}_\Pi \mathbf{z}$, where \mathbf{P}_Π is a partial identity matrix with the k columns of \mathbf{I}_n selected by Π , and \mathbf{z} is the non-zero part of $\boldsymbol{\beta}$. Conditioned on the event that the support Π is successfully recovered, we can reconstruct the amplitude of the signal $\boldsymbol{\beta}$ by first reconstructing the amplitude on the detected support Π via least squares as

$$\hat{\mathbf{z}} = \mathbf{X}_\Pi^\dagger \mathbf{y} = \mathbf{X}_\Pi^\dagger (\mathbf{X}_\Pi \mathbf{z} + \boldsymbol{\eta}) = \mathbf{z} + (\mathbf{X}_\Pi^H \mathbf{X}_\Pi)^{-1} \mathbf{X}_\Pi^T \boldsymbol{\eta},$$

then fill in the zero entries to obtain $\hat{\boldsymbol{\beta}}$. Since

$$\begin{aligned} \|(\mathbf{X}_\Pi^H \mathbf{X}_\Pi)^{-1} \mathbf{X}_\Pi^H \boldsymbol{\eta}\|_2^2 &\leq \|(\mathbf{X}_\Pi^H \mathbf{X}_\Pi)^{-1}\|_2 \|\mathbf{X}_\Pi^H \boldsymbol{\eta}\|_2^2 \\ &\leq 4k \|\mathbf{X}_\Pi^H \boldsymbol{\eta}\|_\infty^2 \leq 8k\sigma^2 \log p, \end{aligned} \quad (2.18)$$

therefore

$$\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_2^2 = \|\hat{\mathbf{z}} - \mathbf{z}\|_2^2 \leq 4(1 + \alpha)k\sigma^2 \log p.$$

We have the following theorem.

Theorem 2.2.6. *Suppose \mathbf{X} satisfies the SCP for any $p \geq 128$, and $\boldsymbol{\eta} \sim \mathcal{N}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$. If the sparsity level of $\boldsymbol{\beta}$ satisfies (2.11) and its nonzero entries satisfy (2.13) for $0 \leq t \leq k-1$ and $\alpha > 0$, then the ℓ_2 norm of the difference between the original signal and the reconstructed signal via a least-squares estimation on the detected support from the OMP algorithm satisfies*

$$\|\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\|_2^2 \leq 4(1 + \alpha)k\sigma^2 \log p$$

with probability at least $1 - k(p^\alpha \pi)^{-1} - 2p^{-2 \log 2} - 4p^{-1}$ using Algorithm 1, and with probability at least $1 - (k+1)(p^\alpha \pi)^{-1} - 2p^{-2 \log 2} - 4p^{-1}$ using Algorithm 2.

We now compare our bound with the performance guarantee of OMP for support recovery provided in [19], which we have modified slightly for complex Gaussian noise. In order to select exactly the correct support with probability at least $1 - (k+1)(p^\alpha \pi)^{-1}$ for the OMP Algorithm 2 with the stopping rule $\delta = \sigma \sqrt{(1 + \alpha) \log p}$, the signal $\boldsymbol{\beta}$ needs to satisfy

$$\text{MAR} > \frac{4(1 + \alpha)}{(1 - (2k - 1)\mu)^2} \cdot \left(\frac{k \log p}{n \text{SNR}} \right), \quad (2.19)$$

therefore the sparsity level of $\boldsymbol{\beta}$ satisfies

$$k < \max_{0 < \theta < 1} \min \left\{ \frac{(1 - \theta)^2 \text{SNR}_{\min}}{4(1 + \alpha) \log p}, \frac{1}{2} + \frac{\theta}{2\mu} \right\}.$$

The first term is the same as that in (2.16), but the second term gives $k \sim \mathcal{O}(\mu^{-1})$, therefore $k \lesssim \mathcal{O}(n^{1/\gamma})$. We achieved a much tighter bound (2.16) by sacrificing the probability of success to $1 - (k+1)(p^\alpha \pi)^{-1} - 2p^{-2 \log 2} - 4p^{-1}$.

We also compare with the performance guarantee of the Sorted One Step Thresholding (SOST) algorithm analyzed in [23], which outputs the index set of the k -largest entries in absolute values of $\mathbf{f} = \mathbf{X}^H \mathbf{y}$. By rephrasing Theorem 4 in [23], in order to

select the correct support with probability at least $1 - 6p^{-1}$, the sparsity level of β satisfies

$$k < \max_{0 < \theta < 1} \min \left\{ \frac{(1 - \theta)^2 \text{SNR}_{\min}}{16 \log p}, \frac{\theta^2}{800 \mu^2 \log p} \cdot \frac{1}{\text{MAR}}, \frac{n}{2 \log p} \right\}. \quad (2.20)$$

Compared with (2.16), the performance of the OMP algorithm is reduced because of the third term, when \mathbf{X} is not a tight frame. On the other hand, the SOST algorithm performs poorly when the MAR of the signal is much smaller than 1, as seen from the second term in (2.20).

In the next two sections, we provide two applications of CS. The first one is asynchronous multi-user detection in random access channels, presented in Section 2.3 [25], where users are distributed in a large space and our job is to find them. The second one is diagnostic grade wireless ECG transmission and monitoring, presented in Section 2.4 [26], where our job is to encode a signal in a large space so that it can be found.

2.3 Demodulation of Mutually Interfering Signals

Demodulation of mutually interfering signals is central to multiaccess communications. It includes the special case of the Random Access Channel (RAC) that arises in modeling control channels in wireless networks where Multi-User Detection (MUD) is used to recover active users. It may be expanded to include demodulation of transmitted symbols. The two biggest impediments are the asynchronous character of random access and the lack of Channel State Information (CSI) at the Base Station (BS). The signature waveforms are obtained by modulating a chip waveform by digital sequence of length L . Our goal is to maximize the number of users N that the network can support, and the number of active users K that the BS can reliably demodulate with or without requiring knowledge of the delays or CSI.

A baseline architecture for demodulation of a sparse superposition is a bank of matched filters, each correlating the received signal with a shift of a signature waveform. The first drawback is the number of required filters which is $N(\tau + 1)$ where τ is the maximum delay. A second drawback is that when the signature waveforms are not orthogonal, the noise will be colored and amplified.

We consider an alternative architecture where the analog signal is sampled directly at the chip rate. This approach does not amplify noise but it does require a high-rate Analog-to-Digital (A/D) converter. We frame the challenge of demodulation as a CS problem where the columns of the measurement matrix are randomly sampled shifts of the digital sequences used to generate the signature waveforms. This is the extension to asynchronous communication of the architecture for synchronous MUD proposed in [27]. This model of asynchronous RAC appears in [28] where compressive demodulation is accomplished through convex optimization; see also [29] for a treatment of synchronous RAC. The drawback of [28] is that convex optimization is difficult to realize to meet the need of real-time decoding. Compared with the result in [27], where the required number of samples is on the order of $K^2 \log N(\tau + 1)$, we

can process N users with a maximum discrete delay τ from a simple matching pursuit algorithm using a number of samples on the order of $K \log N(\tau + 1)$.

2.3.1 System Model

Consider a multiuser system with N users. We assume that users communicate using spread spectrum waveform of the form

$$x_n(t) = \sqrt{P_n} \sum_{l=0}^{L-1} a_{n,l} g(t - lT_c), \quad t \in [0, T), \quad n = 1, \dots, N, \quad (2.21)$$

where $g(t)$ is a unit-energy pulse $\int |g(t)|^2 dt = 1$, T is the symbol duration, T_c is the chip duration, P_n denotes the transmit power of the n th user, and the spreading sequence

$$\tilde{\mathbf{a}}_n = [a_{n,0} \quad \cdots \quad a_{n,L-1}]^T, \quad n = 1, \dots, N \quad (2.22)$$

is the L -length (real or complex-valued) codeword assigned to the n th user. The signal at the receiver is given by

$$y(t) = \sum_{n=1}^N g_n \sqrt{P_n} \delta_{\{n \in \mathcal{I}\}} b_n x_n(t - \tau_n) + w(t), \quad (2.23)$$

where $g_n \in \mathbb{C}$ and $\tau_n \in \mathbb{R}_+$ are the channel fading coefficient and the delay associated with the n th user, respectively. We assume binary phase-shift keying (BPSK) transmission, where $b_n \in \{-1, 1\}$ is the transmitted symbol of the n th user, and $w(t)$ is a complex additive white Gaussian noise (AWGN) introduced by the receiver circuitry. Denote by \mathcal{I} the set of active users. The Dirac function $\delta_{\{x\}} = 1$ if x is true otherwise it is equal to zero.

Define the individual discrete delays $\tau'_n \triangleq \lfloor \tau_n / T_c \rfloor \in \mathbb{Z}_+$, and the maximum discrete delay $\tau \triangleq \max_n \tau'_n \in \mathbb{Z}_+$. While the values of τ'_n are unknown, τ is assumed to be known by the transmitters and receivers.

We extend the vectors $\tilde{\mathbf{a}}_n$ periodically. That is, taking \mathbf{a}_n as the last $P = L - \tau - 1$ symbols of $\tilde{\mathbf{a}}_n$, we have $\tilde{a}_{n,l} = a_{n,P-\tau-1+l}$ for $l = 1, \dots, \tau + 1$. As a result, any length P sub-sequence of the vectors $\tilde{\mathbf{a}}_n$ will be a cyclic shift of \mathbf{a}_n .

We assume the codewords are of a reasonable length relative to the delays such that $P > M$. The receiver starts sampling from the $\tau + 1$ sample, so that all active users' waveforms have arrived. Then the receiver takes M compressive measurements which come from uniformly random subsampling of the received sequence or its DFT. As a result, the output data vector can be written as

$$\mathbf{y} = \mathbf{I}_\Omega \mathbf{A} \mathbf{R} \mathbf{b} + \mathbf{w} \triangleq \mathbf{H} \mathbf{R} \mathbf{b} + \mathbf{w}, \quad (2.24)$$

where $\mathbf{y} \in \mathbb{C}^{M \times 1}$, $\mathbf{H} = \mathbf{I}_\Omega \mathbf{A} \in \mathbb{C}^{M \times N(\tau+1)}$, where M is the number of samples, $\mathbf{A} \in \mathbb{C}^{P \times N(\tau+1)}$, and the noise is Gaussian distributed with zero mean and variance

$\sigma^2 \mathbf{I}_{M \times M}$. The subsampling matrix is \mathbf{I}_Ω , where Ω denotes indices of samples. The columns of the matrix \mathbf{A} have a block structure as $\mathbf{A} = [\mathbf{A}_1 \ \cdots \ \mathbf{A}_N]$, with each block $\mathbf{A}_n \in \mathbb{R}^{P \times (\tau+1)}$ consisting of a circulant-shifted codeword as

$$\mathbf{A}_n = [\mathcal{T}_0 \tilde{\mathbf{a}}_n \ \mathcal{T}_1 \tilde{\mathbf{a}}_n \ \cdots \ \mathcal{T}_\tau \tilde{\mathbf{a}}_n], \quad (2.25)$$

where the notation \mathcal{T}_k denotes circulant shift the matrix by k . The vector $\mathbf{b} \in \mathbb{C}^{N(\tau+1)}$ contains the transmitted symbols; it is a concatenation of N vectors \mathbf{b}'_n , each of length $\tau + 1$, with only one non-zero entry at the location of τ'_n :

$$b'_{n,m} = b_n \delta_{\{m=\tau'_n+1\}}.$$

The diagonal matrix \mathbf{R} has entries containing the transmitted power, channel gain, and the symbols:

$$R_{mm} = g_n \sqrt{P_n} \delta_{\{m=(n-1)(\tau+1)+\tau'_n\}}, \quad n = 1, \dots, N, \quad m = 1, \dots, N(\tau + 1).$$

We assume the support of active users \mathcal{I} is a uniform random K -subset of $\{1, \dots, N\}$.

2.3.2 Demodulation and Codeword Construction

Demodulation is accomplished by a matching pursuit procedure (Algorithm 3) that takes explicit account of the block structure of the measurement matrix \mathbf{A} and corresponding block sparsity of the received signal. The key feature is that after selecting one column from a block the demodulator ignores the other columns in that block, since they correspond to alternative shifts of the same signature waveform.

Algorithm 3 Matching Pursuit Detector for Asynchronous MUD

- 1: Input: matrices \mathbf{H} and \mathbf{R} , signal vector \mathbf{y} , number of active users K
 - 2: Output: active user set \mathcal{I} , transmitted symbols $b_n, n \in \mathcal{I}$
 - 3: Initialize: $\mathcal{I}_0 :=$ empty set, $\hat{\mathbf{b}}_0 := \mathbf{0}$, $\mathbf{v}_0 := \mathbf{y}$,
 $\mathcal{H}_0 = \{1, \dots, N(\tau + 1)\}$
 - 4: **for** $j = 0 \rightarrow K - 1$ **do**
 - 5: Compute: $\mathbf{f} := \mathbf{H}^H \mathbf{v}_j$
 - 6: Find $i = \operatorname{argmax}_{n \in \mathcal{H}_j} |f_n|$
 - 7: Detect active users: $\mathcal{I}_{j+1} = \mathcal{I}_j \cup \{[i/(\tau + 1)]\}$
 - 8: Update: $\mathcal{H}_{j+1} = \mathcal{H}_j \setminus \{[i/(\tau + 1)](\tau + 1) + 1, \dots, [i/(\tau + 1)](\tau + 1)\}$
 - 9: Detect symbols: $[\hat{\mathbf{b}}_{j+1}]_i = \operatorname{sgn}(r_i f_i)$, and $[\hat{\mathbf{b}}_{j+1}]_n = [\hat{\mathbf{b}}_j]_n$ for $n \neq i$.
 - 10: Update residual: $\mathbf{v}_{j+1} = \mathbf{v}_j - \mathbf{H} \mathbf{R} \mathbf{b}_{j+1}$
 - 11: **end for**
 - 12: $\hat{\mathcal{I}} = \mathcal{I}_K, \hat{\mathbf{b}} = \hat{\mathbf{b}}_K$
-

The performance guarantee of Algorithm 3 can be stated in terms of the coherence properties of \mathbf{H} , which shows with a number of samples on the order of $K \log N(\tau+1)$, we can detect N users with a maximum discrete delay τ with high probability. The readers are referred to [25] for details.

Here we describe how to construct the signature sequences $\tilde{\mathbf{a}}_n$ from Gabor frames. Define $\mathbf{g} \in \mathbb{C}^P$ be a seed vector with each entry $|g_i|^2 = 1/M$ and let $\mathbf{T}(\mathbf{g}) \in \mathbb{C}^{P \times P}$ be the circulant matrix generated from \mathbf{g} as $\mathbf{T}(\mathbf{g}) = [\mathcal{T}_0 \mathbf{g} \ \cdots \ \mathcal{T}_\tau \mathbf{g}]$. Its eigen-decomposition can be written as

$$\mathbf{T}(\mathbf{g}) = \mathbf{F} \text{diag}(\mathbf{F}^H \mathbf{g}) \mathbf{F}^H \triangleq \mathbf{F} \text{diag}(\hat{\mathbf{g}}) \mathbf{F}^H,$$

where $\mathbf{F} = \frac{1}{\sqrt{P}}[\boldsymbol{\omega}_0, \boldsymbol{\omega}_1, \dots, \boldsymbol{\omega}_{P-1}]$ is the DFT matrix with columns

$$\boldsymbol{\omega}_m = [e^{j2\pi \frac{m}{P} \cdot 0}, e^{j2\pi \frac{m}{P} \cdot 1}, \dots, e^{j2\pi \frac{m}{P} \cdot (P-1)}]^T.$$

We define corresponding diagonal matrices $\mathbf{W}_m = \text{diag}[\boldsymbol{\omega}_m]$, for $m = 0, 1, \dots, P-1$. Then the Gabor frame generated from \mathbf{g} is an $P \times P^2$ block matrix of the form

$$\Phi = [\mathbf{W}_0 \mathbf{T}(\mathbf{g}), \mathbf{W}_1 \mathbf{T}(\mathbf{g}), \dots, \mathbf{W}_{P-1} \mathbf{T}(\mathbf{g})]. \quad (2.26)$$

where each column has norm $\sqrt{P/M}$. If we apply the DFT to the Gabor frame Φ , to obtain $\hat{\Phi} = \mathbf{F}^H \Phi$, then the order of time-shift and frequency modulation is reversed; therefore $\hat{\Phi}$ is composed of circulant matrices with proper ordering of columns. We index the P^2 columns using $P \times P$ by setting $m = Pq + \ell$. The matrix Φ_ℓ is obtained by keeping all columns with $r = \ell \pmod{P}$, so Φ_ℓ can be written as $\Phi_\ell = \sqrt{P} \cdot \text{diag}(\mathbf{S}^\ell \mathbf{g}) \mathbf{F}$, where \mathbf{S} is the right-shift matrix by one, and $\hat{\Phi}_\ell = \mathbf{F} \Phi_\ell = \sqrt{P} \mathbf{T}(\mathbf{W}_\ell \hat{\mathbf{g}})$ is a circulant matrix. We use $[\Phi_1, \dots, \Phi_{P-1}]$ as the matrix \mathbf{A} .

At the receiver, a partial DFT is applied to the received symbol, so $\mathbf{I}_\Omega = \mathbf{F}_\Omega$ is a partial DFT matrix, and the resulted matrix $\mathbf{H} = \Phi_\Omega$ is a subsampled Gabor frame defined in (2.26), with unit-norm columns. The Gabor frame is known to satisfy the coherence property [23] and [25] shows that this is also true for the subsampled Gabor frames.

The maximum discrete delay τ this Gabor frame construction can support is $P-1$, where $\mathbf{W}_\ell \hat{\mathbf{g}}$, $\ell = 1, \dots, P$ can be assigned as signature sequences ($\tilde{\mathbf{a}}_n$'s) to a user, so the maximum number of total user should satisfy $N \leq P$. In general, if $\tau < P-1$, we can split Φ_ℓ into blocks to support multiple users, and send $\mathcal{T}_{d(\tau+1)} \mathbf{W}_\ell \hat{\mathbf{g}}$ as signature sequences for $d = 0, \dots, \lfloor P/\tau \rfloor$ and $\ell = 1, \dots, P$, so the maximum number of total user satisfies $N \leq P \lfloor P/\tau \rfloor$.

Remark: For simplicity we only consider a flat fading channel model so that there is a single non-zero entry in each block of the matrix \mathbf{A} , but we could easily extend to multi-path fast-fading channel models, where there are multiple non-zero entries in each block.

2.3.3 Numerical Examples

We now provide numerical experiments to validate the performance. Let the seed vector for the Gabor frame be either an Alltop sequence (which meets the Welch bound) with length $P = 127$, or a random uniform vector with length $P = 128$. The

channel gain and transmitted power is assumed to be known, and with $R_{mm} = 1$ for all $m = 1, \dots, N(\tau + 1)$.

The active users are selected first by uniformly choosing at random from 1 to P , and then, for each active user, the delay is chosen uniformly at random with the maximum chip delay as P . First, we fix the number of active users, namely $K = 2$ or $K = 5$, and apply the Matching Pursuit Decoder described in Algorithm 3 for noise-free case or noisy case where the AWGN noise is $\text{SNR} = 20\text{dB}$ per measurement. The partial DFT matrix is applied with randomly selected rows and the number of Monte Carlo runs is 10,000. Fig. 2.1 shows the MUD error rate with respect to the number of measurements. We attribute superior performance of the Gabor frame determined by the Alltop sequence to its optimal coherence, while for random Gabor frames the coherence properties are only satisfied in expectation. A different perspective is found in Fig. 2.2 which describes the phase transition for detection of K users using M measurements.

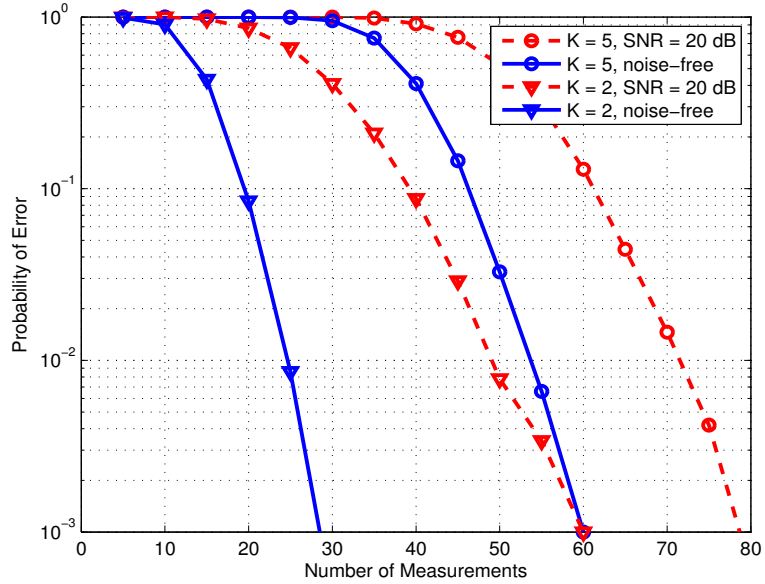
Finally, we consider a maximum delay that is relatively small, for example $\tau = 16$ when $P = 128$ for a random Gabor frame. We transmit the first sequence within the block of the circulant matrix, resulting in a total number of $P^2/\tau = 1024$ users, and Fig. 2.3 shows the MUD error rate with respect to the number of random measurements.

To sum up, we have provided a new architecture for compressive demodulation of mutually interfering signals in this section. The advantage over standard MUD is that, while only using a number of samples on the order of $K \log N(\tau + 1)$, we can process N users with a maximum discrete delay τ from a simple matching pursuit algorithm, compared to $N(\tau + 1)$ samples for standard MUD. The architecture also supports blind MUD by assigning multiple waveforms to a given user and transmitting information by the choice of waveform.

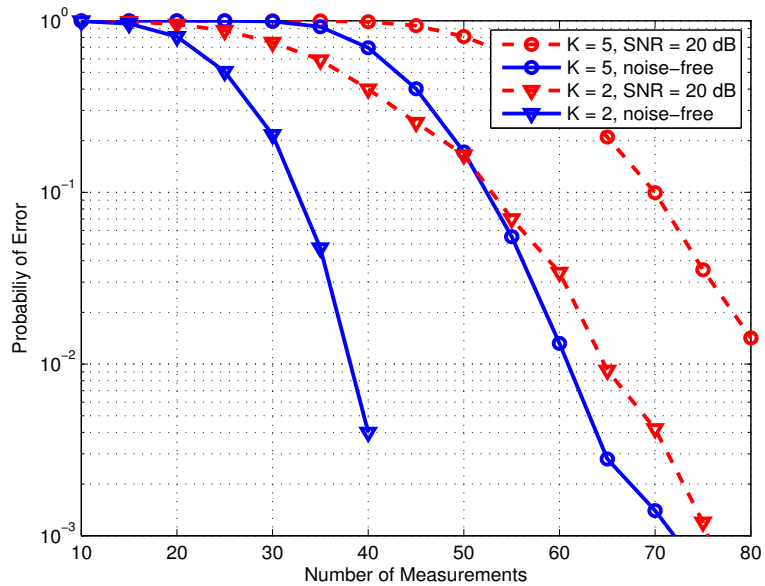
2.4 Diagnostic Wireless ECG Monitoring

For remote monitoring of ECG, it is extremely important to maintain the clinical integrity of the signals. Continuous monitoring of ECG is widely used in many clinical settings, including Intensive Care Units (ICUs), post-operative monitoring, emergency care and in ambulatory settings such as Holter monitoring. As interpretation of continuous ECG requires analysis of as many as 105 cardiac cycles per patient per day, there has been a need for tools to perform automated labeling and classification of ECG. American National Standard Institute (ANSI) and Association for the Advancement of Medical Instrumentation (AAMI) have established standards for automated tools such as EC57 [30] that are recognized by the FDA in United States. The tools for implementing these protocols, along with ECG databases and annotations by experts for normal and abnormal ECG beats are available in the public domain at [31], [32]. We refer to annotations by experts as ground truth.

Commercial ECG machines that provide automated labeling and classification benchmark their classification performance against the ground truth. While EC57 does not specify minimum performance requirements for metrics such as Sensitivity



(a) Alltop Gabor Frame



(b) Random Gabor Frame

Figure 2.1: Multi-user detection error rate with respect to the number of measurements using a Gabor frame generated from (a) an Alltop sequence with length $P = 127$, and (b) a random uniform vector with length $P = 128$ for different active users and SNR, where the maximum chip delay is P . The Alltop Gabor frame has better performance due to its optimal coherence properties.

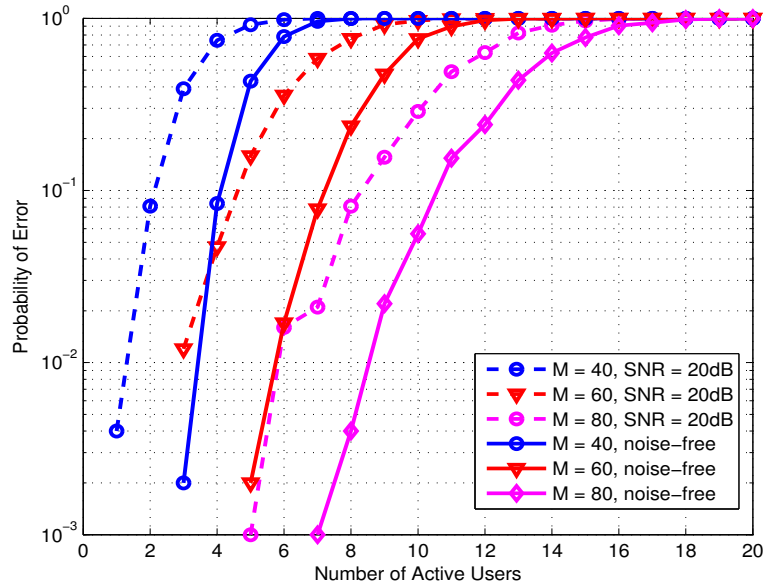


Figure 2.2: Multi-user detection error rate with respect to the number of active users using an Alltop Gabor frame with $P = 127$ for fixed number of measurements.

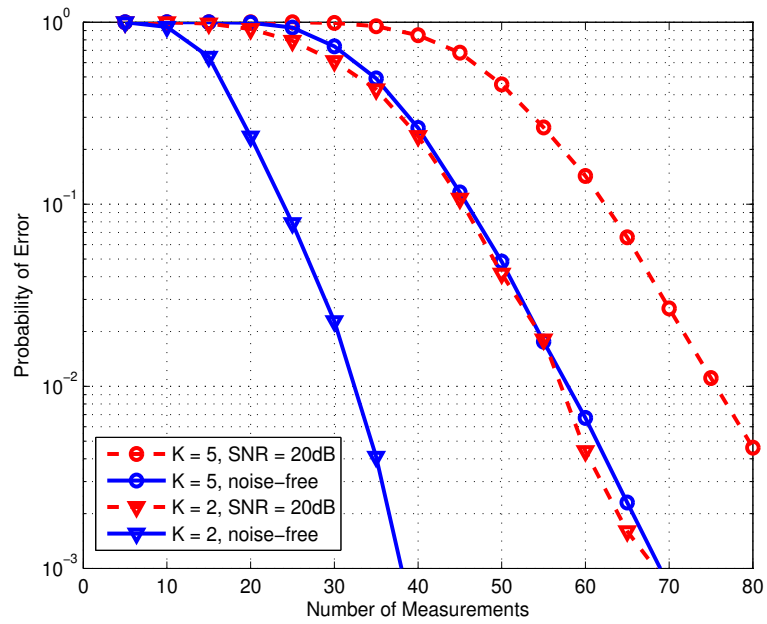


Figure 2.3: Multi-user detection error rate as a function of the number K of active users for a fixed number of measurements M .

(Se – percentage of true events detected) and Positive Predictivity (+P – percentage of detected events that are true), it mandates that such performance metrics with databases in [30] are disclosed. From a clinician’s perspective, a wireless continuous ECG monitoring system should provide diagnostic utility similar to that of a wired system in current standard of care, while enabling non-intrusive form factors for use in free living conditions. This means that there should be no statistically significant degradation in performance due to the wireless link. We observe that 2% or more degradation in classification accuracy from a wired baseline performance is clinically significant and about 1% Packet Loss Rate (PLR) can cause significant degradation in performance.

Packet losses occur in wireless networks due to fading, interference, congestion, system loading, etc. Popular choices for radios in Body Area Networks (BAN) such as Bluetooth and Zigbee operate in the crowded 2.4GHz band, along with IEEE 802.11. In [33], the authors investigated the interference of 802.11 traffic presented to ZigBee nodes in BAN and found 33% – 56% packet loss rate, depending upon the network setup. Another study [34] based on Zigbee reported packet losses as much as 50% in a clinical trial involving remote ECG monitoring. Note that important events like QRS-complex in ECG signals occur over a short period of time, thus packet losses can result in significant loss of clinically relevant data. In this section, we evaluate the proposed packet loss mitigation approach using compressive sensing with MIT-BIH and AHA databases using a state-of-the-art commercial ECG arrhythmia classification software and show that the performance does not degrade even at high packet loss rates [26].

2.4.1 CS Approaches at Sensor and Receiver

We first demonstrate the sparsity of ECG signals in the wavelet domain. Fig. 2.4 shows a sample of raw ECG signals in time domain and its wavelet transform using Daubechies-4 (D4) as mother wavelet. It is straightforward to validate the sparsity of ECG signals in the wavelet domain, from the fast decay of the ordered coefficients in amplitude in red. More formally, denote the ECG signal in a frame by $\mathbf{s} \in \mathbb{R}^N$, where N is the length of a frame, and the wavelet transform by the matrix $\mathbf{W} \in \mathbb{R}^{N \times N}$, then \mathbf{s} can be written as

$$\mathbf{s} = \mathbf{W}\mathbf{x}, \tag{2.27}$$

where \mathbf{x} is the sparse representation of \mathbf{s} in the wavelet domain.

Sensor Signal Processing: We project the the ECG signal in the time domain to a random space prior to transmission, by multiplying it with a sensing matrix $\mathbf{H}_s \in \mathbb{R}^{N \times N}$ whose entries are assumed to be i.i.d. Bernoulli entries from $\{0, 1\}$. Hence the encoded ECG signal $\mathbf{r} \in \mathbb{R}^N$ can be written as

$$\mathbf{r} = \mathbf{H}_s\mathbf{s}. \tag{2.28}$$

Notice there is no compression at the sensor side. The encoded ECG signal \mathbf{r} is then split into smaller packets and transmitted wirelessly.

In practice, each of the elements h_{ij} of the matrix \mathbf{H}_s is generated from a Linear Feedback Shift Register (LFSR) sequence and quantized into a one-bit value in $\{0, 1\}$.

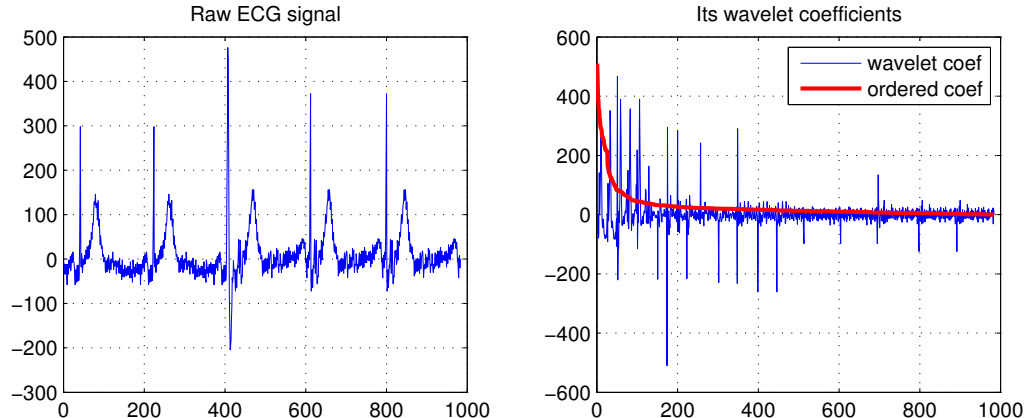


Figure 2.4: A sample of raw ECG signal in time domain and in the wavelet domain.

Thus, the matrix multiplication is implemented as additions only, corresponding to the state of the quantized LFSR output. It is also important to keep in mind bit-growth due to the CS encoding operation. Suppose that the incoming x_i are b_1 bits each. If the k th row of the matrix \mathbf{H}_s contains D ones, then the output r_k will require $b_2 = b_1 + \log_2(D)$ bits if full resolution is to be maintained. To reduce the memory required to store and the bandwidth required to transmit \mathbf{r} , one approach is to implement a sparse matrix where the number of ones per row is small. In this work, the input \mathbf{x} is 16 bits/sample and the memory budget for the output \mathbf{r} is 20 bits/sample. Therefore, we construct a matrix with a maximum of 16 ones per row to allow for 4 bits of expansion.

For a real-time implementation, we utilize a double randomization scheme where each column c_k of the matrix \mathbf{H}_s is a one-bit quantized output of a $\log_2 N$ LFSR sequence starting at a seed s_k . It is well known that a b -bit LFSR sequence is cyclic with periodicity $2^b - 1$ and is unique for a given starting seed s . Thus, by selecting an LFSR sequence with a maximum length of $N - 1$ and a quantization threshold, we are able to control the density of ones per row. The quantization threshold is selected based on the desired ones-density per row. In order to make the columns statistically independent, the master LFSR sequence that provides the quantized $\{0, 1\}$ values also point to a starting seed s_k for the k th column.

Receiver Signal Processing: Notice that the channel randomly drops packets, we could represent the effect of packet loss via a random partial identity matrix $\mathbf{H}_c \in \mathbb{R}^{L \times N}$, which is obtained by removing rows of the identity matrix corresponding to lost packets, and L is the number of received samples. Then the received ECG signal \mathbf{y} over a frame can be represented as

$$\mathbf{y} = \mathbf{H}_c \mathbf{r} + \mathbf{n} = \mathbf{H}_c \mathbf{H}_s \mathbf{s} + \mathbf{n} \triangleq \mathbf{H} \mathbf{s} + \mathbf{n}, \quad (2.29)$$

where \mathbf{n} denotes possible noise. Due to the randomness in \mathbf{H}_c and \mathbf{H}_s , the matrix $\mathbf{H} = \mathbf{H}_c \mathbf{H}_s$ satisfies the RIP with high probability. On the receiver side, we implemented a fast implementations of batch OMP algorithms [35]. Fig. 2.5 shows a

six-second segment of ECG recording from the record 203 of MIT-BIH database. The matrix \mathbf{H}_s is a 128×128 matrix and each ECG frame is transmitted in 8 packets. The middle row shows the data in the random space, with missing segments corresponding to packet losses. It can be seen from the reference (top row) and test (bottom row) annotations that there are two locations where an atrial premature beat (A) was mis-categorized as a normal sinus rhythm (N). All the remaining beats were correctly classified in this segment.

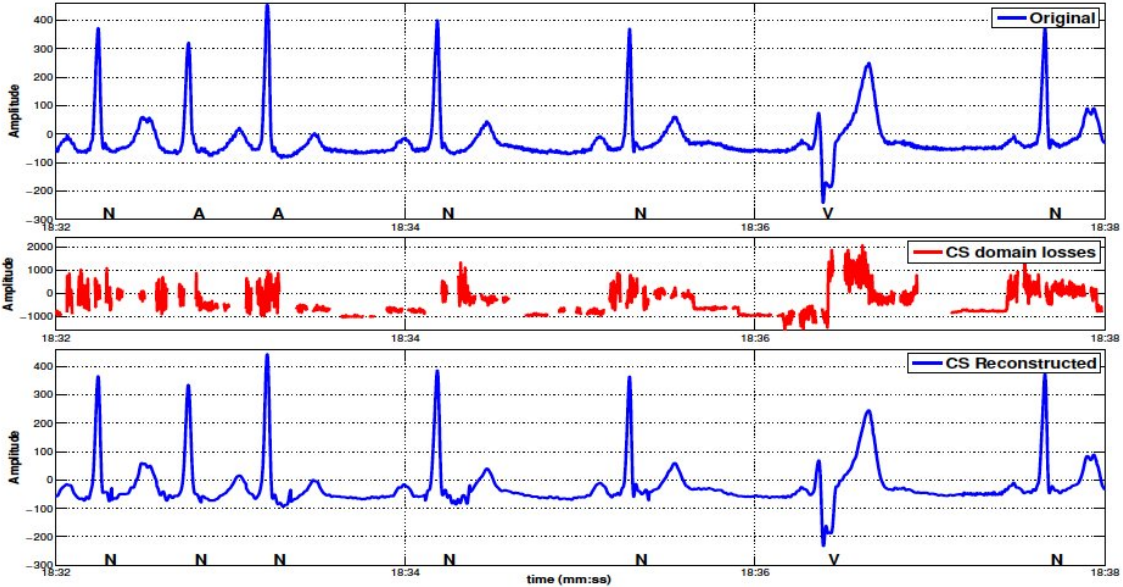


Figure 2.5: ECG Reconstruction using the proposed packet loss mitigation framework.

2.4.2 Validation for Diagnostic Wireless ECG

As described earlier, ANSI/AAMI specification EC57 [30] provides a framework for validating the performance of automated software tools that classify and label large amounts of ECG waveform data resulting from continuous monitoring. In this section, we apply current standard of care protocols in wired settings to wireless ECG monitoring. The databases we consider are MIT-BIH Arrhythmia Database (48 records of 30 minutes each) and AHA database for ventricular arrhythmia detectors (80 records of 35 minutes each). Overall, these databases contain a wide variety of cardiac rhythms comprising nearly 250,000 normal and 22,000 abnormal heart beats from multiple subjects. All of the records in the databases come with ground truth of annotations by cardiologists. Commercial ECG machines that provide automated labeling and classification benchmark their performance against this ground truth. The metrics for beat classification performance are Sensitivity (Se) and Positive Predictivity (+P),

defined as follows [30]:

$$\begin{aligned} \text{Se} &= \text{TP} = (\text{TP} + \text{FN}) \\ +\text{P} &= \text{TP} = (\text{TP} + \text{FP}) \end{aligned}$$

where,

- A correctly detected event is called a true positive (TP);
- An erroneously rejected (missed) event is called a false negative (FN);
- An erroneously detected non-event is called a false positive (FP);
- A correctly rejected non-event is called a true negative (TN).

We implemented arrhythmia analysis and beat classification using the Mortara algorithm [36]. We re-sampled each record at 500 Hz and scaled to 2.5 V/LSB to meet the specifications for ECG data as input to the Mortara arrhythmia analysis library. The arrhythmia analysis library was compiled into an executable to read the ECG records and output measurements including heart rate, ST values, QRS amplitudes, etc. along with beat and event classification such as normal sinus rhythm, pre-ventricular, ventricular fibrillation, asystole, bigeminy, pause, etc. The algorithm processes from 1 to 8 leads and can detect QRS complexes as long as at least one lead is valid. The output from Mortara is formatted such that it can be used directly with EC57 tools for comparison with the ground truth annotations by cardiologists provided in the databases. EC57 requires testing and disclosure of the algorithms sensitivity and positive predictivity along with Root Mean Square (RMS) heart rate error. The comparison may start after 5 minutes from the beginning of the record. For a beat to be correctly classified, the algorithm must identify the beat with correct classification within 150 ms of the actual event.

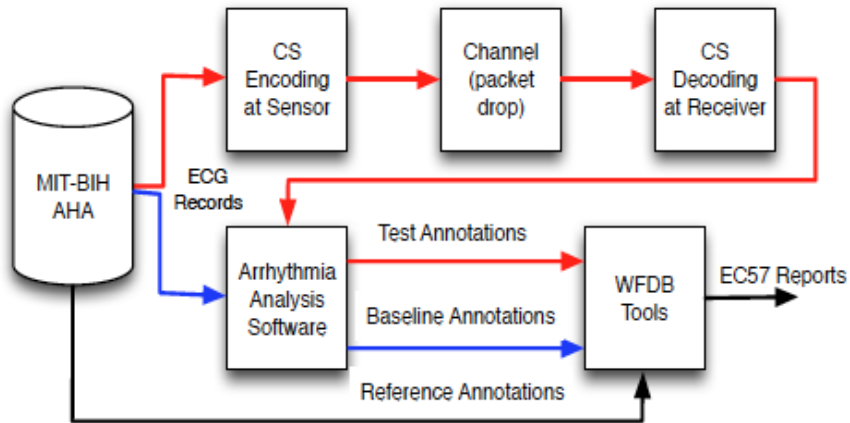


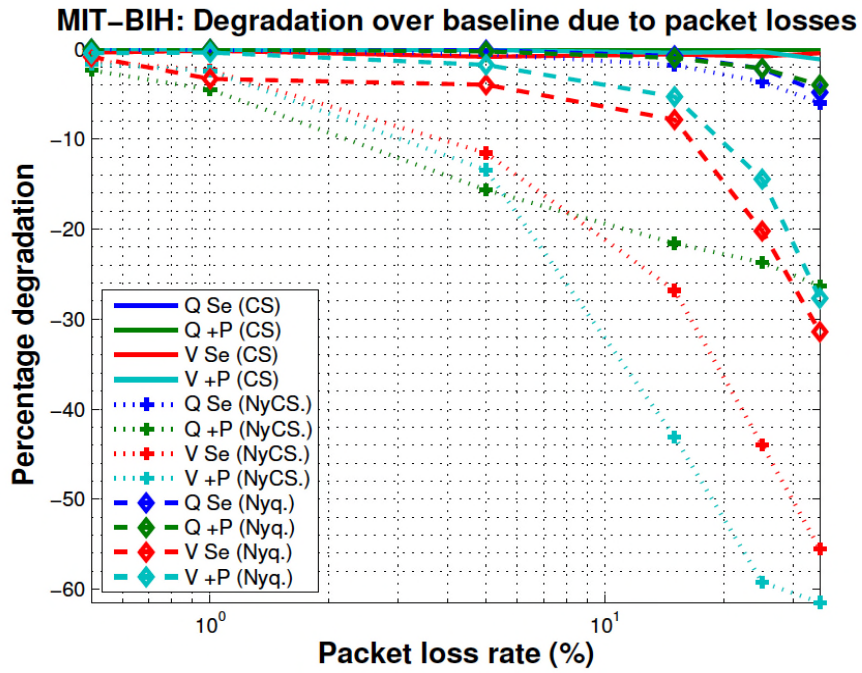
Figure 2.6: Experimental setup for ECG validation.

Fig. 2.6 depicts the experimental setup used to validate the packet loss mitigation proposed in this study. The black path labeled “Reference Annotations” represents

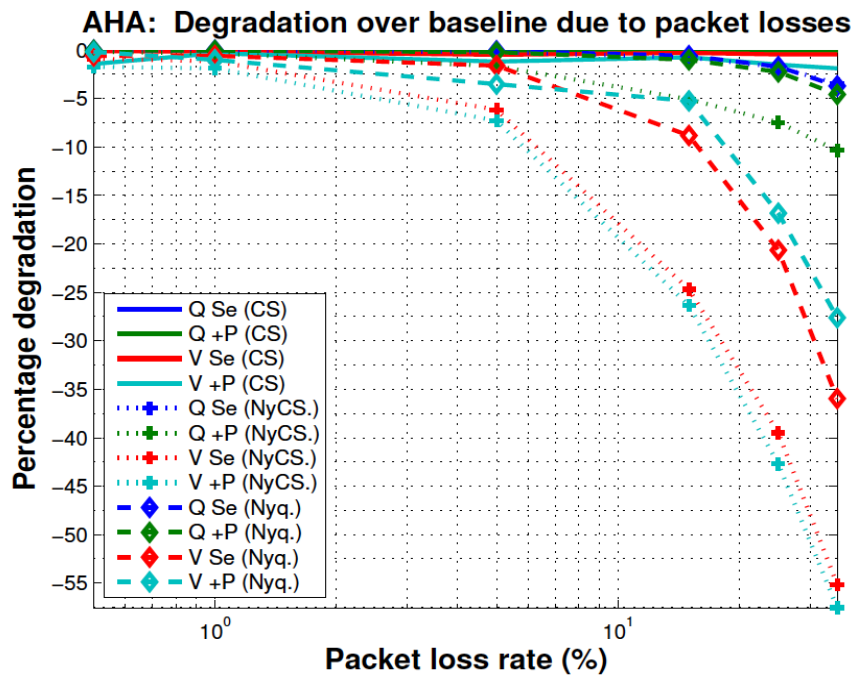
annotations of the ECG waveforms by cardiologists. The blue path labeled “Baseline Annotations” represents the current gold standard in a wireline setting. Each cardiac database was analyzed to provide a baseline performance measure to confirm that introduction of CS did not affect the system performance and to confirm that the system performance of the test bed used for this work provides results that match those from prior EC57 compliance tests. The red path labeled “Test Annotations” represents wireless case in this study. At the sensor, we used a sparse sensing matrix \mathbf{H}_s of size 128×128 as described in Section 2.4.1. Note that increasing the dimensions of measurement matrices provides better reconstruction accuracy, at the expense of increase in encoder complexity and additional latency. It is essential from low power perspective that the application layer is optimized for a given radio in BAN. We experimented with 32 packets per frame, corresponding to 4 ECG samples per packet. A bursty channel-error model was used to drop packets at loss rates of 0.5%, 1%, 5%, 15%, 25% and 35%. The received data was reconstructed into the Nyquist (time) domain and provided to the Mortara arrhythmia analysis algorithm to generate annotations.

We evaluate the proposed packet loss mitigation approach with MIT-BIH and AHA databases using a state-of-the-art commercial ECG arrhythmia classification software and show that the performance does not degrade even at high packet loss rates. Fig. 2.7 (a) and (b) present degradation in beat classification as a function of packet loss rate for MIT-BIH and AHA databases, respectively. The solid lines represent the CS based packet loss mitigation approach (CS), the dashed lines represent the case with no random projections at the sensor but with sparse reconstruction at the receiver (NyCS) and the dotted lines represent the Nyquist domain data (NyQ), respectively. The legends Q and V correspond to normal QRS sinus rhythms and abnormal VEB rhythms, respectively. The legends Se and +P correspond to sensitivity and positive predictivity, respectively.

We observe that 2% or more degradation in classification accuracy from a wired baseline performance is clinically significant. From Fig. 2.7, it can be seen that performance degrades monotonically for both NyCS and NyQ, compared with CS. This is particularly true for sensitivity and positive predictivity of abnormal rhythms (V), compared with normal sinus rhythms (Q). The positive predictivity degradation is also severe, suggesting more false positives as packet loss rate increases. Without some method of packet loss mitigation, a 1% packet loss rate can cause clinically significant degradation. While the high packet loss conditions studied here are corner cases, we believe that typical loss rates of around 5% are typical for BAN modems in the crowded 2.4 GHz band. As a reference, packet loss rates of 1 – 3% are commonly used to evaluate voice quality in 3G standards.



(a) MIT-BIH database



(b) AHA database

Figure 2.7: The diagnostic degradation in (a) MIT-BIH database and (b) AHA database.

Chapter 3

Sensitivity of Compressive Sensing to Basis Mismatch

3.1 Motivation

In a great number of fields of engineering and applied science the problem confronting the designer is to *invert an image*, acquired from a sensor suite, for the underlying field that produced the image. And typically the desired resolution for the underlying field exceeds the temporal or spatial resolution of the image itself. Here we give *image* its most general meaning to encompass a times series, a space series, a space-time series, a 2-D image, and so on. Similarly, we give *field* its most general meaning to encompass complex-exponential modes, radiating modes, coded modulations, multipath components, and the like. Certainly this interpretation includes the problem of identifying field elements from electromagnetic and acoustic images, multipath components in wireless communication, radiating sources in radar and sonar, and light sources in optical imaging and spectrometry.

Broadly speaking there are two main (classical) principles for inverting the kinds of images that are measured in speech, communication, radar, sonar, and optics. The first principle is one of matched filtering, wherein a sequence of test images is matched to the measured image. The test images are generated by scanning a prototype image (e.g., a waveform or a steering vector) through frequency, wavenumber, doppler, and/or delay. In time series analysis, this amounts to classical spectrum analysis to identify the frequency modes, and the corresponding mode amplitudes or powers, of the signal [37], [38]. In phased-array processing, it amounts to spectrum analysis in frequency and wavenumber to identify the frequency-wavenumber coordinates of source radiations impinging on the array [39]–[42]. In Space-Time Adaptive Processing (STAP) for radar and sonar, it amounts to spectrum analysis in delay, frequency, and wavenumber to reconstruct the radar/sonar field [43], [44]. The second principle is one of parameter estimation in a separable linear model, wherein a sparse *modal* representation for the field is posited and estimates of linear parameters (complex amplitudes of modes) and nonlinear mode parameters (frequency, wavenumber, delay, and/or doppler) are extracted, usually based on maximum likelihood, or some

variation on linear prediction, using ℓ_2 minimization (see, e.g., [38], [45], [46]). There is a comprehensive literature in electrical engineering, physics, and chemistry on the performance and limitations of these two classical principles (see, e.g., [37]–[49]). One important limitation is that any subsampling of the measured image has consequences for resolution (or bias) and for variability (or variance).

The recent advent of compressive sensing theory has revolutionized our view of imaging, as it demonstrates that subsampling has manageable consequences for image inversion, provided that the image is sparse in an *a priori known* basis (see, e.g., [50]–[61]). For imaging problems in spectrum analysis (estimating complex exponential modes), and passive and active radar/sonar (estimating Doppler and angle of arrival), this basis is usually *taken* to be a Fourier basis (actually a DFT basis) constructed for resolution of $2\pi/N$, with N a window length, array length, or pulse-to-pulse processing length. Several articles (see, e.g., [62]–[68]) consider the use of compressed sensing theory for discrete radar/sonar imaging, and sensor array processing, when the targets are taken to be on a regular grid in delay, Doppler, and wavenumber, and study this theory as a new high resolution imaging principle. But no matter how large the size N of the grid is, the actual field will not place its sources on the center of the grid points $\{2\pi n/N\}$ in frequency or wavenumber, or on the center of the grid points in delay-Doppler-wavenumber. This means the image is actually *not sparse* in the DFT basis or the basis defined by the grid. In fact any target that lies between two cells of a discretely-resolved range-doppler plane or frequency-wavenumber plane will spill non-zero values into all cells, with the amplitude of the spillage following a Dirichlet kernel, decaying as $1/f$, where f is frequency or wavenumber. This spillage turns a sparse representation into an incompressible one in the DFT basis. Grid misalignment problems also arise in many other applications including channel estimation [69, 70]. These observations raise the following question: What is the sensitivity of CS for image inversion to mismatch between the assumed basis for sparsity and the actual basis in which the image is sparse? We aim to answer this question and examine its consequences for problems in spectrum analysis, beamforming, modal analysis, and radar/sonar imaging.

In order to frame our question more precisely, let us begin with two models for a measured image $\mathbf{s} \in \mathbb{C}^N$. In the mathematical model to be *assumed* in the CS procedure, the image is composed as $\mathbf{s} = \Psi_0 \mathbf{x}$, where the basis $\Psi_0 \in \mathbb{C}^{N \times N}$ is known, and is typically a gridded imaging matrix (e.g., the N -point DFT matrix), and $\mathbf{x} \in \mathbb{C}^N$ is a sparse or compressible vector of field parameters that compose the image as a linear combination of columns of Ψ_0 . But, as a matter of fact, the image \mathbf{s} is composed by the physics as $\mathbf{s} = \Psi_1 \boldsymbol{\theta}$, where the basis $\Psi_1 \in \mathbb{C}^{N \times N}$ is determined by a point spread function, a Green’s function, or an impulse response, and the field parameter vector $\boldsymbol{\theta}$ is sparse. Typically Ψ_1 is determined by frequency, wavenumber, delay, and/or doppler parameters that are *unknown* a priori. More importantly, these parameters do *not* lie exactly on the gridding points of Ψ_0 , e.g., a DFT matrix or an identity matrix. So $\Psi_0 \neq \Psi_1$. We call this *basis mismatch*, and note that it is present in all imaging problems, no matter how large N is, or equivalently no matter how fine-grained the gridding procedure is.

Each of the hypothesized models for the image \mathbf{s} may be inverted for its field parameters:

$$\mathbf{x} = \Psi_0^{-1}\mathbf{s} \quad \text{and} \quad \boldsymbol{\theta} = \Psi_1^{-1}\mathbf{s}. \quad (3.1)$$

These inversions determine the coordinate transformation

$$\mathbf{x} = \Psi\boldsymbol{\theta} \quad \text{and} \quad \boldsymbol{\theta} = \Psi^{-1}\mathbf{x} \quad (3.2)$$

where $\Psi = \Psi_0^{-1}\Psi_1 \in \mathbb{C}^{N \times N}$.

If \mathbf{s} is sparse in Ψ_1 , then the field parameters $\boldsymbol{\theta}$ will be sparse in the identity basis, denoted by \mathbf{I} . The field parameters $\mathbf{x} = \Psi\boldsymbol{\theta}$ in the model $\mathbf{s} = \Psi_0\mathbf{x}$ will be sparse in the Ψ basis, but not in the identity basis. So the question is, “what is the consequence of assuming that \mathbf{x} is sparse in \mathbf{I} , when in fact it is only sparse in an *unknown* basis Ψ , which is determined by the mismatch between Ψ_0 and Ψ_1 ?”

We answer this question by deriving bounds on the ℓ_1 -norm (and also the ℓ_2 -norm) of the error in approximating the presumably sparse parameter vector \mathbf{x} from its CS measurements. We start by analyzing the effect of basis mismatch on the best k -term approximation error $\mathbf{x} - \mathbf{x}_k$, which is central to arguing for the accuracy of the basis pursuit solution. Here, we derive bounds for $\|\mathbf{x} - \mathbf{x}_k\|_1$ in terms of the mismatch level between Ψ and \mathbf{I} and the image dimension N . More specifically, we find a tight upper bound for $\|\mathbf{x} - \mathbf{x}_k\|_1$ and show that this bound grows as $(N - k)\beta\|\boldsymbol{\theta}\|_q$, where β upper bounds the degree of mismatch between the rows of Ψ and \mathbf{I} , and $q \geq 1$. We also establish a tight lower bound on the worst-case ℓ_1 error in the best k -term approximation $\mathbf{x} - \mathbf{x}_k$ and show that it grows as $(N - k)\eta\|\boldsymbol{\theta}\|_q$, with η representing the lower bound on the degree of mismatch between the rows of Ψ and \mathbf{I} . Subsequently, we derive bounds on the image inversion error in estimating the parameter vector \mathbf{x} and the image reconstruction error in estimating the image $\mathbf{s} = \Psi_0\mathbf{x}$ using basis pursuit. All these bounds exhibit similar growth behavior as those for the best k -term approximation error, but they are not tight. Our analysis shows that, in the presence of basis mismatch, exact or near-exact (within noise levels) recovery cannot be guaranteed from bounds that involve best k -term approximation errors and suggests that the basis pursuit recovery may suffer large errors. Our numerical examples demonstrate a considerable performance degradation in recovering \mathbf{x} from CS measurements, when the assumed basis for sparsity is a DFT basis but the actual sparsity basis does not align with the DFT basis. The inaccuracy in field reconstruction persists even when the number of CS measurements is increased to the full image dimension. Comparisons show that classical image inversion approaches, such as reduced rank linear prediction, can provide more reliable reconstructions of the field than basis pursuit with a similar number of measurements in the presence of basis mismatch.

We note that although we present our results in the context of basis pursuit, our mathematical analysis is applicable to any sparse recovery principle that relies on the accuracy of best k -term approximations for performance guarantees. These include greedy recovery algorithms, such as Regularized Orthogonal Matching Pursuit (ROMP) [71] and Compressive Sampling Matching Pursuit (CoSaMP) [72]. The

implication of our results is that, at least for high resolution spectrum analysis, Direction-Of-Arrival estimation, or delay-Doppler imaging, and radar/sonar imaging, where the problem is to identify a small number of modal parameters, extra care may be needed to account for the effects of basis mismatch.

3.2 Problem Formulation

There are two steps in CS, namely compressed recording and inversion (or recovery) for parameters (see, e.g., [50]–[53]). In recording, we make linear measurements of \mathbf{s} as in (3.1), with possible additive noise \mathbf{b} , so the low-dimensional observation \mathbf{y} is *assumed* to be

$$\mathbf{y} = \Phi \mathbf{s} + \mathbf{b} = (\Phi \Psi_0) \mathbf{x} + \mathbf{b} \triangleq \mathbf{A} \mathbf{x} + \mathbf{b} \quad (3.3)$$

where $\Phi \in \mathbb{C}^{M \times N}$ is the CS matrix (typically a matrix with i.i.d. Gaussian or i.i.d. Bernoulli entries), M is the number of measurements, and Ψ_0 is the assumed basis that sparsely composes \mathbf{s} as $\mathbf{s} = \Psi_0 \mathbf{x}$. We define $\mathbf{A} = \Phi \Psi_0 \in \mathbb{C}^{M \times N}$ as the measurement matrix. Without loss of generality, we need only deal with \mathbf{A} and Ψ in the following discussions.

If the measurement matrix \mathbf{A} satisfies the *Restricted Isometry Property* (RIP) with $\delta_{2k}^{\mathbf{A}} < \sqrt{2} - 1$ [11, 14], one can seek the sparsest solution using Basis Pursuit (BP):

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad s.t. \quad \mathbf{y} = \mathbf{A} \mathbf{x}. \quad (3.4)$$

In the noisy case, the problem is modified as:

$$\mathbf{x}^* = \arg \min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad s.t. \quad \|\mathbf{y} - \mathbf{A} \mathbf{x}\|_2 \leq \epsilon \quad (3.5)$$

where $\|\mathbf{b}\|_2 \leq \epsilon$ is the bounded noise. The solution \mathbf{x}^* to (3.4) approximates \mathbf{x} as

$$\|\mathbf{x} - \mathbf{x}^*\|_1 \leq C_0 \|\mathbf{x} - \mathbf{x}_k\|_1 \quad \text{and} \quad \|\mathbf{x} - \mathbf{x}^*\|_2 \leq C_0 k^{-1/2} \|\mathbf{x} - \mathbf{x}_k\|_1, \quad (3.6)$$

where \mathbf{x}_k is the best k -term approximation of \mathbf{x} . For the noisy case (3.5), we have [15, 14]

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C_0 k^{-1/2} \|\mathbf{x} - \mathbf{x}_k\|_1 + C_1 \epsilon. \quad (3.7)$$

The constants C_0 and C_1 are given in (2.6).

In the matched case where the hypothesized basis Ψ_0 coincides with the actual basis Ψ_1 , the mismatched basis $\Psi = \Psi_0^{-1} \Psi_1$ reduces to \mathbf{I} and $\mathbf{x} = \boldsymbol{\theta}$ is sparse in the identity basis \mathbf{I} . Then, for k -sparse $\boldsymbol{\theta}$ the bound $\|\mathbf{x} - \mathbf{x}_k\|_1$ is zero in (3.6) and the solution \mathbf{x}^* is an exact recovery of \mathbf{x} in the noise-free case.

However, in the mismatched case where $\Psi_0 \neq \Psi_1$, $\mathbf{x} = \Psi \boldsymbol{\theta}$ is actually sparse in the Ψ basis, rather than the \mathbf{I} basis. The question is, “what is the consequence of minimizing $\|\mathbf{x}\|_1$ under the constraint $\mathbf{y} = \mathbf{A} \mathbf{x}$, when in fact the correct problem is to minimize $\|\boldsymbol{\theta}\|_1$ under the constraint $\mathbf{y} = \mathbf{A} \Psi \boldsymbol{\theta} = \Phi \Psi_1 \boldsymbol{\theta}$?”

Remark: ROMP [71] and CoSaMP [72] have similar universal performance bounds as BP. For example, under the RIP of the measurement matrix, ROMP can approxi-

mately recover any compressible signal from noisy observations. Given noisy observations $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b}$ with $\|\mathbf{b}\|_2 \leq \epsilon$, ROMP produces a $2k$ -sparse signal approximation \mathbf{x}^* that satisfies the bound

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C\sqrt{\log 2k} \left(\frac{1}{\sqrt{k}} \|\mathbf{x} - \mathbf{x}_k\|_1 + \epsilon \right). \quad (3.8)$$

where $C = 160$ [71]. Similarly, for a given precision parameter ζ , CoSaMP produces a $2k$ -sparse signal approximation \mathbf{x}^* that satisfies

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C \max \left\{ \zeta, \frac{1}{\sqrt{2k}} \|\mathbf{x} - \mathbf{x}_k\|_1 + \epsilon \right\}. \quad (3.9)$$

where $C = 3.42$ [72]. Therefore, our analysis of the best k -term approximation error is also relevant to these algorithms.

3.3 Main Results

3.3.1 Degradation of the Best k -Term Approximation

Let us express the mismatched basis $\mathbf{\Psi} = \mathbf{\Psi}_0^{-1}\mathbf{\Psi}_1 \in \mathbb{C}^{N \times N}$ as $\mathbf{\Psi} = \mathbf{I} + \mathbf{E}$, where \mathbf{E} is a perturbation matrix with respect to the identity basis and $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$. Let \mathbf{x}_k and $\boldsymbol{\theta}_k$ denote the best k -term approximations to \mathbf{x} and $\boldsymbol{\theta}$ respectively. We have the following theorem.

Theorem 3.3.1 (Best k -term approximation error). *Let $\mathbf{\Psi} = \mathbf{I} + \mathbf{E}$ and $\mathbf{x} = \mathbf{\Psi}\boldsymbol{\theta}$. Let $1 \leq p, q \leq \infty$ and $1/p + 1/q = 1$. If the rows $\mathbf{e}_m^T \in \mathbb{C}^{1 \times N}$ of \mathbf{E} are bounded as $\|\mathbf{e}_m\|_p \leq \beta$ for $1 \leq m \leq N$, then*

$$\left| \|\mathbf{x} - \mathbf{x}_k\|_1 - \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 \right| \leq (N - k)\beta\|\boldsymbol{\theta}\|_q. \quad (3.10)$$

The bound is achieved when the entries of \mathbf{E} satisfy

$$e_{mn} = \pm\beta \cdot e^{j(\arg(\theta_m) - \arg(\theta_n))} \cdot (|\theta_n|/\|\boldsymbol{\theta}\|_q)^{q/p} \quad (3.11)$$

for $n \in T_{\boldsymbol{\theta}}$ and $1 \leq m \leq N$.

Proof. See Appendix 7.2.1. □

When $\boldsymbol{\theta}$ is k -sparse in \mathbf{I} , i.e., $\boldsymbol{\theta} = \boldsymbol{\theta}_k$, then (3.10) reduces to

$$\|\mathbf{x} - \mathbf{x}_k\|_1 \leq (N - k)\beta\|\boldsymbol{\theta}\|_q, \quad (3.12)$$

which shows that the upper bound on the ℓ_1 error $\|\mathbf{x} - \mathbf{x}_k\|_1$ is linearly increasing in N , β and $\|\boldsymbol{\theta}\|_q$, and linearly decreasing in k .

Remark: When $p = \infty$ and $q = 1$, the inequality (3.10) corresponds to the case where the entries themselves of the perturbation matrix \mathbf{E} are upper bounded, that

is, when $|e_{mn}| \leq \beta$ for $n \in T_\theta$. When $p = q = 2$, the inequality (3.10) corresponds to the case where the row 2-norms of \mathbf{E} are bounded as $\|\mathbf{e}_m\|_2 \leq \beta$.

Remark: Theorem 3.3.1 can be easily generalized to have ℓ_2 norms on the left-hand-side of (3.10), i.e.,

$$\|\|\mathbf{x} - \mathbf{x}_k\|_2 - \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_2\| \leq (N - k)^{1/2} \beta \|\boldsymbol{\theta}\|_q. \quad (3.13)$$

This follows by taking similar steps as in the proof of Theorem 3.3.1, but using the Minkowski inequality instead of the triangle inequality.

Corollary 3.3.1 (Normalized ℓ_1 error). *Let $\boldsymbol{\Psi} = \mathbf{I} + \mathbf{E}$ and $\mathbf{x} = \boldsymbol{\Psi}\boldsymbol{\theta}$. Let $1 \leq p, q \leq \infty$ and $1/p + 1/q = 1$. If the rows $\mathbf{e}_m^T \in \mathbb{C}^{1 \times N}$ of \mathbf{E} are bounded as $\|\mathbf{e}_m\|_p \leq \beta$ for $1 \leq m \leq N$, then the ℓ_1 norm of \mathbf{x} is bounded as*

$$(1 - N\beta)\|\boldsymbol{\theta}\|_1 \leq \|\mathbf{x}\|_1 \leq (1 + N\beta)\|\boldsymbol{\theta}\|_1 \quad (3.14)$$

and the normalized ℓ_1 error for approximating \mathbf{x} by \mathbf{x}_k is bounded as

$$\frac{1}{1 + N\beta} \frac{\|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1}{\|\boldsymbol{\theta}\|_1} - \frac{(N - k)\beta}{1 + N\beta} \leq \frac{\|\mathbf{x} - \mathbf{x}_k\|_1}{\|\mathbf{x}\|_1} \leq \frac{1}{1 - N\beta} \frac{\|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1}{\|\boldsymbol{\theta}\|_1} + \frac{(N - k)\beta}{1 - N\beta} \quad (3.15)$$

where the upper bound in (3.15) is valid if $N\beta < 1$.

Proof. The inequalities in (3.14) follow from Theorem 3.3.1 by setting $k = 0$ and $q = 1$, and the bound follows by combining (3.10) and (3.14). \square

Remark: In Theorem 3.3.1, we have characterized the mismatch between the assumed basis $\boldsymbol{\Psi}_0$ and the actual basis $\boldsymbol{\Psi}_1$ by considering bounds on the p -norm of the rows of perturbation matrix $\mathbf{E} = \boldsymbol{\Psi}_0^{-1}(\boldsymbol{\Psi}_1 - \boldsymbol{\Psi}_0)$ that captures the deviation of $\boldsymbol{\Psi} = \boldsymbol{\Psi}_0^{-1}\boldsymbol{\Psi}_1$ from the identity matrix \mathbf{I} . But we may think of the incoherence $\mu(\boldsymbol{\Psi}_0, \boldsymbol{\Psi}_1)$ between $\boldsymbol{\Psi}_0$ and $\boldsymbol{\Psi}_1$ as another way of characterizing the degree of mismatch between $\boldsymbol{\Psi}_0$ and $\boldsymbol{\Psi}_1$. And in fact $\mu(\boldsymbol{\Psi}_0, \boldsymbol{\Psi}_1)$ and the ‘‘max norm’’ of $\|\mathbf{E}\|$ are related, as we now show. Let $\boldsymbol{\psi}_i(j)$ denote the j th column of the basis $\boldsymbol{\Psi}_i$, $i = 0, 1$. The incoherence between the bases $\boldsymbol{\Psi}_0$ and $\boldsymbol{\Psi}_1$ is defined as

$$\mu(\boldsymbol{\Psi}_0, \boldsymbol{\Psi}_1) = \max_{i,j} |\langle \boldsymbol{\psi}_0(i), \boldsymbol{\psi}_1(j) \rangle|. \quad (3.16)$$

Let $\|\mathbf{M}\|_{\max} = \max_{i,j} |M_{i,j}|$ denote the max-norm of a matrix \mathbf{M} . Then, we can express $\mu(\boldsymbol{\Psi}_0, \boldsymbol{\Psi}_1)$ as

$$\mu(\boldsymbol{\Psi}_0, \boldsymbol{\Psi}_1) = \|\boldsymbol{\Psi}_0^H \boldsymbol{\Psi}_1\|_{\max} = \|\boldsymbol{\Psi}_0^H \boldsymbol{\Psi}_0 \boldsymbol{\Psi}_0^{-1} \boldsymbol{\Psi}_1\|_{\max} \leq \|\boldsymbol{\Psi}_0^H \boldsymbol{\Psi}_0\|_{\max} \|\boldsymbol{\Psi}\|_{\max} \quad (3.17)$$

where $\boldsymbol{\Psi} = \boldsymbol{\Psi}_0^{-1}\boldsymbol{\Psi}_1$. When $\boldsymbol{\Psi}_0$ is unitary, we have $\mu(\boldsymbol{\Psi}_0, \boldsymbol{\Psi}_1) = \|\boldsymbol{\Psi}\|_{\max}$. Substituting $\boldsymbol{\Psi} = \mathbf{I} + \mathbf{E}$ and applying the triangle inequality gives

$$\mu(\boldsymbol{\Psi}_0, \boldsymbol{\Psi}_1) \leq \|\boldsymbol{\Psi}_0^H \boldsymbol{\Psi}_0\|_{\max} (1 + \|\mathbf{E}\|_{\max}). \quad (3.18)$$

Theorem 3.3.1 establishes an upper bound for the ℓ_1 error $\|\mathbf{x} - \mathbf{x}_k\|_1$ in the best k -term approximation of \mathbf{x} for the case where the degree of mismatch is upper bounded. The following theorem considers the case where the level of mismatch between the bases Ψ_0 and Ψ_1 is lower bounded and establishes a lower bound for the worst-case ℓ_1 error $\|\mathbf{x} - \mathbf{x}_k\|_1$.

Theorem 3.3.2 (Worst-case best k -term approximation error). *Let $\Psi = \mathbf{I} + \mathbf{E}$ and $\mathbf{x} = \Psi\boldsymbol{\theta}$. Let $1 \leq p, q \leq \infty$ and $1/p + 1/q = 1$. If the rows of \mathbf{E} are lower bounded as $\|\mathbf{e}_m\|_p \geq \eta$, then*

$$\max_{\mathbf{E}: \|\mathbf{e}_m\|_p \geq \eta} \|\mathbf{x} - \mathbf{x}_k\|_1 \geq \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 + (N - k)\eta\|\boldsymbol{\theta}\|_q. \quad (3.19)$$

Proof. See Appendix 7.2.2. □

Theorem 3.3.1 shows that the ℓ_1 norm $\|\mathbf{x} - \mathbf{x}_k\|_1$ of the best- k term approximation error will be no worse than the upper bound deduced from (3.10) if the mismatch level between Ψ_0 and Ψ_1 (measured by $\|\mathbf{e}_m\|_p$) is upper bounded by β . In contrast, Theorem 2 shows that if the mismatch level is lower bounded by η then there exists a mismatch scenario where $\|\mathbf{x} - \mathbf{x}_k\|_1$ is no better than the bound in (3.19). Both bounds grow linearly with the grid size N and the mismatch levels. Fig. 3.1 illustrates the interplay between Theorems 3.3.1 and 3.3.2 when $\boldsymbol{\theta}$ is k -sparse, i.e., when $\|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 = 0$. The outer ball has radius $A(\beta)\|\boldsymbol{\theta}\|_q$, with $A(\beta) = (N - k)\beta$ and $q \geq 1$, and corresponds to the upper bound in Theorem 3.3.1. The inner ball, with radius $A(\eta)\|\boldsymbol{\theta}\|_q$, corresponds to Theorem 2. The figure illustrates that there always exists a mismatched basis $\Psi = \mathbf{I} + \mathbf{E}$, with mismatch level $\eta \leq \|\mathbf{e}_m\|_p \leq \beta$, for which the best k -term approximation \mathbf{x}_k lies between the two balls, away from zero, and thus no guarantee can be provided for the performance of basis pursuit.

3.3.2 Confidence Bounds for Image Inversion

We now derive bounds for the ℓ_1 and ℓ_2 norms of the image inversion error $\mathbf{x} - \mathbf{x}^*$. Here we invert the image $\mathbf{s} = \Psi_0\mathbf{x}$ for its field parameter vector \mathbf{x} , using basis pursuit, under the assumption that \mathbf{s} has a sparse representation in Ψ_0 , when in fact this representation is non-sparse or incompressible. These bounds can then be used to find bounds on the ℓ_1 and ℓ_2 norms of the image reconstruction error $\mathbf{s} - \mathbf{s}^*$, which speaks to how well the image \mathbf{s} can be recomposed in the basis Ψ_0 using the mismatched basis pursuit approximation \mathbf{x}^* .

Theorem 3.3.3 (Image inversion error). *Let \mathbf{A} be fixed and satisfy $\delta_{2k}^{\mathbf{A}} < \sqrt{2} - 1$ at the assumed sparsity level k . Let $1 \leq p, q \leq \infty$ and $1/p + 1/q = 1$. If the rows of \mathbf{E} satisfy $\|\mathbf{e}_m\|_p \leq \beta$, then the image inversion error $\|\mathbf{x} - \mathbf{x}^*\|_1$ is bounded as*

$$\|\mathbf{x} - \mathbf{x}^*\|_1 \leq C_0\|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 + C_0(N - k)\beta\|\boldsymbol{\theta}\|_q. \quad (3.20)$$

For noisy recovery (3.5), with $\|\mathbf{b}\|_2 \leq \epsilon$, we have

$$\|\mathbf{x} - \mathbf{x}^*\|_2 \leq C_0k^{-1/2}\|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 + C_0(N - k)k^{-1/2}\beta\|\boldsymbol{\theta}\|_q + C_1\epsilon, \quad (3.21)$$

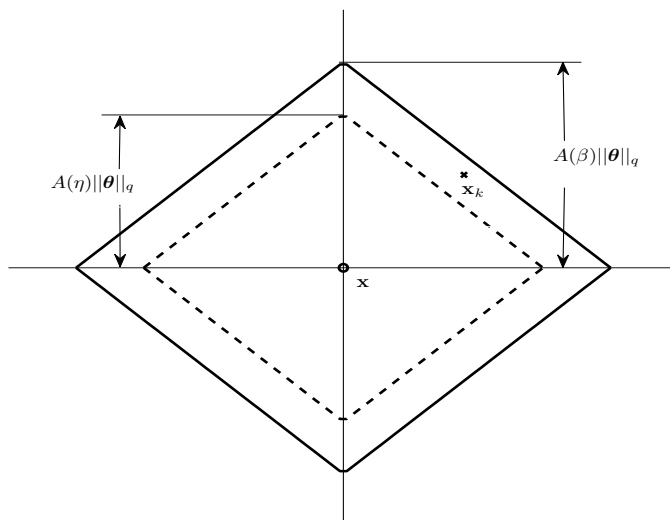


Figure 3.1: Confidence ℓ_1 balls for \mathbf{x}_k when $\eta \leq \|\mathbf{e}_m\|_p \leq \beta$ and $\boldsymbol{\theta}$ is k -sparse.

where C_0 and C_1 are given in (2.6).

Proof. This follows easily by combining Theorem 3.3.1 and the performance bounds of basis pursuit. \square

Theorem 3.3.3 is a direct consequence of Theorem 3.3.1 and therefore the discussion about the connection between the max-norm of the perturbation matrix \mathbf{E} and the incoherence measure $\mu(\boldsymbol{\Psi}_0, \boldsymbol{\Psi}_1)$ also applies here. It shows that the upper bound on $\|\mathbf{x} - \mathbf{x}^*\|_1$ is linearly increasing in the image dimension N and in the mismatch level β , and linearly decreasing in k . The upper bound on the ℓ_1 error in image reconstruction has similar growth behavior since

$$\|\mathbf{s} - \mathbf{s}^*\|_2 = \|\boldsymbol{\Psi}_0 \mathbf{x} - \boldsymbol{\Psi}_0 \mathbf{x}^*\|_2 \leq \|\boldsymbol{\Psi}_0\|_2 \|\mathbf{x} - \mathbf{x}^*\|_2. \quad (3.22)$$

3.3.3 Fourier Imaging and DFT Grid Mismatch

A mismatch case of particular interest arises in Fourier imaging when a sparse signal with arbitrary frequency components is taken to be sparse in a DFT basis. Our objective in this section is to highlight the particularly problematic nature of basis mismatch in this application.

Suppose the sparsity basis $\boldsymbol{\Psi}_0$ in the mathematical model $\mathbf{s} = \boldsymbol{\Psi}_0 \mathbf{x}$, assumed by the compressed sensing procedure, is the unitary N -point DFT basis. Then the l th column of $\boldsymbol{\Psi}_0$ is a Vandermonde vector of the form $\boldsymbol{\psi}_{0,\ell} =$

$[1, e^{j2\pi\ell/N}, \dots, e^{j2\pi\ell(N-1)/N}]^T$ and the basis Ψ_0 is

$$\Psi_0 = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \dots & 1 \\ 1 & e^{j2\pi/N} & \dots & e^{j2\pi(N-1)/N} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{j2\pi(N-1)/N} & \dots & e^{j2\pi(N-1)^2/N} \end{bmatrix}. \quad (3.23)$$

Without loss of generality, let us assume that the ℓ th column $\psi_{1,\ell}$ of the actual sparsity basis Ψ_1 is mismatched to the ℓ th column of Ψ_0 by $\Delta\theta_\ell$ in (normalized) frequency, where $0 \leq \Delta\theta_\ell < \frac{2\pi}{N}$. Then, Ψ_1 is given by

$$\Psi_1 = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \dots & 1 \\ e^{j\Delta\theta_0} & e^{j(\frac{2\pi}{N} + \Delta\theta_1)} & \dots & e^{j(\frac{2\pi(N-1)}{N} + \Delta\theta_{N-1})} \\ \vdots & \vdots & \ddots & \vdots \\ e^{j\Delta\theta_0(N-1)} & e^{j(\frac{2\pi}{N} + \Delta\theta_1)(N-1)} & \dots & e^{j(\frac{2\pi(N-1)}{N} + \Delta\theta_{N-1})(N-1)} \end{bmatrix}. \quad (3.24)$$

The mismatched basis $\Psi = \Psi_0^{-1}\Psi_1$ then can be written as

$$\Psi = \begin{bmatrix} L(\Delta\theta_0 - 0) & L(\Delta\theta_1 - \frac{2\pi(N-1)}{N}) & \dots & L(\Delta\theta_{N-1} - \frac{2\pi}{N}) \\ L(\Delta\theta_0 - \frac{2\pi}{N}) & L(\Delta\theta_1 - 0) & \dots & L(\Delta\theta_{N-1} - \frac{2\pi(N-2)}{N}) \\ \vdots & \vdots & \ddots & \vdots \\ L(\Delta\theta_0 - \frac{2\pi(N-1)}{N}) & L(\Delta\theta_1 - \frac{2\pi(N-2)}{N}) & \dots & L(\Delta\theta_{N-1} - 0) \end{bmatrix} \triangleq \mathbf{I} + \mathbf{E},$$

where $L(\theta)$ is the Dirichlet kernel given by

$$L(\theta) = \frac{1}{N} \sum_{n=0}^{N-1} e^{jn\theta} = \frac{1}{N} e^{j\frac{\theta(N-1)}{2}} \frac{\sin(\theta N/2)}{\sin(\theta/2)}. \quad (3.25)$$

The (m, ℓ) th element of the mismatched basis Ψ is a sample of the Dirichlet kernel $L(\theta)$ at $\theta = \Delta\theta_\ell - \frac{2\pi}{N}(m - \ell)$, where $m = 0, 1, \dots, N - 1$ and $\ell = 0, 1, \dots, N - 1$.

The Dirichlet kernel $L(\theta)$, shown in Fig. 3.2 (ignoring the unimodular phasing term) for $N = 64$, decays slowly as $|L(\theta)| \leq (N\theta/2\pi)^{-1}$ for $|\theta| \leq \pi$, with $L(0) = 1$. This decay behavior follows from the fact that $|\sin(\theta/2)| \geq 2|\theta/2\pi|$ for $|\theta| \leq \pi$, where the equality holds when $|\theta| = \pi$. This means that $(N\theta/2\pi)^{-1}$ is in fact the envelope of $|L(\theta)|$. Therefore, every mismatch between a physical frequency θ_ℓ and the corresponding DFT frequency $2\pi\ell/N$ produces a column in the mismatched basis Ψ for which the entries vanish slowly as each column is traversed. The consequence of this is that the parameter vector \mathbf{x} in the mathematical model $\mathbf{s} = \Psi_0\mathbf{x}$, for which the CS procedure is seeking a sparse solution, is in fact incompressible. This follows from the coordinate transformation $\mathbf{x} = \Psi\boldsymbol{\theta}$ between the true sparse parameter vector $\boldsymbol{\theta}$ and the presumed sparse vector \mathbf{x} . The few nonzero elements of $\boldsymbol{\theta}$ leak in to all locations through Ψ and the slow decay of elements of Ψ makes \mathbf{x} incompressible.

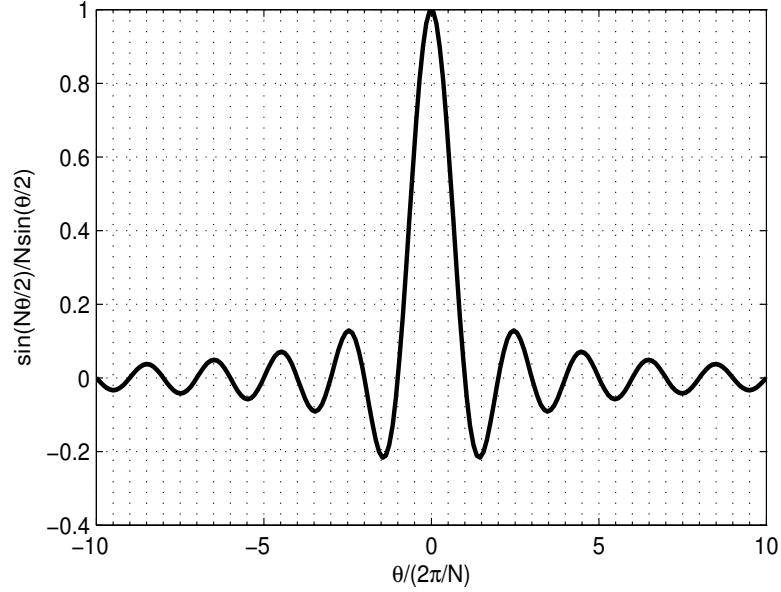


Figure 3.2: The function $\frac{1}{N} \frac{\sin(N\theta/2)}{\sin(\theta/2)}$ versus $\theta/(2\pi/N)$.

In addition to frequency mismatch, the columns of Ψ_1 may also be mismatched to those of Ψ_0 by damping factors $\lambda_\ell \leq 0$. In such a case, the basis Ψ_1 is given by

$$\Psi_1 = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ e^{[\lambda_0+j\Delta\theta_0]} & e^{[\lambda_1+j(\frac{2\pi}{N}+\Delta\theta_1)]} & \cdots & e^{[\lambda_{N-1}+j(\frac{2\pi(N-1)}{N}+\Delta\theta_{N-1})]} \\ \vdots & \vdots & \ddots & \vdots \\ e^{[\lambda_0+j\Delta\theta_0](N-1)} & e^{[\lambda_1+j(\frac{2\pi}{N}+\Delta\theta_1)](N-1)} & \cdots & e^{[\lambda_{N-1}+j(\frac{2\pi(N-1)}{N}+\Delta\theta_{N-1})](N-1)} \end{bmatrix},$$

and the (m, ℓ) th element of the mismatch basis Ψ is

$$\psi_{m,\ell} = \frac{1}{N} \sum_{n=0}^{N-1} e^{n[\lambda_\ell+j(\Delta\theta_\ell-\frac{2\pi(m-\ell)}{N})]}. \quad (3.26)$$

In general, the basis mismatch problem exists in almost all applications and is not limited to Fourier imaging. However, we emphasize Fourier imaging in this chapter as a ubiquitous imaging problem, where basis mismatch seems to have a particularly destructive effect.

3.4 Numerical Examples

We now present three simple numerical examples to demonstrate the effect of basis mismatch on the performance of CS. The first example considers single tone mismatch to the DFT grid, the second considers both frequency and damping mismatch, and

the third example considers a synthetic mismatch where the bound established in Theorem 3.3.1 is achieved.

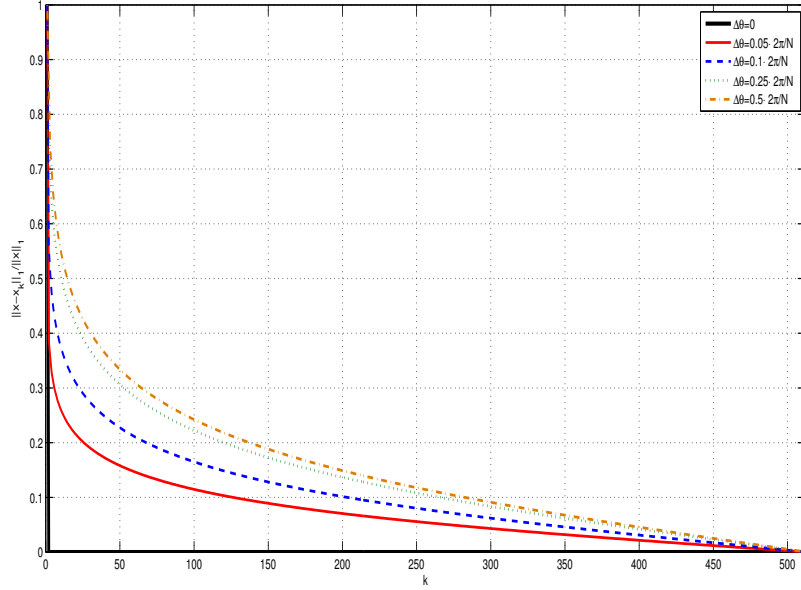
3.4.1 Tone Reconstruction with Basis Pursuit

Here we assume that the image \mathbf{s} is a single tone (a 1-sparse signal) with an unknown frequency. The actual sparsity basis Ψ_1 is a matrix, in which the ℓ th column is mismatched in frequency by $0 \leq \Delta\theta < \frac{2\pi}{N}$ with respect to the ℓ -th column of the N -point DFT matrix, which the compressed sensing procedure takes as the sparsity basis Ψ_0 . The rest of the columns of Ψ_1 and Ψ_0 are assumed to be identical. Damping mismatch is not considered in this example.

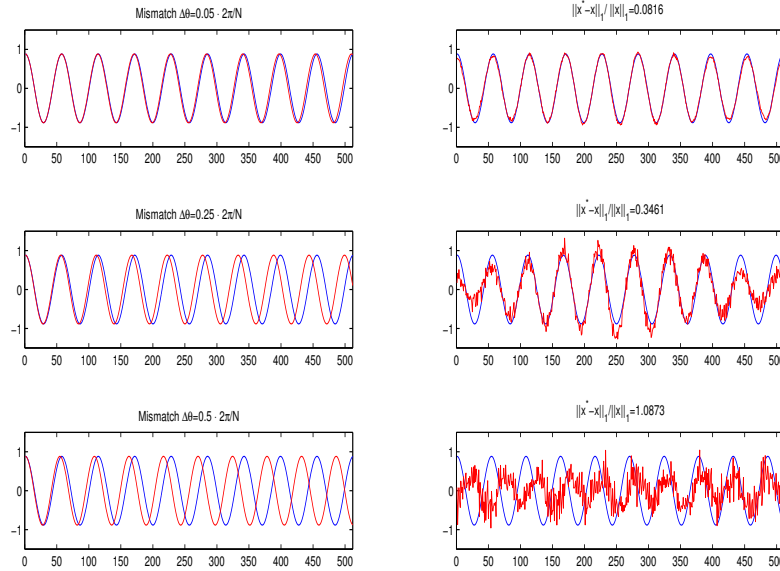
Fig. 3.3 (a) shows the normalized ℓ_1 error $\|\mathbf{x} - \mathbf{x}_k\|_1 / \|\mathbf{x}\|_1$ for different frequency mismatch levels $\Delta\theta$ and different values of k , when the single tone is located off the ($\ell = 10$)-th DFT frequency in the ($N = 512$)-point DFT grid. The plot shows that the normalized best k -term approximation error is considerable even at moderate mismatch levels, and speaks to the fact that, even for a slight mismatch with respect to the DFT mode, the presumably sparse representation is in fact incompressible due to the slow decay of the Dirichlet kernel.

Fig. 3.3 (b) compares the reconstructed tone to the actual tone for different mismatch levels. In the left column, the blue plots show the actual tone and the red plots show the closest tone on the DFT grid to the actual tone, for the corresponding mismatch level. On the right hand side, the red plots show the reconstructed tone $\mathbf{s}^* = \Psi_0 \mathbf{x}^*$, where Ψ_0 is the DFT basis, and the blue plots again show the actual tone. The number of compressed sensing measurements used for reconstruction is $M = 64$. The frequency mismatch $\Delta\theta / (2\pi/N)$ is 0.05 for the plots in the top row, 0.25 for the plots in the middle row, and 0.5 for the plots in the bottom row. The inaccuracy in reconstruction is noticeable for 25% (i.e., $\Delta\theta = 0.25 \frac{2\pi}{N}$) and 50% frequency mismatch.

Fig. 3.4 (a) shows how the normalized image inversion error $\|\mathbf{x} - \mathbf{x}^*\|_1 / \|\mathbf{x}\|_1$ varies with the number of measurements M in $N = 512$ dimensions for different mismatch levels $\Delta\theta = 0.005 \cdot 2\pi/N$, $0.05 \cdot 2\pi/N$ and $0.1 \cdot 2\pi/N$. The measurement matrix \mathbf{A} is generated with random Gaussian entries satisfying RIP conditions, and the experiment takes 50 runs and returns the average error. We observe that without increasing the number of measurements beyond what is required in the mismatch-free case the normalized error can be large, even for a small frequency mismatch $\Delta\theta = 0.005 \cdot 2\pi/N$. This threshold is approximately $3k \log N/k = 27$ in this example. For measurement dimensions smaller than 27 the normalized error is large, but it decreases relatively fast as the number of measurements increases. However, after this threshold the normalized error decays slowly and only when the number of measurements is increased to N the error goes to zero. Fig. 3.4 (b) shows the normalized image inversion error $\|\mathbf{x} - \mathbf{x}^*\|_1 / \|\mathbf{x}\|_1$ versus the frequency mismatch level $\Delta\theta / (2\pi/N)$ for several (M, N) pairs, where M is the number of measurements and N is the grid dimension.

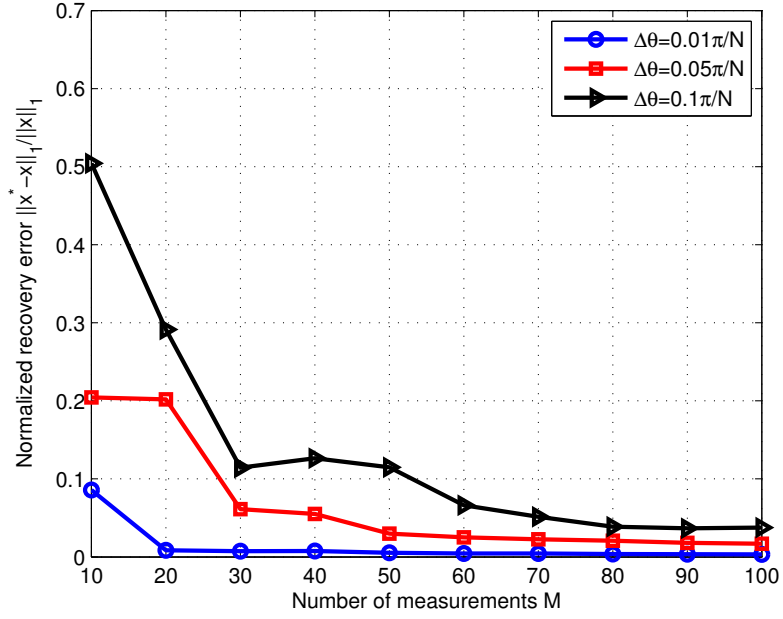


(a)

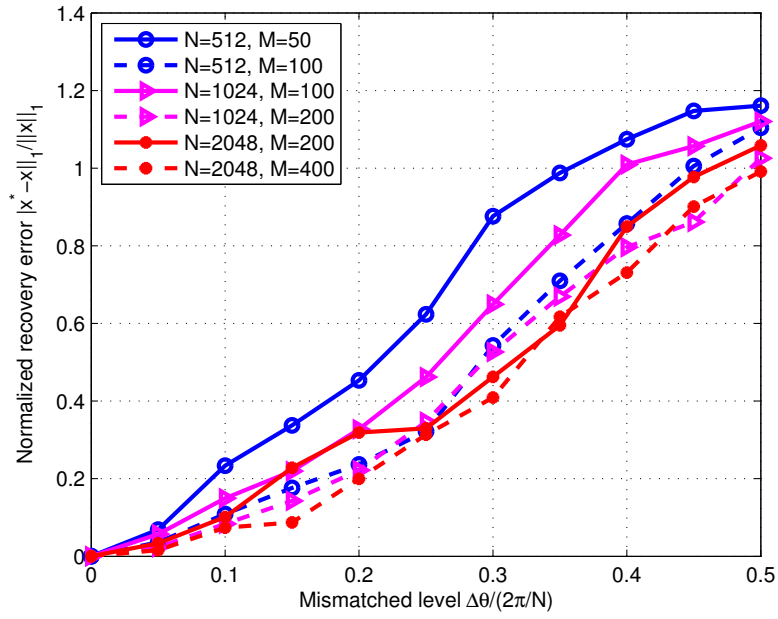


(b)

Figure 3.3: (a) $\|\mathbf{x} - \mathbf{x}_k\|_1 / \|\mathbf{x}\|_1$ versus $\Delta\theta/(2\pi/N)$ for various k . (b) Left column: the actual tone (blue) superimposed on the closest DFT tone; Right column: the reconstructed tone (red) superimposed on the actual tone (blue). The frequency mismatch $\Delta\theta/(2\pi/N)$ is 0.05, 0.25 and 0.5 for the plots in the top, middle and bottom row.



(a)



(b)

Figure 3.4: (a) Normalized image inversion error $\|\mathbf{x} - \mathbf{x}^*\|_1 / \|\mathbf{x}\|_1$ versus the number of measurements M for different frequency mismatch levels $\Delta\theta = 0.005 \cdot 2\pi/N$, $0.05 \cdot 2\pi/N$ and $0.1 \cdot 2\pi/N$, with $N = 512$. (b) $\|\mathbf{x} - \mathbf{x}^*\|_1 / \|\mathbf{x}\|_1$ versus the frequency mismatch level $\Delta\theta / (2\pi/N)$ for different (M, N) pairs.

3.4.2 Modal Analysis

We now give examples to demonstrate the effect of DFT grid mismatch on modal analysis based on CS measurements and compare the results with those obtained using classical image inversion principles, namely standard DFT imaging and linear prediction (LP). The reader is referred to [38], [45], [46] for a description of linear prediction. In all the experiments, the dimension of the image/signal is $N = 64$. The number of measurements M used for inversion is the same for all methods and we report results for $M = N/4 = 16$ to $M = N/2 = 32$ to $M = N/1 = 64$.

We first consider the case where the field we wish to invert for contains only modes that are aligned with the DFT frequencies. This is to demonstrate that CS and LP both provide perfect field recovery when there is no mismatch, as shown in Fig. 3.5(a)-(c). In each subfigure (a) through (c) there are four panels. In the top-left panel the true underlying modes are illustrated with stems whose locations on the unit disc indicate the frequencies of the modes, and whose heights illustrate the mode amplitudes. The phases of the modes are randomly chosen, and not indicated on the figures. The frequencies at which modes are placed, and their amplitudes, are $(2\pi \cdot 9/N, 1)$, $(2\pi \cdot 10/N, 1)$, $(2\pi \cdot 20/N, .5)$, and $(2\pi \cdot 45/N, .2)$. These frequencies are perfectly aligned with the DFT frequencies. No noise is considered for now. We observe that both CS and LP provide perfect recovery. The DFT processing however has leakage according to the Dirichlet kernel unless the measurement dimension is increased to the full dimension $N = 64$. This was of course expected.

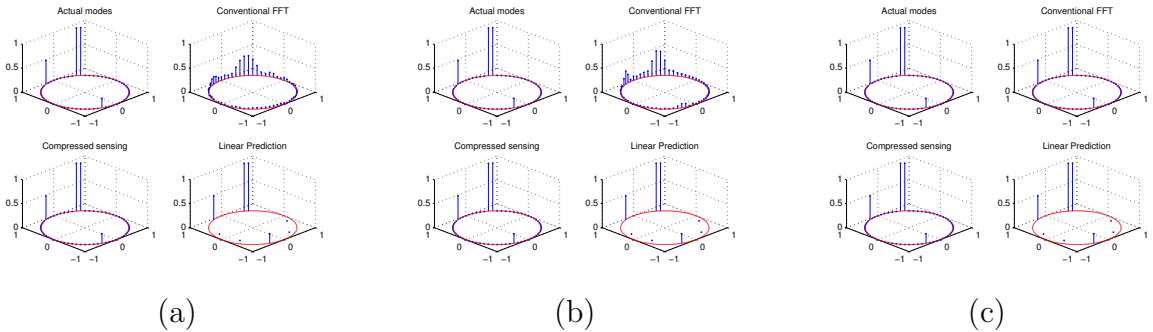


Figure 3.5: Comparison of DFT, CS, and LP inversions in the absence of basis mismatch (a) $M = N/4 = 16$, (b) $M = N/2 = 32$, and (c) $M = N/1 = 64$.

But what is the connection between the circular plots in Fig. 3.5(a)-(c) and the models $\Psi_0 \mathbf{x}$ and $\Psi_1 \boldsymbol{\theta}$? The top-left panel (actual modes) in each subplot is an “illustration” of $(\Psi_1, \boldsymbol{\theta})$, with the locations of the bars on the unit disc corresponding to active modes from Ψ_1 and the heights of the bars corresponding to the values of nonzero entries in $\boldsymbol{\theta}$. The top-right panel (conventional FFT) illustrates $(\Psi_0, \hat{\mathbf{x}})$, where $\hat{\mathbf{x}}$ is the estimate of \mathbf{x} obtained by DFT processing the measurement vector \mathbf{y} . The bottom-left (compressed sensing) illustrates (Ψ_0, \mathbf{x}^*) , where \mathbf{x}^* is the solution to (3.4) (or to (3.5) in the noisy cases to follow). The bottom-right panel (LP) illustrates $(\hat{\Psi}_1, \hat{\boldsymbol{\theta}})$, where $\hat{\Psi}_1$ and $\hat{\boldsymbol{\theta}}$ are, respectively, estimates of Ψ_1 and $\boldsymbol{\theta}$ obtained by LP (or

reduced-rank LP). When there is no mismatch and noise, (Ψ_0, \mathbf{x}^*) and $(\hat{\Psi}_1, \hat{\theta})$ match (Ψ_1, θ) . However, they significantly differ when mismatch is introduced.

We now introduce basis mismatch either by moving some of the modes off the DFT grid or by damping them. For frequency mismatch, the first two modes are moved to $(2\pi \cdot 9.25/N, 1)$ and $(2\pi \cdot 9.75/N, 1)$. For damping mismatch the mode at $(2\pi \cdot 9/N, 1)$ is drawn off the unit circle to radius 0.95, so that the mode is damped as $(0.95)^n$. The rest of the modes are the same as in the mismatch free case. Fig. 3.6 (a)-(f) shows the inversion results for DFT, CS, and LP (order 8) for $M = N/4 = 16$ to $M = N/2 = 32$ to $M = N/1 = 64$. In all cases, DFT and CS result in erroneous inversion. The inaccuracy in inversion persists even when the number of measurements is increased to the full dimension. However, we observe that LP is always exact. These are all noise free cases.

But can the mismatch effect be compensated for by replacing the observation constraint $\mathbf{y} = \mathbf{A}\mathbf{x}$ in basis pursuit with a quadratic constraint $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon$? The plots in Fig. 3.7 suggest that this is not the case. These plots show CS inversions for three different values of ϵ . The matched vs. mismatched modes and mode amplitudes in these plots are exactly the same as those in Fig. 3.6 and the number of measurements is $M = 16$. These results are not an artifact of the quadratic constraint allowed or the choice of rows in the compressed recording matrix, as we have experimented with many values and reported typical plots. For large values of ϵ the inversion returns a zero vector, as seen in the fourth subplot in Fig. 3.7(a),(b). The reason is that when $\epsilon > \|\mathbf{y}\|_2$ the sparsest solution satisfying $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon$ is $\mathbf{x} = 0$.

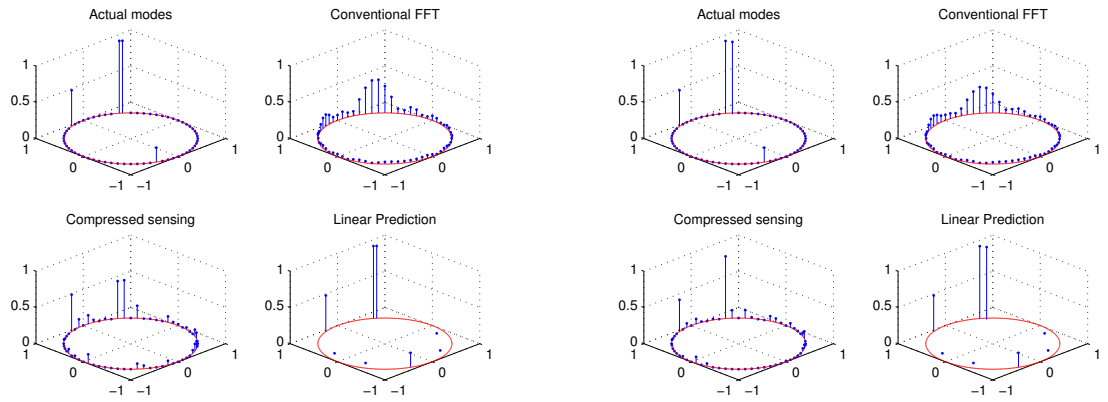
Finally, we consider noisy observations for both mismatched and mismatch-free cases. In the mismatch-free case, the frequencies at which modes are placed, and their amplitudes, are $(2\pi \cdot 9/N, 1)$, $(2\pi \cdot 11/N, 1)$, $(2\pi \cdot 20/N, .5)$, and $(2\pi \cdot 45/N, .2)$. For frequency mismatch, the first two modes are moved to $(2\pi \cdot 9.25/N, 1)$ and $(2\pi \cdot 10.75/N, 1)$. For damping mismatch the mode at $(2\pi \cdot 9/N, 1)$ is drawn off the unit circle to radius 0.95. The number of measurements is $M = N/2 = 32$. The LP order is changed to 16, but rank reduction [45], [46] is applied to reduce the order back to 8 as is typical in noisy cases. The inversion results are shown in Fig. 3.8.

3.4.3 A Synthetic Worst-Case Example

Let $\Psi_0 = \mathbf{I}$ be the sparsity basis assumed in the compressed sensing procedure and let $\Psi_1 = \mathbf{I} + \beta\mathbf{1}\mathbf{1}^T$ be the actual sparsity basis, where $\mathbf{1} = [1, 1, \dots, 1]^T$ is an N -dimensional vector of ones. Then, the mismatched basis is $\Psi = \Psi_0^{-1}\Psi_1 = \mathbf{I} + \beta\mathbf{1}\mathbf{1}^T$. Assume that the true parameter vector θ is a sparse vector in which all nonzero entries are equal to one and the positions of the nonzero entries are uniformly distributed. The coordinate transformation $\mathbf{x} = \Psi\theta$ has a simple form in this example and is given by

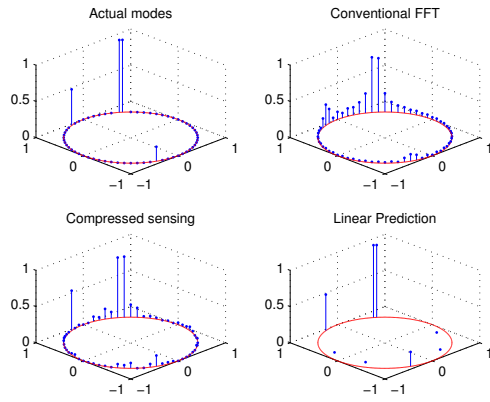
$$\mathbf{x} = (\mathbf{I} + \beta\mathbf{1}\mathbf{1}^T)\theta = \theta + k\beta\mathbf{1}. \quad (3.27)$$

Thus, a nonzero entry of 1 in θ results in a $1 + k\beta$ entry in \mathbf{x} and a zero entry in θ produces a $k\beta$ entry in \mathbf{x} . This example amounts to the worst case mismatch scenario in Theorem 3.3.1, with $q = 1$, where the upper bound is achieved. In this



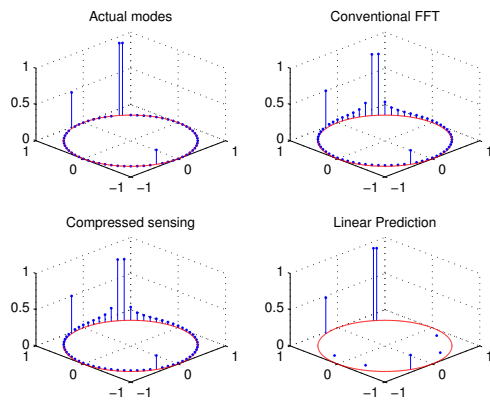
(a) Frequency mismatch

(b) Damping mismatch



(c) Frequency mismatch

(d) Damping mismatch



(e) Frequency mismatch

(f) Damping mismatch

Figure 3.6: Comparison of DFT, CS, and LP inversions in the presence of basis mismatch. Moving from the top to middle to bottom row, the number of measurements changes from $M = N/4 = 16$ to $M = N/2 = 32$ to $M = N/1 = 64$.

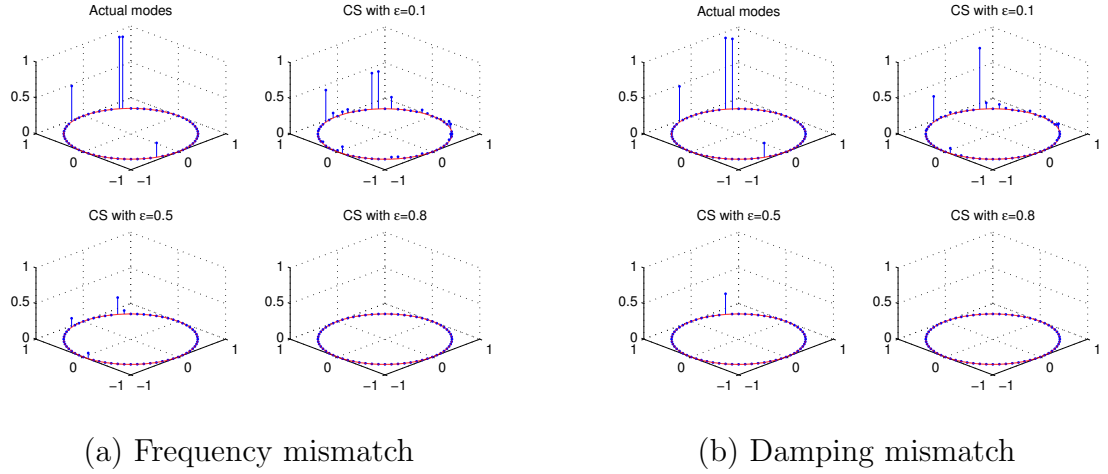


Figure 3.7: CS performance as the quadratic constraint $\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \leq \epsilon$ is relaxed from $\epsilon = 0.1$ to $\epsilon = 0.5$ to $\epsilon = 0.8$; (a) with frequency mismatch (b) with damping mismatch.

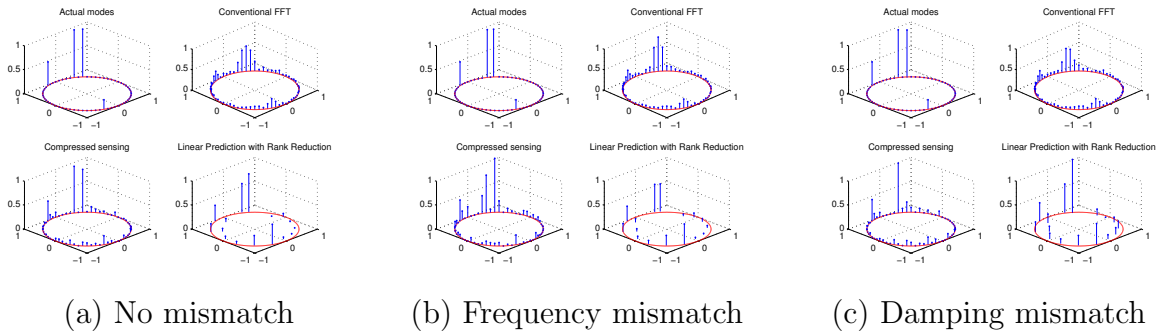


Figure 3.8: Comparison of DFT, CS, and LP inversions with noisy observations: (a) no mismatch, (b) frequency mismatch, and (c) damping mismatch.

case, $\|\mathbf{x}\|_1 = (1 + N\beta)\|\boldsymbol{\theta}\|_q$, and the normalized error bound can be written as

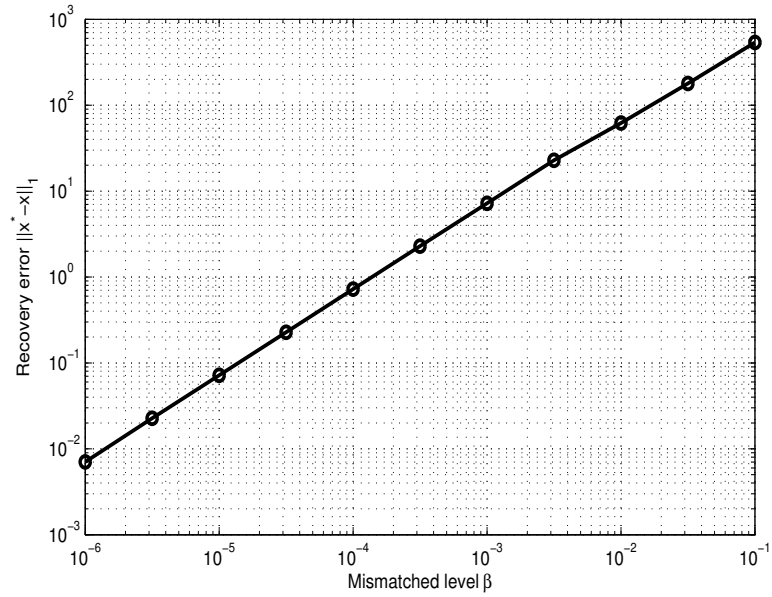
$$\frac{\|\mathbf{x} - \mathbf{x}^*\|_1}{\|\mathbf{x}\|_1} \leq \frac{C_0}{1 + N\beta} \frac{\|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1}{\|\boldsymbol{\theta}\|_1} + C_0 \frac{(N - k)\beta}{1 + N\beta}. \quad (3.28)$$

Fig. 3.9 (a),(b), respectively, show the ℓ_1 error $\|\mathbf{x} - \mathbf{x}^*\|_1$ and the normalized ℓ_1 error $\|\mathbf{x} - \mathbf{x}^*\|_1/\|\mathbf{x}\|_1$ as a function of the mismatch parameter β for dimensions $N = 512$, $M = 64$. We notice that the ℓ_1 error $\|\mathbf{x} - \mathbf{x}^*\|_1$ grows linearly with the increase in the mismatched level β , which agrees with Theorem 3.3.3. We also observe that even for moderate amounts of mismatch the normalized ℓ_1 error is considerable. When the mismatch level β goes to infinity, the normalized error bound converges to $C_0 \left(\frac{N-k}{N}\right)$. In Fig. 3.9 (b), the normalized error curve becomes flat when β is above $\mathcal{O}(10^{-2})$.

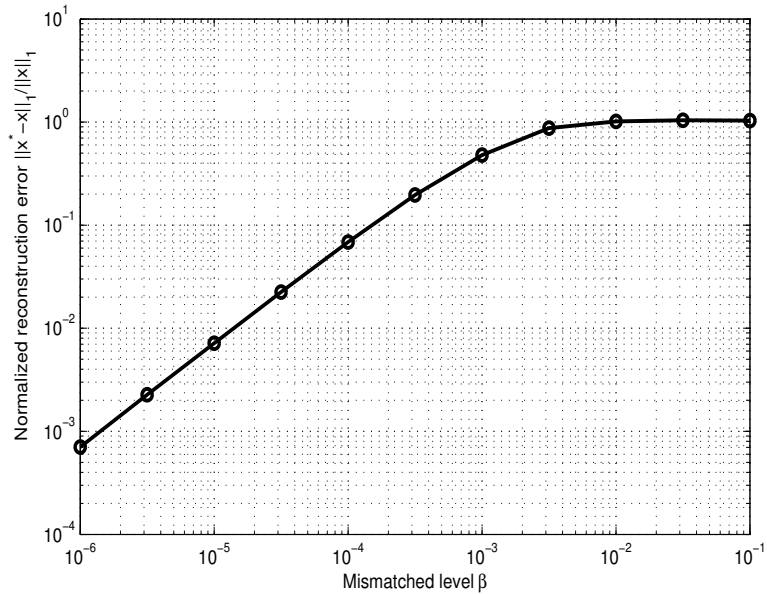
3.5 Conclusions

The theory of CS suggests that compressed recording has manageable consequences for image inversion, provided the image is sparse in an a priori known basis, e.g., a DFT basis or a basis associated with a range-Doppler-wavenumber grid. But no physical field is sparse in the DFT basis or in any a priori known basis defined by a regular grid in delay, doppler, frequency, and/or wavenumber, and there is always mismatch between the mathematical model for sparsity and the physical model for sparsity.

In this chapter, we have investigated the sensitivity of CS (specifically basis pursuit) to mismatch between the assumed basis for sparsity and the actual sparsity basis. Our mathematical analysis and numerical examples indicate that the performance of compressed sensing for approximating a sparse physical field may degrade considerably in the presence of basis mismatch, even when the assumed basis corresponds to a fine-grained discretization of the parameter space. Our analysis suggests that for high resolution spectrum analysis, DOA estimation, or delay-doppler imaging, where the problem is to identify a small number of modal parameters, extra care may be needed to account for the effects of basis mismatch. As a final remark, the parameter estimation problem in Fourier imaging and image inversion might be better handled using more sophisticated sparse models such as a low-rank smooth manifold.



(a)



(b)

Figure 3.9: (a) The recovery error $\|\mathbf{x} - \mathbf{x}^*\|_1$ from basis pursuit versus the mismatched level β for $N = 512$, $M = 64$ and $k = 10$. (b) The normalized recovery error $\|\mathbf{x} - \mathbf{x}^*\|_1 / \|\mathbf{x}\|_1$ from basis pursuit versus the mismatched level β for $N = 512$, $M = 64$ and $k = 10$.

Chapter 4

Low-rank Subspace Estimation from Partial Streaming Data

4.1 Introduction

Many real world data sets exhibit an embedding of low-dimensional structure in a high-dimensional manifold. Smooth manifolds are well approximated by tangent planes, and trees of tangent planes [73]. When the embedding is assumed linear, the underlying low-dimensional structure becomes a linear subspace. Subspace Identification and Tracking (SIT) plays an important role in various signal processing tasks such as online identification of network anomalies [74], moving target localization [75], beamforming [76], and denoising [77]. Conventional SIT algorithms collect full measurements of the data stream at each time, and subsequently update the subspace estimate by utilizing the track record of the stream history in different ways [78, 79].

The theory of Compressed Sensing (CS) [11, 10] and Matrix Completion (MC) [80, 81] have shown that it is possible to infer data structure from highly incomplete observations. Compared with CS, which allows reconstruction of a single vector from only a few attributes by assuming it is sparse in a pre-determined basis or dictionary, MC allows reconstruction of a matrix from a few entries by assuming it is low-rank. A popular method to perform MC is to minimize the nuclear norm of the underlying matrix [80, 81] which requires no prior knowledge of rank, in parallel with ℓ_1 minimization for sparse recovery in CS. Other approaches include greedy algorithms such as OptSpace [82] which requires an estimate of rank for initialization. Identifying the underlying low-rank structure in MC is equivalent to subspace identification in a batch setting. When the number of observed entries is slightly larger than the subspace rank, it has been shown that with high probability, it is possible to test whether a highly incomplete vector of interest lies in a known subspace [83].

In high-dimensional problems, it might be expensive and even impossible to collect data from all dimensions. For example in wireless sensor networks, collecting from all sensors continuously will quickly drain the battery power. Ideally, we would prefer to only collect data from a fixed budget of sensors each time to increase the overall battery life, and still be able to identify the underlying structure. Another example is

in online recommendation systems, where it is impossible to expect rating feedbacks from all users on every product. Therefore it is of growing interest to identify and track a low-dimensional subspace from highly incomplete information of a data stream in an online fashion. In this setting, the estimate of the subspace is updated and tracked across time when new observations become available with low computational cost. The GROUSE algorithm [84] has been proposed for SIT from online partial observations using rank-one updates of the estimated subspace on the Grassmannian manifold. However, performance is limited by the existence of “barriers” in the search path [85] which result in GROUSE being trapped at a local minima. We demonstrate this behavior through numerical examples in Section 4.5 in the context of direction-of-arrival estimation.

In this chapter we further study the problem of SIT given partial observations from a data stream as in GROUSE. Our proposed algorithm is dubbed Parallel Estimation and Tracking by REcursive Least Squares (PETRELS). The underlying low-dimensional subspace is identified by minimizing the geometrically discounted sum of projection residuals on the observed entries per time index, via a recursive procedure with discounting for each row of the subspace matrix in parallel. The missing entries are then reconstructed via least-squares estimation if required. The discounting factor balances the algorithm’s ability to capture long term behavior and changes to that behavior to improve adaptivity. We also benefit from the fact that our optimization of the estimated subspace is on all the possible low-rank subspaces, not restricted to the Grassmannian manifold. We discuss the convergence properties of PETRELS by revealing its connection with the well-known Projection Approximation Subspace Tracking (PAST) algorithm [78] in the full observation scenario, and analyze the convergence in a rank-one subspace update scenario. Finally, we provide numerical examples to measure the impact of the discount factor, estimated rank and number of observed entries. In the context of direction-of-arrival estimation we demonstrate superior performance of PETRELS over GROUSE in terms of separating close-located modes and tracking the changes in the scene. We also compare PETRELS with state of the art batch matrix completion algorithms, showing it as a competitive alternative when the subspace is fixed.

The rest of the chapter is organized as follows. Section 4.2 states the problem and provides background in the context of matrix completion and conventional subspace tracking. Section 4.3 describes the algorithm in detail. We discuss convergence issues of PETRELS in Section 4.4. Section 4.5 shows numerical evaluation of PETRELS.

4.2 Problem Statement and Related Work

4.2.1 Problem Statement

We consider the following problem. At each time t , a vector $\mathbf{x}_t \in \mathbb{R}^M$ is generated as:

$$\mathbf{x}_t = \mathbf{U}_t \mathbf{a}_t + \mathbf{n}_t \in \mathbb{R}^M, \quad (4.1)$$

where the columns of $\mathbf{U}_t \in \mathbb{R}^{M \times r_t}$ span a low-dimensional subspace, the vector $\mathbf{a}_t \in \mathbb{R}^{r_t}$ specifies the linear combination of columns and is Gaussian distributed as $\mathbf{a}_t \sim \mathcal{N}(0, \mathbf{I}_{r_t})$, and \mathbf{n}_t is additive white Gaussian noise distributed as $\mathbf{n}_t \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_M)$. The rank of the underlying subspace r_t is not assumed known exactly and can be slowly changing over time. The entries in the vectors \mathbf{x}_t can be considered as measurements from different sensors in a sensor network, or values of different pixels from a video frame. In practice, the upper bound of the subspace rank is considered known such that $r_t \leq r$ for any t .

We assume only partial entries of the full vector \mathbf{x}_t are observed, as depicted in Fig. 4.1, given by

$$\mathbf{y}_t = \mathbf{p}_t \odot \mathbf{x}_t = \mathbf{P}_t \mathbf{x}_t \in \mathbb{R}^M, \quad (4.2)$$

where \odot denotes point-wise multiplication, $\mathbf{P}_t = \text{diag}[\mathbf{p}_t]$, $\mathbf{p}_t = [p_{1t}, p_{2t}, \dots, p_{Mt}]^T \in \{0, 1\}^M$ with $p_{mt} = 1$ if the m th entry is observed at time t . We denote $\Omega_t = \{m : p_{mt} = 1\}$ as the set of observed entries at time t . In a random observation model, we assume the measurements are taken uniformly at random.

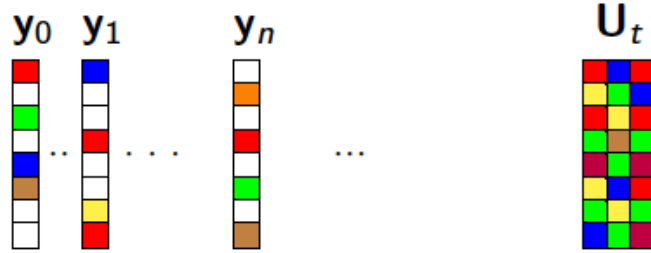


Figure 4.1: Illustration of the measurement model: only a small number of entries in each vector \mathbf{y}_t are observed, and the estimate of the low-rank matrix \mathbf{U}_t is updated online.

We are interested in an online estimate of a low-rank subspace $\mathbf{D}_n \in \mathbb{R}^{M \times r}$ at each time index n , which identifies and tracks the changes in the underlying subspace, from streaming partial observations $(\mathbf{y}_t, \mathbf{p}_t)_{t=1}^n$. The rank of the estimated subspace \mathbf{D}_n is assumed known and fixed throughout the algorithm as r . In practice, we assume is the upper bound of the rank of the underlying subspace. The desired properties for the algorithm include:

- *Low complexity:* each step of the online algorithm at time index n should be adaptive with small complexity compared to running a batch algorithm using history data;
- *Small storage:* The online algorithm should require a storage size that does not grow with the data size;
- *Convergence:* The subspace sequence generated by the online algorithm should converge to the true subspace $\mathbf{U} = \mathbf{U}_n$ if it is constant under the assumption that the sampled covariance matrix of \mathbf{x}_t converges.

- *Adaptivity*: The online algorithm should be able to track the changes of the underlying subspace in a timely fashion.

4.2.2 Conventional Subspace Identification and Tracking

When \mathbf{x}_t 's are fully observed, our problem is equivalent to the classical SIT problem, which is widely studied and has a rich literature in the signal processing community. Here we describe the Projection Approximation Subspace Tracking (PAST) algorithm in detail which is the closest to our proposed algorithm in the conventional scenario. First, consider optimizing the scalar function with respect to a subspace $\mathbf{W} \in \mathbb{R}^{M \times r}$, given by

$$J(\mathbf{W}) = \mathbb{E} \|\mathbf{x}_t - \mathbf{W}\mathbf{W}^T \mathbf{x}_t\|_2^2, \quad (4.3)$$

When $\mathbf{U}_t = \mathbf{U}$ is fixed over time, let $\mathbf{C}_x = \mathbb{E}[\mathbf{x}_t \mathbf{x}_t^T] = \mathbf{U}\mathbf{U}^T + \sigma^2 \mathbf{I}_M$ be the data covariance matrix. It is shown in [78] that the global minima of (4.3) is the only stable stationary point, which is obtained by $\mathbf{W} = \mathbf{U}_r \mathbf{Q}$, where \mathbf{U}_r is composed of the r dominant eigenvectors of \mathbf{C}_x , and $\mathbf{Q} \in \mathbb{C}^{r \times r}$ is a unitary matrix. When the data is noise-free, we can choose $\mathbf{U}_r = \mathbf{U}$. This motivates PAST to optimize the following function at time n without constraining \mathbf{W} to have orthogonal columns:

$$\mathbf{W}_n = \underset{\mathbf{W} \in \mathbb{R}^{M \times r}}{\operatorname{argmin}} \sum_{i=1}^n \alpha^{n-i} \|\mathbf{x}_t - \mathbf{W}\mathbf{W}^T \mathbf{x}_t\|_2^2, \quad (4.4)$$

$$\approx \underset{\mathbf{W} \in \mathbb{R}^{M \times r}}{\operatorname{argmin}} \sum_{i=1}^n \alpha^{n-i} \|\mathbf{x}_t - \mathbf{W}\mathbf{W}_{n-1}^T \mathbf{x}_t\|_2^2, \quad (4.5)$$

where the expectation in (4.3) is replaced by geometrically weighting the previous observations by α in (4.4), which is further approximated by replacing the second \mathbf{W} by its previous estimate in (4.5). Based on (4.5), the subspace \mathbf{W}_n can be found by first estimating the coefficient using the previous subspace estimate as $\mathbf{a}_n = \mathbf{W}_{n-1}^T \mathbf{x}_n$, then update the subspace as

$$\mathbf{W}_n = \underset{\mathbf{W} \in \mathbb{R}^{M \times r}}{\operatorname{argmin}} \sum_{i=1}^n \alpha^{n-i} \|\mathbf{x}_t - \mathbf{W} \mathbf{a}_t\|_2^2. \quad (4.6)$$

Now let $\alpha = 1$ and $\mathbf{R}_n = \sum_{i=1}^n \mathbf{a}_i \mathbf{a}_i^T$. In [86], the asymptotic dynamics of the PAST algorithm is described by the Ordinary Differential Equation (ODE) below and its equilibrium as t goes to infinity:

$$\begin{aligned} \dot{\mathbf{R}} &= \mathbb{E}[\tilde{\mathbf{a}}_n \tilde{\mathbf{a}}_n^T] - \mathbf{R} = \mathbf{W}^T \mathbf{C}_x \mathbf{W} - \mathbf{R}, \\ \dot{\mathbf{W}} &= \mathbb{E}[\mathbf{x}_n (\mathbf{x}_n - \mathbf{W} \tilde{\mathbf{a}}_n)^T] \mathbf{R}^\dagger = (\mathbf{I} - \mathbf{W}\mathbf{W}^T) \mathbf{C}_x \mathbf{W} \mathbf{R}^\dagger, \end{aligned}$$

where $\tilde{\mathbf{a}}_n = \mathbf{W}^T \mathbf{x}_n$, $\mathbf{R} = \mathbf{R}(t)$ and $\mathbf{W} = \mathbf{W}(t)$ are continuous time versions of \mathbf{R}_n and \mathbf{W}_n , and \dagger denotes pseudo-inverse. It is proved in [86] that as t increases, $\mathbf{W}(t)$ converges to the global optima, i.e. to a matrix which spans the eigenvectors of \mathbf{C}_x corresponding to the r largest eigenvalues. In Section 4.4.1 we show that our proposed

PETRELS algorithm becomes essentially equivalent to the PAST algorithm when all entries of the data stream are observed, and can be shown to converge globally.

4.2.3 Matrix Completion

When only partial observations are available and $\mathbf{U}_t = \mathbf{U}$ are fixed, our problem is closely related to the Matrix Completion (MC) problem, which has been extensively studied recently. Assume $\mathbf{X} \in \mathbb{R}^{M \times n}$ is a low-rank matrix, \mathbf{P} is an $M \times n$ mask matrix with 0 at missing entries and 1 at observed entries. Let $\mathbf{Y} = \mathbf{P} \odot \mathbf{X}$ be the observed partial matrix, where \odot denotes point-wise multiplication. Matrix completion aims to solve the following problem:

$$\min_{\mathbf{Z}} \text{rank}(\mathbf{Z}) \quad \text{s.t.} \quad \mathbf{P} \odot (\mathbf{X} - \mathbf{Z}) = \mathbf{0}, \quad (4.7)$$

i.e. to find a low-rank matrix such that the observed entries are satisfied. This problem is combinatorially intractable due to the rank constraint. It has been shown in [80] that by replacing the rank constraint with nuclear norm minimization, (4.7) can be replaced by a convex optimization problem, given the following spectral-regularized MC problem:

$$\min_{\mathbf{Z}} \frac{1}{2} \|\mathbf{P} \odot (\mathbf{X} - \mathbf{Z})\|_F^2 + \mu \|\mathbf{Z}\|_*, \quad (4.8)$$

where $\|\mathbf{Z}\|_*$ is the nuclear norm of \mathbf{Z} , i.e. the sum of singular values of \mathbf{Z} , and $\mu > 0$ is a regularization parameter. Interestingly, the solution of (4.8) is the same as that of (4.7), i.e. provides exact recovery for matrix completion under mild conditions [80]. The nuclear norm [87] of \mathbf{Z} is given by

$$\|\mathbf{Z}\|_* = \min_{\mathbf{U}, \mathbf{V}: \mathbf{Z} = \mathbf{U}\mathbf{V}^T} \frac{1}{2} (\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2) \quad (4.9)$$

where $\mathbf{U} \in \mathbb{C}^{M \times r}$ and $\mathbf{V} \in \mathbb{C}^{n \times r}$. Substituting (4.9) in (4.8) we can rewrite the MC problem as

$$\min_{\mathbf{U}, \mathbf{V}} \|\mathbf{P} \odot (\mathbf{X} - \mathbf{U}\mathbf{V})\|_F^2 + \mu (\|\mathbf{U}\|_F^2 + \|\mathbf{V}\|_F^2). \quad (4.10)$$

Our problem formulation can be viewed as an online way of solving the above batch-setting MC problem, where the columns of \mathbf{X} are drawn randomly and treated as a new measurement at each time index with a fixed underlying subspace $\mathbf{U}_t = \mathbf{U}$. The online algorithm has potential advantages for adapting changes in matrix size and avoiding large matrix manipulations. We will compare the PETRELS algorithm against some of the popular MC algorithms proposed recently in Section 4.5.

4.3 The PETRELS Algorithm

We now describe our proposed Parallel Estimation and Tracking by REcursive Least Squares (PETRELS) algorithm.

4.3.1 PETRELS

We first define the function $f_t(\mathbf{D})$ at each time $t = 1, \dots, n$ for a fixed subspace $\mathbf{D} \in \mathbb{R}^{M \times r}$, which is the total projection residual on the observed entries,

$$f_t(\mathbf{D}) = \min_{\mathbf{a}_t} \|\mathbf{P}_t(\mathbf{x}_t - \mathbf{D}\mathbf{a}_t)\|_2^2, \quad t = 1, \dots, n. \quad (4.11)$$

Here r is the rank of the estimated subspace, which is assumed known and fixed throughout the algorithm¹. We aim to minimize the following loss function at each time n with respect to the underlying subspace:

$$\mathbf{D}_n = \operatorname{argmin}_{\mathbf{D} \in \mathbb{R}^{M \times r}} F_n(\mathbf{D}) = \operatorname{argmin}_{\mathbf{D} \in \mathbb{R}^{M \times r}} \sum_{t=1}^n \lambda^{n-t} f_t(\mathbf{D}), \quad (4.12)$$

where \mathbf{D}_n is the estimated subspace of rank r at time n , and the parameter $0 \ll \lambda \leq 1$ discounts past observations.

To motivate the loss function in (4.12) we note that if $\mathbf{U}_t = \mathbf{U}$ is not changing over time, then the RHS of (4.12) is minimized to zero when \mathbf{D}_n spans the subspace defined by \mathbf{U} . If \mathbf{U}_t is slowly changing, then λ is used to control the memory of the system and maintain tracking ability at time n . For example, by using $\lambda \rightarrow 1$ the algorithm gradually loses its ability to forget the past.

Fixing \mathbf{D} , $f_t(\mathbf{D})$ can be written as

$$f_t(\mathbf{D}) = \mathbf{x}_t^T (\mathbf{P}_t - \mathbf{P}_t \mathbf{D} (\mathbf{D}^T \mathbf{P}_t \mathbf{D})^\dagger \mathbf{D}^T \mathbf{P}_t) \mathbf{x}_t, \quad (4.13)$$

where \dagger denotes matrix pseudo-inverse. Plugging this back to (4.12) the exact optimization problem becomes:

$$\mathbf{D}_n = \operatorname{argmin}_{\mathbf{D} \in \mathbb{R}^{M \times r}} \sum_{t=1}^n \lambda^{n-t} \mathbf{x}_t^T (\mathbf{P}_t - \mathbf{P}_t \mathbf{D} (\mathbf{D}^T \mathbf{P}_t \mathbf{D})^\dagger \mathbf{D}^T \mathbf{P}_t) \mathbf{x}_t,$$

which is difficult to solve over \mathbf{D} and requires storing all previous observations. Instead, we propose PETRELS to approximately solve this optimization problem.

Before developing PETRELS we note that if there are further constraints on the coefficients \mathbf{a}_t 's, a regularization term can be incorporated as:

$$f_t(\mathbf{D}) = \min_{\mathbf{a}_t \in \mathbb{R}^r} \|\mathbf{P}_t(\mathbf{D}\mathbf{a}_t - \mathbf{x}_t)\|_2^2 + \beta \|\mathbf{a}_t\|_p, \quad (4.14)$$

where $p \geq 0$. For example, $p = 1$ enforces a sparse constraint on \mathbf{a}_t , and $p = 2$ enforces a norm constraint on \mathbf{a}_t .

In (4.12) the discount factor λ is fixed, and the influence of past estimates decreases geometrically; a more general online objective function can be given as

$$F_n(\mathbf{D}) = \lambda_n F_{n-1}(\mathbf{D}) + f_n(\mathbf{D}), \quad (4.15)$$

¹The rank may not equal the true subspace dimension.

where the sequence $\{\lambda_n\}$ is used to control the memory and adaptivity of the system in a more flexible way.

Algorithm 4 PETRELS for SIT from Partial Observations

Input: a stream of vectors \mathbf{y}_t and observed pattern \mathbf{P}_t .

Initialization: an $M \times r$ random matrix \mathbf{D}_0 , and $(\mathbf{R}_m^0)^\dagger = \delta \mathbf{I}_r$, $\delta > 0$ for all $m = 1, \dots, M$.

```

1: for  $n = 1, 2, \dots$  do
2:    $\mathbf{a}_n = (\mathbf{D}_{n-1}^T \mathbf{P}_n \mathbf{D}_{n-1})^\dagger \mathbf{D}_{n-1}^T \mathbf{y}_n$ .
3:    $\hat{\mathbf{x}}_n = \mathbf{D}_{n-1} \mathbf{a}_n$ .
4:   for  $m = 1, \dots, M$  do
5:      $\beta_m^n = 1 + \lambda^{-1} \mathbf{a}_n^T (\mathbf{R}_m^{n-1})^\dagger \mathbf{a}_n$ ,
6:      $\mathbf{v}_m^n = \lambda^{-1} (\mathbf{R}_m^{n-1})^\dagger \mathbf{a}_n$ ,
7:      $(\mathbf{R}_m^n)^\dagger = \lambda^{-1} (\mathbf{R}_m^{n-1})^\dagger + p_{mt} (\beta_m^n)^{-1} \mathbf{v}_m^n (\mathbf{v}_m^n)^T$ ,
8:      $\mathbf{d}_m^n = \mathbf{d}_m^{n-1} + p_{mn} (x_{mn} - \mathbf{a}_n^T \mathbf{d}_m^{n-1}) (\mathbf{R}_m^n)^\dagger \mathbf{a}_n$ .
9:   end for
10: end for

```

The proposed PETRELS algorithm, as summarized by Algorithm 4, alternates between coefficient estimation and subspace update at each time n . In particular, the coefficient vector \mathbf{a}_n is estimated by minimizing the projection residual on the previous subspace estimate \mathbf{D}_{n-1} :

$$\mathbf{a}_n = \operatorname{argmin}_{\mathbf{a} \in \mathbb{R}^r} \|\mathbf{P}_n(\mathbf{x}_n - \mathbf{D}_{n-1} \mathbf{a})\|_2^2 = (\mathbf{D}_{n-1}^T \mathbf{P}_n \mathbf{D}_{n-1})^\dagger \mathbf{D}_{n-1}^T \mathbf{y}_n, \quad (4.16)$$

where \mathbf{D}_0 is a random subspace initialization. The full vector \mathbf{x}_n can then be estimated subsequently as:

$$\hat{\mathbf{x}}_n = \mathbf{D}_{n-1} \mathbf{a}_n, \quad (4.17)$$

and the residual error is given as $\mathbf{r}_n = \mathbf{P}_n(\mathbf{x}_n - \hat{\mathbf{x}}_n)$.

The subspace \mathbf{D}_n is then updated by minimizing

$$\mathbf{D}_n = \operatorname{argmin}_{\mathbf{D}} \sum_{t=1}^n \lambda^{n-t} \|\mathbf{P}_t(\mathbf{x}_t - \mathbf{D} \mathbf{a}_t)\|_2^2, \quad (4.18)$$

where \mathbf{a}_t , $t = 1, \dots, n$ are estimates from (4.16). Comparing (4.18) with (4.12), the optimal coefficients are substituted for the previous estimated coefficients. This results in a simpler problem for finding \mathbf{D}_n . The discount factor mitigates the error propagation and compensates for the fact that we used the previous coefficients updated rather than solving (4.12) directly, therefore improving the performance of the algorithm.

The objective function in (4.18) can be equivalently decomposed into a set of smaller problems for each row of $\mathbf{D}_n = [\mathbf{d}_1^n, \mathbf{d}_2^n, \dots, \mathbf{d}_M^n]^T$ as

$$\mathbf{d}_m^n = \underset{\mathbf{d}_m}{\operatorname{argmin}} \sum_{t=1}^n \lambda^{n-t} p_{mt} (x_{mt} - \mathbf{a}_t^T \mathbf{d}_m)^2, \quad (4.19)$$

for $m = 1, \dots, M$. To find the optimal \mathbf{d}_m^n , we equate the derivative of the RHS of (4.19) to zero, resulting in

$$\left(\sum_{t=1}^n \lambda^{n-t} p_{mt} \mathbf{a}_t \mathbf{a}_t^T \right) \mathbf{d}_m^n - \sum_{t=1}^n \lambda^{n-t} p_{mt} x_{mt} \mathbf{a}_t = \mathbf{0}.$$

which can be rewritten as $\mathbf{R}_m^n \mathbf{d}_m^n = \mathbf{s}_m^n$, where $\mathbf{R}_m^n = \sum_{t=1}^n \lambda^{n-t} p_{mt} \mathbf{a}_t \mathbf{a}_t^T$, and $\mathbf{s}_m^n = \sum_{t=1}^n \lambda^{n-t} p_{mt} x_{mt} \mathbf{a}_t$. Therefore, \mathbf{d}_m^n can be found as

$$\mathbf{d}_m^n = (\mathbf{R}_m^n)^\dagger \mathbf{s}_m^n. \quad (4.20)$$

Now we show how (4.20) can be updated recursively. First we rewrite

$$\mathbf{R}_m^n = \lambda \mathbf{R}_m^{n-1} + p_{mn} \mathbf{a}_n \mathbf{a}_n^T, \quad (4.21)$$

$$\mathbf{s}_m^n = \lambda \mathbf{s}_m^{n-1} + p_{mn} x_{mn} \mathbf{a}_n, \quad (4.22)$$

for all $m = 1, \dots, M$. Then we plug (4.21) and (4.22) into (4.20), and get

$$\begin{aligned} \mathbf{R}_m^n \mathbf{d}_m^n &= \lambda \mathbf{s}_m^{n-1} + p_{mn} x_{mn} \mathbf{a}_n \\ &= \lambda \mathbf{R}_m^{n-1} \mathbf{d}_m^{n-1} + p_{mn} x_{mn} \mathbf{a}_n \\ &= \mathbf{R}_m^n \mathbf{d}_m^{n-1} - p_{mn} \mathbf{a}_n \mathbf{a}_n^T \mathbf{d}_m^{n-1} + p_{mn} x_{mn} \mathbf{a}_n \\ &= \mathbf{R}_m^n \mathbf{d}_m^{n-1} + p_{mn} (x_{mn} - \mathbf{a}_n^T \mathbf{d}_m^{n-1}) \mathbf{a}_n, \end{aligned} \quad (4.23)$$

where \mathbf{d}_m^{n-1} is the the row estimate in the previous time $n - 1$. This results in a parallel procedure to update all rows of the subspace matrix \mathbf{D}_n , give as

$$\mathbf{d}_m^n = \mathbf{d}_m^{n-1} + p_{mn} (x_{mn} - \mathbf{a}_n^T \mathbf{d}_m^{n-1}) (\mathbf{R}_m^n)^\dagger \mathbf{a}_n. \quad (4.24)$$

Finally, by the Recursive Least-Squares (RLS) updating formula [47], $(\mathbf{R}_m^n)^\dagger$ can be easily updated without matrix inversion using

$$(\mathbf{R}_m^n)^\dagger = (\lambda \mathbf{R}_m^{n-1} + p_{mn} \mathbf{a}_n \mathbf{a}_n^T)^\dagger = \lambda^{-1} (\mathbf{R}_m^{n-1})^\dagger + p_{mt} \mathbf{G}_m^n; \quad (4.25)$$

where $\mathbf{G}_m^n = (\beta_m^n)^{-1} \mathbf{v}_m^n (\mathbf{v}_m^n)^T$, with β_m^n and \mathbf{v}_m^n given as

$$\beta_m^n = 1 + \lambda^{-1} \mathbf{a}_n^T (\mathbf{R}_m^{n-1})^\dagger \mathbf{a}_n, \quad \mathbf{v}_m^n = \lambda^{-1} (\mathbf{R}_m^{n-1})^\dagger \mathbf{a}_n.$$

To enable the RLS procedure, the matrix $(\mathbf{R}_m^0)^\dagger$ is initialized as a matrix with large entries on the diagonal, which we choose arbitrarily as the identity matrix $(\mathbf{R}_m^0)^\dagger =$

$\delta \mathbf{I}_r$, $\delta > 0$ for all $m = 1, \dots, M$. It is worth-noting that the fast implementation of RLS updating rules is in general very efficient. However, cautions need to be taken since direct application of fast RLS algorithms suffer from numerical instability of finite-precision operations when running for a long time [88].

The PETRELS algorithm can be regarded as a second-order stochastic gradient descent method to solve (4.12) by using \mathbf{d}_m^{n-1} , $m = 1, \dots, M$ as a warm start at time n . Specifically, we can write the gradient of $f_n(\mathbf{D})$ in (4.11) at \mathbf{D}_{n-1} as

$$\left. \frac{\partial f_n(\mathbf{D})}{\partial \mathbf{D}} \right|_{\mathbf{D}=\mathbf{D}_{n-1}} = -2\mathbf{P}_n(\mathbf{x}_n - \mathbf{D}_{n-1}\mathbf{a}_n)\mathbf{a}_n^T, \quad (4.26)$$

where \mathbf{a}_n is given in (4.16). Then the gradient of $F_n(\mathbf{D})$ at \mathbf{D}_{n-1} is given as

$$\left. \frac{\partial F_n(\mathbf{D})}{\partial \mathbf{D}} \right|_{\mathbf{D}=\mathbf{D}_{n-1}} = -2 \sum_{t=1}^n \lambda^{n-t} \mathbf{P}_t(\mathbf{x}_t - \mathbf{D}_{n-1}\mathbf{a}_t)\mathbf{a}_t^T.$$

Therefore the Hessian for each row of \mathbf{D} at \mathbf{d}_m^{n-1} is

$$\mathbf{H}_n(\mathbf{d}_m^{n-1}, \lambda) = \left. \frac{\partial^2 F_n(\mathbf{D})}{\partial \mathbf{d}_m \partial \mathbf{d}_m^T} \right|_{\mathbf{d}_m=\mathbf{d}_m^{n-1}} = 2 \sum_{t=1}^n \lambda^{n-t} p_{mt} \mathbf{a}_t \mathbf{a}_t^T. \quad (4.27)$$

Hence the updating rule for each row \mathbf{d}_m given in (4.24) can be written as

$$\mathbf{d}_m^n = \mathbf{d}_m^{n-1} - \mathbf{H}_n(\mathbf{d}_m^{n-1}, \lambda)^{-1} \frac{\partial f_n(\mathbf{D})}{\partial \mathbf{d}_m^{n-1}}, \quad (4.28)$$

which is equivalent to second-order stochastic gradient descent approach. Compared with first-order algorithms, PETRELS enjoys a faster convergence speed to the stationary point of $F_n(\mathbf{D})$ and gets rid of the problem of tuning the step-size [89].

4.3.2 Comparison with GROUSE

The GROUSE algorithm [84] proposed by Balzano et. al. addresses the same problem of online identification of low-rank subspace from highly incomplete information. The GROUSE method can be viewed as optimizing (4.12) for $\lambda = 1$ at each time n using a first-order stochastic gradient descent on the *orthogonal* Grassmannian defined as $\mathcal{G}_r = \{\mathbf{D} \in \mathbb{R}^{M \times r} : \mathbf{D}^T \mathbf{D} = \mathbf{I}_r\}$ instead of $\mathbb{R}^{M \times r}$. Thus, in the GROUSE algorithm,

$$\mathbf{D}_n = \underset{\mathbf{D} \in \mathcal{G}_r}{\operatorname{argmin}} G_n(\mathbf{D}) = \underset{\mathbf{D} \in \mathcal{G}_r}{\operatorname{argmin}} \sum_{t=1}^n f_t(\mathbf{D}). \quad (4.29)$$

GROUSE updates the subspace estimate along the direction of $\nabla f_t(\mathbf{D})|_{\mathbf{D}=\mathbf{D}_{n-1}}$ on \mathcal{G}_r , given by a fast rank-one update as

$$\mathbf{D}_n = \mathbf{D}_{n-1} - \left[(\cos(\sigma\eta_n) - 1) \frac{\hat{\mathbf{x}}_t}{\|\hat{\mathbf{x}}_t\|_2} + \sin(\sigma\eta_t) \frac{\mathbf{r}_t}{\|\mathbf{r}_t\|_2} \right] \frac{\mathbf{a}_n^T}{\|\mathbf{a}_n\|_2}, \quad (4.30)$$

where $\sigma = \|\hat{\mathbf{x}}_t\|_2 \|\mathbf{r}_t\|_2$, and η_n is the step-size at time n . At each step GROUSE also alternates between coefficient estimation (4.16) and subspace update (4.30). If the step size satisfies

$$\lim_{n \rightarrow \infty} \eta_n = 0 \quad \text{and} \quad \sum_{t=1}^{\infty} \eta_t = \infty,$$

then GROUSE is guaranteed to converge to a stationary point of $G_n(\mathbf{D})$. However, due to the existence of “barriers” in the search path on the Grassmannian [85], GROUSE may be trapped at a local minima as shown in Section 4.5 in the example of direction-of-arrival estimation.

Remark: If we relax the objective function of GROUSE (4.29) to all rank- r subspace $\mathbb{R}^{M \times r}$, given as

$$\mathbf{D}_n = \operatorname{argmin}_{\mathbf{D} \in \mathbb{R}^{M \times r}} \sum_{t=1}^n f_t(\mathbf{D}),$$

which is equivalent to PETRELS without discounting. Therefore following the discussions above, it is possible to use a second-order stochastic gradient descent method to update the underlying subspace, yielding the update rule for each row of \mathbf{D}_n as

$$\mathbf{d}_m^n = \mathbf{d}_m^{n-1} - \gamma_n \mathbf{H}_n(\mathbf{d}_m^{n-1}, \lambda = 1)^{-1} \frac{\partial f_n(\mathbf{D})}{\partial \mathbf{d}_m^{n-1}}, \quad (4.31)$$

where γ_n is the step-size at time n . Compared with the update rule for PETRELS in (4.28), the discount parameter has a similar role as the step-size, but weights the contribution of previous data input geometrically.

4.3.3 Simplified PETRELS

In the subspace update step of PETRELS in (4.18), consider replacing the objective function in (4.12) by

$$\mathbf{D}_n = \operatorname{argmin}_{\mathbf{D}} \hat{F}_n(\mathbf{D}) = \operatorname{argmin}_{\mathbf{D}} \sum_{t=1}^n \lambda^{n-t} \|\hat{\mathbf{x}}_t - \mathbf{D} \mathbf{a}_t\|_2^2, \quad (4.32)$$

where \mathbf{a}_t and $\hat{\mathbf{x}}_t$, $t = 1, \dots, n$ are estimates from earlier steps in (4.16) and (4.17). The only change we made is to remove the partial observation operator from the objective function, and replace it by the full vector estimate. It remains true that $\mathbf{d}_m^n = \operatorname{argmin}_{\mathbf{d}_m} \hat{F}_n(\mathbf{d}_m) = \mathbf{d}_m^{n-1}$ if the corresponding m th entry of \mathbf{x}_n is unobserved,

i.e. $m \notin \Omega_n$, since

$$\begin{aligned}\hat{F}_n(\mathbf{d}_m) &= \sum_{t=1}^{n-1} \lambda^{n-t} \|\hat{x}_{mt} - \mathbf{d}_m^T \mathbf{a}_t\|_2^2 + \|(\mathbf{d}_m^{n-1} - \mathbf{d}_m)^T \mathbf{a}_t\|_2^2, \\ &= \lambda \hat{F}_{n-1}(\mathbf{d}_m) + \|(\mathbf{d}_m^{n-1} - \mathbf{d}_m)^T \mathbf{a}_t\|_2^2 \\ &\geq \lambda \hat{F}_{n-1}(\mathbf{d}_m^{n-1}) = \hat{F}_n(\mathbf{d}_m^{n-1})\end{aligned}$$

is minimized when $\mathbf{d}_m = \mathbf{d}_m^{n-1}$ for $m \notin \Omega_n$.

This modification is equivalent to the original PETRELS in the full observation scenario, but generally leads to a simplified updating rule for \mathbf{R}_m^n , since now the updating formula for all rows \mathbf{d}_m 's is the same, where $\mathbf{R}_m^n = \mathbf{R}_n = \lambda \mathbf{R}_{n-1} + \mathbf{a}_n \mathbf{a}_n^T$ for all m . The row updating formula (4.24) is replaced by

$$\mathbf{D}_n = \mathbf{D}_{n-1} + \mathbf{P}_n(\mathbf{x}_n - \mathbf{D}_{n-1} \mathbf{a}_n) \mathbf{a}_n^T \mathbf{R}_n^\dagger, \quad (4.33)$$

which further saves storage requirement for the PETRELS algorithm from $\mathcal{O}(Mr^2)$, to store all \mathbf{R}_m^n 's, to $\mathcal{O}(r^2)$. In terms of performance, we analyze its convergence in the rank-one scenario in Section 4.4.2, and compare it with PETRELS numerically in Section 4.5.

4.4 Convergence Analysis

4.4.1 PETRELS in the Full Observation Scenario

In the full observation regime, i.e. $\mathbf{y}_n = \mathbf{x}_n$ for all n , the PETRELS algorithm becomes essentially equivalent to the PAST algorithm [78] for SIT, except that the coefficient is estimated as $\mathbf{a}_n = \mathbf{D}_{n-1}^T \mathbf{y}_n = \mathbf{D}_{n-1}^T \mathbf{x}_n$ in the PAST algorithm versus $\mathbf{a}_n = (\mathbf{D}_{n-1}^T \mathbf{D}_{n-1})^{-1} \mathbf{D}_{n-1}^T \mathbf{x}_n$ in PETRELS. Now let $\lambda = 1$. Similar to PAST in [86], the asymptotic dynamics of the PETRELS algorithm can be described by the ODE below,

$$\dot{\mathbf{R}} = \mathbb{E}[\tilde{\mathbf{a}}_n \tilde{\mathbf{a}}_n^T] - \mathbf{R} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{C}_x \mathbf{D} (\mathbf{D}^T \mathbf{D})^{-1} - \mathbf{R}, \quad (4.34)$$

$$\dot{\mathbf{D}} = \mathbb{E}[\mathbf{x}_n (\mathbf{x}_n - \mathbf{D} \tilde{\mathbf{a}}_n)^T] \mathbf{R}^\dagger = (\mathbf{I} - \mathbf{D} (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) \mathbf{C}_x \mathbf{D} (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{R}^{-1}. \quad (4.35)$$

where $\tilde{\mathbf{a}}_n = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{x}_n$, $\mathbf{R} = \mathbf{R}(t)$ and $\mathbf{D} = \mathbf{D}(t)$ are continuous-time versions of \mathbf{R}_n and \mathbf{D}_n . Now let $\tilde{\mathbf{D}} = \mathbf{D} (\mathbf{D}^T \mathbf{D})^{-1/2}$ and $\tilde{\mathbf{R}} = (\mathbf{D}^T \mathbf{D})^{1/2} \mathbf{R} (\mathbf{D}^T \mathbf{D})^{1/2}$, since from (4.35)

$$\mathbf{D}^T \dot{\mathbf{D}} = \mathbf{D}^T (\mathbf{I} - \mathbf{D} (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T) \mathbf{C}_x \mathbf{D} (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{R}^{-1} = \mathbf{0},$$

we have $\frac{d}{dt}(\mathbf{D}^T \mathbf{D}) = \mathbf{D}^T \dot{\mathbf{D}} + \dot{\mathbf{D}}^T \mathbf{D} = \mathbf{0}$ and further $\frac{d}{dt} f(\mathbf{D}^T \mathbf{D}) = \mathbf{0}$ for any function of $\mathbf{D}^T \mathbf{D}$. Hence,

$$\dot{\tilde{\mathbf{D}}} = \dot{\mathbf{D}}(\mathbf{D}^T \mathbf{D})^{-1/2}, \quad (4.36)$$

$$\dot{\tilde{\mathbf{R}}} = (\mathbf{D}^T \mathbf{D})^{1/2} \dot{\mathbf{R}}(\mathbf{D}^T \mathbf{D})^{1/2}. \quad (4.37)$$

Therefore (4.34) and (4.35) can be rewritten as

$$\begin{aligned} \dot{\tilde{\mathbf{R}}} &= \tilde{\mathbf{D}}^T \mathbf{C}_x \tilde{\mathbf{D}} - \tilde{\mathbf{R}}, \\ \dot{\tilde{\mathbf{D}}} &= (\mathbf{I} - \tilde{\mathbf{D}} \tilde{\mathbf{D}}^T) \mathbf{C}_x \tilde{\mathbf{D}} \tilde{\mathbf{R}}^\dagger, \end{aligned}$$

which is equivalent to the ODE of PAST. Hence we conclude that PETRELS will converge to the global optima in the same dynamic as the PAST algorithm.

4.4.2 Simplified PETRELS in Rank-One Scenario

In the partial observation regime, it is straightforward to see that PETRELS converges to local optima from the gradient descent viewpoint. With respect to convergence to global optima, it is difficult to directly prove performance guarantees for PETRELS due to the asynchronous update for each row of the subspace matrix; however, the simplified version of PETRELS, where all rows are jointly updated at each time, is easier to analyze. In particular, when the underlying subspace is rank one, and the estimated rank of the subspace is also one, we are able to show that the simplified PETRELS algorithm is guaranteed to approach a global optima in each step with high probability with respect to random observations in the noise-free setting.

Let the ground truth of the rank-one subspace be denoted by $\mathbf{d} \in \mathbb{C}^M$ up to a scaling factor. Assume $\mathbf{x}_n = c_n \mathbf{d} \in \mathbb{C}^M$ is the full stream vector at the n th time index. The set Ω_n includes indices of observed entries, which are sampled uniformly with replacement, and its complement is denoted by Ω_n^c . Let $\mathbf{d}_n \in \mathbb{C}^M$ be the updated subspace estimate at time n from PETRELS, we group the entries in \mathbf{d} , \mathbf{d}_{n-1} and \mathbf{d}_n as

$$\mathbf{d} = [\mathbf{d}_{\Omega_n}, \mathbf{d}_{\Omega_n^c}], \quad \mathbf{d}_n = [\mathbf{d}_{\Omega_n}^n, \mathbf{d}_{\Omega_n^c}^n], \quad \mathbf{d}_{n-1} = [\mathbf{d}_{\Omega_n}^{n-1}, \mathbf{d}_{\Omega_n^c}^{n-1}].$$

In simplified PETRELS, the coefficient estimate in (4.16) at time n is given as

$$a_n = \frac{(\mathbf{d}_{\Omega_n}^{n-1})^T \mathbf{d}_{\Omega_n}}{\|\mathbf{d}_{\Omega_n}^{n-1}\|_2^2} c_n,$$

then $\mathbf{d}_{\Omega_n}^n$ is given by (4.33) as

$$\mathbf{d}_{\Omega_n}^n = \frac{1}{a_n} \mathbf{x}_{\Omega_n}^n = \frac{\|\mathbf{d}_{\Omega_n}^{n-1}\|_2^2}{(\mathbf{d}_{\Omega_n}^{n-1})^T \mathbf{d}_{\Omega_n}} \mathbf{d}_{\Omega_n}, \quad (4.38)$$

and $\mathbf{d}_{\Omega_n^c}^n = \mathbf{d}_{\Omega_n^c}^{n-1}$ (i.e. the randomness of the data). Note that \mathbf{d}_n does not depend on the coefficient c_n , so that the performance is solely determined by the initialization \mathbf{d}_0 and the random partial observations.

We are interested in the convergence performance under this special case. Define the squared inner product between \mathbf{d} and \mathbf{d}_n as

$$\gamma_n = \frac{(\mathbf{d}^T \mathbf{d}_n)^2}{\|\mathbf{d}\|_2^2 \|\mathbf{d}_n\|_2^2}. \quad (4.39)$$

Note that γ_n is a lower bounded by 0 and upper bounded by 1, and it will converge to 1 when \mathbf{d}_n converges to \mathbf{d} up to a scalar. The coherence of a vector $\mathbf{d} \in \mathbb{C}^M$ is defined as [81]

$$\mu(\mathbf{d}) = \frac{M \|\mathbf{d}\|_\infty^2}{\|\mathbf{d}\|_2^2}, \quad (4.40)$$

and it is obvious that $1 \leq \mu(\mathbf{d}) \leq M$. With these preparations, we now state the following result.

Theorem 4.4.1. *At step n , $\gamma_n \geq \gamma_{n-1}$ with probability at least*

$$1 - \exp\left(-\frac{\gamma_{n-1} |\Omega_n|}{2 [16\mu(\mathbf{d}) + \gamma_{n-1}\mu(\mathbf{d}_{n-1})] \mu(\mathbf{d}_{n-1})}\right)$$

with respect to uniformly sampled measurements.

Proof. See Appendix 7.3.1. □

Remark: To get $\gamma_n \geq \gamma_{n-1}$ with probability at least $1 - \epsilon$, where $\epsilon > 0$, we let

$$\frac{\gamma_{n-1} |\Omega_n|}{2 [16\mu(\mathbf{d}) + \gamma_{n-1}\mu(\mathbf{d}_{n-1})] \mu(\mathbf{d}_{n-1})} = \log\left(\frac{1}{\epsilon}\right),$$

therefore the number of measurements is bounded as

$$\begin{aligned} |\Omega_n| &= 2 \log\left(\frac{1}{\epsilon}\right) \mu(\mathbf{d}_{n-1}) \frac{16\mu(\mathbf{d}) + \gamma_{n-1}\mu(\mathbf{d}_{n-1})}{\gamma_{n-1}} \\ &\leq 34 \log\left(\frac{1}{\epsilon}\right) \frac{\mu(\mathbf{d}_{n-1})}{\gamma_{n-1}} \max\{\mu(\mathbf{d}_{n-1}), \mu(\mathbf{d})\}. \end{aligned}$$

If $\mu(\mathbf{d}_{n-1})$ is close to $\mu(\mathbf{d})$, then the required number of measurements is approximately $34 \log\left(\frac{1}{\epsilon}\right) \frac{\mu(\mathbf{d})^2}{\gamma_{n-1}}$. For an incoherence vector where $\mu(\mathbf{d})$ is close to 1, then we could see the required number of measurements is proportional to $4 \log\left(\frac{1}{\epsilon}\right)$, decrease per step as inverse proportional to γ_{n-1} , which is irrelevant to the actual dimension of the vector M . This tells that it is possible to track and estimate an underlying subspace from highly incomplete measurements.

The convergence in the general case of partial observation is beyond the scope of this chapter and will be left for future investigations.

4.5 Numerical Results

Our numerical results fall into four parts. First we examine the influence of parameters specified in the PETRELS algorithm, such as discount factor, rank estimation, and its robustness to noise level. Next we look at the problem of direction-of-arrival estimation and show the proposed PETRELS algorithm demonstrates performance superior to GROUSE by identifying and tracking all the targets almost perfectly even in low SNR. Thirdly, we compare our approach with matrix completion, and show that PETRELS is at least competitive with state of the art batch algorithms. Finally, we provide numerical simulations for the extensions of the PETRELS algorithm.

4.5.1 Choice of Parameters

At each time t , a vector \mathbf{x}_t is generated as

$$\mathbf{x}_t = \mathbf{D}_{true} \mathbf{a}_t + \mathbf{n}_t, t = 1, 2, \dots \quad (4.41)$$

where \mathbf{D}_{true} is an r -dimensional subspace generated with i.i.d. $\mathcal{N}(0, 1)$ entries, \mathbf{a}_t is an $r \times 1$ vector with i.i.d. $\mathcal{N}(0, 1)$ entries, and \mathbf{n}_t is an $m \times 1$ Gaussian noise vector with i.i.d. $\mathcal{N}(0, \epsilon^2)$ entries. We further fix the signal dimension $m = 500$ and the subspace rank $r_{true} = 10$. We assume that a fixed number of entries in \mathbf{x}_t , denoted by K , are revealed each time. This restriction is not necessary for the algorithm to work as shown in matrix completion simulations, but we make it here in order to get a meaningful estimate of \mathbf{a}_t . Denoting the estimated subspace by $\hat{\mathbf{D}}$, we use the normalized subspace reconstruction error to examine the algorithm performance; this is calculated as $\|\mathcal{P}_{\hat{\mathbf{D}}_{\perp}} \mathbf{D}_{true}\|_F^2 / \|\mathbf{D}_{true}\|_F^2$, where $\mathcal{P}_{\hat{\mathbf{D}}_{\perp}}$ is the projector for the orthogonal subspace $\hat{\mathbf{D}}_{\perp}$.

The choice of discount factor λ plays an important role in how fast the algorithm converges. With $K = 50$, a mere 10% percent of the full dimension, the rank is estimated accurately as $r = 10$ in a noise-free setting where $\epsilon = 0$. We run the algorithm to time $n = 2000$ for the same data, the normalize subspace reconstruction error is minimized when λ is around 0.98 in Fig. 4.2. Hence, we will keep $\lambda = 0.98$ hereafter.

In reality it is almost impossible to accurately estimate the intrinsic rank in advance. Fortunately the convergence rate of our algorithm degrades gracefully as the rank estimation error increases. In Fig. 4.3, the evolution of normalized subspace error is plotted against data stream index, for rank estimation $r = 10, 12, 14, 16, 18$. We only examine the rank over-estimation case here since this is usually the case in applications, and the readers are referred to the next section for examples for the case of rank underestimation.

Taking more measurements per time leads to faster convergence since it is approaching the full information regime, as shown in Fig. 4.4. Theoretically it requires $M \sim \mathcal{O}(r \log r) \approx 23$ measurements to test if an incomplete vector is within a subspace of rank r [83], and the simulation shows our algorithm can work even when M is close to this lower bound.

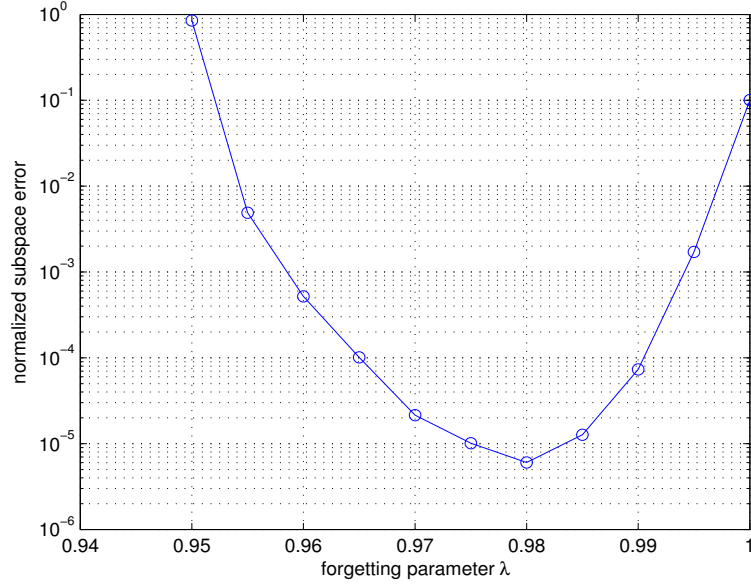


Figure 4.2: The normalized subspace reconstruction error as a function of the discount factor λ after running the algorithm to time $n = 2000$ when 50 out of 500 entries of the signal are observed each time without noise.

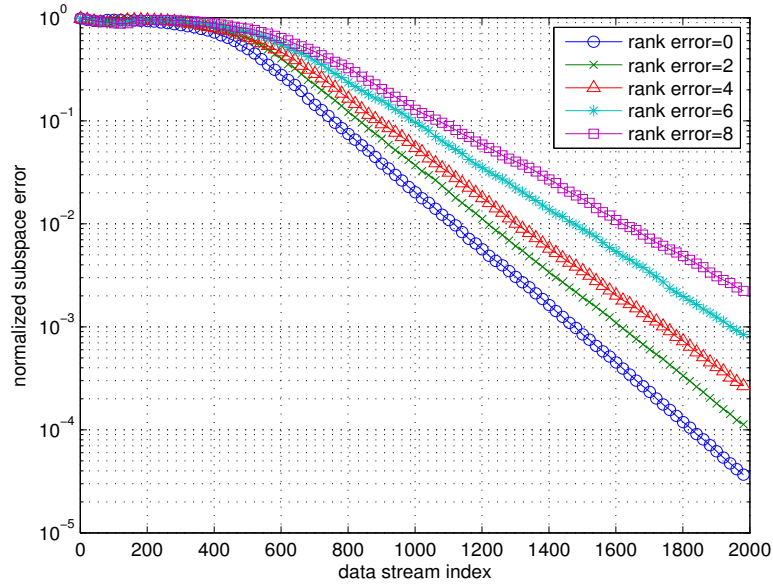


Figure 4.3: The normalized subspace reconstruction error as a function of data stream index when the rank is over-estimated when 50 out of 500 entries of the signal are observed each time without noise.

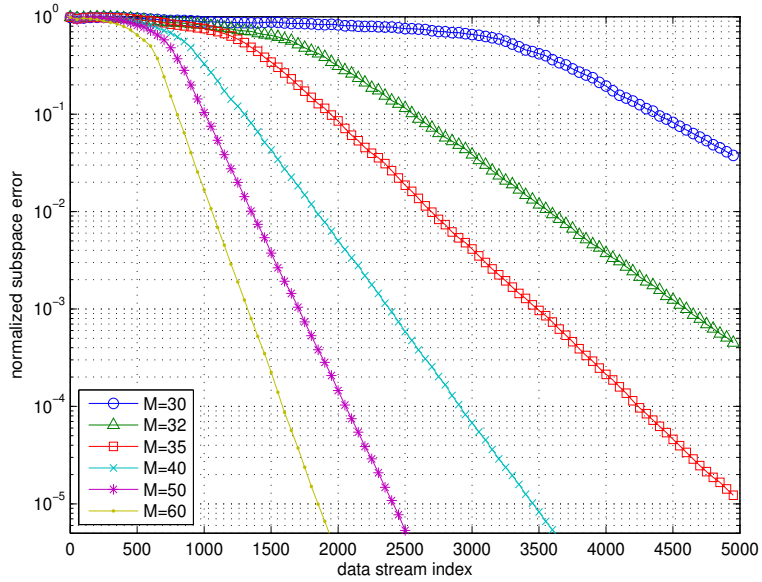


Figure 4.4: The normalized subspace reconstruction error as a function of data stream index when the number of entries observe per time M out of 500 entries are varied with accurate rank estimation and no noise.

Finally the robustness of the algorithm is tested against the noise variance ϵ^2 in Fig. 4.5, where the normalized subspace error is plotted against data stream index for different noise level ϵ .

We now consider a scenario where a subspace of rank $r = 10$ changes abruptly at time index $n = 3000$ and $n = 5000$, and examine the performance of GROUSE [84] and PETRELS in Fig. 4.6 when the rank is over-estimated by 4 and the noise level is $\epsilon = 10^{-3}$. The normalized residual error for data stream and normalized subspace error are shown respectively in Fig. 4.6 (a) and (b). Both PETRELS and GROUSE can successfully track the changed subspace, but PETRELS can track the change faster.

4.5.2 Direction-Of-Arrival Analysis

Given GROUSE [84] as a baseline, we evaluate the resilience of our algorithm to different data models and applications. We use the following example of Direction-Of-Arrival analysis in array processing to compare the performance of these two methods. Assume there are $n = 256$ sensors from a linear array, and the measurements from all sensors at time t are given as

$$\mathbf{x}_t = \mathbf{V}\Sigma\mathbf{a}_t + \mathbf{n}_t, t = 1, 2, \dots \quad (4.42)$$

where $\mathbf{V} \in \mathbb{C}^{n \times p}$ is a Vandermonde matrix given as

$$\mathbf{V} = [\boldsymbol{\alpha}_1(\omega_1), \dots, \boldsymbol{\alpha}_p(\omega_p)], \quad (4.43)$$

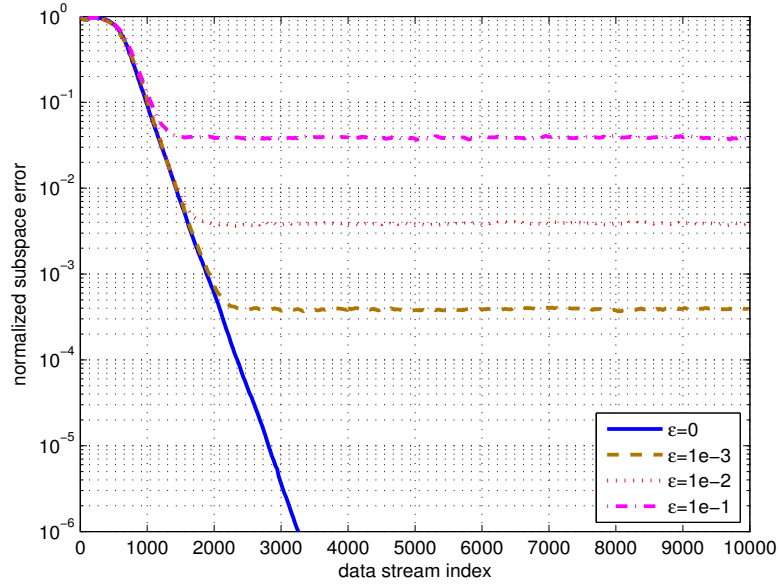


Figure 4.5: The normalized subspace error against data stream index with different noise level ϵ when 50 out of 500 entries of the signal are observed each time with accurate rank estimation $r = 10$.

where $\mathbf{\alpha}_i(\omega_i) = [1, e^{j2\pi\omega_i}, \dots, e^{j2\pi\omega_i(n-1)}]^T$, $0 \leq \omega_i < 1$; $\mathbf{\Sigma} = \text{diag}\{\mathbf{d}\} = \text{diag}\{d_1, \dots, d_p\}$ is a diagonal matrix which characterizes the amplitudes of each mode. The coefficients \mathbf{a}_t are generated with $\mathcal{N}(0, 1)$ entries, and the noise is generated with $\mathcal{N}(0, \epsilon^2)$ entries, where $\epsilon = 0.1$.

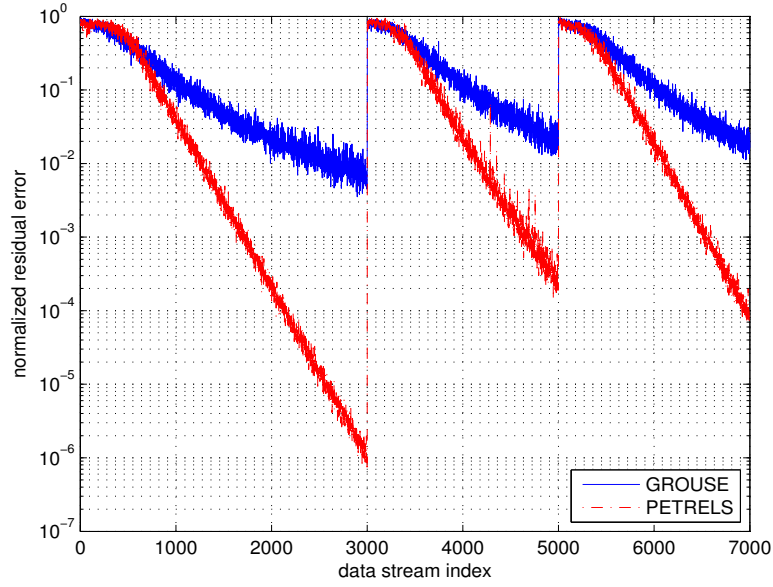
Each time we collect measurements from $K = 30$ random sensors. We are interested in identifying all $\{\omega_i\}_{i=1}^p$ and $\{d_i\}_{i=1}^p$. This can be done by applying the well-known ESPRIT algorithm [41] to the estimated subspace $\hat{\mathbf{D}}$ of rank r , where r is specified a-priori corresponding to the number of modes to be estimated. Specifically, if $\mathbf{D}_1 = \hat{\mathbf{D}}(1 : n - 1)$ and $\mathbf{D}_2 = \hat{\mathbf{D}}(2 : n)$ are the first and the last $n - 1$ rows of $\hat{\mathbf{D}}$, then from the eigenvalues of the matrix $\mathbf{T} = \mathbf{D}_1^\dagger \mathbf{D}_2$, denoted by λ_i , $i = 1, \dots, r$, the set of $\{\omega_i\}_{i=1}^p$ can be recovered as

$$\omega_i = \frac{1}{2\pi} \arg \lambda_i, \quad i = 1, \dots, r. \quad (4.44)$$

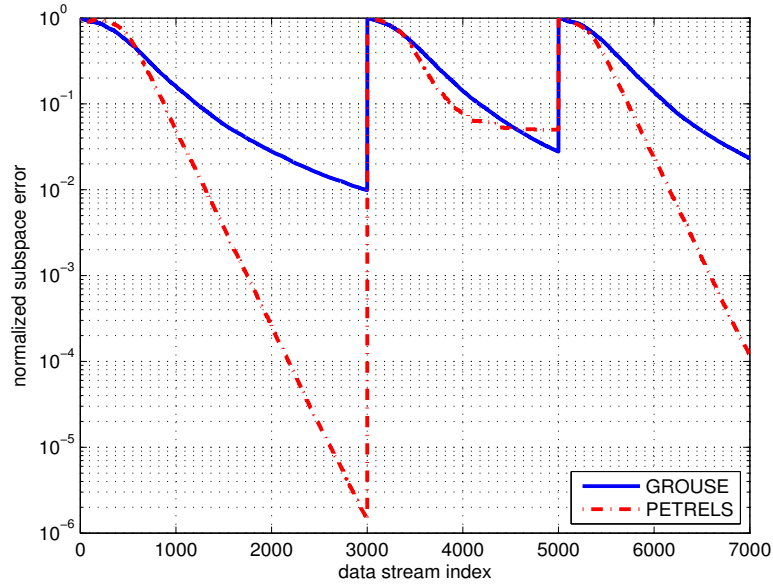
In a dynamic setting when the underlying subspace is varying, the proposed PE-TRELS algorithm does a better job of discarding out-of-date modes and picking up new ones. We divide the running time into 4 parts, and the frequencies and amplitudes are specified as follows:

1. Start with the same frequencies

$$\omega = [0.1769, 0.1992, 0.2116, 0.6776, 0.7599];$$



(a) Normalized residual error



(b) Normalized subspace error

Figure 4.6: The normalized subspace error when the underlying subspace is changing with fixed rank $r = 10$. The rank is over-estimated by 4 and the noise level is $\epsilon = 10^{-3}$, when 50 out of 500 entries of the signal are observed each time for both GROUSE and PETRELS.

and amplitudes

$$d = [0.3, 0.8, 0.5, 1, 0.1].$$

2. Change two modes (only frequencies) at stream index 1000:

$$\omega = [0.1769, 0.1992, \mathbf{0.4116}, 0.6776, \mathbf{0.8599}];$$

and amplitudes

$$d = [0.3, 0.8, 0.5, 1, 0.1].$$

3. Add one new mode at stream index 2000:

$$\omega = [0.1769, 0.1992, 0.4116, 0.6776, 0.8599, \mathbf{0.9513}];$$

and amplitudes

$$d = [0.3, 0.8, 0.5, 1, 0.1, \mathbf{0.6}].$$

4. Delete the weakest mode at stream index 3000:

$$\omega = [0.1769, 0.1992, 0.4116, 0.6776, 0.9513];$$

and amplitudes

$$d = [0.3, 0.8, 0.5, 1, 0.6].$$

Fig. 4.7 shows the ground truth of mode locations and amplitudes for the scenario above. Note that there are three closely located modes and one weak mode in the beginning, which makes the task challenging. We compare the performance of the proposed PETRELS algorithm and GROUSE. The rank specified in both algorithms is $r = 10$, which is the estimated number of modes; in our case it is twice the number of true modes, and the estimated directions at each time for 10 modes are shown against the data stream index in Fig. 4.8. The color shows the amplitude corresponding to the color bar. The proposed PETRELS algorithm identifies all modes correctly. In particular it distinguishes the three closely-spaced modes perfectly in the beginning, and identifies the appearance and disappearance of the later weak mode.

The auxiliary modes are exhibited as “noise” in the scatter plot. With GROUSE the closely spaced nodes are erroneously estimated as one mode, the weak mode is missing, and spurious modes have been introduced. The PETRELS algorithm also fully tracked the later changes in accordance with the entrance and exit of each mode, while GROUSE is not able to react to changes in the data model.

4.5.3 Matrix Completion

We compare performance of the proposed PETRELS algorithm on matrix completion against batch algorithms LMaFit [90], FPCA [91], Singular Value Thresholding (SVT) [92], OptSpace [82] and GROUSE [84]. The low-rank matrix is generated from a matrix factorization model with $\mathbf{X} = \mathbf{UV}^T \in \mathbb{R}^{1000 \times 2000}$, where $\mathbf{U} \in \mathbb{R}^{1000 \times 10}$ and $\mathbf{V} \in \mathbb{R}^{2000 \times 10}$, all entries in \mathbf{U} and \mathbf{V} are generated from standard normal distribution

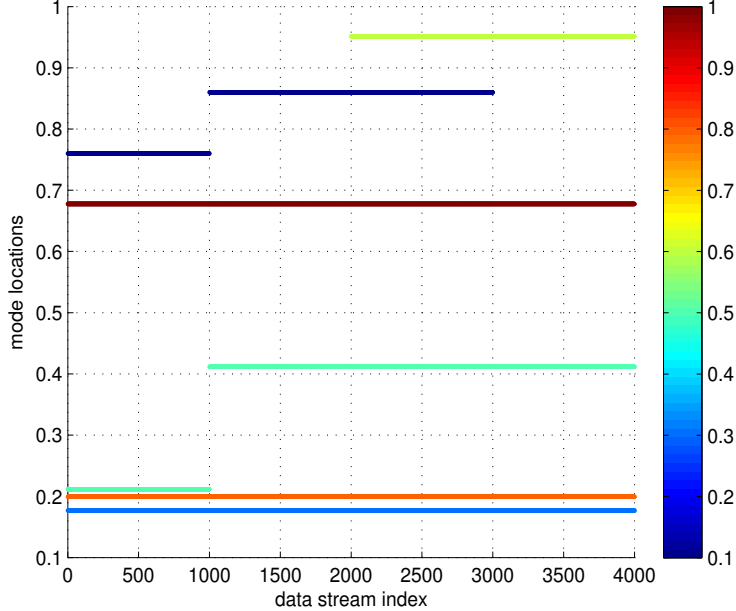


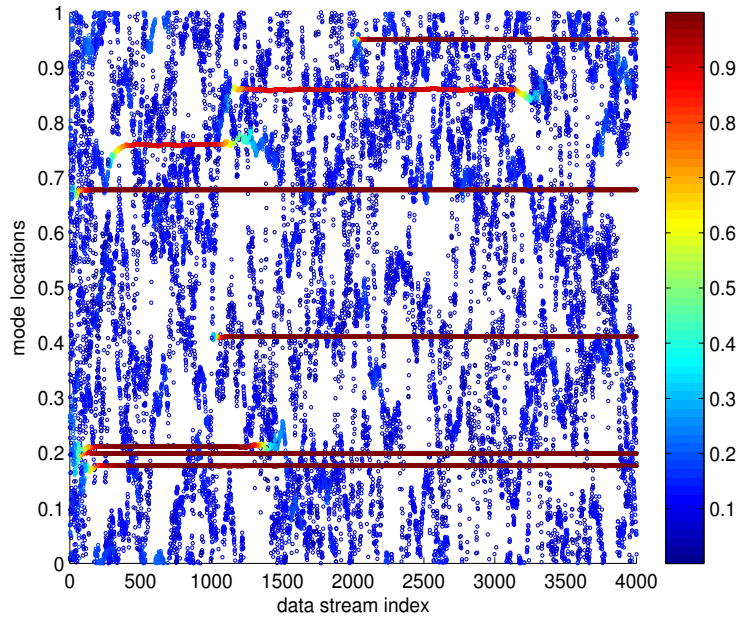
Figure 4.7: Ground truth of the actual mode locations and amplitudes in a dynamic scenario.

$\mathcal{N}(0, 1)$ (Gaussian data) or uniform distribution $\mathcal{U}[0, 1]$ (uniform data). The sampling rate is taken to be 0.05, so only 5% of all entries are revealed.

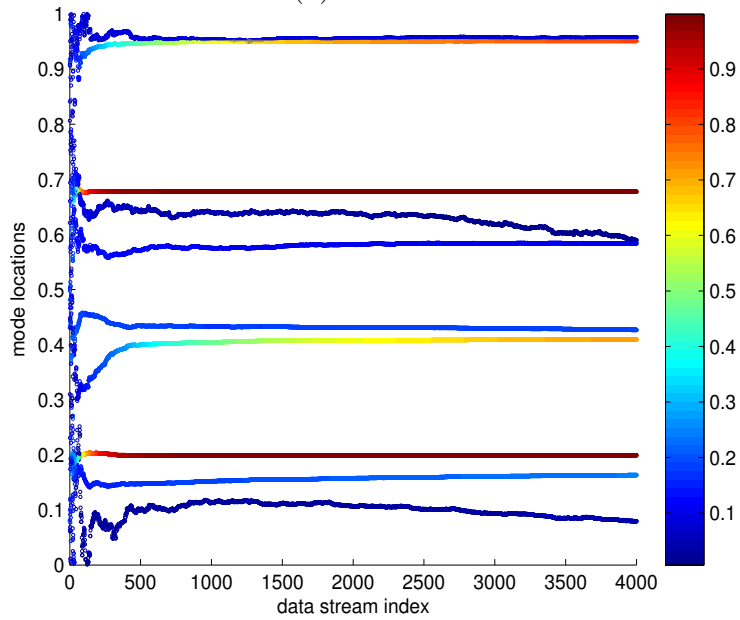
The running time is plotted against the normalized matrix reconstruction error, calculated as $\|\hat{\mathbf{X}} - \mathbf{X}\|_F / \|\mathbf{X}\|_F$, where $\hat{\mathbf{X}}$ is the reconstructed low-rank matrix for Gaussian data and uniform data respectively in Fig. 4.9 (a) and (b). The proposed PETRELS algorithm matches the performance of batch algorithms on Gaussian data and improves upon the accuracy of most algorithms on uniform data, where the Grassmannian-based optimization approach may encounter “barriers” for its convergence. Note that different algorithms have different input parameter requirements. For example, OptSpace needs to specify the tolerance to terminate the iterations, which directly decides the trade-off between accuracy and running time; PETRELS and GROUSE require an initial estimate of the rank. Our simulation here only shows one particular realization and we simply conclude that PETRELS is competitive.

4.5.4 Simplified PETRELS

Under the same simulation setup as for Fig. 4.3 except that the subspace of rank 10 is generated by $\hat{\mathbf{D}}_{true} = \mathbf{D}_{true}\mathbf{\Sigma}$, where $\mathbf{\Sigma}$ is a diagonal matrix with 5 entries from $\mathcal{N}(0, 1)$ and 5 entries from $0.01 \cdot \mathcal{N}(0, 1)$, we examine the performance of the simplified PETRELS algorithm (with $\lambda = 0.9$) in Section 4.3.3 A and the original PETRELS (with $\lambda = 0.98$) algorithm when the rank of the subspace is over-estimated as 12 or under-estimated as 8. When the rank of \mathbf{D} is over-estimated, the change in (4.10) will introduce more errors and converges slower compared with the original



(a) PETRELS



(b) GROUSE

Figure 4.8: Tracking of mode changes in direction-of-arrival estimation using PETRELS and GROUSE algorithms: the estimated directions at each time for 10 modes are shown against the data stream. All changes are identified and tracked successfully by PETRELS, but not by GROUSE.

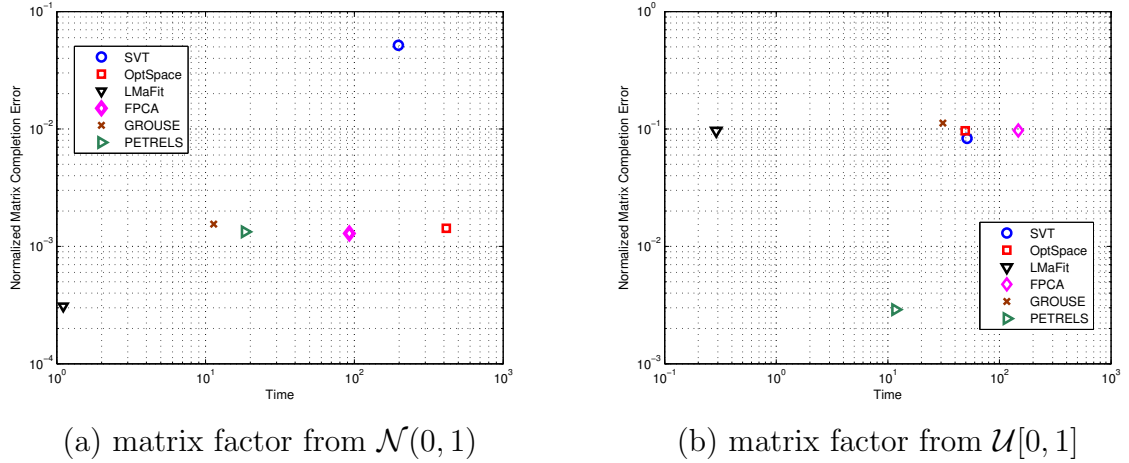


Figure 4.9: Comparison of matrix completion algorithms in terms of speed and accuracy: PETRELS is a competitive alternative for matrix completion tasks.

PETRELS algorithm; however, when the rank of \mathbf{D} is under-estimated, the simplified PETRELS performs better than PETRELS.

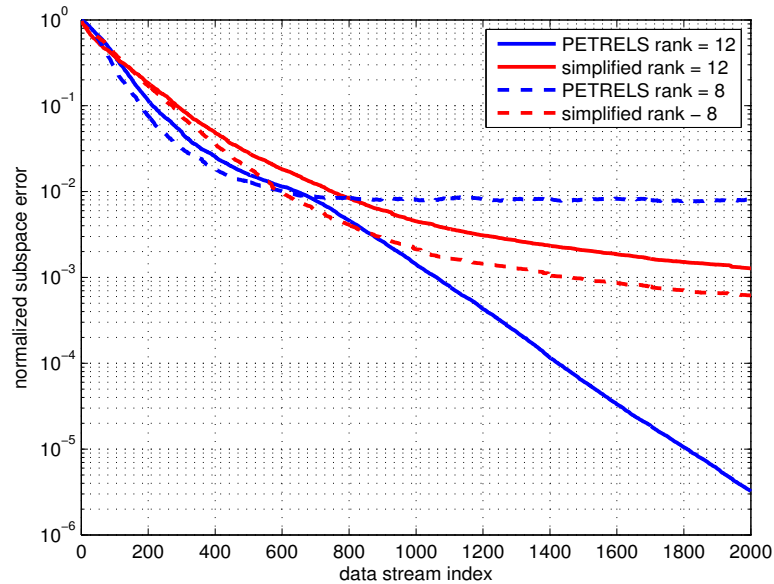


Figure 4.10: The normalized subspace reconstruction error against data stream index when the rank is over-estimated as 12 or under-estimated as 8 for the original PETRELS and modified algorithm.

Chapter 5

Waveform Coordination for Range Sidelobe Suppression in Radar

5.1 Introduction

Phase coding [93] is a common technique in radar for constructing waveforms with an impulse-like autocorrelation function. In this technique, a long pulse is phase coded with a unimodular sequence and the autocorrelation function of the coded waveform is controlled via the autocorrelation function of the unimodular sequence. A key issue in phase coding is the presence of range sidelobes in the ambiguity function of the coded waveforms. Range sidelobes due to a strong reflector can result in masking of nearby weak targets. It is however impossible to design a single unimodular sequence for which the autocorrelation function has no range sidelobes. This has led to the idea of using complementary sets of unimodular sequences [94]–[98] for phase coding.

Perhaps the most famous class of complementary sequences are Golay complementary sequences or Golay pairs introduced by Marcel Golay [94]. Golay complementary sequences have the property that the sum of their autocorrelation functions vanishes at all delays other than zero. Thus, if each sequence is transmitted separately and the autocorrelation functions are added together the output will be free of range sidelobes. In other words, the effective ambiguity function of a Golay pair of phase coded waveforms is free of range sidelobes along the zero-Doppler axis. However, this ideal property is very sensitive to Doppler effect. Off the zero-Doppler axis the ambiguity function of Golay pairs of phase coded waveforms has large range sidelobes, e.g., see [93],[99]. The ambiguity function of a pulse train of Golay complementary waveforms, in which the two waveforms are transmitted alternatively in time over several Pulse Repetition Intervals (PRIs), suffers from the same problem. The sensitivity of Golay complementary waveforms to Doppler has been a major barrier in adoption of these waveforms for radar pulse compression. Various generalizations of complementary waveforms, including multiple complementary waveforms [98], multiphase (or polyphase) complementary waveforms [100], and near-complementary waveforms [99],[101] suffer from a similar problem.

A natural question to ask is whether or not it is possible to construct a *Doppler resilient* pulse train of Golay complementary waveforms, for which the range sidelobes of the pulse train ambiguity function vanish inside a desired Doppler interval. This question was recently considered in [102, 103], where it is shown that by carefully choosing the order in which a Golay pair of phase coded waveforms is transmitted over time we can clear out the range sidelobes of the pulse train ambiguity function along modest (close to zero) Doppler shifts. More specifically, it is shown that range sidelobes along modest Doppler shifts can be significantly suppressed if the transmission of the two waveforms in pair is coordinated according to the locations of 1's and -1 's in a *Prouhet-Thue-Morse (PTM) sequence* defined over $\{1, -1\}$. The resulting pulse train is called a *PTM pulse train* and its ambiguity function has a high-order null along the zero-Doppler axis. This led to the discovery that if the transmission of a Golay pair of phase coded waveforms is coordinated in time according to the entries in a biphasic sequence then the magnitude of the range sidelobes can be controlled by shaping the spectrum of the biphasic sequence [104].

The rest of the chapter is organized as follows. Section 5.2 describes the Golay complementary waveforms. In Section 5.3, we extend the result of [102]–[104] to construct pulse trains of Golay complementary waveforms, for which the range sidelobes of the ambiguity function vanish inside a desired Doppler interval away from zero. This is accomplished by coordinating the transmission of a Golay pair of phase coded waveforms in time according to the 1's and -1 's in a $(2^M, m)$ -PTM sequence. The $(2^M, m)$ -PTM sequence has length $2^M \times m$ and is obtained by repeating each 1 and -1 in the length- 2^M PTM sequence m times, that is by oversampling a length- 2^M PTM sequence by a proper factor m determined by the Doppler region of interests. Finally in Section 5.4, we present a new radar primitive that enables instantaneous radar polarimetry at essentially no increase in signal processing complexity, by the transmission of Golay complementary waveforms across orthogonal polarization channels and multiple antennas over time in an Alamouti fashion to construct a unitary waveform matrix, and discuss its Doppler resilience via generalized PTM sequencing for a four-antenna MIMO radar.

5.2 Golay Complementary Waveforms

Definition 5.2.1. *Two length L unimodular sequences of complex numbers $x(\ell)$ and $y(\ell)$ are Golay complementary if for $k = -(L - 1), \dots, (L - 1)$ the sum of their autocorrelation functions satisfies*

$$C_x(k) + C_y(k) = 2L\delta(k), \quad (5.1)$$

where $C_x(k)$ is the autocorrelation of $x(\ell)$ at lag k and $\delta(k)$ is the Kronecker delta function. Henceforth we may drop the discrete time index ℓ from $x(\ell)$ and $y(\ell)$ and simply use x and y . Each member of the pair (x, y) is called a Golay sequence.

The baseband waveform $s_x(t)$ phase coded by the Golay sequence x is given by

$$s_x(t) = \sum_{\ell=0}^{L-1} x(\ell)\Omega(t - \ell T_c), \quad (5.2)$$

where $\Omega(t)$ is a unit energy pulse shape supported mostly on $(0, T_c)$ and T_c is the chip length, i.e. $\int_0^{T_c} \Omega(t)^2 dt \approx 1$. The ambiguity function $\chi_{s_x}(\tau, \nu)$ of $s_x(t)$ is given by

$$\begin{aligned} \chi_{s_x}(\tau, \nu) &= \int_{-\infty}^{\infty} s_x(t) \overline{s_x(t - \tau)} e^{-j\nu t} dt \\ &= \sum_{k=-(L-1)}^{L-1} A_x(k, \nu T_c) \chi_{\Omega}(\tau - k T_c, \nu) \\ &= A_x(k_1(\tau), \nu T_c) \chi_{\Omega}(\tau - k_1(\tau) T_c, \nu) + A_x(k_2(\tau), \nu T_c) \chi_{\Omega}(\tau - k_2(\tau) T_c, \nu), \end{aligned} \quad (5.3)$$

where $k_1(\tau) = \lceil \frac{\tau}{T_c} \rceil - 1$, $k_2(\tau) = \lceil \frac{\tau}{T_c} \rceil = k_1 + 1$, $\overline{s(t)}$ is the complex conjugate of $s(t)$, $\chi_{\Omega}(\tau, \nu)$ is the ambiguity function of the pulse shape $\Omega(t)$, and $A_x(k, \nu T_c)$ is given by

$$A_x(k, \nu T_c) = \sum_{\ell=-(L-1)}^{L-1} x[\ell] \overline{x[\ell - k]} e^{-j\nu \ell T_c}. \quad (5.4)$$

At $\nu = 0$, $A_x(k, 0) = C_x(k)$ is the autocorrelation of the sequence x . The last equation of (5.3) follows from that $\Omega(t)$ has finite support.

The baseband waveform $s_x(t)$ is modulated by a carrier of frequency ω_c and transmitted. It reflects back from a target after a round-trip delay τ_0 and Doppler shift ν_0 . The received signal is demodulated by ω_c , low-passed and the baseband waveform is given by

$$r_x(t) = A s_x(t - \tau_0) e^{j\nu_0 t} e^{j\phi}, \quad (5.5)$$

where A is the scattering coefficient, $\phi = \omega_c \tau_0$ is the phase shift in the carrier due to propagation delay, $\nu_0 = 2\omega_c v/c$ is the Doppler shift, where v is the velocity of the target. The received signal goes through a matched-filter bank matched to the signal $s_x(t - \tau) e^{j\nu t}$, in which the delay τ and Doppler shift ν may vary. The output is equivalent to sampling the cross-ambiguity function

$$\begin{aligned} \chi_{r_x s_x}(\tau, \nu) &= \int_{-\infty}^{\infty} r_x(t) \overline{s_x(t - \tau)} e^{-j\nu t} dt \\ &= A e^{j\phi} e^{-j(\nu - \nu_0)\tau_0} \chi_{s_x}(\tau - \tau_0, \nu - \nu_0). \end{aligned} \quad (5.6)$$

We can think of the individual delay τ_0 and Doppler shift ν_0 associated with scatterers as specifying points in a two-dimensional image with coordinates of delay τ and Doppler shift ν . When the received signal is processed by a matched filter

to a complex conjugate replica of $s_x(t - \tau_0)e^{j\nu_0 t}$, the output is maximized since the ambiguity function $\chi_{s_x}(\tau, \nu)$ is maximized at $(0, 0)$. From (5.6), it also implies that by using a single pulse it is impossible to achieve the ideal thumbtack output, i.e. a sharp response at the correct delay and Doppler shift, and zero otherwise.

If the complementary waveforms $s_x(t)$ and $s_y(t)$ are transmitted separately in time, with a length- T time interval between the two transmissions. The ambiguity function of the radar waveform $S(t) = s_x(t) + s_y(t - T)$ is given by

$$\chi_S(\tau, \nu) = \chi_{s_x}(\tau, \nu) + e^{-j\nu T} \chi_{s_y}(\tau, \nu) + \chi_{s_x s_y}(\tau + T, \nu) + e^{-j\nu T} \chi_{s_y s_x}(\tau - T, \nu). \quad (5.7)$$

The cross-ambiguity terms only appear at $\pm T$, therefore can be omitted by truncation on the delay-Doppler plane. However, in practice (5.7) will not be computed directly. Instead the two non-offset terms will be computed separately by matching to $s_x(t)$ and $s_y(t)$ and combined. After all, the sum of the match filtered returns gives the ambiguity function

$$\chi_S(\tau, \nu) = \chi_{s_x}(\tau, \nu) + e^{j\nu T} \chi_{s_y}(\tau, \nu). \quad (5.8)$$

Remark: Here, we ignore the range aliasing effects and only focus on the *mainlobe* of the ambiguity function, which corresponds to $\chi_S(\tau, \nu)$ given in (5.8). Range aliasing effects can be accounted for using standard techniques devised for this purpose (e.g. see [93]) and hence will not be further discussed.

Since the chip length T_c is typically very small, the relative Doppler shift over chip intervals is negligible compared to the relative Doppler shift over the PRI T , and hence the ambiguity function $\chi_S(\tau, \nu)$ can be approximated by

$$\chi_S(\tau, \nu) = \sum_{k=-(L-1)}^{L-1} [C_x(k) + e^{j\nu T} C_y(k)] \chi_\Omega(\tau - kT_c, \nu), \quad (5.9)$$

where in this approximation we have replaced $A_x(k, \nu T_c)$ and $A_y(k, \nu T_c)$ with the autocorrelation functions $C_x(k)$ and $C_y(k)$, respectively.

Along the zero-Doppler axis ($\nu = 0$), the ambiguity function $\chi_S(\tau, \nu)$ reduces to the autocorrelation sum

$$\chi_S(\tau, 0) = \sum_{k=-(L-1)}^{L-1} [C_x(k) + C_y(k)] \chi_\Omega(\tau - kT_c, 0) = 2L \chi_\Omega(\tau, 0) \quad (5.10)$$

which is “free” of range sidelobes.¹ Fig. 5.1(a) shows the perfect autocorrelation property of a Golay pair of phase coded waveforms. Off the zero-Doppler axis however, the ambiguity function has large sidelobes in delay (range) as Fig. 5.1(b) shows. The color bar values are in dB.

The range sidelobes persist even when a pulse train is constructed in which $S(t) = s_x(t) + s_y(t - T)$ is transmitted several times, with a pulse repetition interval PRI of T

¹The shape of the autocorrelation function depends on the autocorrelation function $\chi_\Omega(\tau, 0)$ for the pulse shape $\Omega(t)$. The Golay complementary property eliminates range sidelobes caused by replicas of $\chi_\Omega(\tau, 0)$ at nonzero integer delays.

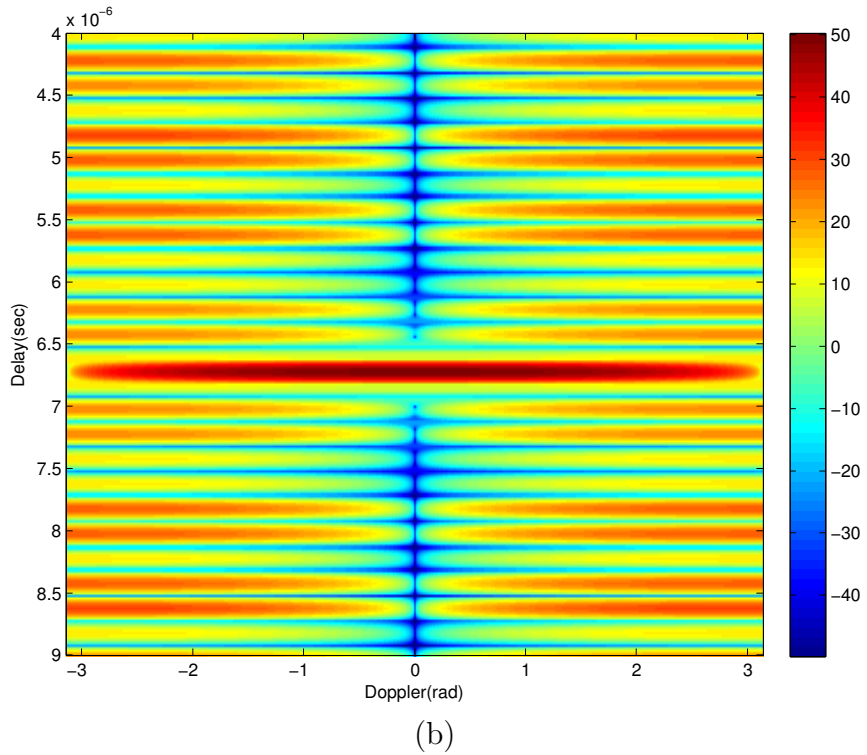
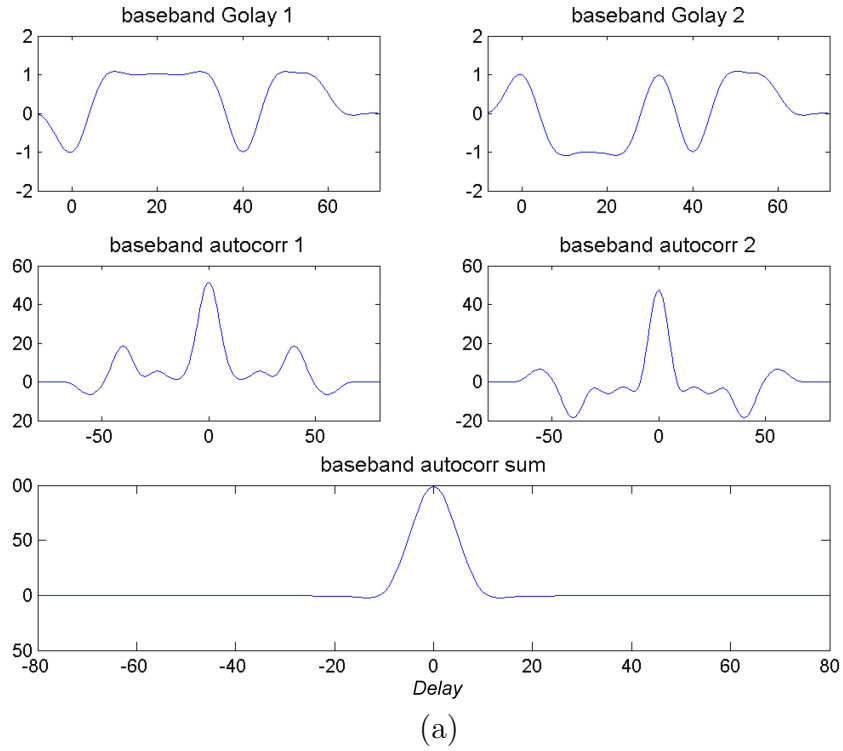


Figure 5.1: (a) The perfect autocorrelation property of a Golay pair of phase-coded waveforms; (b) The ambiguity function of a Golay pair of phase-coded waveforms separated in time.

seconds between consecutive transmissions. In other words, the ambiguity function of a conventional pulse train of Golay complementary waveforms, where the transmitter alternates between $s_x(t)$ and $s_y(t)$ during several PRIs, also has large range sidelobes along nonzero Dopplers. The range sidelobes in the ambiguity function can cause masking of a weak target that is situated near a strong reflector. Fig. 5.2 shows the delay-Doppler map at the output of a radar receiver (matched filter), when an alternating pulse train of Golay complementary waveforms was transmitted over $N = 256$ PRIs. The radar scene contains three stationary reflectors at different ranges and two slow-moving targets, which are 30dB weaker than the stationary reflectors. We notice that the range sidelobes from the strong reflectors make it difficult to resolve the weak targets. The sensitivity of Golay complementary waveforms to Doppler effect has been the main barrier in adopting these waveforms for radar pulse compression.

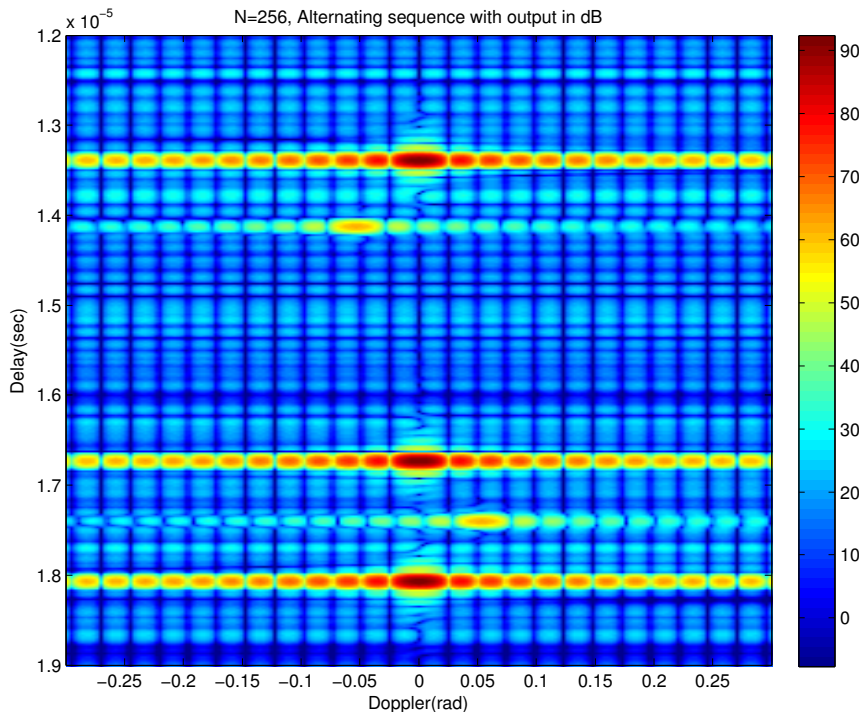


Figure 5.2: Ambiguity function of a length-256 alternating pulse train of Golay complementary waveforms.

5.3 Doppler Resilience of Golay Waveforms

5.3.1 Range Sidelobe Suppression in Small Doppler

It is natural to ask whether or not it is possible to construct a *Doppler resilient* pulse train of Golay complementary waveforms, for which the range sidelobes of the pulse train ambiguity function vanish inside a desired Doppler interval. The developments

in [102] led to the introduction of \mathcal{P} -pulse trains in [104] and the discovery that if the transmission of a Golay pair of phase coded waveforms is coordinated in time according to the 1's and -1 's in a biphasic sequence \mathcal{P} , then magnitude of the range sidelobes can be controlled by shaping the spectrum of \mathcal{P} .

Definition 5.3.1. Consider a biphasic sequence $\mathcal{P} = \{p_n\}_{n=0}^{N-1}$, $p_n \in \{-1, 1\}$ of length N , where N is even. Let 1 represent $s_x(t)$ and let -1 represent $s_y(t)$. The \mathcal{P} -pulse train $Z_{\mathcal{P}}(t)$ of $(s_x(t), s_y(t))$ is defined as

$$Z_{\mathcal{P}}(t) = \frac{1}{2} \sum_{n=0}^{N-1} [(1 + p_n)s_x(t - nT) + (1 - p_n)s_y(t - nT)]. \quad (5.11)$$

The n th entry in the pulse train is $s_x(t)$ if $p_n = 1$ and it is $s_y(t)$ if $p_n = -1$. Consecutive entries in the pulse train are separated in time by a PRI T .

The ambiguity function of the \mathcal{P} -pulse train $Z_{\mathcal{P}}(t)$, after ignoring range aliases which are offset from the zero-delay axis by integer multiples of T by similar discussions in the previous section, is given by

$$\chi_{Z_{\mathcal{P}}}(\tau, \nu) = \frac{1}{2} [\chi_{s_x}(\tau, \nu) + \chi_{s_y}(\tau, \nu)] \sum_{n=0}^{N-1} e^{jn\nu T} + \frac{1}{2} [\chi_{s_x}(\tau, \nu) - \chi_{s_y}(\tau, \nu)] \sum_{n=0}^{N-1} p_n e^{jn\nu T}.$$

We discretize $\chi_{Z_{\mathcal{P}}}(\tau, \nu)$ by letting $\tau = kT_c$ and $\nu = \theta/T$, and ignore the pulse shape ambiguity function, the above equation can be further written as

$$\begin{aligned} \chi_{Z_{\mathcal{P}}}(k, \theta) &= \frac{1}{2} [A_x(k, \theta) + A_y(k, \theta)] \sum_{n=0}^{N-1} e^{jn\theta} + \frac{1}{2} [A_x(k, \theta) - A_y(k, \theta)] \sum_{n=0}^{N-1} p_n e^{jn\theta} \\ &= \frac{1}{2} [C_x(k) + C_y(k)] \sum_{n=0}^{N-1} e^{jn\theta} + \frac{1}{2} [C_x(k) - C_y(k)] \sum_{n=0}^{N-1} p_n e^{jn\theta}, \end{aligned} \quad (5.12)$$

where $\theta = \nu T$ is the relative Doppler shift over a PRI, and we have replaced $A_x(k, \theta)$ and $A_y(k, \theta)$ with $C_x(k)$ and $C_y(k)$ again.

The first term on the RHS of (5.12) is free of range sidelobes due to the complementary property of Golay sequences x and y . The second term represents the range sidelobes, as $C_x(k) - C_y(k)$ is not an impulse. The magnitude of the range sidelobes is proportional to the magnitude of the spectrum $S_{\mathcal{P}}(\theta)$ of the sequence \mathcal{P} , which is given by

$$S_{\mathcal{P}}(\theta) = \sum_{n=0}^{N-1} p_n e^{jn\theta}. \quad (5.13)$$

The question is how to design the sequence \mathcal{P} to suppress the range sidelobes along a desired Doppler interval. One way to accomplish this is to design the sequence \mathcal{P} so that its spectrum $S_{\mathcal{P}}(\theta)$ has a high-order null at a Doppler frequency inside the desired interval. This idea has been explored in [102]–[104], where it is shown that

the spectrum of a PTM sequence of length 2^{M+1} has an M th-order null at $\theta = 0$. The PTM sequence is formally defined as follows.

Definition 5.3.2. [105]-[108] *The Prouhet-Thue-Morse (PTM) sequence $\mathcal{P} = (p_k)_{k \geq 0}$ over $\{-1, 1\}$ is defined by the following recursions:*

1. $p_0 = 1$
2. $p_{2k} = p_k$
3. $p_{2k+1} = \bar{p}_k = -p_k$

for all $k > 0$.

Example 5.3.1. *The PTM sequence of length 8 is given as*

$$\mathcal{P} = (p_k)_{k=0}^7 = +1 \ -1 \ -1 \ +1 \ -1 \ +1 \ +1 \ -1.$$

The corresponding pulse train of Golay complementary waveforms is given by

$$\begin{aligned} Z_{\mathcal{P}}(t) = & s_x(t) + s_y(t - T) + s_y(t - 2T) + s_x(t - 3T) \\ & + s_y(t - 4T) + s_x(t - 5T) + s_x(t - 6T) + s_y(t - 7T). \end{aligned}$$

The ambiguity function of $Z_{\mathcal{P}}(t)$ has a second-order null along the zero-Doppler axis.

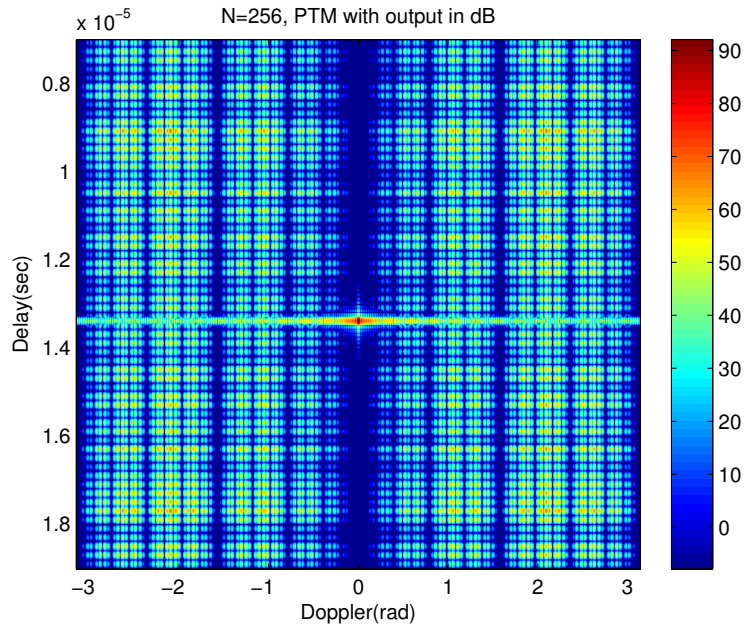
The PTM sequence was originally discovered by Prouhet [105] for partitioning a set of integers into disjoint subsets, where the sums of the elements raised to some integer power in the two subsets are equal.² Prouhet was interested in the following problem. Given M , is it possible to partition the set of integers $\mathcal{S} = \{0, 1, \dots, N-1\}$ into two disjoint subsets $\mathcal{S}_0 \subset \mathcal{S}$ and $\mathcal{S}_1 \subset \mathcal{S}$ such that

$$\sum_{n \in \mathcal{S}_0} n^m - \sum_{n \in \mathcal{S}_1} n^m = 0, \quad (5.14)$$

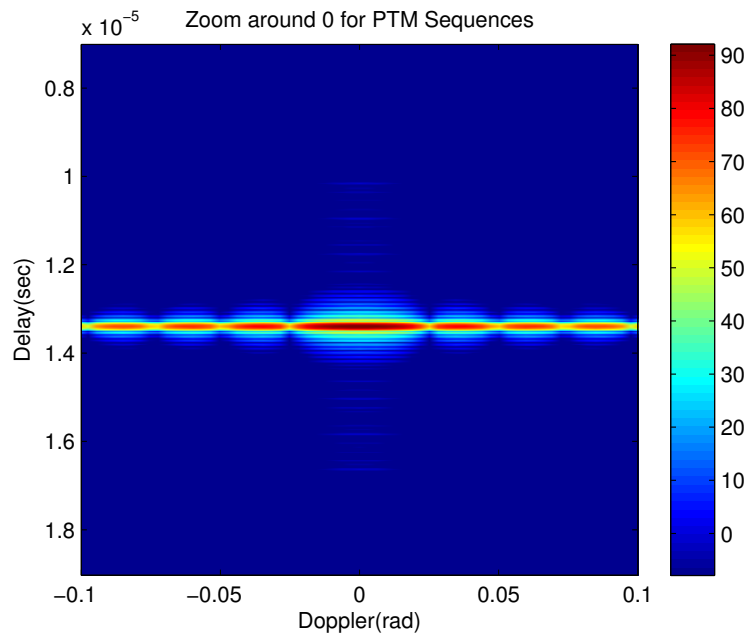
for all $0 \leq m \leq M$? Prouhet proved that this is possible only when $N = 2^{M+1}$ and that the partition is identified by the PTM sequence. In other words, the solution is $\mathcal{S}_0 = \{n | 0 \leq n \leq 2^{M+1} - 1, p_n = 1\}$ and $\mathcal{S}_1 = \{n | 0 \leq n \leq 2^{M+1} - 1, p_n = -1\}$, where $\{p_n\}_{n=0}^{2^{M+1}-1}$ is the length- 2^{M+1} PTM sequence.

Fig. 5.3 (a) shows the ambiguity function of a length- $(N = 2^8)$ PTM pulse train of Golay complementary waveforms, which has a seventh-order null at zero-Doppler. The horizontal axis is Doppler shift in rad and the vertical axis is delay in sec. The magnitude of the pulse train ambiguity function is color coded and presented in dB scale. A zoom in around zero-Doppler is shown in Fig. 5.3 (b). We notice that the range sidelobes inside the Doppler interval $[-0.1, 0.1]$ rad have been cleared out. They are at least 80 dB below the peak of the ambiguity function.

²However, Prouhet did not explicitly write the sequence. This was left to Thue [106] and Morse [107], who each rediscovered the PTM sequence independently.



(a)



(b)

Figure 5.3: Ambiguity function of a length- $(N = 2^8)$ PTM pulse train of Golay complementary waveforms: (a) the entire Doppler band (b) Doppler band $[-0.1, 0.1]$ rad

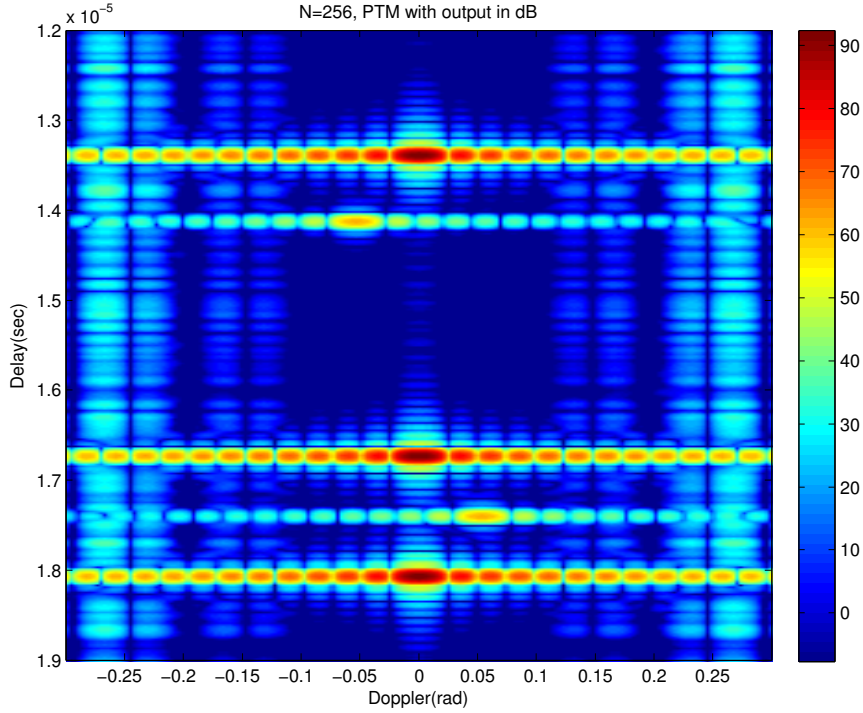


Figure 5.4: The PTM pulse train of Golay complementary waveforms can bring out weak targets which would have otherwise been masked by the range sidelobes of nearby strong reflectors.

Fig. 5.4 shows the effect of range sidelobe suppression in bringing out weak targets in the presence of strong reflectors for the five target scenario discussed earlier in Fig. 5.2. This example demonstrates the value of PTM pulse trains for radar imaging. We note that the Golay complementary sequences used for phase coding in obtaining figures in this chapter are of size 64 and the pulse shape is a raised-cosine pulse. The chip length is $T_c = 100$ nsec, the carrier frequency is 17 GHz (corresponding to a surveillance radar), and the PRI is $T = 50$ μ sec.

5.3.2 Range Sidelobes Suppression in Higher Doppler

We now consider the design of biphasic sequences whose spectra have high-order nulls at Doppler frequencies other than zero. Consider the Taylor expansion of the spectrum $S_{\mathcal{P}}(\theta)$ around $\theta = \theta_0$:

$$S_{\mathcal{P}}(\theta) = \sum_{t=0}^{\infty} \frac{1}{t!} f_{\mathcal{P}}^{(t)}(\theta_0) (\theta - \theta_0)^t \quad (5.15)$$

where the coefficients $f_{\mathcal{P}}^{(t)}(\theta_0)$ are given by

$$f_{\mathcal{P}}^{(t)}(\theta_0) = \left[\frac{d^t}{d\theta^t} S_{\mathcal{P}}(\theta) \right]_{\theta=\theta_0} = j^t \sum_{n=0}^{N-1} n^t p_n e^{jn\theta_0}, \quad t = 0, 1, 2, \dots \quad (5.16)$$

We wish to zero-force all the derivatives $f_{\mathcal{P}}^{(t)}(\theta_0)$ up to order M , that is we wish to design the sequence \mathcal{P} so that

$$f_{\mathcal{P}}^{(t)}(\theta_0) = 0, \quad \text{for all } t = 0, 1, \dots, M. \quad (5.17)$$

We consider rational Doppler shifts $\theta_0 = 2\pi l/m$, where l and $m \neq 1$ are co-prime integers. We assume the length of \mathcal{P} is $N = mq$ for some integer q . We have the following theorem.

Theorem 5.3.3 (Oversampled PTM Sequencing). *Let $\mathcal{P} = \{p_n\}_{n=0}^{2^M m - 1}$ be a $(2^M, m)$ -PTM sequence, that is to say that $\{p_{rm+i}\}_{r=0}^{2^M - 1}$, $i = 0, \dots, m-1$ is a PTM sequence of length 2^M , then the spectrum $S_{\mathcal{P}}(\theta)$ of \mathcal{P} has M th-order nulls at all $\theta_0 = 2\pi l/m$ where l and $m \neq 1$ are co-prime integers.*

Proof. See Appendix 7.4.1 □

Remark: For the special case $m = 2$ and $\theta_0 = \pi$, $f_{\mathcal{P}}^{(t)}(\theta_0) = j^t \sum_{n=0}^{N-1} n^t (-1)^n p_n$. It follows from [102] that $\{(-1)^n p_n\}$ has to be a PTM sequence of length 2^{M+1} in order to zero-force up to the M th-order Taylor expansion at zero Doppler. We denote $\{\tilde{p}_n\} = \{(-1)^n p_n\}$, and henceforth $p_n = (-1)^n \tilde{p}_n$. This result coincides with the $(2^M, 2)$ -PTM sequence as in Theorem 5.3.3. From this point, the result in [102] can be treated as a special case of the $(2^M, m)$ -PTM sequence.

We have the following corollary.

Corollary 5.3.2. *Let \mathcal{P} be the $(2^M, m)$ -PTM sequence, then the spectrum $S_{\mathcal{P}}(\theta)$ of \mathcal{P} has*

1. *an $(M-1)$ th-order null at $\theta_0 = 0$.*
2. *$(M-h-1)$ th-order nulls at all $\theta_0 = 2\pi l/(2^h m)$, where l and $m \neq 1$ are co-prime, and $1 \leq h \leq M-1$.*

Proof. See Appendix 7.4.2. □

Example 5.3.3. *The spectrum of the $(2^3, 2)$ -PTM sequence, i.e.,*

$$\mathcal{P} = \{+1, +1, -1, -1, -1, -1, +1, +1, -1, -1, +1, +1, +1, +1, -1, -1\},$$

has a third-order null at $\theta_0 = \pi$, a second-order null at $\theta_0 = 0$, first-order nulls at $\theta_0 = \pi/2$ and $\theta_0 = 3\pi/2$, and zeroth-order nulls at $\theta_0 = \pi/4$ and $\theta_0 = 3\pi/4$. This spectrum is shown in Fig. 5.5 in solid black line. The spectrum of the $(2^2, 2)$ -PTM sequence is also shown. This spectrum has a second-order null at $\theta_0 = \pi$ rad, a first-order null at $\theta_0 = 0$, and zeroth-order nulls at $\theta_0 = \pi/2$ and $\theta_0 = 3\pi/2$.

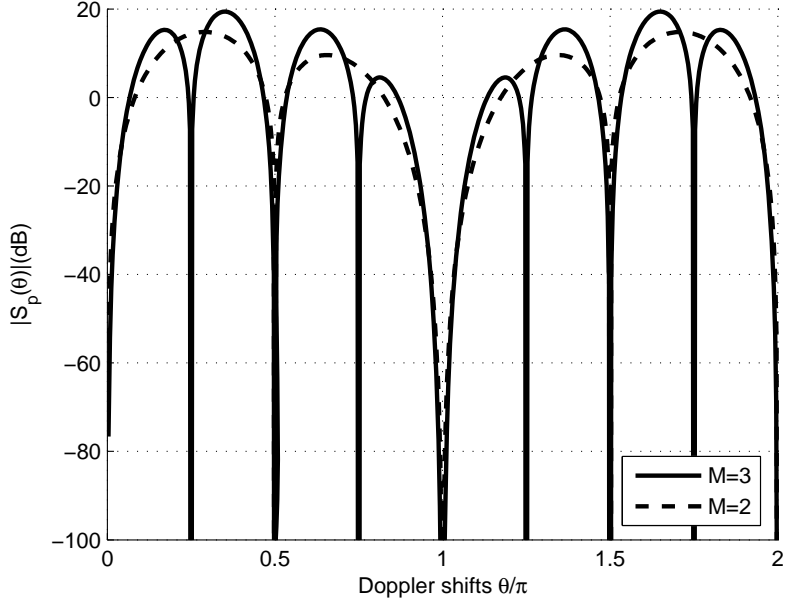


Figure 5.5: The spectra of $(2^3, 2)$ - and $(2^2, 2)$ -PTM sequences.

For certain values of m , it is possible to achieve an M th-order null with a sequence of length shorter than $2^M \times m$, as the following theorem shows.

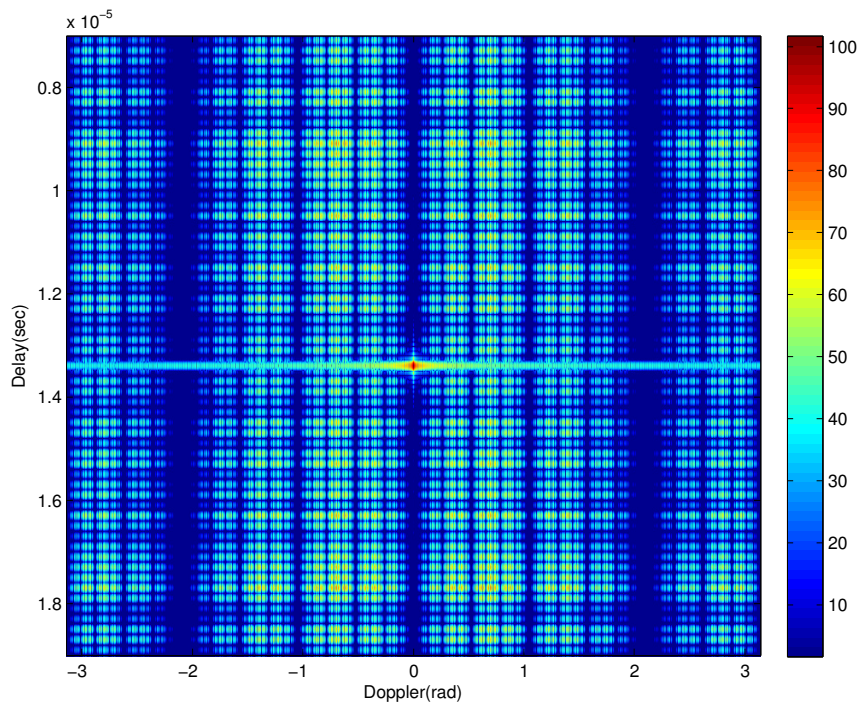
Theorem 5.3.4. *Let $m \equiv 2 \pmod{4}$, and $\tilde{\mathcal{P}} = \{\tilde{p}_n\}_{n=0}^{2^{M-1}m-1}$ be the $(2^M, m/2)$ -PTM sequence, then $\mathcal{P} = \{p_n\}_{n=0}^{2^{M-1}m-1}$, where $p_n = (-1)^n \tilde{p}_n$ has M th-order nulls at all $\theta_0 = 2\pi l/m$ where l and $m \neq 1$ are co-prime integers.*

Proof. See Appendix 7.4.3. □

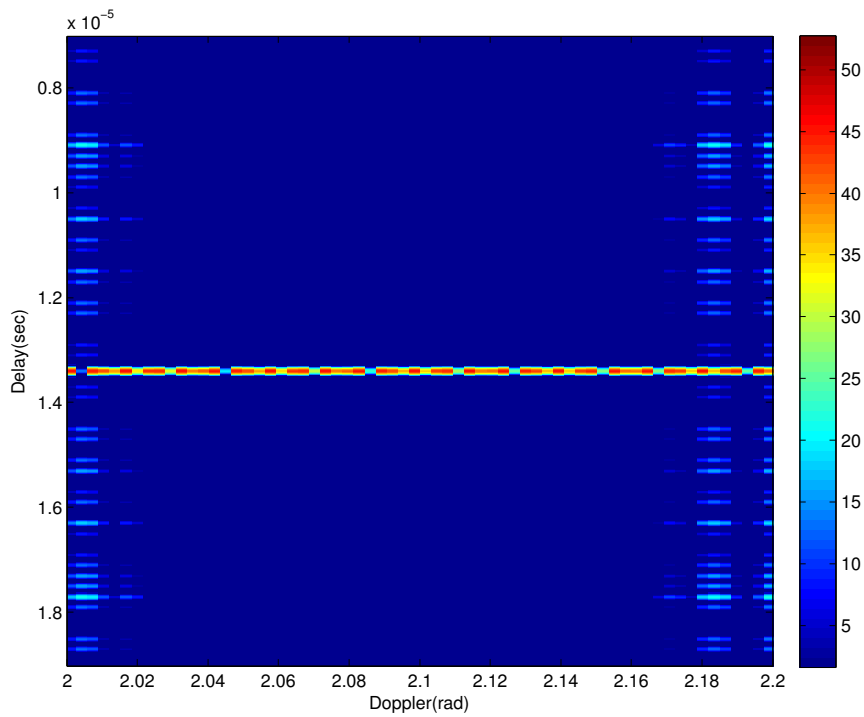
Fig. 5.6 (a) shows the ambiguity function of a $(2^8, 3)$ -PTM sequence of Golay complementary waveforms. The color bar values are in dB. This ambiguity function has an eighth-order null at $\theta_0 = \pm 2\pi/3$, a seventh-order null at zero Doppler, sixth-order nulls at $\theta_0 = \pm\pi/3$, and so on. A zoom in around $\theta_0 = 2\pi/3$ is provided in Fig. 5.6 (b) to demonstrate that range sidelobes in this Doppler region are significantly suppressed. The range sidelobes in this region are at least 80 dB below the peak of the ambiguity function.

5.4 Instantaneous Radar Polarimetry

Fully polarimetric radar systems are capable of transmitting and receiving on two orthogonal polarizations simultaneously. The combined signal then has an electric field vector that is modulated both in direction and amplitude by the waveforms on the two polarization channels as seen in Fig. 5.7, and the receiver is used to obtain both polarization components of the reflected waveform. The use of two orthogonal



(a)



(b)

Figure 5.6: Ambiguity function of the $(2^8, 3)$ -PTM pulse train of Golay complementary waveforms: (a) the entire Doppler band (b) Zoom in around $\theta_0 = 2\pi/3$.

polarizations increases the degrees of freedom available for target detection and can result in significant improvement in detection performance.

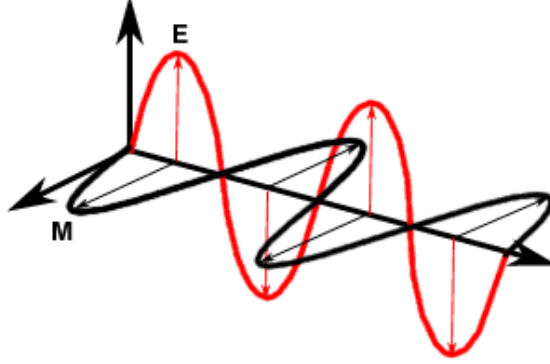


Figure 5.7: Two orthogonal polarizations of the electromagnetic field.

In general the reflection properties that apply to each polarization component are also different, and indeed reflection can change the direction of polarization. Thus, polarimetric radars are able to obtain the scattering tensor of a target

$$\mathbf{\Sigma} = \begin{pmatrix} \sigma_{VV} & \sigma_{VH} \\ \sigma_{HV} & \sigma_{HH} \end{pmatrix}, \quad (5.18)$$

where σ_{VH} denotes the target scattering coefficient into the vertical polarization channel due to a horizontally polarized incident field. Target detection is enhanced by concurrent rather than serial access to the cross-polarization components of the scattering tensor, which varies more rapidly in standard radar models used in target detection and tracking [109, 110] than in models used in remote sensing or synthetic aperture radar [111, 112]. In fact what is measured is the combination of three matrices

$$\mathbf{H} = \begin{pmatrix} h_{VV} & h_{VH} \\ h_{HV} & h_{HH} \end{pmatrix} = \mathbf{C}_{\text{Rx}} \mathbf{\Sigma} \mathbf{C}_{\text{Tx}}, \quad (5.19)$$

where \mathbf{C}_{Rx} and \mathbf{C}_{Tx} correspond to the polarization coupling properties of the transmit and receive antennas, whereas $\mathbf{\Sigma}$ results from the target. In most radar systems the transmit and receive antennas are common, and so the matrices \mathbf{C}_{Tx} and \mathbf{C}_{Rx} are conjugate. The cross-coupling terms in the antenna polarization matrices are clearly frequency and antenna geometry dependent but for the linearly polarized case this value is typically no better than about -20dB.

In this section, we follow [113],[114] to describe a new approach to radar polarimetry that uses orthogonal polarization modes to provide essentially independent channels for viewing a target, and achieve diversity gain. Unlike conventional radar polarimetry, where polarized waveforms are transmitted sequentially and processed non-coherently, the approach in [113],[114] allows for *Instantaneous Radar Polarime-*

try (IRP), where polarization modes are combined coherently on a pulse-by-pulse basis. IRP enables detection based on full polarimetric properties of the target and hence can provide better discrimination against clutter. When compared to a radar system with a singly-polarized transmitter and a singly-polarized receiver instantaneous radar polarimetry can achieve the same detection performance (same false alarm and detection probabilities) with a substantially smaller transmit energy, or alternatively it can detect at substantially greater ranges for a given transmit energy. The aim in this section is to present the main idea for waveform transmission in IRP, its Doppler resilience via PTM sequencing, and the generalization to a four-antenna radar system.

5.4.1 Unitary Waveform Matrices

Let us employ both polarization modes to transmit four phase-coded waveforms w_H^1 , w_V^1 , w_H^2 , w_V^2 . On each polarization mode we transmit two phase-coded waveforms separated by a PRI of length T . We employ Alamouti coding [115] to coordinate the transmission of waveforms over the V and H channels; that is, we define

$$w_H^2 = \widetilde{w_V^1} \quad \text{and} \quad w_V^2 = -\widetilde{w_H^1}, \quad (5.20)$$

where $\widetilde{\cdot}$ denotes complex conjugate time-reversal. After discretizing (at chip intervals) and converting time-indexed sequences to z -transform domain, we can write the radar receive equation in matrix form as

$$\mathbf{R}(z) = z^{-d} \mathbf{H} \mathbf{W}(z) + \mathbf{Z}(z), \quad (5.21)$$

where $\mathbf{R}(z)$ is the 2×2 radar measurement matrix at the receiver, \mathbf{H} is the 2×2 scattering matrix in (5.19) and $\mathbf{Z}(z)$ is a noise matrix. The entries of \mathbf{H} are taken to be constant (fixed draw from a random vector) since they correspond to a fixed range (delay d) and a fixed time. For now, we limit our analysis to zero-Doppler axis and will postpone the treatment of Doppler effect. The Alamouti waveform matrix $\mathbf{W}(z)$ is given by

$$\mathbf{W}(z) = \begin{pmatrix} w_V^1(z) & -\widetilde{w_H^1}(z) \\ w_H^1(z) & \widetilde{w_V^1}(z) \end{pmatrix}, \quad (5.22)$$

where $\widetilde{w}(z) = z^L \overline{w}(z^{-1})$ for a length L sequence w .

If we require the matrix $\mathbf{W}(z)$ to be unitary, that is

$$\mathbf{W}(z) \widetilde{\mathbf{W}}(z)^H = \begin{pmatrix} w_V^1(z) & -\widetilde{w_H^1}(z) \\ w_H^1(z) & \widetilde{w_V^1}(z) \end{pmatrix} \begin{pmatrix} \widetilde{w_V^1}(z) & \widetilde{w_H^1}(z) \\ -w_H^1(z) & w_V^1(z) \end{pmatrix} = 2Lz^{-L} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (5.23)$$

then it is easy to estimate the scattering matrix \mathbf{H} by post-multiplying (5.21) with $\widetilde{\mathbf{W}}(z)^H$. The unitary condition is equivalent to

$$|w_V^1(z)|^2 + |w_H^1(z)|^2 = 2L. \quad (5.24)$$

This is the same condition as the Golay complementary condition and is satisfied by choosing $w_V(z) = X(z)$ and $w_H(z) = Y(z)$, where $X(z)$ and $Y(z)$ are z -transforms of the Golay complementary sequences x and y , respectively. This shows that by properly coordinating the transmission of Golay complementary waveforms across polarizations and over time we can make the four channels HH, VV, VH, and HV available at the receiver (with delay L) using only linear processing. The four matched filters (in z -domain) are given by

$$\mathbf{Q}(z) = \begin{pmatrix} m_1(z)r_V(z) & m_2(z)r_V(z) \\ m_1(z)r_H(z) & m_2(z)r_H(z) \end{pmatrix} \quad (5.25)$$

where $m_1(z) = \tilde{w}_V^1(z)z^{-L} - \tilde{w}_H^1(z)$ and $m_2(z) = \tilde{w}_H^1(z)z^{-L} + \tilde{w}_V^1(z)$.

Remark: The above description suggests that a radar image will be available only on every second pulse, since two PRIs are required to form an image. However, after the transmission of the first pulse, images can be made available at every PRI. This is done by reversing the roles of the waveforms transmitted on the two pulses. Thus, in the analysis, the matrix $\mathbf{W}(z)$ in (5.21) is replaced by

$$\mathbf{V}(z) = \begin{pmatrix} -\tilde{w}_H^1(z) & w_V^1(z) \\ \tilde{w}_V^1(z) & w_H^1(z) \end{pmatrix}, \quad (5.26)$$

which is still unitary due to the interplay between Golay property and Alamouti coding. Moreover, the processing involved is essentially invariant from pulse to pulse: the return pulse in each of the channels is correlated against the transmit pulse on that channel. This yields an estimate of the scattering matrix on each pulse.

5.4.2 Doppler Resilient IRP

In the previous section, we restricted our analysis to zero-Doppler axis. Off the zero-Doppler axis, a relative Doppler shift of θ exists between consecutive waveforms on each polarization channel and the radar measurement equation changes to

$$\mathbf{R}(z) = z^{-d}\mathbf{H}\mathbf{W}(z)\mathbf{D}(\theta) + \mathbf{Z}(z), \quad (5.27)$$

where $\mathbf{D}(\theta) = \text{diag}(1, e^{j\theta})$. Consequently, the unitary property of $\mathbf{W}(z)$ can no longer be used to estimate \mathbf{H} , due to the fact that $\mathbf{W}(z)\mathbf{D}(\theta)\widetilde{\mathbf{W}}(z)$ is not a factor of identity. In other words, off the zero-Doppler axis the four elements in $\mathbf{W}(z)\mathbf{D}(\theta)\widetilde{\mathbf{W}}(z)$ have range sidelobes.

The matrix $\mathbf{W}(z)\mathbf{D}(\theta)\widetilde{\mathbf{W}}(z)$ can be viewed as a *matrix-valued ambiguity function* for IRP. In Section 4.3 and 4.4, we showed that the range sidelobes of the ambiguity function of a pulse train of Golay waveforms can be suppressed (in a desired Doppler band) by carefully selecting the order in which the Golay complementary waveforms are transmitted. This result can be extended to IRP. In fact, the range sidelobes of the IRP matrix-valued ambiguity function can be cleared out (inside a desired Doppler

interval) by coordinating the transmission of the following Alamouti matrices

$$\mathbf{X}_1(z) = \begin{pmatrix} X(z) \\ Y(z) \end{pmatrix} \quad \text{and} \quad \mathbf{X}_{-1}(z) = \begin{pmatrix} -\tilde{Y}(z) \\ \tilde{X}(z) \end{pmatrix} \quad (5.28)$$

according to the 1's and -1's in the PTM sequence or its oversampled versions (depending on the Doppler interval of interest) where 1 represents $\mathbf{X}_1(z)$ and -1 represents $\mathbf{X}_{-1}(z)$.

5.4.3 Four-Antenna IRP

The above IRP can be also equivalently employed on the two antennas in a 2×2 Multiple-Input-Multiple-Output (MIMO) radar system, by identifying each polarization with an antenna. Now we present its generalization to a 4×4 radar system, where on each antenna we transmit a phase-coded waveform separated by a PRI of length T . The waveform matrix transmitted over four PRIs is constructed as

$$\mathbf{W}_4(z) = \begin{pmatrix} \mathbf{W}(z) & -\widetilde{\mathbf{W}}(z) \\ \mathbf{W}(z) & \widetilde{\mathbf{W}}(z) \end{pmatrix}, \quad (5.29)$$

where $\mathbf{W}(z)$ is defined in (5.22). We write $\mathbf{W}_4(z)$ explicitly as

$$\mathbf{W}_4(z) = \begin{pmatrix} w_V^1(z) & -\tilde{w}_H^1(z) & -\tilde{w}_V^1(z) & w_H^1(z) \\ w_H^1(z) & \tilde{w}_V^1(z) & -\tilde{w}_H^1(z) & -w_V^1(z) \\ w_V^1(z) & -\tilde{w}_H^1(z) & \tilde{w}_V^1(z) & -w_H^1(z) \\ w_H^1(z) & \tilde{w}_V^1(z) & \tilde{w}_H^1(z) & w_V^1(z) \end{pmatrix}, \quad (5.30)$$

and it is easy to see $\mathbf{W}_4(z)$ is also unitary since

$$\mathbf{W}_4(z)\widetilde{\mathbf{W}}_4(z)^H = \begin{pmatrix} \mathbf{W}(z) & -\widetilde{\mathbf{W}}(z) \\ \mathbf{W}(z) & \widetilde{\mathbf{W}}(z) \end{pmatrix} \begin{pmatrix} \widetilde{\mathbf{W}}(z) & \widetilde{\mathbf{W}}(z) \\ -\mathbf{W}(z) & \mathbf{W}(z) \end{pmatrix} = 4Lz^{-L}I_4, \quad (5.31)$$

where we make use of (5.23) and I_4 is the 4×4 identity matrix. This suggests we could use similar technique to coordinate the transmission of Golay complementary waveforms across antennas and time according to (5.30) and achieve IRP.

Similarly, the matrix $\mathbf{W}_4(z)\mathbf{D}_4(\theta)\widetilde{\mathbf{W}}_4(z)$ can be viewed as a matrix-valued ambiguity function for the 4×4 IRP, where $\mathbf{D}_4(\theta) = \text{diag}(1, e^{j\theta}, e^{j2\theta}, e^{j3\theta})$. We could clear out the range sidelobes of the IRP matrix-valued ambiguity function inside a desired Doppler interval by coordinating the transmission of the following four matrices

$$\mathbf{X}_0(z) = \begin{pmatrix} X(z) \\ Y(z) \\ X(z) \\ Y(z) \end{pmatrix}, \mathbf{X}_1(z) = \begin{pmatrix} -\tilde{Y}(z) \\ \tilde{X}(z) \\ -\tilde{Y}(z) \\ \tilde{X}(z) \end{pmatrix}, \mathbf{X}_2(z) = \begin{pmatrix} -\tilde{X}(z) \\ -\tilde{Y}(z) \\ \tilde{X}(z) \\ \tilde{Y}(z) \end{pmatrix}, \mathbf{X}_3(z) = \begin{pmatrix} Y(z) \\ -X(z) \\ -Y(z) \\ X(z) \end{pmatrix}$$

according to the generalized Prouhet sequence $\mathcal{P} = \{p_n\}_{n=0}^{N-1}$ such that $p_n \in \{0, 1, 2, 3\}$ or its oversampled versions (depending on the Doppler interval of interest) where $p_n = i$ represents transmitting $\mathbf{X}_i(z)$ in the n th PRI.

The sequence \mathcal{P} is the answer to the more general Prouhet's problem of finding a partition of $[0, 4^{M+1} - 1]$ into 4 sets I_0, I_1, I_2 and I_3 such that

$$\sum_{i \in I_0} i^k = \sum_{i \in I_1} i^k = \sum_{i \in I_2} i^k = \sum_{i \in I_3} i^k \quad \text{for } k = 0, 1, \dots, M.$$

Prouhet gave the following solution: define the sequence $\mathcal{P} = (T_4(n))_{n \geq 0}$ by $T_4(n) = s_4(n) \bmod 4$, where $s_4(n)$ equals the sum of the digits in the base-4 representation of the integer n , then $I_j = \{0 \leq i \leq 4^{M+1} - 1 : T_4(i) = j\}$.

Example 5.4.1. Let \mathcal{P} be the generalized Prouhet sequence, we coordinate the transmission of $\mathbf{X}_0(z), \dots, \mathbf{X}_3(z)$ according to \mathcal{P} . For example, Let $N = 4^2 - 1 = 15$, \mathcal{P} and are the waveforms are transmitted in order as

\mathcal{P}	0	1	2	3	1	2	3	0
$\mathbf{X}(z)$	\mathbf{X}_0	\mathbf{X}_1	\mathbf{X}_2	\mathbf{X}_3	\mathbf{X}_1	\mathbf{X}_2	\mathbf{X}_3	\mathbf{X}_0
\mathcal{P}	2	3	0	1	3	0	1	2
$\mathbf{X}(z)$	\mathbf{X}_2	\mathbf{X}_3	\mathbf{X}_0	\mathbf{X}_1	\mathbf{X}_3	\mathbf{X}_0	\mathbf{X}_1	\mathbf{X}_2

to create a second order null at zero Doppler.

Chapter 6

Optimal Training Sequences in Multi-User MIMO-OFDM Systems

6.1 Introduction

Information-theoretic analysis by Foschini [116] and by Telatar [117] has shown that multiple antennas at the transmitter and receiver enable high-rate wireless communication. Space-time codes, introduced by Tarokh et al. [118], improve the reliability of communication over fading channels by correlating signals across different transmit antennas. Orthogonal Frequency Division Multiplexing (OFDM) [119] is widely adopted in broadband communications standards for its efficient implementation, high spectral efficiency, and robustness to Inter-Symbol Interference (ISI). However, there are two main drawbacks in OFDM; the first is high Peak-to-Average Power Ratio (PAPR) which results in larger backoff with nonlinear amplifiers, and the second is high sensitivity to frequency errors and phase noise. The focus in this chapter is training sequence design for the combination of Multiple-Input-Multiple-Output (MIMO) systems and OFDM technology (see [120] and references therein), and we aim to make this combination more attractive by reducing the overhead that is necessary for channel estimation.

Current multi-user MIMO-OFDM systems [121] support multiple users by assigning each time/frequency slot to only one user. For example, in OFDMA systems (adopted in the WiMAX [122] and LTE standards [123]), different users are assigned different subcarriers within the same OFDMA symbol. A different method of separating users is through the random-access CSMA/CA medium access control (MAC) protocol used in WLAN standards, e.g. IEEE 802.11n. Both methods require that users not overlap in either time or frequency and this restriction results in a significant loss in spectral efficiency. The introduction of multiple receive antennas at the base station means that it is possible to improve spectral efficiency by allowing users to overlap while maintaining decodability, as in the recently-proposed Coordinated MultiPoint transmission (CoMP) techniques in the LTE-Advanced standard [124].

Accurate Channel State Information (CSI) is required at the receiver for coherent detection and is typically acquired by sending known training sequences from the

transmit antennas and inferring channel parameters from the received signals. Various OFDM channel estimation schemes [125]-[127] have been proposed for Single-Input Single-Output (SISO) systems. However channel estimation is more challenging in a multi-user MIMO-OFDM system because there are more link parameters to calculate, and their estimation is complicated by interference between different transmissions. The direct approach is to invert a large matrix that describes cross-antenna interference at each OFDM tone [128]. Complexity can be reduced by exploiting the correlation between adjacent subchannels [129]. It is also possible to develop solutions in the time domain [130] where the challenge is to estimate time of arrivals. Here it is possible to reduce complexity by exploiting the power-delay profiles of the typical urban and hilly terrain propagation models. MIMO Channel estimation schemes were investigated in [131] for single-carrier single-user systems in the context of GSM-EDGE.

Linear Least-Squares (LLS) channel estimation is of great practical importance since it does not require prior knowledge of the channel statistics and enjoys low implementation complexity. We consider frequency-selective block-fading channels where the Time Domain (TD) representation requires fewer parameters than the Frequency Domain (FD) representation. Our focus is on the design of (optimal) training sequences for Multi-User MIMO OFDM systems that minimize the mean squared error of time-domain LLS channel estimation. The design of optimal training sequences for single-user MIMO-OFDM systems is investigated in [132] and [133]. The Fourier methods used in [132] provide some control over PAPR and some resilience to frequency offsets. The construction of optimal training sequences for multi-user MIMO-OFDM systems has been investigated in both the time domain [134] and the frequency domain [135], but these designs do not easily extend to multiple OFDM training symbols. It is also possible to take advantage of the similarities between communications and radar signal processing, where the path gains and delays are the range/Doppler coordinates of a scattering source and the problem is to estimate them. The unitary filter bank developed for Instantaneous Radar Polarimetry [136] supports frequency domain LLS channel estimation in a 2×2 MIMO OFDM system [137] and is able to suppress interference over two OFDM symbols with linear complexity. This example is a special case of a more general construction of filter banks for the analysis of acoustic surface waves [138, 98]. A limitation of these methods is that the number of OFDM training symbols is at least the number of transmit antennas.

In contrast, our framework supports the design of optimal training sequences for an arbitrary number of transmit antennas and an arbitrary number of training symbols. It provides the first general solution to the channel estimation problem for Multi-User MIMO-OFDM systems where Spatial Division Multiple Access (SDMA) is employed to increase the spectral efficiency. The optimality of our designs holds irrespective of the number of transmit antennas per user, the number of OFDM sub-carriers, the channel delay spread, and the number of users *provided* that the number of tones dedicated to estimation exceeds the product of the number of transmit antennas and the worst case delay spread. Not only does our design algorithm generate training sequences that minimize mean squared channel estimation error, but the designs have additional properties that make them very attractive from several implementation

perspectives: 1) Individual training sequences can be drawn from standard signal constellations, 2) Low PAPR, and 3) Low channel estimation complexity without sacrificing optimality.

The rest of this chapter is organized as follows. The uplink Multi-User MIMO-OFDM communication system model is described in Section 6.2. The design of optimal training sequences is given for one and multiple training symbol scenarios separately in Section 6.3. Simulation results are presented in Section 6.4.

6.2 System Model

We consider the uplink of a Multi-User MIMO-OFDM system, as shown in Fig. 6.1. We denote the Discrete Fourier Transform (DFT) size by N and the number of users by L ($L \geq 1$) where the i th user is equipped with M_i transmit antennas, $0 \leq i \leq L - 1$. Therefore, the total number of transmit antennas among all users is given by $M = \sum_{i=0}^{L-1} M_i$.

We assume that the channel is quasi-static and remains constant over K successive OFDM training symbols. The channel from the j th transmit antenna of the i th user to the Base Station (BST) can be represented either in TD or FD. Let the Channel Frequency Response (CFR) be $\mathbf{H}_{i,j} = [H_{i,j}(0), \dots, H_{i,j}(N - 1)]^T$ where $H_{i,j}(k)$, $0 \leq k \leq N - 1$, is the frequency response at the k th subcarrier. However, the Channel Impulse Response (CIR) in TD is represented by a much smaller number of parameters. We assume that the maximal memory over all CIRs is ν_{\max} , and write the CIR as $\mathbf{h}_{i,j} = [h_{i,j}(0), \dots, h_{i,j}(\nu_{\max})]^T$. Estimating the CIR instead of the CFR leads to the reduction of the number of unknowns from MN to $M(\nu_{\max} + 1)$. Hence, a more accurate channel estimate is attainable using the same amount of training. Furthermore, the CFR can be reconstructed from the CIR as follows

$$H_{i,j}(k) = \frac{1}{\sqrt{N}} \sum_{t=0}^{\nu_{\max}} h_{i,j}(t) e^{-j \frac{2\pi}{N} tk}. \quad (6.1)$$

At the j th ($0 \leq j \leq M_i - 1$) transmit antenna of the i th ($0 \leq i \leq L - 1$) user, an OFDM symbol $\mathbf{X}_{i,j}$ of size N is given by $\mathbf{X}_{i,j} = [X_{i,j}(0), \dots, X_{i,j}(N - 1)]^T$. Let $\mathbf{x}_{i,j} = [x_{i,j}(0), x_{i,j}(1), \dots, x_{i,j}(N - 1)]^T$ be the Inverse Discrete Fourier Transform (IDFT) of $\mathbf{X}_{i,j}$. We use a Cyclic-Prefix (CP) of length L_p for the guard interval in the OFDM system so that

$$\tilde{\mathbf{x}}_{i,j} = [x(N - L_p + 1), \dots, x(N - 1), x(0), \dots, x(N - 1)]^T \quad (6.2)$$

where L_p is chosen to be greater than the channel memory, i.e. $L_p \geq (\nu_{\max} + 1)$. Finally, $\tilde{\mathbf{x}}_{i,j}$ goes through Parallel-to-Serial (P/S) conversion and is modulated to the carrier frequency for transmission.

At the base station, all users are synchronized with the BST, where the received signal is down-converted to baseband and passed through a Serial-to-Parallel (S/P) converter. Then, the CP is removed and the Fast Fourier Transform (FFT) is applied.

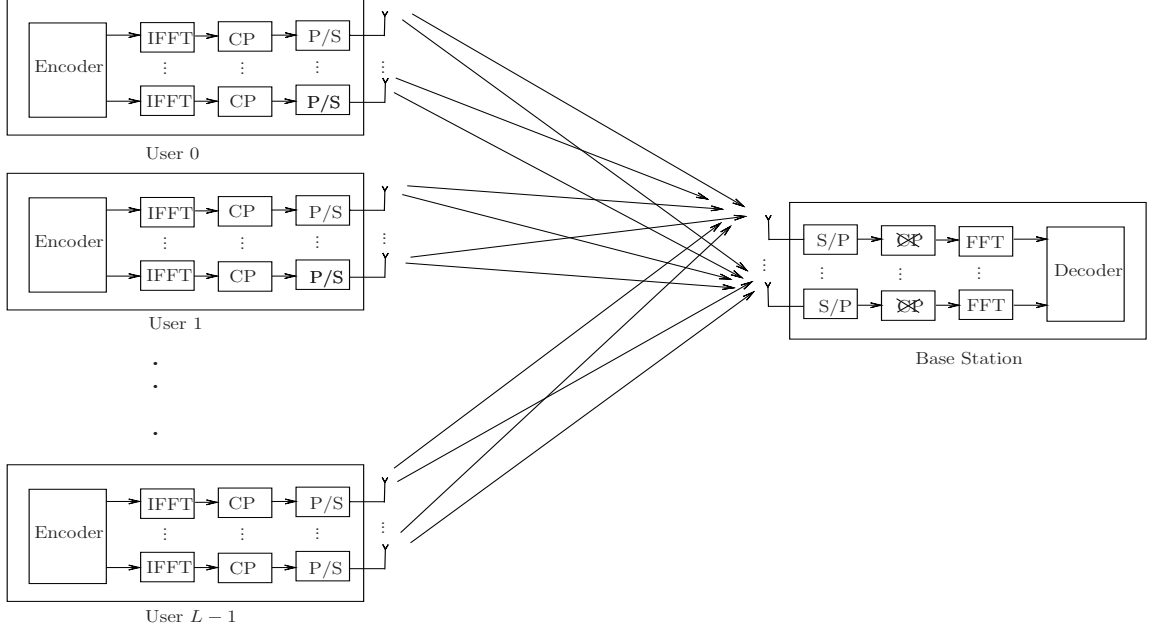


Figure 6.1: The uplink of a Multi-User MIMO-OFDM communication system.

The received OFDM symbol $\mathbf{Y} = [Y(0), \dots, Y(N-1)]^T$ in one symbol time can be written as

$$\mathbf{Y} = \sum_{i=0}^{L-1} \sum_{j=0}^{M_i-1} \text{diag}(\mathbf{H}_{i,j}) \mathbf{X}_{i,j} + \mathbf{N}, \quad (6.3)$$

where $\mathbf{N} \sim \mathcal{N}(\mathbf{0}_{N \times 1}, \sigma^2 \mathbf{I}_N)$ is Additive White Gaussian Noise (AWGN). We consider the mapping $(i, j) \mapsto m : m = \sum_{s=0}^i M_s + j - M_i$, $0 \leq m \leq M-1$, and re-label $\mathbf{H}_{i,j}$ and $\mathbf{X}_{i,j}$ as \mathbf{H}_m and \mathbf{X}_m , respectively. The label can be inverted easily as

$$i = \underset{0 \leq i^* \leq L-1}{\text{argmin}} i^* \quad \text{s.t.} \quad m \leq \sum_{s=0}^{i^*} M_s, \quad j = m + M_i - \sum_{s=0}^i M_s. \quad (6.4)$$

Then, equation (6.3) can be written as

$$\mathbf{Y} = \sum_{m=0}^{M-1} \text{diag}(\mathbf{H}_m) \mathbf{X}_m + \mathbf{N}. \quad (6.5)$$

Remark: The development of the algorithm requires labeling of transmit antennas among all users, and that both the BST and all the users are aware of that labeling.

6.3 Main Results

We first consider TD LLS channel estimation when only one training symbol is allowed by leveraging the channel representation in TD. A general approach when $K \geq 2$ training symbol is given by further incorporating space-time code structure into the design. Then, a special construction utilizing Quaternions is given when $K = 2$. Finally, an alternative scheme using equally-spaced pilots instead of the whole symbol for training is given under some mild conditions. In the following sections, we assume one receive antenna, since the same channel estimation scheme can be applied at all receive antennas without loss of generality.

6.3.1 One OFDM Training Symbol

Since there are fewer parameters to be estimated in the TD, we apply the IDFT of size N to (6.5), and get

$$\begin{aligned} \mathbf{y} &= \sum_{m=0}^{M-1} \mathbf{S}_m \mathbf{h}_m + \mathbf{n} \\ &= [\mathbf{S}_0 \ \mathbf{S}_1 \ \cdots \ \mathbf{S}_{M-1}] [\mathbf{h}_0^H \ \mathbf{h}_1^H \ \cdots \ \mathbf{h}_{M-1}^H]^H + \mathbf{n} \\ &\triangleq \mathbf{S} \mathbf{h} + \mathbf{n}, \end{aligned} \tag{6.6}$$

where $\mathbf{y} \in \mathbb{C}^N$, $\mathbf{h}_m \in \mathbb{C}^{\nu_{\max}+1}$, $0 \leq m \leq M-1$, and $\mathbf{S}_m \in \mathbb{C}^{N \times (\nu_{\max}+1)}$ is the circulant training matrix constructed from the corresponding training sequence transmitted over the m th antenna.

Let $\mathbf{F} = [\mathbf{f}_0, \cdots, \mathbf{f}_{N-1}]$ be the DFT matrix of size N with \mathbf{f}_i denoting its i^{th} column, and let $\mathbf{F}_0 = [\mathbf{f}_0, \cdots, \mathbf{f}_{\nu_{\max}}]$ be composed of the first $(\nu_{\max} + 1)$ columns of \mathbf{F} . Then, \mathbf{S}_m can be written as

$$\mathbf{S}_m = \mathbf{F}^H \mathbf{D}_m \mathbf{F}_0, \tag{6.7}$$

where $\mathbf{D}_m = \text{diag}(X_m(0), \cdots, X_m(N-1))$. The matrix $\mathbf{S} \in \mathbb{C}^{N \times M(\nu_{\max}+1)}$ defined in (6.6) is formed by horizontally concatenating the matrices \mathbf{S}_m , $0 \leq m \leq M-1$. To enable LLS channel estimation, the following condition on dimensionality has to be satisfied [139]

$$N \geq M(\nu_{\max} + 1) \quad \text{or}, \quad M \leq \frac{N}{(\nu_{\max} + 1)}. \tag{6.8}$$

To minimize the variance of the channel estimation error, the matrix \mathbf{S} is required to satisfy $\mathbf{S}^H \mathbf{S} = c \mathbf{I}_{M(\nu_{\max}+1)}$, and this requires that

$$\mathbf{S}_m^H \mathbf{S}_n = c \delta_{mn} \mathbf{I}_{(\nu_{\max}+1)}, \quad 0 \leq m, n \leq M-1. \tag{6.9}$$

Given (6.7), the optimality condition becomes

$$\mathbf{F}_0^H \mathbf{D}_m^H \mathbf{D}_n \mathbf{F}_0 = c \delta_{mn} \mathbf{I}_{(\nu_{\max}+1)}, \quad 0 \leq m, n \leq M-1. \tag{6.10}$$

Next, let \mathbf{F}_m be composed of $(\nu_{\max} + 1)$ consecutive columns of \mathbf{F} starting at index $m(\nu_{\max} + 1)$, i.e.

$$\mathbf{F}_m = [\mathbf{f}_{m(\nu_{\max}+1)}, \dots, \mathbf{f}_{(m+1)(\nu_{\max}+1)-1}] = \mathbf{\Lambda}_m \mathbf{F}_0, \quad 0 \leq m \leq M - 1, \quad (6.11)$$

where

$$\mathbf{\Lambda}_m = \text{diag} \left(1, e^{j \frac{2\pi(\nu_{\max}+1)}{N} m}, \dots, e^{j \frac{2\pi(\nu_{\max}+1)(N-1)}{N} m} \right). \quad (6.12)$$

It can be easily shown that $\mathbf{F}_m^H \mathbf{F}_n = \delta_{mn} \mathbf{I}_{(\nu_{\max}+1)}$. Now we present a general approach which gives a family of optimal training sequences. As a starting point, we choose the FD training sequence as an arbitrary constant-amplitude sequence \mathbf{X} . Let $\mathbf{D} = \text{diag}(X(0), \dots, X(N-1))$, then $\mathbf{D}^H \mathbf{D} = c \mathbf{I}_N$ where c is determined by the signal constellation and/or transmit power constraints. The FD training sequence at the m th transmit antenna is given by

$$\mathbf{X}_m = \mathbf{\Lambda}_m \mathbf{X}, \quad 0 \leq m \leq M - 1. \quad (6.13)$$

Equivalently, $\mathbf{D}_m = \mathbf{\Lambda}_m \mathbf{D} = \mathbf{D} \mathbf{\Lambda}_m$, $0 \leq m \leq M - 1$. Furthermore, we have the following theorem.

Theorem 6.3.1. *The choice of FD training sequences in (6.13) is optimal for a single training OFDM symbol.*

Proof. It is enough to show that (6.9) holds. Since

$$\mathbf{S}_m = \mathbf{F}^H \mathbf{D}_m \mathbf{F}_0 = \mathbf{F}^H \mathbf{D} \mathbf{\Lambda}_m \mathbf{F}_0 = \mathbf{F}^H \mathbf{D} \mathbf{F}_m, \quad (6.14)$$

it follows that

$$\mathbf{S}_m^H \mathbf{S}_n = (\mathbf{F}^H \mathbf{D} \mathbf{F}_m)^H \mathbf{F}^H \mathbf{D} \mathbf{F}_n = \mathbf{F}_m^H \mathbf{D}^H \mathbf{D} \mathbf{F}_n = c \delta_{mn} \mathbf{I}_{(\nu_{\max}+1)}. \quad (6.15)$$

□

The LLS estimate (LLSE) of \mathbf{h} is given as $\hat{\mathbf{h}} = \frac{1}{c} \mathbf{S}^H \mathbf{y}$, where each CIR can be estimated as $\hat{\mathbf{h}}_m = \frac{1}{c} \mathbf{S}_m^H \mathbf{y}$. Then, the CFR is given by

$$\hat{\mathbf{H}}_m = \frac{1}{c} \mathbf{F} \mathbf{S}_m^H \mathbf{y} = \frac{1}{c} (\mathbf{F} \mathbf{F}_0^H) \mathbf{D}_m^H \mathbf{F} \mathbf{y}, \quad 0 \leq m \leq M - 1. \quad (6.16)$$

The resulting channel estimation error variance is given by

$$\sigma_e^2 = \sigma^2 \text{Tr}((\mathbf{S}^H \mathbf{S})^{-1}) = \frac{M(\nu_{\max} + 1)}{c} \sigma^2. \quad (6.17)$$

6.3.2 K OFDM Training Symbols with $K \geq 2$

The major limitation of using only one training OFDM symbol is that the total number of transmit antennas is limited by $\frac{N}{(\nu_{\max}+1)}$. When the channel is quasi-static

over $K \geq 2$ OFDM training symbols it is possible to increase the number of admissible transmit antennas and reduce MMSE by a factor of K .

Denoting the received TD OFDM symbol in the t th symbol time by \mathbf{y}_t , $0 \leq t \leq K - 1$, we express the received symbol block \mathbf{y} as

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{K-1} \end{bmatrix} = \begin{bmatrix} \mathbf{S}_{00} & \mathbf{S}_{01} & \cdots & \mathbf{S}_{0,M-1} \\ \mathbf{S}_{10} & \mathbf{S}_{11} & \cdots & \mathbf{S}_{1,M-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{S}_{K-1,0} & \mathbf{S}_{K-1,1} & \cdots & \mathbf{S}_{K-1,M-1} \end{bmatrix} \begin{bmatrix} \mathbf{h}_0 \\ \mathbf{h}_1 \\ \vdots \\ \mathbf{h}_{M-1} \end{bmatrix} + \mathbf{n}$$

$$\triangleq \mathbf{S}\mathbf{h} + \mathbf{n}, \quad (6.18)$$

where $\mathbf{S}_{tm} = \mathbf{F}^H \mathbf{D}_{tm} \mathbf{F}_0$, and the matrices \mathbf{D}_{tm} 's are diagonal with the FD training sequences appearing on their main diagonals. Least-square estimation is possible when the following dimensionality condition for the matrix $\mathbf{S} \in \mathbb{C}^{KN \times M(\nu_{\max} + 1)}$ holds

$$KN \geq M(\nu_{\max} + 1), \text{ or, } \sum_{i=0}^{L-1} M_i = M \leq \frac{KN}{(\nu_{\max} + 1)}. \quad (6.19)$$

For \mathbf{S} to be optimal, it has to satisfy $\mathbf{S}^H \mathbf{S} = \tilde{c} \mathbf{I}_{M(\nu_{\max} + 1)}$ for some \tilde{c} . We extend our previous approach by constructing a unitary matrix of higher dimension with the space-time code structure. Let the matrix $\mathbf{\Sigma} \in \mathbb{C}^{KN \times KN}$ be constructed as a Kronecker product $\mathbf{\Sigma} = \mathbf{U} \otimes \mathbf{V}$ where $\mathbf{U} = [U_{tq}] \in \mathbb{C}^{K \times K}$ is a unitary matrix and $\mathbf{V} \in \mathbb{C}^{N \times N}$ is a diagonal matrix satisfying $\mathbf{V}^H \mathbf{V} = \tilde{c} \mathbf{I}_N$. Therefore the matrix $\mathbf{\Sigma}$ satisfies

$$\mathbf{\Sigma}^H \mathbf{\Sigma} = \mathbf{U}^H \mathbf{U} \otimes \mathbf{V}^H \mathbf{V} = \tilde{c} \mathbf{I}_{KN}. \quad (6.20)$$

We give the following general design of optimal training sequences. For $0 \leq m \leq M - 1$, let $p = \lfloor \frac{m}{K} \rfloor$, $0 \leq p \leq \lfloor \frac{M-1}{K} \rfloor$ and $q = m - Kp \in \{0, \dots, K - 1\}$. For the m th transmit antenna, its FD training sequence matrix at the t th OFDM training symbol is given by

$$\mathbf{D}_{tm} = \mathbf{\Sigma}_{tq} \mathbf{\Lambda}_p, \quad \text{if } m = Kp + q, \quad 0 \leq m \leq M - 1, \quad (6.21)$$

where $\mathbf{\Sigma}_{tq} = U_{tq} \mathbf{V}$ is the $N \times N$ diagonal matrix located at the (t, q) block of $\mathbf{\Sigma}$.

The bijection $\pi : m \mapsto \{p, q\}$ groups the antennas into K classes depending on the equivalence of the residue q . For two antennas *not* in the same class, their training sequences can be proved orthogonal over *any* OFDM training symbol. For two antennas in the same class, their training sequences can be proved orthogonal over *all* K OFDM training symbols. We give the detailed proof below.

Theorem 6.3.2 (K Training Symbols). *The training sequences in (6.21) are optimal for K training OFDM symbols.*

Proof. See Appendix 7.5.1. □

Finally, the LLSE of \mathbf{h} is given by $\hat{\mathbf{h}} = \frac{1}{\tilde{c}} \sum_{t=0}^{K-1} \mathbf{S}_t^H \mathbf{y}_t$ where each CIR can be estimated as $\hat{\mathbf{h}}_m = \frac{1}{\tilde{c}} \sum_{t=0}^{K-1} \mathbf{S}_{tm}^H \mathbf{y}_t$. Then, the CFR at the m th transmit antenna is

given by

$$\hat{\mathbf{H}}_m = \frac{1}{\tilde{c}} \mathbf{F} \hat{\mathbf{h}}_m = \frac{1}{\tilde{c}} (\mathbf{F} \mathbf{F}_0^H) \sum_{t=0}^{K-1} \mathbf{D}_{tm}^H \mathbf{F} \mathbf{y}_t. \quad (6.22)$$

Let $\tilde{c} = Kc$, the resulting channel estimation error variance is given by

$$\sigma_e^2 = \sigma^2 \text{Tr} \left(\sum_{t=0}^{K-1} \mathbf{S}_t^H \mathbf{S}_t \right)^{-1} = \frac{M(\nu_{\max} + 1)}{Kc} \sigma^2.$$

6.3.3 Special case when $K = 2$

When $K = 2$, like the Alamouti Space-Time Block Code (STBC), our construction of training sequences makes use of Hamilton's Biquaternions. We choose two FD training sequences \mathbf{X} and \mathbf{Z} where the sum of their squared amplitudes is constant, i.e.

$$\mathbf{D}_X^H \mathbf{D}_X + \mathbf{D}_Z^H \mathbf{D}_Z = \tilde{c} \mathbf{I}_N. \quad (6.23)$$

where $\mathbf{D}_X = \text{diag}(X(0), \dots, X(N-1))$, and $\mathbf{D}_Z = \text{diag}(Z(0), \dots, Z(N-1))$.

For $0 \leq m \leq M-1$, let $p = \lfloor \frac{m}{2} \rfloor$, $0 \leq p \leq \lfloor \frac{M-1}{2} \rfloor$ and $q = m - 2p \in \{0, 1\}$. Let $\mathbf{X}_p = \mathbf{\Lambda}_p \mathbf{X}$ and $\mathbf{Z}_p = \mathbf{\Lambda}_p \mathbf{Z}$, $0 \leq p \leq \lfloor \frac{M-1}{2} \rfloor$, where $\mathbf{\Lambda}_p$ is defined in (6.12). The diagonal FD training matrices of the m th antenna in the 0th and 1st training symbols are given by \mathbf{D}_{0m} and \mathbf{D}_{1m} respectively:

$$\mathbf{D}_{0m} = \begin{cases} \mathbf{\Lambda}_p \mathbf{D}_X, & \text{if } m = 2p \\ \mathbf{\Lambda}_p^H \mathbf{D}_Z^H, & \text{if } m = 2p + 1 \end{cases}, \quad \mathbf{D}_{1m} = \begin{cases} \mathbf{\Lambda}_p \mathbf{D}_Z, & \text{if } m = 2p \\ -\mathbf{\Lambda}_p^H \mathbf{D}_X^H, & \text{if } m = 2p + 1 \end{cases}. \quad (6.24)$$

Theorem 6.3.3 (Alamouti Signaling). *The FD training sequences in (6.24) are optimal for two training OFDM symbols.*

Proof. See Appendix 7.5.2. □

If all the users employ two transmit antennas and Alamouti code, their training sequences in two symbol intervals are assigned according to (6.24), which can be generated simply using the same Alamouti code generator, thus greatly reduce the training assignment complexity.

6.3.4 Discussions

The PAPR of the training sequence $S(n)$, $0 \leq n \leq N-1$, is given by

$$\text{PAPR} = \frac{\max_n |S(n)|^2}{\frac{1}{N} \sum_{n=0}^{N-1} |S(n)|^2}. \quad (6.25)$$

The transform operator $\mathbf{\Lambda}_m$ between different FD training sequences can be viewed as a frequency modulation, which is equivalent to circulant shift of the training sequence in the TD. Hence, we have the following proposition.

Proposition 6.3.1. *All TD training sequences of \mathbf{X}_m 's in (6.13) have the same PAPR.*

This property is important when designing the training sequences. As long as the PAPR of \mathbf{X} is low, all training symbols will have low PAPR. Another merit of our design is that if $\frac{N}{(\nu_{\max}+1)} = 2^k$, for some integer k , and if \mathbf{X} is chosen from a 2^k -phase shift keying (PSK) constellation, then the transform $\mathbf{\Lambda}_m$ guarantees that all FD training sequences $\{\mathbf{X}_m, 0 \leq m \leq M-1\}$ will belong to the same 2^k -PSK constellation, which is very easy to generate.

One possible choice for \mathbf{X} is a Constant-Amplitude-Zero-Auto-Correlation (CAZAC) sequence [140], which is a complex-valued sequence with constant amplitude and zero autocorrelation at nonzero lags. One example of a CAZAC sequence of length N is the chirp sequence given by

$$X(k) = \begin{cases} \sqrt{c} \exp(j \frac{\pi u k^2}{N}), & \text{if } N \text{ is even} \\ \sqrt{c} \exp(j \frac{\pi u k(k+1)}{N}), & \text{if } N \text{ is odd} \end{cases}, 0 \leq k \leq N-1. \quad (6.26)$$

where u is any integer relatively prime¹ to N . A disadvantage of this and other CAZAC sequences is that the entries are not restricted to a standard signal constellation. An alternative is provided by Golay complementary sequences [94] which only assume values from $\{-\sqrt{c}, \sqrt{c}\}$. A third possibility is the flat sequence (impulsive in TD) $\{\mathbf{X} : X(k) = \sqrt{c}, \text{ for all } k\}$. These three choices have different PAPRs, as summarized in Table 6.1. Given the above discussion, it is possible to generate a family of optimal training sequences with low PAPR from a standard signal constellation.

	Chirp-based	Golay-based	TD Impulsive
PAPR	0 dB	≤ 3 dB	10 dB

Table 6.1: PAPR Comparison of Three Training Sequence Candidates

It is also possible to exploit alternative design to reduce the number of pilots per training symbol. Assume the number of subcarriers N can be decomposed as $N = N_p T$ where $N_p \geq (\nu_{\max} + 1)$, then it is possible to use N_p equally-spaced pilots in each training symbol instead of the whole symbol. At the m th antenna, the training sequence is given by diagonal matrix $\hat{\mathbf{D}}_m \in \mathbb{C}^{N_p \times N_p}$ and the pilot locations are $\{sT\}_{s=0}^{N_p-1}$. Consider the one training symbol scenario without loss of generality. Instead of taking IDFT of (6.5) of length N , we now take IDFT of (6.5) *only at pilot tones* of length N_p , and get

$$\hat{\mathbf{y}} = \sum_{m=0}^{M-1} \hat{\mathbf{S}}_m \mathbf{h}_m + \mathbf{n}, \quad (6.27)$$

where $\hat{\mathbf{y}} \in \mathbb{C}^{N_p}$, $\hat{\mathbf{S}}_m = \hat{\mathbf{F}}^H \hat{\mathbf{D}}_m \hat{\mathbf{F}}_0 \in \mathbb{C}^{N_p \times (\nu_{\max}+1)}$, $\hat{\mathbf{F}}$ is the DFT matrix of size N_p , $\hat{\mathbf{F}}_0 = [\hat{f}_{st} = e^{j \frac{2\pi st}{N_p}}] \in \mathbb{C}^{N_p \times (\nu_{\max}+1)}$ is the submatrix of \mathbf{F}_0 at rows corresponding to pilot frequencies. It is obvious that $\hat{\mathbf{F}}_0$ is also the first $(\nu_{\max} + 1)$ columns of the

¹Two integers are said to be relatively prime if their greatest common divisor is 1.

DFT matrix $\hat{\mathbf{F}}$, and $\hat{\mathbf{F}}_0^H \hat{\mathbf{F}}_0 = \frac{1}{T} \mathbf{I}_{\nu_{\max}+1}$. Therefore, it is clear that we can follow the same framework in both single and multiple training symbol scenarios, by replacing N by N_p in both the dimensionality conditions and design parameters at the cost of increasing the MMSE by a factor of T .

6.4 Numerical Results

We consider uplink transmission in a Multi-User MIMO-OFDM system with $N = 64$ and $\nu_{\max} = 15$. Each BST is equipped with two co-located receive antennas and two users are each equipped with two transmit antennas over which the Alamouti STBC is employed. Each user employs a non-systematic rate-1/2 convolutional code with octal generator (133, 171) and constraint length 7, which are further QPSK modulated. All channel paths are assumed to have uncorrelated and identically-distributed CIRs with 8 zero-mean complex Gaussian taps following an exponentially-decaying power delay profile with a 3 dB decay per tap. K OFDM training symbols are transmitted over each transmit antenna for channel estimation as described in Section 6.3. The CIR estimates are used for detection of the OFDM data symbols through the joint Linear Minimum-Mean-Square-Error (LMMSE) technique [141] where the received signals from the two receive antennas are processed jointly to separate the two users.

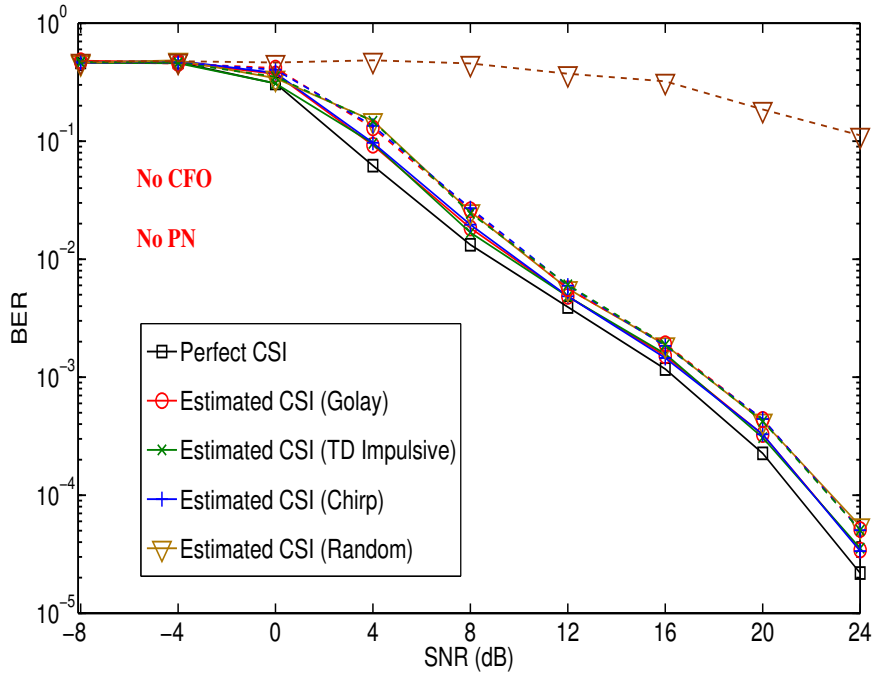


Figure 6.2: BER versus SNR for $K = 1$ (dashed) and 2 (solid) training OFDM symbols from optimal training sequences (generated from Golay, TD impulsive and chirp) and non-optimal random sequence compared with perfect CSI.

Using these parameters, the dimensionality condition in (6.19) is met with $K \geq 1$. In Fig. 6.2, the Bit Error Rate (BER) performances of three FD training sequences

proposed in Section 3.4 (namely: Chirp, Golay, and TD Impulsive) with $K = 1$ and 2 are compared with the perfect CSI case. All training sequences can be generated from standard QPSK constellation except chirp sequences. In Fig. 6.2, all users are assumed to have perfect frequency synchronization with the receiver. All training sequences achieve roughly the same BER performance with SNR losses of 1.5 and 0.7 dB for $K = 1$ and 2, respectively compared with the perfect CSI case. The performance of a random BPSK sequence *not* satisfying the optimality condition is also shown for comparison. The performance of the random sequence is inferior to that of the other sequences satisfying the optimality condition; especially with $K = 1$ training symbol where the number of equations equals the number of unknowns making the channel estimate unreliable when the optimality condition is not satisfied. From another perspective, our optimally-designed training sequences with $K = 1$ training symbol achieve comparable performance to that of the random sequence with $K = 2$ training symbols, i.e. with 50% less training overhead. This is besides the additional complexity needed to invert the matrix $\mathbf{S}^H \mathbf{S}$ which is not a scaled identity in the case of non-optimal sequences.

We further examine the performance of the optimal training sequences when Phase Noise (PN) is present at the transmitter and receivers in Fig. 6.3 when its variance $\sigma_{pn}^2 = 0, 10^{-5}$ and 10^{-4} . Although all the optimal training sequences obtain MMSE, their PAPR and robustness to phase noise and carrier frequency offset is different. The readers are referred to [8] which explores this tradeoff in detail.

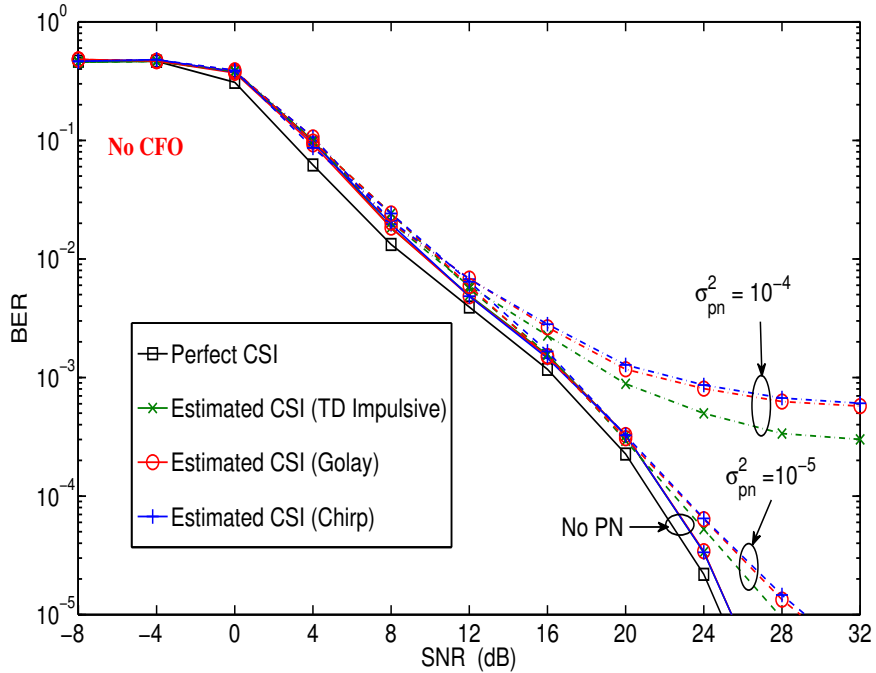


Figure 6.3: BER versus SNR for different PN levels $\sigma_{pn}^2 = 0$ (solid), 10^{-5} (dashed), and 10^{-4} (dash-dotted) with $K = 2$ training OFDM symbols.

Chapter 7

Appendix

7.1 Proofs in Chapter 2

7.1.1 Preparations for Proofs

The Statistical Orthogonality Condition (StOC) for a measurement matrix \mathbf{X} is first introduced in [23] and defined as below.

Definition 7.1.1. *Let $\bar{\Pi} = (\pi_1, \dots, \pi_p)$ be a random permutation of $\{1, \dots, p\}$, and define $\Pi = (\pi_1, \dots, \pi_k)$ and $\Pi^c = (\pi_{k+1}, \dots, \pi_p)$ for any $k \leq p$. Then the matrix \mathbf{X} is said to satisfy the (k, ϵ, δ) -StOC, if there exist $\epsilon, \delta \in [0, 1)$ such that the inequalities*

$$\|(\mathbf{X}_{\Pi}^H \mathbf{X}_{\Pi} - \mathbf{I})\mathbf{z}\|_{\infty} \leq \epsilon \|\mathbf{z}\|_2, \quad (7.1)$$

$$\|\mathbf{X}_{\Pi^c}^H \mathbf{X}_{\Pi} \mathbf{z}\|_{\infty} \leq \epsilon \|\mathbf{z}\|_2, \quad (7.2)$$

hold for every fixed $\mathbf{z} \in \mathbb{C}^k$ with probability exceeding $1 - \delta$, with respect to $\bar{\Pi}$.

We have the following proposition rephrased from [23] stating that the StOC is satisfied with high probability if X satisfies the strong coherence property.

Proposition 7.1.2 ([23]). *If the matrix \mathbf{X} satisfies the strong coherence property, then it satisfies (k, ϵ, δ) -StOC for $k \leq n/(2 \log p)$, with $\epsilon = 10\mu\sqrt{2 \log p}$ and $\delta \leq 4p^{-1}$.*

If (7.1) and (7.2) hold for a realization of permutation $\bar{\Pi}$, then for $t \leq k$, let $\Pi_t = (\pi_1, \dots, \pi_t)$ and $\Pi_t^c = (\pi_{t+1}, \dots, \pi_k)$, so that $\Pi_t \cup \Pi_t^c = \Pi$ and $\Pi_t \cap \Pi_t^c = \emptyset$. For every $\mathbf{z} \in \mathbb{C}^t$, we have

$$\left\| \begin{bmatrix} \mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi_t} - \mathbf{I}_t & \mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi_t^c} \\ \mathbf{X}_{\Pi_t^c}^H \mathbf{X}_{\Pi_t} & \mathbf{X}_{\Pi_t^c}^H \mathbf{X}_{\Pi_t^c} - \mathbf{I}_{k-t} \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{0}_{k-t} \end{bmatrix} \right\|_{\infty} \leq \epsilon \|\mathbf{z}\|_2,$$

from (7.1), therefore

$$\|(\mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi_t} - \mathbf{I}_t)\mathbf{z}\|_{\infty} \leq \epsilon \|\mathbf{z}\|_2, \quad \text{and} \quad \|\mathbf{X}_{\Pi_t^c}^H \mathbf{X}_{\Pi_t} \mathbf{z}\|_{\infty} \leq \epsilon \|\mathbf{z}\|_2.$$

Moreover, from (7.2) we have

$$\|\mathbf{X}_{\Pi^c}^H \mathbf{X}_{\Pi_t} \mathbf{z}\|_\infty = \left\| \begin{bmatrix} \mathbf{X}_{\Pi^c}^H \mathbf{X}_{\Pi_t} & \mathbf{X}_{\Pi^c}^H \mathbf{X}_{\Pi_t^c} \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \mathbf{0}_{k-t} \end{bmatrix} \right\|_\infty \leq \epsilon \|\mathbf{z}\|_2.$$

We also need the following proposition that shows a random submatrix of \mathbf{X} is well-conditioned with high probability, which is mainly due to Tropp [142], and first presented in the form below by Candès and Plan [143].

Proposition 7.1.3 ([142, 143]). *Let $\bar{\Pi} = (\pi_1, \dots, \pi_p)$ be a random permutation of $\{1, \dots, p\}$, and define $\Pi = (\pi_1, \dots, \pi_k)$ for any $k \leq p$. Then for $q = 2 \log p$ and $k \leq p/(4\|\mathbf{X}\|_2^2)$, we have*

$$\left(\mathbb{E} \left[\|\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}} - \mathbf{I}\|_2^q \right]\right)^{1/q} \leq 2^{1/q} \left(30\mu \log p + 13 \sqrt{\frac{2k\|\mathbf{X}\|_2^2 \log p}{p}} \right). \quad (7.3)$$

with respect to the random permutation $\bar{\Pi}$.

The following proposition [143] states a probabilistic bound on the extreme singular values of a random submatrix of \mathbf{X} , by applying Markov's inequality

$$\Pr \left(\|\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}} - \mathbf{I}\|_2 \geq 1/2 \right) \leq 2^q \mathbb{E} \left[\|\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}} - \mathbf{I}\|_2^q \right]$$

to Proposition 7.1.3.

Proposition 7.1.4 ([143]). *Let $\bar{\Pi} = (\pi_1, \dots, \pi_p)$ be a random permutation of $\{1, \dots, p\}$, and define $\Pi = (\pi_1, \dots, \pi_k)$ for any $k \leq p$. Suppose that $\mu(\mathbf{X}) \leq 1/(240 \log p)$ and $k \leq p/(c_2^2 \|\mathbf{X}\|_2^2 \log p)$ for numerical constant $c_2 = 104\sqrt{2}$, then we have*

$$\Pr \left(\|\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}} - \mathbf{I}\|_2 \geq 1/2 \right) \leq 2p^{-2 \log 2}.$$

Notice that $\|\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}} - \mathbf{I}\|_2 = \max\{\lambda_{\max}(\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}}) - 1, 1 - \lambda_{\min}(\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}})\}$, where $\lambda_{\max}(\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}})$ and $\lambda_{\min}(\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}})$ are the maximum and minimum eigenvalues of $\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}}$, i.e. all the eigenvalues of $\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}}$ are bounded in $[1/2, 3/2]$. If for a realization of permutation $\bar{\Pi}$, $\|\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}} - \mathbf{I}\|_2 \geq 1/2$, we have

$$\|(\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}})^{-1}\|_2 \leq 2 \quad \text{and} \quad \|\mathbf{X}_{\bar{\Pi}}(\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}})^{-1}\|_2 \leq \sqrt{2}.$$

Moreover, for $t \leq k$ and $\Pi_t = (\pi_1, \dots, \pi_t)$, we have $\|\mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi_t} - \mathbf{I}_t\|_2 \leq 1/2$, since eigenvalues of $\mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi_t}$ are majorized by eigenvalues of $\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}}$.

Finally, we bound the ℓ_∞ norm of correlated Gaussian noise in the below form. Let $\mathbf{P} \in \mathbb{C}^{n \times n}$ be a projection matrix such that $\mathbf{P}^2 = \mathbf{P}$. Since $\boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_n)$ is i.i.d. complex Gaussian noise, $\mathbf{X}^H \mathbf{P} \boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{X}^H \mathbf{P} \mathbf{X})$ is also Gaussian distributed, but is correlated with covariance matrix $\sigma^2 \mathbf{X}^H \mathbf{P} \mathbf{X}$. We want to bound $\Pr(\|\mathbf{X}^H \mathbf{P} \boldsymbol{\eta}\|_\infty \geq \tau)$ for some $\tau > 0$. First, we need the Sidak's lemma [144] below.

Lemma 7.1.5 (Sidak's lemma). *Let $[X_1, \dots, X_n]$ be a vector of random multivariate normal variables with zero means, arbitrary variances $\sigma_1^2, \dots, \sigma_n^2$ and an arbitrary*

correlation matrix. Then, for any positive numbers c_1, \dots, c_n , we have

$$\Pr(|X_1| \leq c_1, \dots, |X_n| \leq c_n) \geq \prod_{i=1}^n \Pr(|X_i| \leq c_i).$$

Since $\mathbf{X}^H \mathbf{P} \boldsymbol{\eta} \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{X}^H \mathbf{P} \mathbf{X})$, then each $\mathbf{x}_i^H \mathbf{P} \boldsymbol{\eta} \sim \mathcal{CN}(0, \sigma_i^2)$, where $\sigma_i^2 = \sigma^2 \mathbf{x}_i^H \mathbf{P} \mathbf{x}_i \leq \sigma^2$. Then

$$\Pr(|\mathbf{x}_i^H \mathbf{P} \boldsymbol{\eta}| \leq \tau) = 1 - \frac{1}{\pi} e^{-\tau^2/\sigma_i^2} \geq 1 - \frac{1}{\pi} e^{-\tau^2/\sigma^2}.$$

Following Sidak's lemma, for $\tau > 0$ we have

$$\Pr(\|\mathbf{X}^H \mathbf{P} \boldsymbol{\eta}\|_\infty \leq \tau) \geq \prod_{i=1}^p \Pr(|\mathbf{x}_i^H \mathbf{P} \boldsymbol{\eta}| \leq \tau) \geq (1 - \frac{1}{\pi} e^{-\tau^2/\sigma^2})^p \geq 1 - \frac{p}{\pi} e^{-\tau^2/\sigma^2},$$

provided the RHS is greater than zero. We have the proposition below.

Proposition 7.1.6. *Let $\boldsymbol{\eta}$ be a random vector with i.i.d. $\mathcal{CN}(0, \sigma^2)$ entries, \mathbf{P} be a projection matrix, and \mathbf{X} be a unit-column matrix, then for $\tau > 0$ we have*

$$\Pr(\|\mathbf{X}^H \mathbf{P} \boldsymbol{\eta}\|_\infty \leq \tau) \geq 1 - \frac{p}{\pi} e^{-\tau^2/\sigma^2},$$

provided the RHS is greater than zero.

Now let $\tau = \sigma \sqrt{(1 + \alpha) \log p}$ for $\alpha > 0$, we have

$$\Pr\{\|\mathbf{X}^H \mathbf{P} \boldsymbol{\eta}\|_\infty \leq \sqrt{\sigma^2(1 + \alpha) \log p}\} \geq 1 - (p^\alpha \pi)^{-1}.$$

7.1.2 Proof of Theorem 2.2.1

We first write the data vector $\boldsymbol{\beta}$ as $\boldsymbol{\beta} = \mathbf{P}_\Pi \mathbf{z}$, where $\mathbf{z} \in \mathbb{C}^k$ is a deterministic vector, and $\mathbf{P}_\Pi \in \mathbb{R}^{p \times k}$ is a partial identity matrix composed of columns indexed by Π . Then the measurement vector can be written as

$$\mathbf{y} = \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\eta} = \mathbf{X} \mathbf{P}_\Pi \mathbf{z} + \boldsymbol{\eta} = \mathbf{X}_\Pi \mathbf{z} + \boldsymbol{\eta},$$

where \mathbf{X}_Π denotes the submatrix of \mathbf{X} composed of columns indexed by Π .

We note that in OMP, the residual \mathbf{r}_t , $t = 0, \dots, k - 1$ is orthogonal to the selected columns in previous iterations, so in each iteration a new column will be selected. Define a subset Π_t which contains t variables that are selected at the t th iteration, and $\mathbf{P}_t = \mathbf{X}_{\Pi_t} (\mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi_t})^{-1} \mathbf{X}_{\Pi_t}^H$ is the projection matrix onto the linear subspace spanned by the columns of \mathbf{X}_{Π_t} , and we assume $\mathbf{P}_0 = \mathbf{0}$.

We want to prove $\Pi_t \subseteq \Pi$ by induction. First at $t = 0$, $\Pi_t = \emptyset \subset \Pi$. Assume at iteration t , $\Pi_t \subseteq \Pi$, then the residual \mathbf{r}_t can be written as

$$\mathbf{r}_t = (\mathbf{I} - \mathbf{P}_t) \mathbf{y} = (\mathbf{I} - \mathbf{P}_t) \mathbf{X}_\Pi \mathbf{z} + (\mathbf{I} - \mathbf{P}_t) \boldsymbol{\eta} \triangleq \mathbf{s}_t + \mathbf{n}_t.$$

Let $M_{\bar{\Pi}}^t = \|\mathbf{X}_{\bar{\Pi}}^H \mathbf{s}_t\|_{\infty}$, $M_{\Pi^c}^t = \|\mathbf{X}_{\Pi^c}^H \mathbf{s}_t\|_{\infty}$ and $N_t = \|\mathbf{X}^H \mathbf{n}_t\|_{\infty}$, then a sufficient condition for $\Pi_{t+1} \subseteq \Pi$, i.e. for OMP to select a correct variable at the next iteration is that

$$M_{\bar{\Pi}}^t - M_{\Pi^c}^t > 2N_t, \quad (7.4)$$

since $\|\mathbf{X}_{\bar{\Pi}}^H \mathbf{r}_t\|_{\infty} \geq M_{\bar{\Pi}}^t - N_t > M_{\Pi^c}^t + N_t \geq \|\mathbf{X}_{\Pi^c}^H \mathbf{r}_t\|_{\infty}$.

Define the event $\mathcal{G}_1 = \{\mathbf{X} \text{ satisfies the } (k, \epsilon, \delta)\text{-StOC}\}$ that happens with probability at least $1 - 4p^{-1}$ with respect to $\bar{\Pi}$ from Proposition 7.1.2, and the event $\mathcal{G}_2 = \{\|\mathbf{X}_{\bar{\Pi}}^H \mathbf{X}_{\bar{\Pi}} - \mathbf{I}\|_2 \leq 1/2\}$ that happens with probability at least $1 - 2p^{-2 \log 2}$ with respect to $\bar{\Pi}$ from Proposition 7.1.4. Let the event $\mathcal{G} = \mathcal{G}_1 \cap \mathcal{G}_2$. From the above discussions the event \mathcal{G} holds with probability at least $1 - 4p^{-1} - 2p^{-2 \log 2}$ with respect to $\bar{\Pi}$.

Now we bound $M_{\bar{\Pi}}^t$ and $M_{\Pi^c}^t$ under the event \mathcal{G} . Let $\Pi_t^c = \Pi \setminus \Pi_t$ be the set of yet to be selected indices of the support of β , and $\beta_{\Pi_t^c} = \mathbf{z}_{\Pi_t^c}$ be the corresponding coefficients. Since $(\mathbf{I} - \mathbf{P}_t)\mathbf{X}_{\Pi} \mathbf{z} \in \mathcal{R}(\mathbf{X}_{\Pi_t^c})$ belongs to the linear subspace spanned by the columns of $\mathbf{X}_{\Pi_t^c}$, we can find a vector \mathbf{w} of dimension $(k - t)$ such that $\mathbf{X}_{\Pi_t^c} \mathbf{w} = (\mathbf{I} - \mathbf{P}_t)\mathbf{X}_{\Pi} \mathbf{z}$, where the vector \mathbf{w} can be written as

$$\mathbf{w} = (\mathbf{X}_{\Pi_t^c}^H \mathbf{X}_{\Pi_t^c})^{-1} \mathbf{X}_{\Pi_t^c}^H (\mathbf{I} - \mathbf{P}_t) \mathbf{X}_{\Pi} \mathbf{z}_{\Pi_t^c} = \mathbf{z}_{\Pi_t^c} - (\mathbf{X}_{\Pi_t^c}^H \mathbf{X}_{\Pi_t^c})^{-1} \mathbf{X}_{\Pi_t^c}^H \mathbf{P}_t \mathbf{X}_{\Pi} \mathbf{z}_{\Pi_t^c}.$$

We use the following lemma from [19].

Lemma 7.1.7 ([19]). *The minimum and maximum eigenvalue of $\mathbf{X}_{\Pi_t^c}^H (\mathbf{I} - \mathbf{P}_t) \mathbf{X}_{\Pi_t^c}$ is bounded as*

$$\begin{aligned} \lambda_{\min}(\mathbf{X}_{\Pi_t^c}^H (\mathbf{I} - \mathbf{P}_t) \mathbf{X}_{\Pi_t^c}) &\geq \lambda_{\min}(\mathbf{X}_{\Pi}^H \mathbf{X}_{\Pi}), \\ \lambda_{\max}(\mathbf{X}_{\Pi_t^c}^H (\mathbf{I} - \mathbf{P}_t) \mathbf{X}_{\Pi_t^c}) &\leq \lambda_{\max}(\mathbf{X}_{\Pi}^H \mathbf{X}_{\Pi}). \end{aligned}$$

The readers are referred to [19] for the proof. Since we have

$$\|\mathbf{w}\|_2 \leq \|(\mathbf{X}_{\Pi_t^c}^H \mathbf{X}_{\Pi_t^c})^{-1}\|_2 \|\mathbf{X}_{\Pi_t^c}^H (\mathbf{I} - \mathbf{P}_t) \mathbf{X}_{\Pi} \mathbf{z}_{\Pi_t^c}\|_2 \quad (7.5)$$

$$\leq 2 \|\mathbf{X}_{\Pi}^H \mathbf{X}_{\Pi}\|_2 \|\mathbf{z}_{\Pi_t^c}\|_2 \leq 3 \|\mathbf{z}_{\Pi_t^c}\|_2, \quad (7.6)$$

where (7.5) follows from Lemma 7.1.7, and (7.6) follows from Proposition 7.1.4. Also,

$$\begin{aligned} \|\mathbf{X}_{\Pi_t^c}^H \mathbf{P}_t \mathbf{X}_{\Pi} \mathbf{z}_{\Pi_t^c}\|_{\infty} &= \|\mathbf{X}_{\Pi_t^c}^H \mathbf{X}_{\Pi_t} (\mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi_t})^{-1} \mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi} \mathbf{z}_{\Pi_t^c}\|_{\infty} \\ &\leq \epsilon \|(\mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi_t})^{-1} \mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi} \mathbf{z}_{\Pi_t^c}\|_2 \\ &\leq \epsilon \|(\mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi_t})^{-1}\|_2 \|\mathbf{X}_{\Pi_t}^H \mathbf{X}_{\Pi}\|_2 \|\mathbf{z}_{\Pi_t^c}\|_2 \leq \epsilon \|\mathbf{z}_{\Pi_t^c}\|_2, \end{aligned}$$

therefore $M_{\bar{\Pi}}^t$ can be bounded as

$$\begin{aligned} M_{\bar{\Pi}}^t &= \|\mathbf{X}_{\Pi_t^c}^H \mathbf{X}_{\Pi_t^c} \mathbf{z}_{\Pi_t^c} - \mathbf{X}_{\Pi_t^c}^H \mathbf{P}_t \mathbf{X}_{\Pi} \mathbf{z}_{\Pi_t^c}\|_{\infty} \\ &\geq \|\mathbf{z}_{\Pi_t^c}\|_{\infty} - \|(\mathbf{X}_{\Pi_t^c}^H \mathbf{X}_{\Pi_t^c} - \mathbf{I}) \mathbf{z}_{\Pi_t^c}\|_{\infty} - \|\mathbf{X}_{\Pi_t^c}^H \mathbf{P}_t \mathbf{X}_{\Pi} \mathbf{z}_{\Pi_t^c}\|_{\infty} \\ &\geq \|\mathbf{z}_{\Pi_t^c}\|_{\infty} - 2\epsilon \|\mathbf{z}_{\Pi_t^c}\|_2. \end{aligned} \quad (7.7)$$

where (7.7) follows from (7.1). Next, $M_{\Pi^c}^t$ can be bounded as

$$M_{\Pi^c}^t = \|\mathbf{X}_{\Pi^c}^H(\mathbf{I} - \mathbf{P}_t)\mathbf{X}_{\Pi}z\|_\infty = \|\mathbf{X}_{\Pi^c}^H\mathbf{X}_{\Pi^c}\mathbf{w}\|_\infty \leq \epsilon\|\mathbf{w}\|_2 \leq 3\epsilon\|\mathbf{z}_{\Pi^c}\|_2. \quad (7.8)$$

where (7.8) follows from (7.2).

Conditioned on the event \mathcal{G} , for each \mathbf{P}_t , since $\mathbf{I} - \mathbf{P}_t$ is also a projection matrix, define the event

$$\mathcal{H}_t = \{N_t \leq \sigma\sqrt{(1+\alpha)\log p}\}, \quad t = 0, \dots, k-1 \quad (7.9)$$

then from Proposition 7.1.6, it happens with probability at least $1 - (p^\alpha\pi)^{-1}$ with respect to η . We further define the event $\mathcal{H} = \bigcap_{t=0}^{k-1} \mathcal{H}_t$, then from the union bound $\Pr(\mathcal{H}|\mathcal{G}) = \Pr(\mathcal{H}) \geq 1 - k(p^\alpha\pi)^{-1}$. Similarly, for the event $\mathcal{H}' = \bigcap_{t=0}^k \mathcal{H}_t$, then from the union bound $\Pr(\mathcal{H}'|\mathcal{G}) = \Pr(\mathcal{H}') \geq 1 - (k+1)(p^\alpha\pi)^{-1}$.

Now define the event $\mathcal{I} = \mathcal{G} \cap \mathcal{H}$, from the above discussions we have $\Pr(\mathcal{I}) \geq 1 - k(p^\alpha\pi)^{-1} - 2p^{-2\log 2} - 4p^{-1}$. Now we are in order to analyze the OMP performance under the event \mathcal{G} . We want to prove $\Pi_t \subseteq \Pi$ by induction.

Plug in the above bounds (7.7), (7.8) and (7.9) into (7.4), it is sufficient that at the t th iteration

$$\|\mathbf{z}_{\Pi_t^c}\|_\infty > 5\epsilon\|\mathbf{z}_{\Pi_t^c}\|_2 + 2\sigma\sqrt{(1+\alpha)\log p}. \quad (7.10)$$

Note that $\|\mathbf{z}_{\Pi_t^c}\|_\infty \geq |\beta|_{(t+1)}$, $\|\mathbf{z}_{\Pi_t^c}\|_2 \leq \sqrt{k-t}|\beta|_{(t+1)}$, (7.10) is satisfied by the condition in (2.13) for $0 \leq t \leq k-1$, therefore a correct variable is selected at the t th iteration, $\Pi_t \subseteq \Pi$. Since the sparsity level of $\boldsymbol{\beta}$ is k , the OMP algorithm in Algorithm 1 successfully finds the support of $\boldsymbol{\beta}$ in k iterations under the event \mathcal{I} , and we've proved Theorem 2.2.1.

7.1.3 Proof of Theorem 2.2.2

Now we define the event $\mathcal{I}' = \mathcal{H}' \cap \mathcal{G}$, where $\mathcal{I}' \subset \mathcal{I}$ and happens with probability at least $1 - (k+1)(p^\alpha\pi)^{-1} - 2p^{-2\log 2} - 4p^{-1}$. Conditioned on the event \mathcal{I}' , in order to prove Theorem 2.2.2, we need to further show that $\|\mathbf{X}^H\mathbf{r}_t\|_\infty > \delta$ for $0 \leq t \leq k-1$ so that the algorithm doesn't stop early, and $\|\mathbf{X}^H\mathbf{r}_k\|_\infty \leq \delta$ so that the algorithm stops at the k th iteration. While the latter is obvious from the definition of \mathcal{H}_k , for the first inequality we have

$$\begin{aligned} \|\mathbf{X}^H\mathbf{r}_t\|_\infty &\geq M_{\Pi}^t - N_t \\ &\geq \|\mathbf{z}_{\Pi_t^c}\|_\infty - 2\epsilon\|\mathbf{z}_{\Pi_t^c}\|_2 - \sigma\sqrt{(1+\alpha)\log p} \end{aligned} \quad (7.11)$$

$$\begin{aligned} &> 3\epsilon\|\mathbf{z}_{\Pi_t^c}\|_2 + \sigma\sqrt{(1+\alpha)\log p} \\ &\geq \sigma\sqrt{(1+\alpha)\log p} = \delta, \end{aligned} \quad (7.12)$$

where (7.11) follows from (7.7) and (7.9), and (7.12) follows from (7.10).

7.1.4 Proof of Corollary 2.2.4

The proof of Corollary 2.2.4 is straightforward by early-terminating the induction procedure at the k' th iteration.

7.1.5 Proof of Corollary 2.2.5

Again we prove by induction. First at $t = 0$, $\Pi_t = \emptyset \subset \Pi$. Assume at iteration t , the OMP algorithm has successfully detected the t largest entries of $|\boldsymbol{\beta}|$. For $i \in \Pi_t^c$ that corresponds to the $t + 1$ th largest entry of $|\boldsymbol{\beta}|$, we have

$$|\mathbf{x}_i^H \mathbf{r}_t| \geq |z_i| - 2\epsilon \|\mathbf{z}_{\Pi_t^c}\|_2 - \sigma \sqrt{(1 + \alpha) \log p} = |\beta|_{(t+1)} - 2\epsilon \|\mathbf{z}_{\Pi_t^c}\|_2 - \sigma \sqrt{(1 + \alpha) \log p}$$

from a simple variation of (7.7). On the other hand, for $j \in \Pi_t^c$ that corresponds to the rest undetected entries of β , we have

$$|\mathbf{x}_j^H \mathbf{r}_t| \leq |z_j| + 2\epsilon \|\mathbf{z}_{\Pi_t^c}\|_2 + \sigma \sqrt{(1 + \alpha) \log p} \leq |\beta|_{(t+2)} + 2\epsilon \|\mathbf{z}_{\Pi_t^c}\|_2 + \sigma \sqrt{(1 + \alpha) \log p}.$$

To detect the $(t + 1)$ th largest entries it is sufficient to have

$$|\beta|_{(t+1)} - |\beta|_{(t+2)} \geq 4\epsilon \sqrt{k - t} |\beta|_{(t+1)} + 2\sigma \sqrt{(1 + \alpha) \log p}.$$

This is satisfied when (2.17) holds by simply plugging it into the above equation.

7.2 Proofs in Chapter 3

7.2.1 Proof of Theorem 3.3.1

Proof. Without loss of generality, let $|\theta_1| \geq |\theta_2| \geq \dots \geq |\theta_N|$. We have

$$\mathbf{x} = \boldsymbol{\Psi} \boldsymbol{\theta} = (\mathbf{I} + \mathbf{E}) \boldsymbol{\theta} = \boldsymbol{\theta} + \mathbf{E} \boldsymbol{\theta}, \quad (7.13)$$

where the i th entry of \mathbf{x} is given as $x_i = \theta_i + \mathbf{e}_i^T \boldsymbol{\theta}$. By the triangle inequality and Hölder's inequality we have

$$|\mathbf{e}_m^T \boldsymbol{\theta}| \leq \sum_{n=1}^N |e_{mn} \theta_n| \leq \|\mathbf{e}_m\|_p \|\boldsymbol{\theta}\|_q \leq \beta \|\boldsymbol{\theta}\|_q \quad (7.14)$$

for all $1 \leq m \leq N$. Also, by the triangle inequality we have

$$|\theta_m| - |\mathbf{e}_m^T \boldsymbol{\theta}| \leq |\theta_m + \mathbf{e}_m^T \boldsymbol{\theta}| \leq |\theta_m| + |\mathbf{e}_m^T \boldsymbol{\theta}|. \quad (7.15)$$

By combining the above two inequalities, we obtain

$$|\theta_m| - \beta \|\boldsymbol{\theta}\|_q \leq |\theta_m + \mathbf{e}_m^T \boldsymbol{\theta}| \leq |\theta_m| + \beta \|\boldsymbol{\theta}\|_q. \quad (7.16)$$

Let Σ_{N-k} denote any subset of $\{1, \dots, N\}$ with cardinality $(N - k)$. To get the upper bound, we can write

$$\begin{aligned} \|\mathbf{x} - \mathbf{x}_k\|_1 &= \min_{\Sigma_{N-k}} \sum_{m \in \Sigma_{N-k}} |x_m| \\ &= \min_{\Sigma_{N-k}} \sum_{m \in \Sigma_{N-k}} |\theta_m + \mathbf{e}_m^T \boldsymbol{\theta}| \\ &\leq \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 + (N - k)\beta \|\boldsymbol{\theta}\|_q \end{aligned} \quad (7.17)$$

where (7.17) follows from (7.16). The upper bound is achieved when (3.11) holds with the positive sign, given by

$$\begin{aligned} &\min_{\Sigma_{N-k}} \sum_{m \in \Sigma_{N-k}} |\theta_m + \mathbf{e}_m^T \boldsymbol{\theta}| \\ &= \min_{\Sigma_{N-k}} \sum_{m \in \Sigma_{N-k}} \left| \theta_m + e^{j \arg(\theta_m)} \beta \left(\sum_{n=1}^N \left(\frac{|\theta_n|}{\|\boldsymbol{\theta}\|_q} \right)^{q/p} e^{-j \arg(\theta_n)} \cdot \theta_n \right) \right| \\ &= \min_{\Sigma_{N-k}} \sum_{m \in \Sigma_{N-k}} |\theta_m| + (N - k)\beta \left(\sum_{n=1}^N \left(\frac{|\theta_n|}{\|\boldsymbol{\theta}\|_q} \right)^{q/p} |\theta_n| \right) \\ &= \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 + (N - k)\beta \frac{\sum_{n=1}^N |\theta_n|^q}{\|\boldsymbol{\theta}\|_q^{q/p}} \end{aligned} \quad (7.18)$$

$$= \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 + (N - k)\beta \|\boldsymbol{\theta}\|_q, \quad (7.19)$$

where in writing (7.18) and (7.19) we have used the fact that $1/p + 1/q = 1$. To get the lower bound, note that

$$\begin{aligned} \sum_{m \in \Sigma_{N-k}} |x_m| &= \sum_{m \in \Sigma_{N-k}} |\theta_m + \mathbf{e}_m^T \boldsymbol{\theta}| \\ &\geq \sum_{m \in \Sigma_{N-k}} |\theta_m| - (N - k)\beta \|\boldsymbol{\theta}\|_q \\ &\geq \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 - (N - k)\beta \|\boldsymbol{\theta}\|_q \end{aligned} \quad (7.20)$$

for all Σ_{N-k} . Then,

$$\|\mathbf{x} - \mathbf{x}_k\|_1 = \min_{\Sigma_{N-k}} \sum_{m \in \Sigma_{N-k}} |x_m| \geq \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 - (N - k)\beta \|\boldsymbol{\theta}\|_q, \quad (7.21)$$

and it is easy to see that the bound is achieved when (3.11) holds with the negative sign. □

7.2.2 Proof of Theorem 3.3.2

Proof. By taking $\beta = \eta$ in Theorem 3.3.1 and letting \mathbf{E} be the perturbation matrix with $\|\mathbf{e}_m\|_p = \beta = \eta$, we have

$$\|\mathbf{x} - \mathbf{x}_k\|_1 = \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 + (N - k)\eta\|\boldsymbol{\theta}\|_q \quad (7.22)$$

for $\mathbf{x} = (\mathbf{I} + \mathbf{E})\boldsymbol{\theta}$. Since $\mathbf{E} \in \{\tilde{\mathbf{E}} : \|\tilde{\mathbf{e}}_i\|_p \geq \eta\}$, then

$$\max_{\tilde{\mathbf{E}}: \|\tilde{\mathbf{e}}_i\|_p \geq \eta} \|\mathbf{x} - \mathbf{x}_k\|_1 \geq \|\boldsymbol{\theta} - \boldsymbol{\theta}_k\|_1 + (N - k)\eta\|\boldsymbol{\theta}\|_q. \quad (7.23)$$

□

7.3 Proofs in Chapter 4

7.3.1 Proof of Theorem 4.4.1

Proof. We first define $\beta_{\Omega_n}^n = \frac{\|\mathbf{d}_{\Omega_n}\|_2^2 \|\mathbf{d}_{\Omega_n}^n\|_2^2}{(\mathbf{d}_{\Omega_n}^T \mathbf{d}_{\Omega_n}^n)^2}$. Obviously $\beta_{\Omega_n}^n \geq 1$ for all n . With this we have $\|\mathbf{d}_{\Omega_n}^n\|_2^2 = \beta_{\Omega_n}^{n-1} \|\mathbf{d}_{\Omega_n}^{n-1}\|_2^2$, and $\mathbf{d}_{\Omega_n}^T \mathbf{d}_{\Omega_n}^n = \beta_{\Omega_n}^{n-1} \mathbf{d}_{\Omega_n}^T \mathbf{d}_{\Omega_n}^{n-1}$. Combining with $\mathbf{d}_{\Omega_n^c}^n = \mathbf{d}_{\Omega_n^c}^{n-1}$, γ_n can be written as

$$\begin{aligned} \gamma_n &= \frac{(\mathbf{d}_{\Omega_n}^T \mathbf{d}_{\Omega_n}^n + \mathbf{d}_{\Omega_n^c}^T \mathbf{d}_{\Omega_n^c}^n)^2}{\|\mathbf{d}\|_2^2 (\|\mathbf{d}_{\Omega_n}^n\|_2^2 + \|\mathbf{d}_{\Omega_n^c}^n\|_2^2)} \\ &= \frac{(\beta_{\Omega_n}^{n-1} \mathbf{d}_{\Omega_n}^T \mathbf{d}_{\Omega_n}^{n-1} + \mathbf{d}_{\Omega_n^c}^T \mathbf{d}_{\Omega_n^c}^{n-1})^2}{\|\mathbf{d}\|_2^2 (\beta_{\Omega_n}^{n-1} \|\mathbf{d}_{\Omega_n}^{n-1}\|_2^2 + \|\mathbf{d}_{\Omega_n^c}^{n-1}\|_2^2)} \\ &= \gamma^{n-1} \frac{\left[(\beta_{\Omega_n}^{n-1} - 1) \frac{\mathbf{d}_{\Omega_n}^T \mathbf{d}_{\Omega_n}^{n-1}}{\mathbf{d}^T \mathbf{d}^{n-1}} + 1 \right]^2}{(\beta_{\Omega_n}^{n-1} - 1) \frac{\|\mathbf{d}_{\Omega_n}^{n-1}\|_2^2}{\|\mathbf{d}^{n-1}\|_2^2} + 1}, \end{aligned} \quad (7.24)$$

Next we define

$$q_i^{n-1} = \frac{1}{\mathbf{d}^T \mathbf{d}^{n-1}} d_i d_i^{n-1} \quad \text{and} \quad p_i^{n-1} = \frac{1}{\|\mathbf{d}^{n-1}\|_2^2} (d_i^{n-1})^2$$

for $1 \leq i \leq M$, then $|q_i^{n-1}| \leq \frac{1}{M} \left(\frac{\mu(\mathbf{d})\mu(\mathbf{d}_{n-1})}{\gamma_{n-1}} \right)^{1/2}$ and $0 \leq p_i^{n-1} \leq \frac{\mu(\mathbf{d}_{n-1})}{M}$, and we have

$$\begin{aligned} \gamma_n &= \gamma_{n-1} \frac{\left[(\beta_{\Omega_n}^{n-1} - 1) \sum_{i=1}^{|\Omega_n|} q_{\Omega_n(i)}^{n-1} + 1 \right]^2}{(\beta_{\Omega_n}^{n-1} - 1) \sum_{i=1}^{|\Omega_n|} p_{\Omega_n(i)}^{n-1} + 1} \\ &= \gamma_{n-1} \frac{\left[(\beta_{\Omega_n}^{n-1} - 1) \sum_{i=1}^{|\Omega_n|} (p_{\Omega_n(i)}^{n-1} + q_{\Omega_n(i)}^{n-1} - p_{\Omega_n(i)}^{n-1}) + 1 \right]^2}{(\beta_{\Omega_n}^{n-1} - 1) \sum_{i=1}^M p_{\Omega_n(i)}^{n-1} + 1} \\ &\geq \gamma_{n-1} \left[(\beta_{\Omega_n}^{n-1} - 1) \sum_{i=1}^{|\Omega_n|} (2q_{\Omega_n(i)}^{n-1} - p_{\Omega_n(i)}^{n-1}) + 1 \right]. \end{aligned}$$

By applying the Hoeffding's inequality [145] to

$$X_i = 2q_{\Omega_n(i)}^{n-1} - p_{\Omega_n(i)}^{n-1}, \quad i = 1, \dots, |\Omega_n|,$$

where X_i 's are independent random variables satisfying

$$\begin{aligned} X_i &\leq 2|q_{\Omega_n(i)}^{n-1}| - p_{\Omega_n(i)}^{n-1} \leq \frac{2}{M} \left(\frac{\mu(\mathbf{d})\mu(\mathbf{d}_{n-1})}{\gamma_{n-1}} \right)^{\frac{1}{2}}, \\ X_i &\geq -2|q_{\Omega_n(i)}^{n-1}| - p_{\Omega_n(i)}^{n-1} \geq -\frac{2}{M} \left(\frac{\mu(\mathbf{d})\mu(\mathbf{d}_{n-1})}{\gamma_{n-1}} \right)^{\frac{1}{2}} - \frac{\mu(\mathbf{d}_{n-1})}{M}, \end{aligned}$$

and $\mathbb{E}[\sum_{i=1}^{|\Omega_n|} X_i] = \mathbb{E}[\sum_{i=1}^M \mathbf{1}_{\{i \in \Omega_n\}} (2q_i^{n-1} - p_i^{n-1})] = \frac{|\Omega_n|}{M}$, given $\mathbf{1}_{\{\mathcal{G}\}}$ is the indicator function of event \mathcal{G} , we have for $t > 0$,

$$\begin{aligned} \Pr \left(\sum_{i=1}^{|\Omega_n|} X_i - \frac{|\Omega_n|}{M} \leq -t \right) &\leq \exp \left(-\frac{t^2 M^2 \gamma_{n-1}}{[4\mu(\mathbf{d})^{1/2} + (\gamma_{n-1}\mu(\mathbf{d}_{n-1}))^{1/2}]^2 \mu(\mathbf{d}_{n-1}) |\Omega_n|} \right) \\ &\leq \exp \left(-\frac{t^2 M^2 \gamma_{n-1}}{2[16\mu(\mathbf{d}) + \gamma_{n-1}\mu(\mathbf{d}_{n-1})] \mu(\mathbf{d}_{n-1}) |\Omega_n|} \right), \end{aligned}$$

where the second inequality follows from $(a+b)^2 \leq 2(a^2 + b^2)$. Now let $t = \frac{|\Omega_n|}{M}$, we have

$$\begin{aligned} \Pr \left(\sum_{i=1}^{|\Omega_n|} X_i \leq 0 \right) &\leq \exp \left(-\frac{\gamma_{n-1} |\Omega_n|}{2[16\mu(\mathbf{d}) + \gamma_{n-1}\mu(\mathbf{d}_{n-1})] \mu(\mathbf{d}_{n-1})} \right) \\ &\leq \exp \left(-\frac{\gamma_{n-1} |\Omega_n|}{2[16\mu(\mathbf{d}) + \mu(\mathbf{d}_{n-1})] \mu(\mathbf{d}_{n-1})} \right). \end{aligned}$$

Since

$$\begin{aligned} \Pr(\gamma_n \geq \gamma_{n-1}) &\geq \Pr\left(\left(\beta_{\Omega_n}^{n-1} - 1\right) \sum_{i=1}^{|\Omega_n|} (2q_{\Omega_n(i)}^{n-1} - p_{\Omega_n(i)}^{n-1}) \geq 0\right) \\ &\geq \Pr\left(\sum_{i=1}^{|\Omega_n|} (2q_{\Omega_n(i)}^{n-1} - p_{\Omega_n(i)}^{n-1}) \geq 0\right), \end{aligned}$$

therefore we have $\gamma_n \geq \gamma_{n-1}$ with probability at least $1 - \exp\left(-\frac{\gamma_{n-1}|\Omega_n|}{2[16\mu(\mathbf{d}) + \gamma_{n-1}\mu(\mathbf{d}_{n-1})]\mu(\mathbf{d}_{n-1})}\right)$. \square

7.4 Proofs in Chapter 5

7.4.1 Proof of Theorem 5.3.3

Proof. If we express $0 \leq n \leq N - 1$ as $n = rm + i$, where $0 \leq r \leq q - 1$ and $0 \leq i \leq m - 1$, then using binomial expansion for $n^t = (rm + i)^t$ we can write $f_{\mathcal{P}}^{(t)}(\theta_0)$, $\theta_0 = 2\pi l/m$, where l and $m \neq 1$ are co-prime integers, as

$$\begin{aligned} f_{\mathcal{P}}^{(t)}(\theta_0) &= j^t \sum_{n=0}^{N-1} n^t p_n e^{jn\theta_0} \\ &= j^t \sum_{r=0}^{q-1} \sum_{i=0}^{m-1} (rm + i)^t p_{rm+i} e^{j(rm+i)\frac{2\pi l}{m}} \\ &= j^t \sum_{r=0}^{q-1} \sum_{i=0}^{m-1} \sum_{u=0}^t \binom{t}{u} i^{t-u} (rm)^u p_{rm+i} e^{j\frac{2\pi li}{m}} \\ &= j^t \sum_{u=0}^t \binom{t}{u} m^u \sum_{i=0}^{m-1} i^{t-u} e^{j\frac{2\pi li}{m}} \left[\sum_{r=0}^{q-1} r^u p_{rm+i} \right]. \end{aligned} \quad (7.25)$$

Define a length- q sequence $\{b_r\}_{r=0}^{q-1}$ as $b_r = p_{rm+i}$, $0 \leq r \leq q - 1$. If $\{b_r\}_{r=0}^{q-1}$ satisfies

$$\sum_{r=0}^{q-1} r^u b_r = 0, \quad \text{for all } 0 \leq u < t, \quad (7.26)$$

then the coefficient $f_{\mathcal{P}}^{(t)}(\theta_0)$ will be zero. From (5.14), it follows that the zero-forcing condition in (7.26) will be satisfied if $\{b_r\}_{r=0}^{q-1}$ is the PTM sequence of length 2^t . We note that $f_{\mathcal{P}}^{(M)}(\theta_0)$ is automatically zero as $\sum_{i=0}^{m-1} e^{j\frac{2\pi li}{m}} = 0$. Therefore, to zero-force the derivatives $f_{\mathcal{P}}^{(t)}(\theta_0)$ for all $t \leq M$, it is sufficient to select $\mathcal{P} = \{p_n\}_{n=0}^{2^M m - 1}$ such that each $\{p_{rm+i}\}_{r=0}^{q-1}$, $i = 0, \dots, m - 1$ is the length- 2^M PTM sequence. We call such a sequence a $(2^M, m)$ -PTM sequence. The $(2^M, m)$ -PTM sequence has length $2^M \times m$ and is constructed from the length- 2^M PTM sequence by repeating each 1 and -1

in the PTM sequence m times, that is oversampling the PTM sequence by a factor m . \square

7.4.2 Proof of Corollary 5.3.2

Proof. The proof for $\theta_0 = 0$ is straightforward. For $\theta_0 = 2\pi l/(2^h m)$, we have

$$f_{\mathcal{P}}^{(t)}(\theta_0) = \sum_{u=0}^t \binom{t}{u} (2^h m)^u \sum_{i=0}^{2^h m-1} i^{t-u} e^{j \frac{2\pi l i}{2^h m}} \left[\sum_{r=0}^{\frac{q}{2^h}-1} r^u p_{2^h m r+i} \right] \quad (7.27)$$

where $q = 2^M$. The corollary follows from the fact that downsampling a PTM sequence by a power of 2 produces a PTM sequence of shorter length. \square

7.4.3 Proof of Theorem 5.3.4

Proof. Let $m' = m/2$ and $0 \leq t \leq M$, the coefficient $f_{\mathcal{P}}^{(t)}(\theta_0)$ is given by

$$\begin{aligned} f_{\mathcal{P}}^{(t)}(\theta_0) &= j^t \sum_{n=0}^{2^{M-1}m-1} n^t p_n e^{jn\theta_0} \\ &= j^t \sum_{n=0}^{2^M m'-1} n^t (-1)^n \tilde{p}_n e^{j \frac{2\pi l n}{m}} \\ &= j^t \sum_{n=0}^{2^M m'-1} n^t (-1)^n \tilde{p}_n (-1)^n e^{j \frac{\pi(m+2l)n}{m}} \\ &= j^t \sum_{n=0}^{2^M m'-1} n^t \tilde{p}_n e^{j \frac{\pi(m'+l)n}{m'}} \\ &= f_{\tilde{\mathcal{P}}}^{(t)}(\tilde{\theta}_0) \end{aligned}$$

where $\tilde{\theta}_0 = (m' + l)\pi/m'$. The last equation equals to zero following from Theorem 5.3.3, therefore \mathcal{P} has M th-order nulls at θ_0 . \square

7.5 Proofs in Chapter 6

7.5.1 Proof of Theorem 6.3.2

Proof. It is enough to show that

$$\sum_{t=0}^{K-1} \mathbf{S}_{tm}^H \mathbf{S}_{tn} = \mathbf{F}_0^H \left(\sum_{t=0}^{K-1} \mathbf{D}_{tm}^H \mathbf{D}_{tn} \right) \mathbf{F}_0 = \tilde{c} \delta_{mn} \mathbf{I}_{(\nu_{\max}+1)} \quad (7.28)$$

It is obvious that when $m = n$, the above equation holds. When $m \neq n$, we write $m = Kp_1 + q_1$ and $n = Kp_2 + q_2$, then $\mathbf{D}_{tm} = \Sigma_{tq_1} \Lambda_{p_1}$ and $\mathbf{D}_{tn} = \Sigma_{tq_2} \Lambda_{p_2}$ for $0 \leq t \leq K - 1$. The proof can be splitted into two cases:

- When $q_1 = q_2 = q \in \{0, \dots, K - 1\}$ but $p_1 \neq p_2$, we have

$$\sum_{t=0}^{K-1} \mathbf{S}_{tm}^H \mathbf{S}_{tn} = \mathbf{F}_{p_1}^H \left(\sum_{t=0}^{K-1} \Sigma_{tq}^H \Sigma_{tq} \right) \mathbf{F}_{p_2} = \tilde{c} \mathbf{F}_{p_1}^H \mathbf{F}_{p_2} = \mathbf{0}_{(\nu_{\max}+1)}.$$

- When $q_1 \neq q_2$, we have

$$\sum_{t=0}^{K-1} \mathbf{D}_{tm}^H \mathbf{D}_{tn} = \Lambda_{p_1}^H \left(\sum_{t=0}^{K-1} \Sigma_{t,q_1}^H \Sigma_{t,q_2} \right) \Lambda_{p_2} = \mathbf{0}_N. \quad (7.29)$$

Now, Eq. (7.28) follows trivially. \square

7.5.2 Proof of Theorem 6.3.3

Proof. It is enough to show that

$$\mathbf{S}_{0m}^H \mathbf{S}_{0n} + \mathbf{S}_{1m}^H \mathbf{S}_{1n} = \mathbf{F}_0^H (\mathbf{D}_{0m}^H \mathbf{D}_{0n} + \mathbf{D}_{1m}^H \mathbf{D}_{1n}) \mathbf{F}_0 = \tilde{c} \delta_{mn} \mathbf{I}_{(\nu_{\max}+1)}. \quad (7.30)$$

It is obvious that the above equation holds when $m = n$. When $m \neq n$, we write $m = 2p_1 + q_1$ and $n = 2p_2 + q_2$ and consider two cases

- $q_1 = q_2$ but $p_1 \neq p_2$. Without loss of generality, we assume $q_1 = q_2 = 0$ and get

$$\begin{aligned} \mathbf{S}_{0m}^H \mathbf{S}_{0n} &= \mathbf{F}_0^H \mathbf{D}_{0m}^H \mathbf{D}_{0n} \mathbf{F}_0 = \mathbf{F}_{p_1}^H \mathbf{D}_X^H \mathbf{D}_X \mathbf{F}_{p_2}, \\ \mathbf{S}_{1m}^H \mathbf{S}_{1n} &= \mathbf{F}_0^H \mathbf{D}_{1m}^H \mathbf{D}_{1n} \mathbf{F}_0 = \mathbf{F}_{p_1}^H \mathbf{D}_Z^H \mathbf{D}_Z \mathbf{F}_{p_2}. \end{aligned}$$

Hence we have,

$$\mathbf{S}_{0m}^H \mathbf{S}_{0n} + \mathbf{S}_{1m}^H \mathbf{S}_{1n} = \mathbf{F}_{p_1}^H (\mathbf{D}_X^H \mathbf{D}_X + \mathbf{D}_Z^H \mathbf{D}_Z) \mathbf{F}_{p_2} = \tilde{c} \mathbf{F}_{p_1}^H \mathbf{F}_{p_2} = \mathbf{0}_{(\nu_{\max}+1)}. \quad (7.31)$$

- $q_1 \neq q_2$. Without loss of generality, we assume $q_1 = 0$ and $q_2 = 1$ and get

$$\begin{aligned} \mathbf{D}_{0m}^H \mathbf{D}_{0n} + \mathbf{D}_{1m}^H \mathbf{D}_{1n} &= (\Lambda_{p_1} \mathbf{D}_X)^H (\Lambda_{p_2}^H \mathbf{D}_Z^H) + (\Lambda_{p_1} \mathbf{D}_Z)^H (-\Lambda_{p_2}^H \mathbf{D}_X^H) \\ &= \Lambda_{p_1}^H (\mathbf{D}_X^H \mathbf{D}_Z^H - \mathbf{D}_Z^H \mathbf{D}_X^H) \Lambda_{p_2}^H = \mathbf{0}_{(\nu_{\max}+1)}. \end{aligned}$$

therefore (7.30) follows directly. \square

Bibliography

- [1] Y. Chi, L. L. Scharf, A. Pezeshki, and R. Calderbank, “Sensitivity to basis mismatch of compressed sensing for spectrum analysis and beamforming,” in *Proc. 6th U.S./Australia Joint Workshop on Defence Applications of Signal Processing (DASP)*, Lihue, HI, Sep. 26-28 2009.
- [2] Y. Chi, A. Pezeshki, L. L. Scharf, and R. Calderbank, “Sensitivity to basis mismatch in compressed sensing,” in *Proc. 2010 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Dallas, TX, Mar. 2010.
- [3] Y. Chi, L. L. Scharf, A. Pezeshki, and R. Calderbank, “Sensitivity to basis mismatch in compressed sensing,” *IEEE Trans. Signal Processing*, vol. 59, no. 5, pp. 2182–2195, 2011.
- [4] Y. Chi, Y. Eldar, and R. Calderbank, “PETRELS: Subspace estimation and tracking from partial observations,” in *Proc. 2012 IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*.
- [5] —, “PETRELS: Subspace estimation and tracking from partial observations,” *IEEE Trans. Signal Processing*, submitted.
- [6] Y. Chi, A. Pezeshki, R. Calderbank, and S. Howard, “Range sidelobe suppression in a desired doppler interval,” in *Proc. IEEE Waveform Diversity and Design Conference*, Orlando, FL, Feb. 2010.
- [7] Y. Chi, A. Pezeshki, and R. Calderbank, “Complementary waveforms for sidelobe suppression and radar polarimetry,” in *Applications and Methods of Waveform Diversity*, V. Amuso, S. Blunt, E. Mokole, R. Schneible, and M. Wicks, Eds. SciTech Publishing, 2010.
- [8] Y. Chi, A. Gomaa, N. Al-Dhahir, and A. R. Calderbank, “Training signal design and tradeoffs for spectrally-efficient multi-user mimo-ofdm systems.” *IEEE Trans. on Wireless Communications*, vol. 10, no. 7, pp. 2234–2245, 2011.
- [9] —, in *Proc. 2011 European Signal Processing Conference*, Barcelona, Spain, 2011.
- [10] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inform. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

- [11] E. J. Candés and T. Tao, “Decoding by linear programming,” *IEEE Trans. Inform. Theory*, vol. 51, pp. 4203–4215, Dec. 2005.
- [12] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM Rev.*, vol. 43, no. 1, pp. 129–159, Jan. 2001.
- [13] J. A. Tropp, “Greed is good: algorithmic results for sparse approximation,” *IEEE Trans. Inform. Theory*, vol. 50, no. 10, pp. 2231–2242, Oct. 2004.
- [14] E. J. Candés, “The restricted isometry property and its implications for compressed sensing,” *Académie des Sciences*, vol. 1, no. 346, pp. 589–592, 2008.
- [15] E. J. Candés, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inform. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.
- [16] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [17] M. A. Davenport and M. B. Wakin, “Analysis of orthogonal matching pursuit using the restricted isometry property,” *IEEE Trans. Info. Theory*, vol. 56, no. 9, pp. 4395–4401, Sep. 2010.
- [18] T. Cai, L. Wang, and G. Xu, “Stable recovery of sparse signals and an oracle inequality,” *IEEE Trans. Info. Theory*, vol. 56, no. 7, pp. 3516–3522, Jul. 2010.
- [19] T. Cai and L. Wang, “Orthogonal matching pursuit for sparse signal recovery with noise,” *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 1–26, 2011.
- [20] L. R. Welch, “Lower bounds on the maximum cross correlation of signals,” *IEEE Trans. on Info. Theory*, vol. 20, no. 3, pp. 397–399, May 1974.
- [21] S. Kunis and H. Rauhut, “Random sampling of sparse trigonometric polynomials, ii. orthogonal matching pursuit versus basis pursuit,” *Foundations of Computational Mathematics*, vol. 8, no. 6, pp. 737–763, 2007.
- [22] W. Bajwa, R. Calderbank, and D. Mixon, “Two are better than one: Fundamental parameters of frame coherence,” *Appl. Comput. Harmon. Anal.*, 2012.
- [23] W. U. Bajwa, R. Calderbank, and S. Jafarpour, “Why gabor frames? two fundamental measures of coherence and their role in model selection,” *J. Commun. Netw.*, vol. 12, no. 4, pp. 289–307, Aug. 2010.
- [24] R. Calderbank and S. Jafarpour, “Reed muller sensing matrices and the lasso,” in *Proceedings of the 6th international conference on Sequences and their applications*, 2010, pp. 442–463.

- [25] Y. Xie, Y. Chi, L. Applebaum, and R. Calderbank, “Compressive demodulation of mutually interfering signals,” in *2012 Statistical Signal Processing Workshop*, Ann Arbor, MI, Aug. 2012.
- [26] H. Garudadri, Y. Chi, S. Baker, S. Majumdar, P. Baheti, and D. Ballard, “Diagnostic grade wireless ECG monitoring,” *Proc. IEEE Engineering in Medicine and Biology Society*, 2011.
- [27] Y. Xie, Y. Eldar, and A. Goldsmith, “Reduced-dimension multiuser detection,” *submitted to IEEE Trans. Information Theory and arXived*, Dec. 2011.
- [28] L. Applebaum, W. U. Bajwa, M. F. Duarte, and R. Calderbank, “Asynchronous code-division random access using convex optimization,” *Physical Communication*, vol. In Press, 2011.
- [29] A. K. Fletcher, S. Rangan, and V. K. Goyal, “On-off random access channels: A compressed sensing framework,” *submitted to IEEE Trans. Information Theory*, 2010.
- [30] “ANSI/AAMI EC57, Testing and reporting performance results of cardiac rhythm and ST-segment measurement algorithms,” *Health Physics*, vol. Suppl 3, pp. 1–23, 2003.
- [31] “Physionet: the research resource for complex physiologic signals.” [Online]. Available: <http://www.physionet.org/>
- [32] G. B. Moody, *Evaluating ECG Analyzers*. Cambridge, MA: Harvard-MIT Division of Health Sciences and Technology.
- [33] J. Hou, B. Chang, D.-K. Cho, and M. Gerla, “Minimizing 802.11 interference on zigbee medical sensors,” in *BodyNets 2009 conference*, April 2009.
- [34] V. Shnayder, B. Chen, K. Lorincz, T. R. F. Fulford-Hones, and M. Welsh, “Sensor networks for medical care,” *Harvard University Technical Report TR-08-05*, 2005.
- [35] R. Rubinstein, M. Zibulevsky, and M. Elad, “Efficient implementation of the K-SVD algorithm using batch orthogonal matching pursuit,” *Technion Technical Report*, 2008.
- [36] J. Morganroth and I. Gussak, *Cardiac Safety of Noncardiac Drugs: Practical Guidelines for Clinical Research and Drug Development*. Humana Press, 2010.
- [37] C. T. Mullis and L. L. Scharf, “Quadratic estimators of the power spectrum,” in *Advances in Spectrum Estimation*, S. Haykin, Ed. Prentice Hall, 1990, vol. 1, ch. 1, pp. 1–57.
- [38] L. L. Scharf, *Statistical Signal Processing*. MA: Addison-Wesley, 1991.

- [39] H. L. Van Trees, *Optimum Array Processing*. Wiley Interscience, 2002.
- [40] H. Karim and M. Viberg, “Two decades of array signal processing research: The parametric approach,” *IEEE Signal Process. Mag.*, vol. 13, no. 4, pp. 67–94, Jul. 1996.
- [41] R. Roy and T. Kailath, “ESPRIT—Estimation of signal parameters via rotational invariance techniques,” *IEEE Trans. on Acoustics, Speech, Signal Processing*, vol. 37, no. 7, pp. 984–995, Jul. 1989.
- [42] R. Roy, A. Paulraj, and T. Kailath, “ESPRIT—A subspace rotation approach to estimation of parameters of cisoids in noise,” *IEEE Trans. on Acoustics, Speech, Signal Processing*, vol. 34, no. 5, pp. 1340–1344, Oct. 1986.
- [43] J. Ward, “Maximum likelihood angle and velocity estimation with space-time adaptive processing radar,” in *Conf. Rec. 1996 Asilomar Conf. Signals, Sys., Comput.*, Pacific Grove, CA, Nov. 1996, pp. 1265–1267.
- [44] R. Klemm, *Space-Time Adaptive Processing*. IEEE Press, 1998.
- [45] D. W. Tufts and R. Kumaresan, “Singular value decomposition and improved frequency estimation using linear prediction,” *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 30, no. 4, pp. 671–675, Aug. 1982.
- [46] ———, “Estimation of frequencies of multiple sinusoids: making linear prediction perform like maximum likelihood,” *Proc. IEEE.*, vol. 70, pp. 975–989, 1982.
- [47] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 3rd ed. Baltimore, MD: John Hopkins Univ. Press, 1996.
- [48] H. L. Van Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering and Tracking*. IEEE Press, 2007.
- [49] L. T. McWhorter and L. L. Scharf, “Cramer-Rao bounds for deterministic modal analysis,” *IEEE Trans. Signal Processing*, vol. 41, no. 5, pp. 1847–1862, May 1993.
- [50] E. J. Candés, “Compressive sampling,” *Proc. Int. Congress Math.*, vol. 3, pp. 1433–1452, 2006.
- [51] R. Baraniuk, “Compressive sensing,” *IEEE Signal Process. Mag.*, vol. 24, no. 4, pp. 118–121, 2007.
- [52] J. Romberg, “Imaging via compressive sampling.” *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 14–20, Mar. 2008.
- [53] E. J. Candés and M. Wakin, “An introduction to compressive sampling,” *IEEE Signal Process. Mag.*, vol. 25, no. 2, pp. 21–30, Mar. 2008.

- [54] D. L. Donoho and M. Elad, “Optimally sparse representation in general (non-orthogonal) dictionaries via ℓ_1 minimization,” *Proc. Natl. Acad. Sci.*, vol. 100, pp. 2197–2202, 2002.
- [55] D. Donoho and J. Tanner, “Neighborliness of randomly-projected simplices in high dimensions,” *Proc. Natl. Acad. Sci.*, vol. 102, no. 27, pp. 9452–9457, 2005.
- [56] E. J. Candés and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?” *IEEE Trans. Inform. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [57] M. Elad, “Optimized projections for compressed sensing,” *IEEE Trans. Signal Proc.*, vol. 55, no. 12, pp. 5695–5702, Dec. 2007.
- [58] M. Stojnic, W. Xu, and B. Hassibi, “Compressed sensing of approximately sparse signals,” in *Proc. IEEE Int. Symp. Inform. Theory*, Toronto, CA., July 6-11 2008, pp. 2182–2186.
- [59] R. Baraniuk, M. Davenport, R. DeVore, and M. Wakin, “A simple proof of the restricted isometry property for random matrices,” *Constructive Approximation*, vol. 28, no. 3, pp. 253–263, Dec. 2008.
- [60] S. Jafarpour, W. Xu, B. Hassibi, and R. Calderbank, “Efficient and robust compressed sensing using optimized expander graphs,” *IEEE Trans. Inform. Theory*, vol. 55, no. 9, pp. 4299–4308, Sep. 2009.
- [61] R. Calderbank, S. Howard, and S. Jafarpour, “Construction of a large class of deterministic matrices that satisfy a statistical isometry property,” *IEEE J. Selected Topics in Signal Processing, Special Issue on Compressed Sensing*, vol. 4, no. 2, pp. 358–374, Apr. 2010.
- [62] R. Baraniuk and P. Steeghs, “Compressive radar imaging,” in *Proc. 2007 IEEE Radar Conf.*, Waltham, Massachusetts, Apr. 2007, pp. 128–133.
- [63] M. A. Herman and T. Strohmer, “High-resolution radar via compressed sensing,” *IEEE Trans. Signal Processing*, vol. 57, no. 6, pp. 2275–2284, Jun. 2009.
- [64] —, “Compressed sensing radar,” in *Proc. IEEE Radar Conf.*, Rome, Italy, May 2008, pp. 1–6.
- [65] —, “Compressed sensing radar,” in *Proc. IEEE Int. Conf. Acoustic, Speech, and Signal Processing (ICASSP)*, Las Vegas, NV, Apr. 2008, pp. 1509–1512.
- [66] A. Gurbuz, J. McClellan, and V. Cevher, “A compressive beamforming method,” in *Proc. Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, Las Vegas, NV, Apr. 2008, pp. 2617–2620.

- [67] V. Cevher, A. Gurbuz, J. McClellan, and R. Chellappa, “Compressive wireless arrays for bearing estimation of sparse sources in angle domain,” in *Proc. IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, Nevada, Apr. 2008, pp. 2497–2500.
- [68] A. Fannjiang, P. Yan, and T. Strohmer, “Compressed remote sensing of sparse objects,” *arXiv:0904.3999v2*, May 2009.
- [69] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, “Compressed channel sensing: A new approach to estimating sparse multipath channels,” *Proc. IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.
- [70] W. U. Bajwa, J. Haupt, G. Raz, and R. Nowak, “Compressed channel sensing,” in *Proc. 42nd Annu. Conf. Information Sciences and Systems*, Princeton, NJ, Mar. 19–21 2008, pp. 5–10.
- [71] D. Needell and R. Vershynin, “Signal recovery from inaccurate and incomplete measurements via regularized orthogonal matching pursuit,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 2, pp. 310–316, Apr. 2010.
- [72] D. Needell and J. Tropp, “CoSaMP: Iterative signal recovery from incomplete and inaccurate samples,” *Applied and Computational Harmonic Analysis*, vol. 26, pp. 301–321, 2008.
- [73] W. Allard, G. Chen, and M. Maggioni, “Multiscale Geometric Methods for Data Sets II: Geometric Wavelets,” *Appl. Comp. Harm. Anal.*, vol. 32, no. 3, pp. 435–462, May 2012.
- [74] T. Ahmed, M. Coates, and A. Lakhina, “Multivariate online anomaly detection using kernel recursive least squares,” *Proc. 26th IEEE International Conference on Computer Communications*, pp. 625–633, 2007.
- [75] S. Shahbazpanahi, S. Valaee, and M. H. Bastani, “Distributed source localization using ESPRIT algorithm,” *IEEE Transactions on Signal Processing*, vol. 49, no. 10, p. 21692178, 2001.
- [76] R. Kumaresan and D. Tufts, “Estimating the angles of arrival of multiple plane waves,” *IEEE Trans. On Aerospace And Electronic Systems*, vol. AES-19, no. 1, pp. 134–139, 1983.
- [77] A. H. Sayed, *Fundamentals of Adaptive Filtering*. Wiley, NY, 2003.
- [78] B. Yang, “Projection approximation subspace tracking,” *IEEE Transactions on Signal Processing*, vol. 43, no. 1, pp. 95–107, 1995.
- [79] K. Crammer, “Online tracking of linear subspaces,” In *Proc. COLT 2006*, vol. 4005, pp. 438–452, 2006.

- [80] E. J. Candés and T. Tao, “The power of convex relaxation: Near-optimal matrix completion,” *IEEE Trans. Inform. Theory*, vol. 56, no. 5, pp. 2053–2080, 2009.
- [81] E. J. Candes and B. Recht, “Exact matrix completion via convex optimization,” *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2008.
- [82] R. H. Keshavan, A. Montanari, and S. Oh, “Matrix completion from noisy entries,” *Journal of Machine Learning Research*, pp. 2057–2078, 2010.
- [83] L. Balzano, B. Recht, and R. Nowak, “High-dimensional matched subspace detection when data are missing,” in *Proc. ISIT*, June 2010.
- [84] L. Balzano, R. Nowak, and B. Recht, “Online identification and tracking of subspaces from highly incomplete information,” *Proc. Allerton 2010*, 2010.
- [85] W. Dai, O. Milenkovic, and E. Kerman, “Subspace Evolution and Transfer (SET) for Low-Rank Matrix Completion,” *IEEE Trans. Signal Processing*, vol. 59, no. 7, pp. 3120–3132, 2011.
- [86] B. Yang, “Asymptotic convergence analysis of the projection approximation subspace tracking algorithm,” *Signal Processing*, vol. 50, pp. 123–136, 1996.
- [87] R. Mazumder, T. Hastie, and R. Tibshirani, “Spectral regularization algorithms for learning large incomplete matrices,” *Journal of Machine Learning Research*, vol. 11, pp. 1–26, 2009.
- [88] J. Cioffi, “Limited-precision effects in adaptive filtering,” *IEEE Transactions on Circuits and Systems*, vol. 34, no. 7, pp. 821–833, 1987.
- [89] L. Bottou and O. Bousquet, “The tradeoffs of large scale learning,” in *Advances in Neural Information Processing Systems*, 2008, vol. 20, pp. 161–168.
- [90] Z. Wen, W. Yin, and Y. Zhang, “Solving a low-rank factorization model for matrix completion by a non-linear successive over-relaxation algorithm,” *Rice CAAM Tech Report TR10-07*, 2010.
- [91] S. Ma, D. Goldfarb, and L. Chen, “Fixed point and Bregman iterative methods for matrix rank minimization,” *Mathematical Programming*, vol. 1, no. 1, pp. 1–27, 2009.
- [92] J.-F. Cai, E. J. Candes, and Z. Shen, “A singular value thresholding algorithm for matrix completion,” *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1–28, 2008.
- [93] N. Levanon and E. Mozeson, *Radar Signals*. New York: Wiley, 2004.
- [94] M. J. E. Golay, “Complementary series,” *IRE Trans. Inform. Theory*, vol. 7, no. 2, pp. 82–87, April 1961.

- [95] R. J. Turyn, "Ambiguity functions of complementary sequences," *IEEE Trans. Inform. Theory*, vol. 9, pp. 46–47, Jan. 1963.
- [96] G. R. Welty, "Quaternary codes for pulsed radar," *IRE Trans. Inform. Theory*, vol. IT-6, no. 3, pp. 400–408, June 1960.
- [97] Y. Taki, M. Miyakawa, M. Hatori, and S. Namba, "Even-shift orthogonal sequences," *IEEE Trans. Inform. Theory*, vol. IT-15, no. 2, pp. 295–300, Mar. 1969.
- [98] C. C. Tseng and C. L. Liu, "Complementary sets of sequences," *IEEE Trans. Inform. Theory*, vol. IT-18, no. 5, pp. 644–652, Sept. 1972.
- [99] J. Guey and M. R. Bell, "Diversity waveform sets for delay-Doppler imaging," *IEEE Trans. Inform. Theory*, vol. 44, no. 4, pp. 1504–1522, Jul. 1998.
- [100] R. Sivaswami, "Multiphase complementary codes," *IEEE Trans. Inform. Theory*, vol. IT-24, no. 3, pp. 546–552, Sept. 1978.
- [101] —, "Self-clutter cancellation and ambiguity properties of subcomplementary sequences," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-18, no. 2, pp. 163–181, Mar. 1982.
- [102] A. Pezeshki, A. R. Calderbank, W. Moran, and S. D. Howard, "Doppler resilient Golay complementary waveforms," *IEEE Trans. Inform. Theory*, to appear, Sep. 2008.
- [103] —, "Doppler resilient Golay complementary pairs for radar," in *Proc. Stat. Signal Proc. Workshop*, Madison, WI, Aug. 2007.
- [104] S. Suvorova, S. D. Howard, W. Moran, R. Calderbank, and A. Pezeshki, "Doppler resilience, Reed-Muller codes, and complementary waveforms," in *Forty-first Asilomar Conf. Signals, Syst., Comput.*, Nov. 2007, pp. 1839–1843.
- [105] E. Prouhet, "Mèmoire sur quelques relations entre les puissances des nombres," *C. R. Acad. Sci. Paris Sèr.*, vol. I 33, p. 225, 1851.
- [106] A. Thue, "Über unendliche Zeichenreihen," *Norske vid. Selsk. Skr. Mat. Nat. Kl.*, vol. 7, pp. 1–22, 1906.
- [107] M. Morse, "Recurrent geodesics on a surface of negative curvature," *Trans. Amer. Math Soc.*, vol. 22, pp. 84–100, 1921.
- [108] J. P. Allouche and J. Shallit, "The ubiquitous Prouhet-Thue-Morse sequence," in *Sequences and their applications, Proc. SETA'98*, T. H. C. Ding and H. Niederreiter, Eds. Springer Verlag, 1999, pp. 1–16.
- [109] H. L. van Trees, *Detection, Estimation and Modulation Theory, Part III*. New York: Wiley, 1971.

- [110] M. I. Skolnik, Ed., *Radar Handbook*, 2nd ed. McGraw-Hill, 1990.
- [111] R. K. Moore, “Ground echo,” in *in Radar Handbook*, M. I. Skolnik, Ed. New York: McGraw-Hill, 2008.
- [112] R. Sullivan, “Synthetic aperture radar,” in *in Radar Handbook*, M. I. Skolnik, Ed. New York: McGraw-Hill, 2008.
- [113] S. D. Howard, A. R. Calderbank, and W. Moran, “A simple polarization diversity technique for radar detection,” in *Proc. Second Int. Conf. Waveform Diversity and Design*, HI, Jan. 22-27 2006.
- [114] —, “A simple signal processing architecture for instantaneous radar polarimetry,” *IEEE Trans. Inform. Theory*, vol. 53, pp. 1282–1289, Apr. 2007.
- [115] S. Alamouti, “A simple transmit diversity technique for wireless communications,” *IEEE J. Select. Areas Commun.*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
- [116] G. J. Foschini, “Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas,” *Bell Labs Technical Journal*, vol. 1, pp. 41–59, 1996.
- [117] E. Telatar, “Capacity of multi-antenna gaussian channels,” *European Transactions on Telecommunications*, 1999.
- [118] V. Tarokh, N. Seshadri, and A. Calderbank, “Space-time codes for high data rate wireless communication: performance criterion and code construction,” *IEEE Transactions on Information Theory*, 1998.
- [119] A. R. S. Bahai, B. R. Saltzberg, and M. Ergen, *Multi-carrier digital communications theory and applications of OFDM*. New York, NY: Springer, 2004.
- [120] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. New York, NY: Cambridge University Press, 2005.
- [121] M. Jiang and L. Hanzo, “Multiuser MIMO-OFDM for next-generation wireless systems,” *Proceedings of the IEEE*, vol. 95, no. 7, pp. 1430–1469, Jul. 2007.
- [122] “IEEE standard for local and metropolitan area networks part 16: Air interface for fixed and mobile broadband wireless access systems,” *IEEE Std 802.16e-2005*, 2006.
- [123] “Physical channels and modulation,” *3GPP TS 36.211, Ver 8.6.0*, March 2009.
- [124] S. Parkvall, E. Dahlman, A. Furuskar, Y. Jading, M. Olsson, S. Wanstedt, and K. Zangi, “LTE-advanced - evolving LTE towards IMT-advanced,” in *IEEE Vehicular Technology Conference*, 2008.

- [125] Y. G. Li, "Pilot-symbol-aided channel estimation for OFDM in wireless systems," *IEEE Trans. Veh. Technol.*, vol. 49, pp. 1207–1215, 2000.
- [126] L. Deneire, P. Vandenameele, L. van der Perre, B. Gyselinckx, and M. Engels, "A low-complexity ML channel estimator for OFDM," *IEEE Trans. Commun.*, vol. 51, pp. 67 – 75, 2003.
- [127] O. Edfors, M. Sandell, J.-J. V. de Beek, S. K. Wilson, and P. O. Böesson, "OFDM channel estimation by singular value decomposition," *IEEE Trans. Commun.*, vol. 46, pp. 931–939, 1998.
- [128] Y. G. Li, N. Seshadri, and S. Ariyavisitakul, "Channel estimation for OFDM systems with transmitter diversity in mobile wireless channels," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 461 – 471, 1999.
- [129] H. Minn, D. I. Kim, and V. Bhargava, "A reduced complexity channel estimation for OFDM systems with transmit diversity in mobile wireless channels," *IEEE Transactions on Communications*, 2002.
- [130] Z. J. Wang, Z. Han, and K. J. R. Liu, "A MIMO-OFDM channel estimation approach using time of arrivals," *IEEE Trans. On Wireless Communications*, vol. 4, pp. 1207 – 1213, May 2005.
- [131] C. Fragouli, N. Al-Dhahir, and W. Turin, "Training-based channel estimation for multiple-antenna broadband transmissions," *IEEE Transactions on Wireless Communications*, 2003.
- [132] H. Minn and N. Al-Dhahir, "Optimal training signals for MIMO OFDM channel estimation," *IEEE Trans. Wireless Comm.*, vol. 5, no. 5, pp. 1158–1168, April 2006.
- [133] Y. G. Li, "Simplified channel estimation for OFDM systems with multiple transmit antennas," *IEEE Trans. on Wireless Communications*, vol. 1, pp. 67 – 75, 2002.
- [134] Y. Zeng, A. R. Leyman, S. Ma, and T. Ng, "Optimal pilot and fast algorithm for MIMO-OFDM channel estimation," in *Proc. of IEEE International Conference on Information, Communications and Signal Processing (ICICSP)*, 2005.
- [135] F. Horlin and L. V. der Perre, "Optimal training sequences for low complexity ML multi-channel estimation in multi-user MIMO OFDM-based communications," in *Proceeding of IEEE International Conference on Communications (ICC)*, vol. 4, Jun. 2004.
- [136] S. D. Howard, A. R. Calderbank, and W. Moran, "A simple signal processing architecture for instantaneous radar polarimetry," *IEEE Trans. Inform. Theory*, vol. 53, no. 4, pp. 1282–1289, Apr. 2007.

- [137] M. D. Zoltowski, T. R. Qureshi, and R. Calderbank, “Complementary codes based channel estimation for MIMO-OFDM systems,” in *Proc. 46th Annual Allerton Conf. Communication, Control, and Computing*, Monticello, IL, Sep. 2008, pp. 133–138.
- [138] C. C. Tseng, “Signal multiplexing in surface-wave delay lines using orthogonal pairs of Golay’s complementary sequences,” *IEEE Trans. on Sonics and Ultrasonics*, 1971.
- [139] S. M. Kay, *Fundamentals Of Statistical Signal Processing: Estimation Theory*. New Jersey: Prentice Hall, 1993.
- [140] D. Chu, “Polyphase codes with good periodic correlation properties,” *IEEE Trans. on Info. Theory*, vol. IT-18, 1972.
- [141] A. Naguib, N. Seshadri, and A. Calderbank, “Increasing data rate over wireless channels,” *IEEE Signal Processing Magazine*, vol. 17, 2000.
- [142] J. A. Tropp, “On the conditioning of random subdictionaries,” *Applied and Computational Harmonic Analysis*, vol. 25, no. 1, pp. 1–24, 2008.
- [143] E. J. Candés and Y. Plan, “Near-ideal model selection by ℓ_1 minimization,” *Annals of Statistics*, vol. 37, no. 5A, pp. 2145–2177, 2009.
- [144] Z. Sidak, “Rectangular confidence regions for the means of multivariate normal distributions,” *J. of Amer. Stat. Asso.*, vol. 12, no. 4, pp. 626–633, Jun. 1967.
- [145] W. Hoeffding, “Probability inequalities for sums of bounded random variables,” *Journal of the American Statistical Association*, vol. 58, no. 301, pp. pp. 13–30, 1963.