

April 2020

COMPRESSIVE PARAMETER ESTIMATION VIA APPROXIMATE MESSAGE PASSING

Shermin Hamzehei
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/masters_theses_2



Part of the [Signal Processing Commons](#)

Recommended Citation

Hamzehei, Shermin, "COMPRESSIVE PARAMETER ESTIMATION VIA APPROXIMATE MESSAGE PASSING" (2020). *Masters Theses*. 877.
<https://doi.org/10.7275/15995042> https://scholarworks.umass.edu/masters_theses_2/877

This Open Access Thesis is brought to you for free and open access by the Dissertations and Theses at ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

**COMPRESSIVE PARAMETER ESTIMATION VIA
APPROXIMATE MESSAGE PASSING**

Presented

by

SHERMIN HAMZEHEI

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

MASTER OF SCIENCE IN ELECTRICAL AND COMPUTER ENGINEERING

February 2020

Electrical and Computer Engineering

COMPRESSIVE PARAMETER ESTIMATION VIA APPROXIMATE MESSAGE PASSING

Presented

by

SHERMIN HAMZEHEI

Approved as to style and content by:

Marco F. Duarte, Chair

Dennis Goeckel, Member

Hossein Pishro-Nik, Member

Christopher V. Hollot,
Electrical and Computer Engineering

ABSTRACT

COMPRESSIVE PARAMETER ESTIMATION VIA APPROXIMATE MESSAGE PASSING

FEBRUARY 2020

SHERMIN HAMZEHEI

B.S., SHARIF UNIVERSITY OF TECHNOLOGY

M.S.E.C.E., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Marco F. Duarte

The literature on compressive parameter estimation has been mostly focused on the use of sparsity dictionaries that encode a discretized sampling of the parameter space; these dictionaries, however, suffer from coherence issues that must be controlled for successful estimation. To bypass such issues with discretization, we propose the use of statistical parameter estimation methods within the Approximate Message Passing (AMP) algorithm for signal recovery. Our method leverages the use of custom denoisers in place of the usual thresholding steps (which act as denoisers for sparse signals) in AMP. We introduce the design of analog denoisers that are based on statistical parameter estimation algorithms, and we focus on two commonly used examples: frequency estimation and bearing estimation, coupled with the Root MUSIC estimation algorithm. We first analyze the performance of the proposed analog denoiser for signal recovery, and then link the performance in signal estimation to that of parameter estimation. Numerical experiments show significant improvements

in estimation performance versus previously proposed approaches for compressive parameter estimation.

TABLE OF CONTENTS

| | Page |
|--|------------|
| ABSTRACT | iii |
| LIST OF FIGURES | vii |
| CHAPTER | |
| 1. INTRODUCTION | 1 |
| 1.1 Contributions | 2 |
| 1.2 Outline | 3 |
| 2. OVERVIEW | 4 |
| 2.1 Notation | 4 |
| 2.2 Bases | 4 |
| 2.3 Sparsity | 4 |
| 2.4 Sensing Matrices | 5 |
| 2.5 Finding Sparse Representation | 5 |
| 2.6 Algorithmic Performance | 6 |
| 2.6.1 The Restricted Isometry Property | 7 |
| 2.7 Measurement Bounds | 9 |
| 2.8 Approximate Message Passing | 9 |
| 2.9 Denoising-Based Approximate Message Passing (D-AMP) | 10 |
| 2.10 Approximate Onsager Correction via Monte Carlo Method | 11 |
| 2.11 Parametric Signal Models and Parameter Estimation | 12 |
| 2.11.1 Frequency Estimation | 12 |
| 2.11.2 Bearing Estimation | 13 |
| 2.12 Statistical Parametric Estimation Algorithms | 14 |

| | |
|--|-----------|
| 3. PROPOSED RESEARCH | 16 |
| 3.1 Analog Denoisers and Their Analysis | 16 |
| 3.1.1 Signal Generation | 17 |
| 3.1.2 Denoiser Properties | 18 |
| 3.1.2.1 Denoiser Properness | 18 |
| 3.1.2.2 Denoiser Monotonicity | 20 |
| 3.1.3 State evolution | 22 |
| 3.1.4 D-AMP Performance Analysis | 26 |
| 3.1.4.1 Noiseless Measurements | 27 |
| 3.1.4.2 Noisy Measurements | 29 |
| 3.2 Analysis of parameter estimation performance | 32 |
| 3.3 Numerical Evaluation of AMP+MUSIC | 38 |
| 3.3.1 Signal recovery | 38 |
| 3.3.1.1 Frequency-sparse Signals | 38 |
| 3.3.1.2 Bearing-sparse Signals | 40 |
| 3.3.2 Parameter Estimation | 40 |
| 3.3.2.1 Frequency Estimation | 40 |
| 3.3.2.2 Bearing Estimation | 42 |
| 4. CONCLUSION AND FUTURE WORK | 44 |
| BIBLIOGRAPHY | 46 |

LIST OF FIGURES

| Figure | Page |
|--------|---|
| 3.1 | Dependence of the denoiser level κ on signal generation parameters for frequency-sparse signals. (a) Signal length N . (b) Sparsity level K . (c) Window size W . (d) Minimum frequency separation S 19 |
| 3.2 | Dependence of the denoiser level κ on signal generation parameters for bearing-sparse signals. (a) Sparsity level (number of transmitters) K . (b) Number of snapshots Q 20 |
| 3.3 | Dependence of the denoiser monotonicity on signal generation parameters for frequency-sparse signals. (a) Signal length N . (b) Sparsity level K . (c) Window size W . (d) Minimum frequency separation S 22 |
| 3.4 | Dependence of the denoiser monotonicity on signal generation parameters for bearing-sparse signals. (a) Sparsity level (number of transmitters) K . (b) Number of snapshots Q 23 |
| 3.5 | Evaluation of iteration-wise AMP+MUSIC recovery MSE prediction by the state evolution sequence with different measurement SNRs and undersampling ratios δ for frequency-sparse signals. (a) Noiseless measurements. (b) SNR = 30dB. (c) SNR = 20 dB. (d) $\delta = 0.8$ 25 |
| 3.6 | Evaluation of iteration-wise AMP+MUSIC recovery MSE prediction by the state evolution sequence with different measurement SNRs and undersampling ratios δ for bearing-sparse signals. (a) Noiseless measurements. (b) SNR = 30dB. (c) SNR = 20 dB. (d) $\delta = 0.8$ 26 |
| 3.7 | Convergence of state evolution sequence in frequency-sparse signals. (a) SNR = 30dB. (b) SNR = 20dB. 27 |
| 3.8 | Convergence of state evolution sequence in bearing-sparse signals. (a) Noiseless measurements. (b) SNR = 20dB. 27 |
| 3.9 | Effect of different signal generation and denoiser parameters on the terms of Proposition 3.1.1 for frequency-sparse signals. (a) Signal length N . (b) Sparsity level K . (c) Window size W . (d) Minimum frequency separation S 30 |

| | | |
|------|---|----|
| 3.10 | <i>Effect of different signal generation parameters on Proposition 3.1.1 for bearing-sparse signals. (a) Sparsity level K. (b) Number of snapshots Q.</i> | 31 |
| 3.11 | <i>Effect of different signal generation, denoiser, and CS parameters on the terms of Proposition 3.1.2 for frequency-sparse signals. (a) Signal length N. (b) Sparsity level K. (c) Window size W. (d) Minimum frequency separation S. (e) Undersampling ratio δ.</i> | 33 |
| 3.12 | <i>Effect of different signal generation parameters on the terms of Proposition 3.1.2 for bearing-sparse signals. (a) Sparsity level K. (b) Number of snapshots Q.</i> | 34 |
| 3.13 | <i>(a) Performance of the analog denoiser for bearing-sparse signals as a function of the measurement noise variance. (b) Performance of Root MUSIC frequency estimation as a function of the measurement noise variance. (c) Performance of Root MUSIC bearing estimation as a function of the analog denoiser error for bearing-sparse signals. Different colors correspond to different signals in the experiment.</i> | 35 |
| 3.14 | <i>Histogram for the absolute difference between the empirical average bearing estimation error from Root MUSIC and the error predicted by (3.11) from [29] for two sets of numbers of transmitters K. (a) $K = 10$, $K = 20$, and $K = 30$. (b) $K = 10$.</i> | 37 |
| 3.15 | <i>Performance of AMP+MUSIC for frequency-sparse signal recovery. (a) Average normalized mean squared error. (b) Phase transition plot.</i> | 39 |
| 3.16 | <i>Performance of AMP+MUSIC for bearing-sparse signal recovery. (a) Average normalized mean squared error. (b) Phase transition plot.</i> | 41 |
| 3.17 | <i>Numerical comparison of algorithms for compressive frequency estimation. (a) Average frequency estimation error as a function of the number of measurements M. (b) Phase transition plot for frequency estimation.</i> | 42 |
| 3.18 | <i>Numerical comparison of algorithms for compressive bearing estimation. (a) Average bearing estimation error as a function of the number of measurements for different sparsity levels. (b) Phase transition plot for bearing estimation.</i> | 43 |

CHAPTER 1

INTRODUCTION

Compressive sensing (CS) allows the recovery of sparse signals measured at a rate lower than Nyquist [2, 4, 9]. Most existing work in the field of CS focuses on recovering the signal from a few measurements, with more recent extensions to parameter estimation using sparse models. In such models, the signals are recovered by exploiting the fact that they can be described as sparse in some basis or frame. Some examples of such sparse parameter estimation models include frequency estimation, localization, and bearing estimation [5, 13–21, 27, 30, 32, 34, 35].

Multiple methods have been proposed for signal recovery from CS measurements, with iterative methods becoming particularly popular. The Approximate Message Passing (AMP) algorithm [12], which is a variation of the iterative thresholding algorithm, has emerged as a flexible framework for recovery that can leverage models beyond standard sparsity. This iterative algorithm produces an additive white Gaussian noise (AWGN)-polluted version of the original sparse signal at each iteration by integrating a so-called Onsager correction term, enabling the simple application of denoising algorithms (e.g., thresholders for sparse signals). Thus, AMP is flexible enough to expand beyond the standard sparsity model by leveraging denoisers suitable for the class of signals of interest [10, 25].

Many algorithms proposed for compressive parameter estimation leverage a prescribed grid for parameters specified by the sparsity basis; therefore, the continuous parameter domain is replaced by a fixed parameter sample grid [33]. The grid selection has a high importance for such *on-grid* methods, and they often suffer from

poor estimation accuracy when the parameters do not belong to the grid. In *off-grid* methods, refinements to the sparse model output allow for arbitrary parameter estimates, although a grid is still necessary to perform estimation. Such methods most often leverage nonconvex interpolation and optimization, and thus can only provide local convergence guarantees. *Gridless* methods avoid these issues by directly considering the continuous parameter space, and often rely on statistical signal models that provide closed-form solutions for the parameter estimates. Such methods usually have limited applicability to settings featuring measurements from uniformly timed samples, or uniform or sparse linear arrays.

In this research, we propose the integration of statistical gridless methods for parameter estimation with synthesis models based on the signal parametrization to provide *analog denoisers*, which can then be integrated into the AMP algorithm to obtain a framework for compressive parameter estimation. Our approach is enabled by the aforementioned flexibility of AMP, as well as subsequently proposed approach for the numerical computation of the Onsager correction using a Monte-Carlo approximation [25]. One of the significant advantages of our proposed algorithm is that it can be applied to arbitrary parameter estimation problems, as long as there exists a known one-to-one mapping between the observed signal and the parameters of interest and certain technical conditions hold for the analog denoiser operating on the signal of interest.

1.1 Contributions

The contributions of this thesis can be summarized as follows. First, we introduce the design of analog denoisers that leverage parametric models to be integrated with AMP for the purpose of compressive parameter estimation. More specifically, the analog denoiser is used within the AMP algorithm to “denoise” an AWGN-polluted version of the original signal, which simultaneously provides estimates for the pa-

rameters of interest. Second, we study the necessary conditions for a denoiser to be compatible with the AMP algorithm for signal recovery. More precisely, we investigate these essential conditions provided in [25] while focusing on signal recovery rather than parameter estimation for two examples of parametric signal models: frequency-sparse and bearing-sparse signals. Third, we link parameter estimation to signal recovery, study the accuracy of the estimation and recovery in both examples, and provide a numerical analysis for the expected signal recovery and parameter estimation error. Finally, we present numerical results on the performance of both signal recovery and parameter estimation based on our combination of AMP and analog denoisers.

1.2 Outline

This thesis is organized as follows.

Chapter 2 introduces notation and provides essential background, including sparsity, sensing matrices, algorithmic performance, approximate message passing, and statistical parametric estimation algorithms.

Chapter 3 contains the proposed work. In this chapter we first introduce analog denoisers with two running examples for frequency estimation and bearing estimation. We also present numerical simulations that verify the crucial properties of the denoiser for signal recovery in both cases of frequency-sparse and bearing-sparse signals. We then discuss the relationship between signal estimation and parameter estimation. Finally, we provide numerical results for the proposed approach for both signal recovery and parameter estimation from CS measurements, including a comparison to existing algorithms for compressive parameter estimation.

Finally, **Chapter 4** includes the summary of the thesis and suggests some future work.

CHAPTER 2

OVERVIEW

2.1 Notation

We will adopt the convention throughout this thesis that all vectors are written in lower-case bold type and are column vectors, and all matrices are written in upper-case bold type. Moreover, for a vector \mathbf{x} , x_i refers to the i th element of this vector and $x_i(\beta)$ implies that this element is a function of parameter β , while for recursive algorithms, \mathbf{x}^i refers to the vector \mathbf{x} in iteration i .

2.2 Bases

A set $\{\phi_i\}_{i=1}^n$ is called a basis for \mathbb{R}^n if the vectors in the set span \mathbb{R}^n and are linearly independent. This implies that each vector in the space can be uniquely represented as a linear combination of these basis vectors. In other words, for each $\mathbf{x} \in \mathbb{R}^n$, there exist unique coefficients $\{c_i\}_{i=1}^n$ such that

$$\mathbf{x} = \sum_{i=1}^n c_i \phi_i.$$

2.3 Sparsity

Mathematically, we say that a signal \mathbf{x} is K -sparse if it has at most K nonzero elements, i.e., $\|\mathbf{x}\|_0 \leq K$. We also let Σ_K indicate the set of all K -sparse signals

$$\Sigma_K = \{\mathbf{x} : \|\mathbf{x}\|_0 \leq K\}.$$

In many cases, the signals are not sparse themselves, but have a sparse representation in some basis Φ . In such cases, $\mathbf{x} = \Phi\mathbf{c}$, where \mathbf{c} is K -sparse, i.e., $\|\mathbf{c}\|_0 \leq K$.

Compressible signals are the signals that can be represented accurately using $K \ll N$ coefficients. If we define

$$\sigma_K(\mathbf{x})_p = \min_{\hat{\mathbf{x}} \in \Sigma_K} \|\mathbf{x} - \hat{\mathbf{x}}\|_p,$$

for compressible signals, there exist constants C and r such that

$$\sigma_K(\mathbf{x})_2 \leq CK^{-r}.$$

2.4 Sensing Matrices

Let us assume that the signal of interest \mathbf{x} is real valued with finite length N . For this signal, we consider M linear measurements

$$\mathbf{y} = \mathbf{A}\mathbf{x}.$$

Here, $\mathbf{y} \in \mathbb{R}^M$ and \mathbf{A} , called the *sensing matrix*, is an $M \times N$ matrix, mapping \mathbb{R}^N to \mathbb{R}^M , where M is typically much smaller than N . Note that in standard CS the rows of \mathbf{A} are fixed.

For sparse or compressible data, it is possible to design the matrices \mathbf{A} with $M \ll N$ that ensure the accurate recovery of the original signal using different algorithms.

2.5 Finding Sparse Representation

In compressive sensing (CS), the goal is to recover the signal of interest from a few measurements. More precisely, we consider a sparse signal $\mathbf{x} \in \mathbb{R}^N$ and a measurement vector $\mathbf{y} \in \mathbb{R}^M$, where $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$. Here \mathbf{n} represents the measurement

noise. If $M \ll N$, the matrix $\mathbf{A} \in \mathbb{R}^{M \times N}$ is severely underdetermined. Therefore, in order to recover \mathbf{x} , we search for the sparsest signal $\hat{\mathbf{x}}$ such that it best satisfies the measurements,

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{x}\|_0, \text{ s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}. \quad (2.1)$$

Note that since ℓ_0 -norm is defined as the number of nonzero elements in a vector, this method is called *ℓ_0 -norm minimization*. Since (2.1) is non-convex, Chen et al. [6] suggested the use of ℓ_1 -norm as a convex relaxation of (2.1);

$$\hat{\mathbf{x}} = \operatorname{argmin}_{\mathbf{x} \in \mathbb{R}^N} \|\mathbf{x}\|_1, \text{ s.t. } \mathbf{y} = \mathbf{A}\mathbf{x}. \quad (2.2)$$

Since this ℓ_1 -minimization equation, known as *basis pursuit (BP)*, is convex, this algorithm can be implemented as a linear program, leading to polynomial computational complexity in the signal length.

Note that when dealing with large signals, the linear programs, such as (2.2), are extremely computationally complex; therefore, iterative algorithms are proposed to solve for the sparse signal \mathbf{x} . Some examples of such iterative algorithms are matching pursuit, orthogonal matching pursuit, iterative hard-thresholding, compressive sampling matching pursuit, iterative soft-thresholding, and approximate message passing (which will be studied in Section 2.8).

2.6 Algorithmic Performance

In this section, we provide some definitions and theories which are important in studying the guarantees for the performance of the recovery algorithms.

Definition 2.6.1. *The coherence of a matrix \mathbf{A} , $\mu(\mathbf{A})$, is the largest absolute inner product between any two column a_i , a_j of \mathbf{A} :*

$$\mu(\mathbf{A}) = \max_{1 \leq i < j \leq N} \frac{|\langle a_i, a_j \rangle|}{\|a_i\|_2 \|a_j\|_2} \quad (2.3)$$

It can be shown that for the coherence of a matrix the inequality $\mu(\mathbf{A}) \in \left[\sqrt{\frac{N-M}{M(N-1)}}, 1 \right]$ always holds; the lower bound is the Welch bound [citations, p.26](#) . Also, for $M \ll N$, we approximately have $\mu(\mathbf{A}) \geq \frac{1}{\sqrt{M}}$.

Theorem 2.6.1. *If*

$$K < \frac{1}{2} \left(1 + \frac{1}{\mu(\mathbf{A})} \right), \quad (2.4)$$

then for each measurement vector $\mathbf{y} \in \mathbb{R}^M$ there exists at most one K -sparse signal \mathbf{x} such that $\mathbf{y} = \mathbf{A}\mathbf{x}$.

Definition 2.6.2. *The spark of a given matrix \mathbf{A} is the smallest number of columns of \mathbf{A} that are linearly dependent.*

Theorem 2.6.2. *For any vector $\mathbf{y} \in \mathbb{R}^M$, there exists at most one K -sparse signal \mathbf{x} , such that $\mathbf{y} = \mathbf{A}\mathbf{x}$ if and only if $\text{spark}(\mathbf{A}) > 2K$.*

2.6.1 The Restricted Isometry Property

In order to establish guarantees for recovery, even when the signal is contaminated by noise or other corruptions, Candès [4] introduced the following restricted isometry condition on matrices \mathbf{A} .

Definition 2.6.3. A matrix \mathbf{A} satisfies the restricted isometry property (RIP) of order K if there exists a $\delta_K \in (0, 1)$ such that

$$(1 - \delta_K)\|\mathbf{x}\|_2^2 \leq \|\mathbf{Ax}\|_2^2 \leq (1 + \delta_K)\|\mathbf{x}\|_2^2 \quad (2.5)$$

for all K -sparse signals \mathbf{x} .

Note that if \mathbf{A} satisfies the RIP of order K with constant δ_K , then for any $K' < K$, \mathbf{A} satisfies the RIP of order K' with constant $\delta_{K'} \leq \delta_K$.

It can be shown that if a matrix \mathbf{A} satisfies the RIP, then this is sufficient for a variety of algorithms to be able to successfully recover a sparse signal from noisy measurements.

Definition 2.6.4. Let $\mathbf{A} : \mathbb{R}^N \rightarrow \mathbb{R}^M$ denote a sensing matrix and $\Delta : \mathbb{R}^M \rightarrow \mathbb{R}^N$ be a recovery algorithm. We say that the pair (\mathbf{A}, Δ) is C -stable if for any K -sparse signal \mathbf{x} and any $\mathbf{e} \in \mathbb{R}^M$, we have

$$\|\Delta(\mathbf{Ax} + \mathbf{e} - \mathbf{x})\|_2 \leq C\|\mathbf{e}\|_2. \quad (2.6)$$

In words, this definition implies that adding a small amount of noise to the measurements cannot lead to an arbitrarily large impact on the recovered signal.

Theorem 2.6.3. If the pair (\mathbf{A}, Δ) is C -stable, then

$$\frac{1}{C}\|\mathbf{x}\|_2 \leq \|\mathbf{Ax}\|_2 \quad (2.7)$$

for all K -sparse signals \mathbf{x} .

2.7 Measurement Bounds

We now consider how many measurements are necessary to achieve the RIP. If we only focus on the dimensions of the problems, M , N , and K , we can establish a simple lower bound for the number of measurements M .

Theorem 2.7.1. *If $\mathbf{A} \in \mathbb{R}^{M \times N}$ satisfies the RIP of order $2K$ with constant $\delta_K \in (0, \frac{1}{2}]$, then*

$$M \geq CK \log\left(\frac{N}{K}\right)$$

where $C = 1/2 \log(\sqrt{24} + 1) \approx 0.28$.

2.8 Approximate Message Passing

While many algorithms have previously been proposed to recover the signal \mathbf{x} given \mathbf{y} and \mathbf{A} , greedy algorithms have been shown to have high performance and to be fast [3, 7, 26].

As a matter of fact, we search K -sparse vectors \mathbf{x} that satisfy $\mathbf{y} \approx \mathbf{A}\mathbf{x}$. One recovery method is to solve the convex problem. The

Donoho et al. proposed a modification of the traditional iterative thresholding algorithm that adds an ‘‘Onsager’’ correction term [12]. The resulting first-order approximate message passing algorithm (AMP) proceeds as follows: starting from $\mathbf{x}^0 = \mathbf{0}$ and $\mathbf{z}^0 = \mathbf{0}$, at iteration t we set

$$\mathbf{x}^t = \eta_t(\mathbf{A}^H \mathbf{z}^{t-1} + \mathbf{x}^{t-1}), \quad (2.8)$$

$$\mathbf{z}^t = \mathbf{y} - \mathbf{A}\mathbf{x}^t + \frac{1}{\delta} \mathbf{z}^{t-1} \langle \eta'_t(\mathbf{A}\mathbf{z}^{t-1} + \mathbf{x}^{t-1}) \rangle, \quad (2.9)$$

where $\eta'_t(\mathbf{s}) = \frac{\partial}{\partial \mathbf{s}} \eta_t(\mathbf{s})$ is the gradient of the thresholding function $\eta_t(\mathbf{s})$, $\delta = \frac{M}{N}$ is the normalized measurement ratio, and for a length- N vector \mathbf{u} we denote $\langle \mathbf{u} \rangle = \frac{1}{N} \mathbf{1}^T \mathbf{u}$;

the term $\langle \eta'_t(\cdot) \rangle$ is often referred to as the *divergence* of the thresholding function. The Onsager correction term at the end of (2.9) plays a crucial role throughout the AMP algorithm, causing it to achieve a sparsity-undersampling tradeoff matching the theoretical tradeoff for reconstruction via ℓ_1 -norm minimization [12].

2.9 Denoising-Based Approximate Message Passing (D-AMP)

The power of the Onsager correction is that it shapes the distribution of the estimate $\mathbf{A}^H \mathbf{z}^t + \mathbf{x}^t$ to resemble that of $\mathbf{x} + \mathbf{n}_t$, where \mathbf{n}_t is AWGN. AMP then solves the problem of sparse signal denoising using the function η_t , which can be replaced by an arbitrary denoising function D_σ suitable for the signal class of interest [10]. Note that hard and soft thresholders are optimal AWGN denoisers for sparse and approximately sparse signals [8], while there exist multiple denoisers that allow for the AMP recovery framework to be applied beyond the standard class of sparse signals [10, 25].¹

For example, denoising-based AMP (D-AMP) [25] assumes that \mathbf{x} belongs to a class of signals $\mathcal{C} \subseteq \mathbb{R}^N$ for which a family of denoisers $\{D_{\sigma,\tau}\}$ exists. Here, σ is the standard deviation of the added noise to the signal that the denoiser is tuned to and τ denotes additional parameters for the denoiser, if any, and is dropped from the notation when possible. Each denoiser $D_{\sigma,\tau}$ can be applied to $\mathbf{x}_w = \mathbf{x} + \sigma\epsilon$ with $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ to return an estimate $\hat{\mathbf{x}}$ that should be closer to \mathbf{x} than \mathbf{x}_w . Even though the denoiser must satisfy several conditions (detailed in the sequel), these conditions do not change between signal classes, making the derivations in the sequel applicable to a wide variety of signal classes and a broad set of denoisers.

For high-dimensional problems, D-AMP modifies the AMP iterations (2.8-2.9) by replacing the thresholding step with a denoising step, making it applicable to signal classes beyond sparsity; we set $\hat{\sigma}^0 = \|\mathbf{y}\|_2 / \sqrt{N}$ and iterate:

¹Although this idea was originally introduced in [10], [25] considers a finite-length setting and a more tractable class of signals; therefore, we focus on the latter formulation.

$$\mathbf{x}^t = D_{\hat{\sigma}^{t-1}}(\mathbf{A}^H \mathbf{z}^{t-1} + \mathbf{x}^{t-1}), \quad (2.10)$$

$$\mathbf{z}^t = \mathbf{y} - \mathbf{A}\mathbf{x}^t + \frac{1}{M} \mathbf{z}^{t-1} \langle D'_{\hat{\sigma}^{t-1}}(\mathbf{x}^{t-1} + \mathbf{A}^H \mathbf{z}^{t-1}) \rangle, \quad (2.11)$$

$$(\hat{\sigma}^t)^2 = \frac{\|\mathbf{z}^t\|_2^2}{M}. \quad (2.12)$$

As with AMP (2.8-2.9), at each iteration of D-AMP (2.10-2.12) one can write $\mathbf{x}^t + \mathbf{A}^H \mathbf{z}^t \approx \mathbf{x} + \mathbf{n}_t$ with $\mathbf{n}_t \sim \mathcal{N}(\mathbf{0}, (\hat{\sigma}^t)^2 \mathbf{I})$, where $(\hat{\sigma}^t)^2$ approximates the AWGN variance at iteration t .

2.10 Approximate Onsager Correction via Monte Carlo Method

High-performance denoisers are usually data-dependent, and thus it might not be possible to explicitly indicate the input-output relationship for such denoisers. As a result, the computation of the divergence $\langle D'_{\hat{\sigma}^{t-1}}(\mathbf{x}^{t-1} + \mathbf{A}^H \mathbf{z}^{t-1}) \rangle$ in the Onsager correction might not be straightforward. Luckily, Metzler et al. have proposed an approximation of this term, making it feasible to use arbitrary denoisers in D-AMP [25]. Their proposed method relies on a Monte Carlo estimate of the denoiser divergence: given an arbitrary denoiser $D_\sigma(\mathbf{x})$, we can estimate its divergence as

$$\langle D'_{\sigma,\tau}(\mathbf{x}) \rangle \approx \frac{1}{B} \sum_{i=1}^B (\mathbf{b}^i)^* \left(\frac{D_{\sigma,\tau}(\mathbf{x} + \zeta \mathbf{b}^i) - D_{\sigma,\tau}(\mathbf{x})}{\zeta} \right),$$

where $\mathbf{b}^i \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ are i.i.d. random vectors and ζ is chosen to be a sufficiently small constant in the last term. Note that according to the weak law of large numbers, as $B \rightarrow \infty$ this estimation converges to the true value of divergence. In practice, since the signal lengths are generally large, a small number of trials B suffices for accurate approximation [25].

2.11 Parametric Signal Models and Parameter Estimation

Rather than studying signals that are sparse in a basis or a dictionary, we focus on low-dimensional parametric signal models in this research. In such models, the signal of interest \mathbf{x} corresponds to a linear combination of components generated from a parametric form, where the number of components (i.e., the number of degrees of freedom for the signal) is assumed to be significantly smaller than the signal length.

More specifically, we assume that the signal of interest has K parameters $\{\theta_k\}_{k=1}^K \in \mathbb{R}$ and can be represented as

$$\mathbf{x} = \sum_{k=1}^K a_k \Psi(\theta_k). \quad (2.13)$$

Here, Ψ is the mapping function from parameter space to signal space, and $\{a_k\}_{k=1}^K \in \mathbb{C}$ are component amplitudes. We refer to (2.13) as the *synthesis* equation. The goal of parameter estimation algorithms is to obtain estimates $\{\hat{\theta}_k\}_{k=1}^K$ from observations of \mathbf{x} .

In this research, we focus on two specific examples of parametric signal models and parameter estimation.

2.11.1 Frequency Estimation

We concentrate on parametric signals that are linear combination of a few complex sinusoids. Such *frequency-sparse signals* have a number of component frequencies K that is much smaller than their length N . Assuming $\{f_k\}_{k=1}^K$ and $\{a_k\}_{k=1}^K$ to be the sinusoid frequencies and their corresponding amplitudes, respectively, such signals can be represented using (2.13) with $\Psi(f_k) = \mathbf{v}(f_k)$, defined as

$$\mathbf{v}(f) = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & e^{\frac{j2\pi f}{N}} & e^{\frac{j2\pi 2f}{N}} & \dots & e^{\frac{j2\pi(N-1)f}{N}} \end{bmatrix}^T. \quad (2.14)$$

The frequency estimation problem is well known in the signal processing literature as the line spectral estimation problem. Unfortunately, standard on-grid approaches for recovering these parameters suffer from frequency mismatch. For example, the DFT² basis preserves the sparsity of the discrete-time Fourier transform for frequency-sparse signals only when the signal components have *integer frequencies*, i.e., containing an integer number of signal periods. Otherwise, the situation is decidedly more complicated due to the spectral leakage induced by windowing (sinc convolution) [14]. While many popular estimation algorithms exist to overcome this problem [31], our focus in this research will be on the Root MUSIC gridless method, cf. Section 2.12.

2.11.2 Bearing Estimation

Assume that an array of P sensors (usually microphones or antennas) can record one or multiple targets transmitting a signal to the array at specific bearing angles [23]. Moreover, assume that the p^{th} antenna is located at the coordinates (u_p, v_p) and that the antennas are configured as a uniform linear array (i.e., $u_p = u_0 + pd_x$, where d_x is the array inter-element spacing, and $v_p = 0$ for all p). The $P \times 1$ array snapshot vector $\mathbf{x}(q) = [x_1(q) \ x_2(q) \ \dots \ x_P(q)]^T$, containing observations from all antennas at each time $q = 1, \dots, Q$, can be modeled as $\mathbf{x}(q) = \mathbf{S}(\theta)\mathbf{a}(q) + \mathbf{n}(q)$ (dubbed a snapshot), where $\theta = [\theta_1 \ \dots \ \theta_K]^T$ is the $K \times 1$ vector of the bearings, $\mathbf{S}(\theta) = [\mathbf{s}(\theta_1) \ \dots \ \mathbf{s}(\theta_K)]$ is the $P \times K$ signal steering matrix, $\mathbf{a}(q) = [a_1(q) \ \dots \ a_K(q)]^T$ is the $K \times 1$ vector collecting the scalar amplitudes of the received transmissions at time q , and $\mathbf{n}(q)$ is the $P \times 1$ vector of the antenna array AWGN at time q . Each $P \times 1$ steering vector can be expressed as

²Recall that the Discrete Fourier Transform (DFT) of a length- N signal \mathbf{x} is defined as $\mathbf{X}[l] = \sum_{n=1}^N \mathbf{x}[n]e^{-j2\pi ln/N}$, $1 \leq l \leq N$, with inverse transformation $\mathbf{x}[n] = \frac{1}{N} \sum_{l=1}^N \mathbf{X}[l]e^{j2\pi ln/N}$, $1 \leq n \leq N$.

$$\mathbf{s}(\theta) = \begin{bmatrix} \exp(-j\frac{P-1}{2}\frac{2\pi}{\lambda}d_x \sin \theta) \\ \exp(-j\frac{P-3}{2}\frac{2\pi}{\lambda}d_x \sin \theta) \\ \vdots \\ \exp(j\frac{P-1}{2}\frac{2\pi}{\lambda}d_x \sin \theta) \end{bmatrix} = \begin{bmatrix} z^{-(P-1)/2} \\ z^{-(P-3)/2} \\ \vdots \\ z^{(P-1)/2} \end{bmatrix}, \quad (2.15)$$

where $z = \exp(j(2\pi/\lambda)d_x \sin \theta)$ and λ is the signal wavelength. We collect the multiple snapshot observations into the matrix equation $\mathbf{X} = \mathbf{S}(\theta)\mathbf{A} + \mathbf{N}$, with $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_Q]$, $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_Q]$, and $\mathbf{N} = [\mathbf{n}_1 \dots \mathbf{n}_Q]$.³

It is clear from (2.15) that the steering vectors $\mathbf{s}(\theta_k)$ correspond to uniformly sampled complex exponentials with frequencies $f_k = \frac{d_x}{\lambda} \sin \theta_k$. Thus, the angles $\{\theta_k\}_{k=1}^K$ can be obtained by identifying the frequencies for the complex exponential components of the received (noisy) snapshots $\mathbf{x}(q)$ and then solving for θ_k from f_k . As a result, bearing estimation is similar to frequency estimation, and the same statistical parameter estimation algorithms can be applied to both problems.

2.12 Statistical Parametric Estimation Algorithms

For the K -sparse signal \mathbf{x} from (2.13), a statistical parameter estimation method considers a set of observations $\mathbf{s} = \mathbf{x} + \mathbf{n}$, where $\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \sigma_n^2 \mathbf{I})$ denotes an AWGN vector. One can rewrite (2.13) as $\mathbf{x} = \mathbf{\Gamma}\mathbf{a}$, where $\mathbf{\Gamma}$ is a matrix with columns $\mathbf{\Gamma} = \begin{bmatrix} \mathbf{v}(f_1) \dots \mathbf{v}(f_K) \end{bmatrix}$, and the coefficient vector $\mathbf{a} = \begin{bmatrix} a_1 \dots a_K \end{bmatrix}^T$. The autocorrelation matrix \mathbf{R}_{ss} is given by

$$\mathbf{R}_{ss} = \mathbb{E}[\mathbf{s}\mathbf{s}^H] = \mathbf{R}_{xx} + \mathbf{R}_{nn} = \mathbf{\Gamma}\bar{\mathbf{A}}^2\mathbf{\Gamma}^H + \sigma_n^2\mathbf{I} \quad (2.16)$$

³Note that our focus on Gaussian CS matrices Φ in this research requires the measurement of a vectorized version of \mathbf{X} , which in practice requires all antennas to communicate their measurements to a central location for aggregation. The algorithms proposed in this research have also been tailored for distributed compressive bearing estimation, with successful results [22].

where $\bar{\mathbf{A}} = \text{diag}(\mathbb{E}(a_1), \dots, \mathbb{E}(a_K))$ denotes the matrix diagonalization of the expected value of the vector a . Since for distinct frequencies $\{f_k\}_{k=1}^K$, $\text{rank}(\mathbf{\Gamma}\bar{\mathbf{A}}^2\mathbf{\Gamma}^H) = K$, it is easy to see that \mathbf{R}_{xx} has K nonzero eigenvalues $\{\tilde{\lambda}_i\}_{i=1}^K$ (sorted by magnitude), with all other eigenvalues equal to zero. Thus, for the sorted eigenvalues $\{\lambda_i\}_{i=1}^N$ of \mathbf{R}_{ss} , we have

$$\lambda_i = \begin{cases} \tilde{\lambda}_i + \sigma_n^2, & i \leq K, \\ \sigma_n^2, & K < i \leq N. \end{cases}$$

Defining \mathbf{G} as the matrix containing the column eigenvectors for the $N - K$ smallest eigenvalues of \mathbf{R}_{ss} , we have that $\mathbf{R}_{ss}\mathbf{G} = \sigma_n^2\mathbf{G}$, as the corresponding eigenvalues are all σ_n^2 ; thus, plugging in (2.16), we have that $\mathbf{\Gamma}\bar{\mathbf{A}}^2\mathbf{\Gamma}^H\mathbf{G} + \sigma_n^2\mathbf{G} = \sigma_n^2\mathbf{G}$, which in turn implies that $\mathbf{\Gamma}^H\mathbf{G} = 0$. Thus, the frequencies $\{f_k\}_{k=1}^K$ are the only solutions to $\mathbf{v}(f)^H\mathbf{G}\mathbf{G}^H\mathbf{v}(f) = 0$.

In order to determine the component frequencies, the Root MUSIC algorithm [1] searches for the roots of polynomial $\mathbf{p}^H(z)\mathbf{G}\mathbf{G}^H\mathbf{p}(z)$ for $z \in \mathbb{C}$, $|z| = 1$, where $\mathbf{p}(z) = \begin{bmatrix} 1 & z & z^2 & \dots & z^{N-1} \end{bmatrix}^T$. The frequencies can then be established through the relationship $f = (\angle z)/(2\pi)$.

In practice, Root MUSIC uses the sampled windowed autocorrelation matrix $\hat{\mathbf{R}}_{xx} = \frac{1}{L} \sum_{i=1}^L \hat{\mathbf{x}}_i \hat{\mathbf{x}}_i^T$, where $\hat{\mathbf{x}}_i = \begin{bmatrix} x_i & \dots & x_{i+W-1} \end{bmatrix}^T$ denotes the i^{th} window from \mathbf{x} of length W (an algorithm parameter), with each window providing a realization of the parametric signal, and $L = N - W$ denotes the number of windows present in x . Note that the window size should follow $W \in [K + 1, N - (K + 1)]$. In the bearing estimation case, one can consider the multiple snapshots as individual signal realizations instead of using windowed versions of the signals.

CHAPTER 3

PROPOSED RESEARCH

3.1 Analog Denoisers and Their Analysis

While most existing extensions to AMP have focused on sparse signal models, D-AMP extends its applicability to any signal class for which suitable denoisers exist. Thus, we create *analog denoisers* $\hat{\mathbf{x}} = D_{\sigma, \beta}(\mathbf{s})$ for parametric signal classes by merging the statistical estimation and synthesis steps:

$$\{\hat{\theta}_k, \hat{a}_k\}_{k=1}^K = \Omega(\mathbf{s}, \beta), \quad (3.1)$$

$$\hat{\mathbf{x}} = \sum_{k=1}^K a_k \Psi(\theta_k), \quad (3.2)$$

where $\Omega(\mathbf{x}, \beta)$ refers to a statistical parameter estimator and β are its parameters. In terms of the parameter estimation examples mentioned before, the analog denoiser for frequency estimation can be written as

$$\{\hat{f}_k, \hat{a}_k\}_{k=1}^K = \text{MUSIC}(\mathbf{s}, \beta)$$

$$\hat{\mathbf{x}} = \Gamma(\{\hat{f}\}_{k=1}^K) \hat{\mathbf{a}},$$

with parameters $\beta = \{K, W\}$. Additionally, the analog denoiser for bearing estimation can be written as

$$\{\hat{f}_k, \hat{\mathbf{a}}^k\}_{k=1}^K = \text{MUSIC}(\mathbf{X}, \beta)$$

$$\{\hat{\theta}_k\}_{k=1}^K = \left\{ \sin^{-1}(\lambda \hat{f}_k / d_x) \right\}_{k=1}^K$$

$$\hat{\mathbf{X}} = \mathbf{S}(\hat{\theta}) \hat{\mathbf{A}},$$

with parameters $\beta = \{K\}$. Here, the snapshots are contained in \mathbf{X} . The above steps estimate the bearings $\{\hat{\theta}_k\}_{k=1}^K$ and the corresponding amplitude (column) vectors $\{\hat{\mathbf{a}}^k\}_{k=1}^K \in \mathbb{R}^Q$, and $\hat{\mathbf{A}} = \begin{bmatrix} \hat{\mathbf{a}}^1 & \hat{\mathbf{a}}^2 & \dots & \hat{\mathbf{a}}^K \end{bmatrix}^T$. At each iteration of D-AMP, we leverage the above concatenation of the parametric bearing estimation step and the signal synthesizer as an analog denoiser $\hat{\mathbf{X}} = D_{\sigma,\beta}(\mathbf{X})$, noting that the estimates of the parameters are obtained as a byproduct of the analog denoising process in each iteration. Note also that due to the knowledge of the signal sparsity the noise variance is not needed as a parameter. The Onsager correction term for these analog denoisers can be estimated using the numerical scheme described in Section 2.10. We refer to the integration of D-AMP with the proposed Root MUSIC-based analog denoisers as AMP+MUSIC in the sequel.

The remainder of this section focuses on the analysis of the proposed analog denoisers, reviewing a set of necessary conditions on a denoiser for D-AMP to provide successful signal recovery [25] and evaluating these conditions using several numerical experiments. We assume that the signal of interest \mathbf{x} belongs to the class of signals $\mathcal{C} \subseteq \mathbb{R}^N$ defined by the synthesis equation (2.13).

3.1.1 Signal Generation

Throughout this section, we setup numerical experiments while considering frequency-sparse and bearing-sparse signals as follows, unless otherwise stated. For frequency-sparse signals, we choose signals \mathbf{x} with amplitudes and frequencies chosen uniformly at random from $[0, 1)$ and $[0, N)$, respectively. We set the default values for the signal length $N = 512$, sparsity level $K = \lfloor \frac{N}{100} \rfloor = 5$, and window size $W = \lfloor \frac{N}{4} \rfloor = 128$. Additionally, the minimum separation S , defined as the minimum allowed distance between any two adjacent frequencies, is set to zero so that the frequencies are selected arbitrarily, i.e., $S = 0$.

For bearing-sparse signals, we set the number of transmitters $K = 30$, the number of snapshots $Q = 100$, and the number of receivers $P = 128$. The wavelength of the signal λ and the separation between the receiver sensors d_x are set to be 1 and 0.5, respectively. In order to generate the antenna observations, we choose the transmitter locations uniformly at random in $[-\frac{\pi}{2}, \frac{\pi}{2})$ and their amplitudes uniformly at random from $[0, 1)$.

3.1.2 Denoiser Properties

Metzler et al. [25] have stated that for a family of denoisers D_σ to be suitable within the D-AMP algorithm, it must satisfy several conditions. We investigate these conditions to evaluate the analog denoisers described in Section 3.1 for frequency-sparse and bearing-sparse signals.

3.1.2.1 Denoiser Properness

Let us quantify the performance of analog denoisers for analytical purposes.

Definition 3.1.1. [25] *A family of denoisers $D_{\sigma,\beta}$ for the class of signals \mathcal{C} is proper with level $\kappa \in (0, 1)$ if*

$$\sup_{\mathbf{x} \in \mathcal{C}} \frac{\mathbb{E}_\epsilon [\|D_{\sigma,\beta}(\mathbf{x} + \sigma\epsilon) - \mathbf{x}\|_2^2]}{N} < \kappa\sigma^2 \quad (3.3)$$

for every $\sigma > 0$, where the expectation is with respect to $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$.

We perform experiments to study the level of our analog denoisers with varying parameter values and signal generation setups. For frequency-sparse signal experiments, 100 signals are generated at random, while we only generate 40 signals for experiments with bearing-sparse signals. Moreover, in order to compute the expectation, we have used the sample mean over 50 independent noise trials for each variance of the measurement noise. Our computation of the level κ considers noise variances

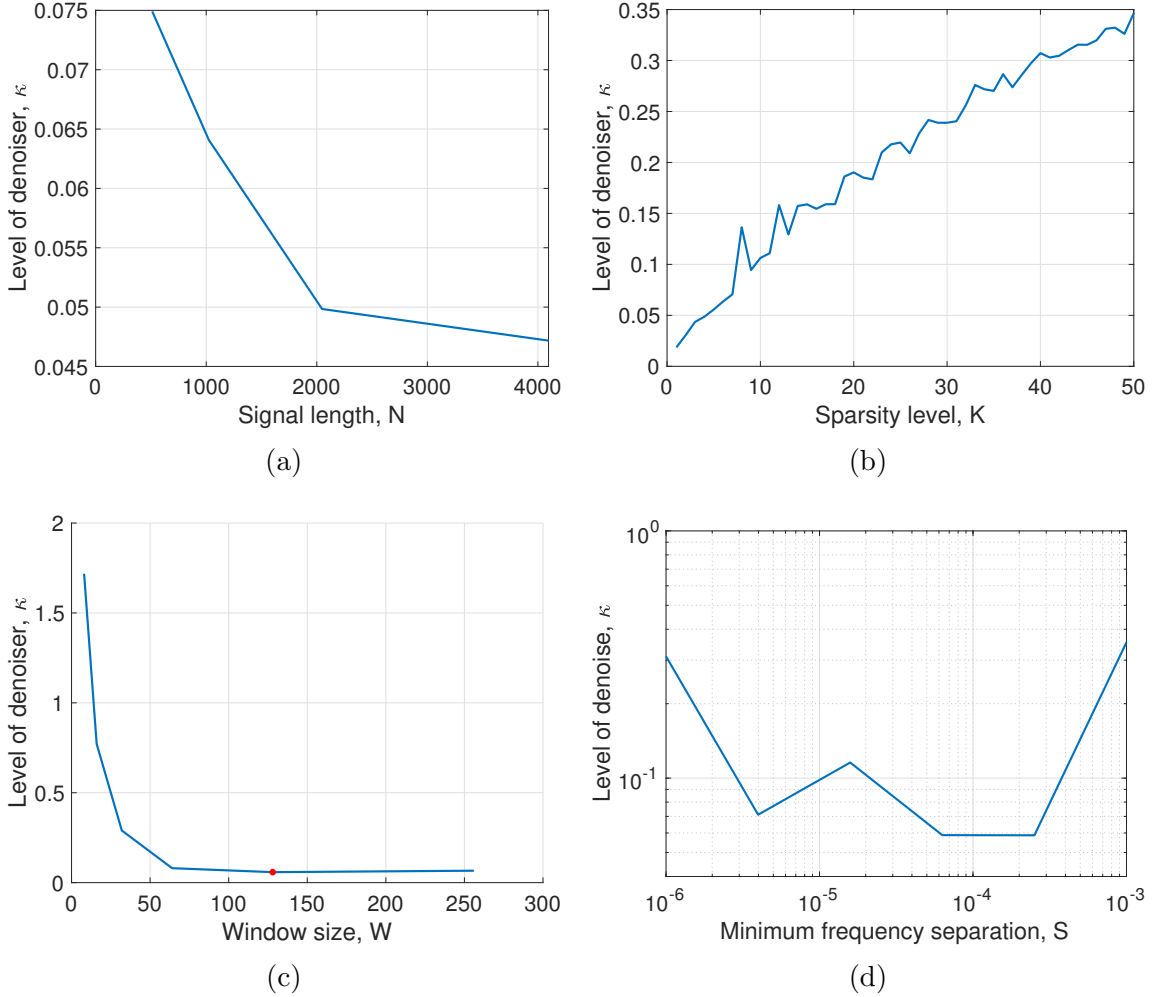


Figure 3.1: *Dependence of the denoiser level κ on signal generation parameters for frequency-sparse signals. (a) Signal length N . (b) Sparsity level K . (c) Window size W . (d) Minimum frequency separation S .*

σ corresponding to signal-to-noise ratios between 0 and 20 dB. Figures 3.1 and 3.2 show the impact of different signal generation parameters on the level of our analog denoisers; note that while studying the effect of each parameter, all the others are set to their default values listed in Section 3.1.1.

The results shown in Figures 3.1 and 3.2 are intuitively reasonable: the level κ is roughly proportional to the signal sparsity K , while increasing the signal length in frequency-sparse signals leads to a reduction of denoiser level. Considering the

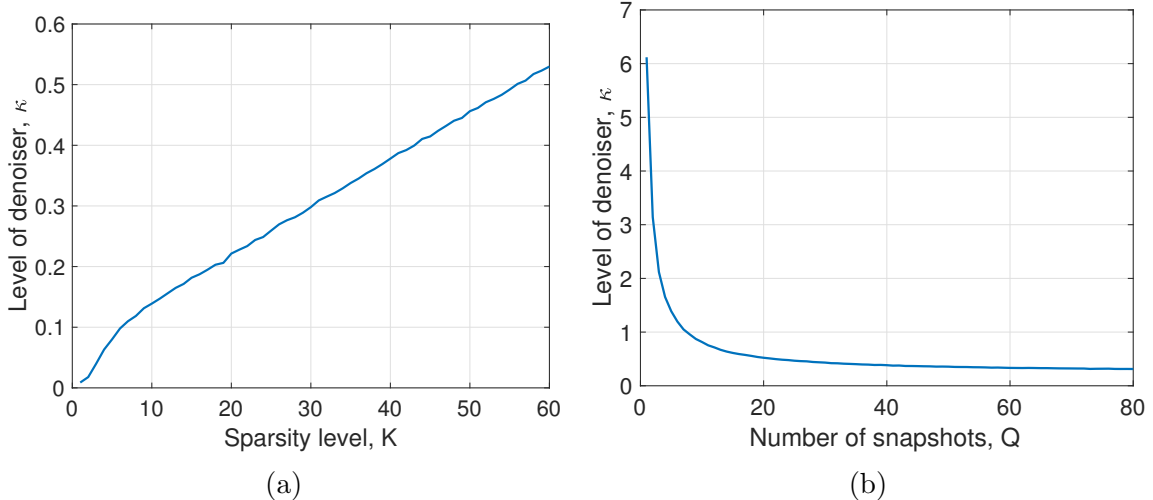


Figure 3.2: *Dependence of the denoiser level κ on signal generation parameters for bearing-sparse signals. (a) Sparsity level (number of transmitters) K . (b) Number of snapshots Q .*

window size W , we observe that the denoiser level converges to a low value once W is sufficiently small. For bearing-sparse signals, one can observe that increasing the number of snapshots leads to a level reduction. This is intuitive given that as the number of snapshots Q grows, the accuracy of the autocorrelation matrix estimate increases.

While the level of denoiser should be in the range $(0, 1)$ for D-AMP use, we observe in some cases for bearing-sparse signals that $\kappa > 1$. Having such large level values means that the noise in the output is greater than that in the input; as a result, the iterations of D-AMP would not converge. For example, in Figure 3.2 $Q \leq 7$ results in $\kappa > 1$; thus, we cannot handle bearing-sparse signals involving fewer than seven snapshots with our analog denoiser. High levels are often alleviated by reducing the complexity of the signal class \mathcal{C} .

3.1.2.2 Denoiser Monotonicity

The next denoiser condition uses the denoiser risk function, written as

$$R_{D_{\sigma,\beta}}(\sigma^2, \mathbf{x}) \triangleq \frac{\mathbb{E}_{\epsilon}(\|D_{\sigma,\beta}(\mathbf{x} + \sigma\epsilon) - \mathbf{x}\|_2^2)}{N}. \quad (3.4)$$

Definition 3.1.2. [25] *A denoiser is monotone if for every signal $\mathbf{x} \in \mathcal{C}$, $R_{D_{\sigma,\beta}}(\sigma^2, \mathbf{x})$ is a non-decreasing function of σ^2 .*

Note that denoiser monotonicity along with the level of the denoiser being below the line $R_{D_{\sigma,\beta}}(\sigma^2, \mathbf{x}) = \sigma^2$ ensures convergence of D-AMP. Monotonicity may also occur for certain ranges of values for σ , and the performance of D-AMP will be dependent on the region of the risk function being observed in its iterations.

We perform numerical experiments to illustrate the risk as a function of the noise variance for multiple signal model and denoiser parameter values. Figures 3.3 and 3.4 study the effect of several parameters on the value of the risk as a function of the measurement noise for frequency and bearing-sparse signals, respectively. In each case the signal of interest \mathbf{x} is fixed across all experiments, and the risk expectation is computed over 50 random realizations of the noise vector $\epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Not surprisingly, Figure 3.3 shows that the slope of the risk function grows as the sparsity increases. Hence, due to the larger level of distortion present at the output of the denoiser, the convergence is slower for larger K .

For frequency-sparse signals, the denoiser level that can be handled depends on the window size parameter W , due to the fact that the covariance matrix can only be estimated accurately for large windows. We see that the analog denoiser is monotonic for all studied parameters, except for high levels of noise. In other words, as long as the measurement noise is sufficiently small, the denoiser shows monotonic behavior and it is guaranteed that the D-AMP will succeed in signal recovery. Figure 3.4 also indicates that the analog denoiser for bearing-sparse signals is monotonic for all parameter values.

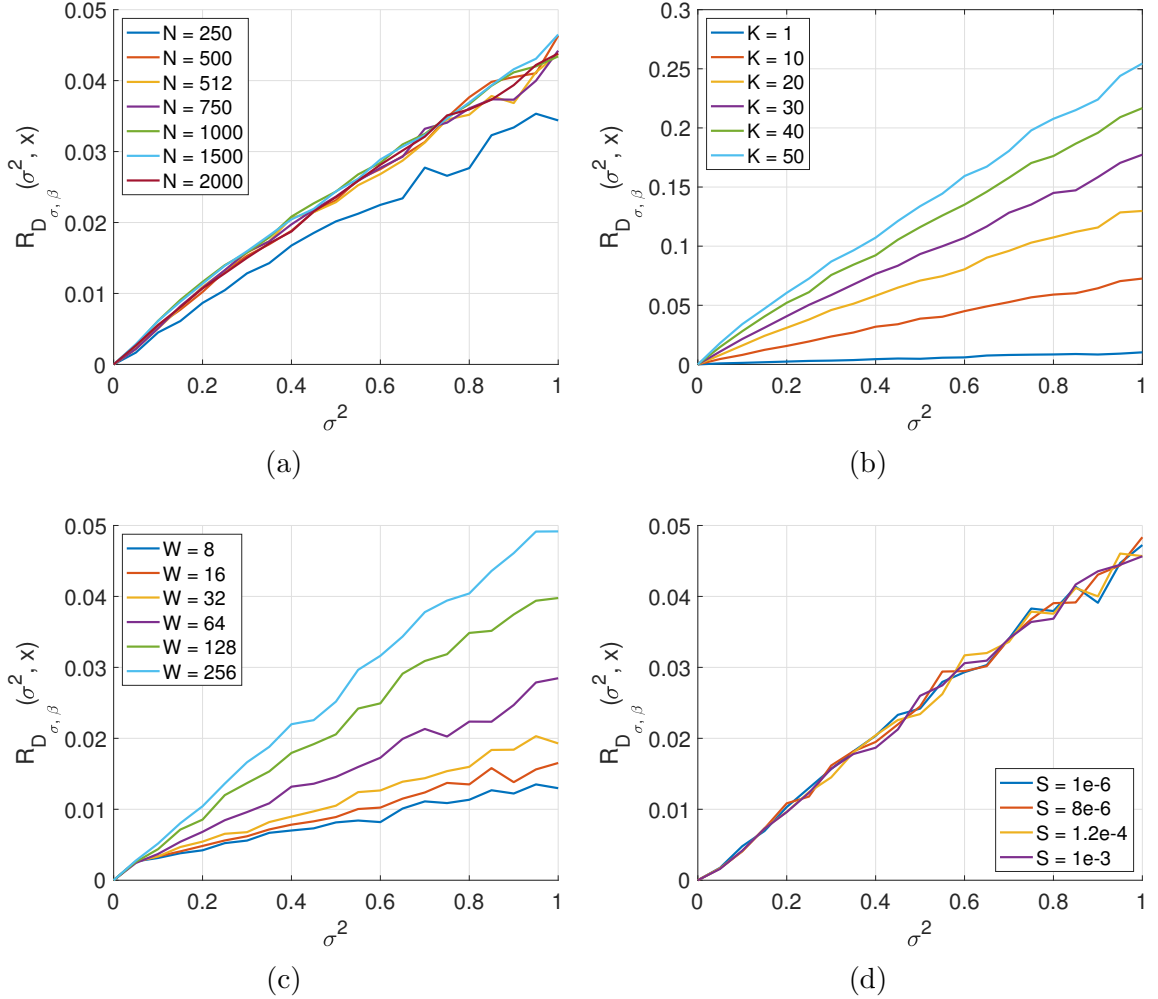


Figure 3.3: *Dependence of the denoiser monotonicity on signal generation parameters for frequency-sparse signals. (a) Signal length N . (b) Sparsity level K . (c) Window size W . (d) Minimum frequency separation S .*

In summary, these experiments indicate that our analog denoisers lead to noise reduction for a wide range of the tested parameters, given the fact that the slope of the risk function is less than one.

3.1.3 State evolution

A key ingredient in the analysis of D-AMP is state evolution [25]. Defining \mathbf{x} as in Section 2.12, starting from $\theta_0(\mathbf{x}, \delta, \sigma^2) = \frac{\|\mathbf{x}\|_2^2}{N}$, state evolution generates a sequence of numbers through the following iterations:

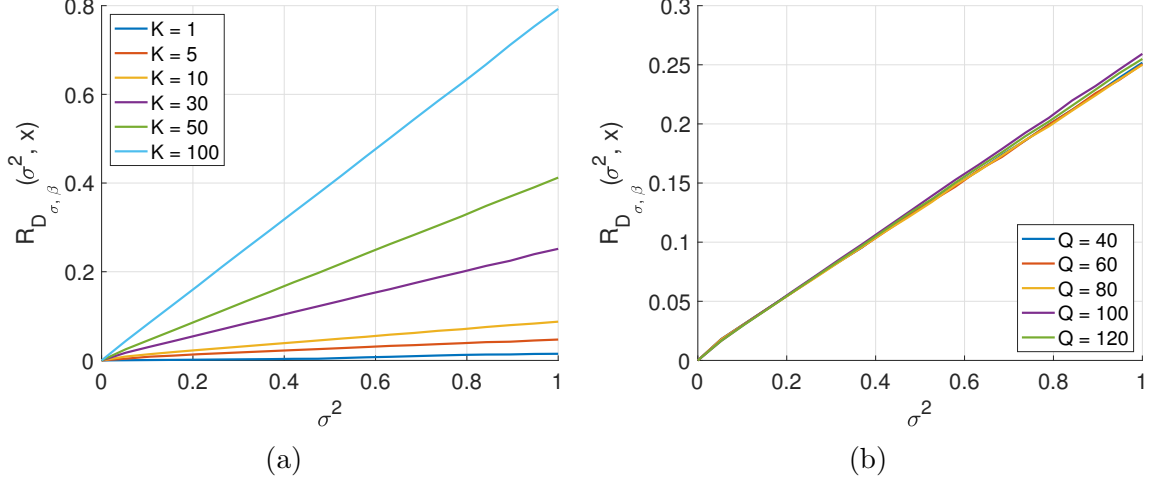


Figure 3.4: *Dependence of the denoiser monotonicity on signal generation parameters for bearing-sparse signals. (a) Sparsity level (number of transmitters) K . (b) Number of snapshots Q .*

$$(\sigma^t)^2 = \frac{\theta^t(\mathbf{x}, \delta, \sigma^2)}{\delta} + \sigma^2, \quad (3.5)$$

$$\theta^{t+1}(\mathbf{x}, \delta, \sigma^2) = R_{D_{\sigma, \beta}}((\sigma^t)^2, \mathbf{x}), \quad (3.6)$$

where δ and σ^2 are the undersampling ratio ($\delta = \frac{M}{N}$) and the variance of the measurement noise, respectively. State evolution is very useful in our analysis because it has been found to accurately predict the performance of D-AMP signal recovery across its iterations when \mathbf{A} is a Gaussian matrix.

Finding 3.1.1. [25] *If the D-AMP algorithm starts from $\mathbf{x}^0 = \mathbf{0}$, for large values of M and N , state evolution predicts the mean squared error of D-AMP, i.e.,*

$$\theta^t(\mathbf{x}, \delta, \sigma^2) \approx \frac{1}{N} \|\mathbf{x}^t - \mathbf{x}\|_2^2.$$

In other words, the state evolution is a predictive tool that attempts to characterize mean squared error (MSE) performance in a theoretical framework. We will now compare this prediction with the experimental performance of AMP+MUSIC.

Our experiments for frequency-sparse signals compare the state evolution sequence from our analog denoisers and the signal recovery MSE from AMP+MUSIC. To perform this comparison, we consider one signal and evaluate the signal recovery error prediction given by the state evolution sequence, according to Finding 3.1.1. We then compare this predicted value to the empirical recovery error obtained by applying the AMP+MUSIC algorithm to the signal on an iteration-by-iteration basis. In order to compute the Onsager correction term, we use the Monte Carlo method of Section 2.10 with $B = 2$ random draws and $\zeta = \frac{1}{\sqrt{N}}$. The experiments are then repeated for different values of the undersampling ratio δ and the measurement noise variance σ^2 ; the results are shown in Figure 3.5. We see that the state evolution sequence predicts the real MSE accurately in the noiseless setting. Additionally, increasing the number of measurements while keeping the signal length constant (i.e., increasing the undersampling ratio δ) results in smaller signal recovery error as well as improved predictions from Finding 3.1.1. However, increasing the measurement noise variance leads to larger signal recovery error.

The state evolution sequence is also a good estimator of the signal recovery error for bearing-sparse signals. The Monte Carlo parameters used to estimate the Onsager correction are $B = 2$ and $\zeta = \frac{1}{\sqrt{QP}}$. Figure 3.6 compares the state evolution sequence and the numerical MSE for different measurement noise variances for bearing-sparse signals. We observe that the state evolution sequence predicts the real MSE with high accuracy. Moreover, as expected, increasing the undersampling ratio δ results in smaller signal recovery error. We observe that for an SNR of 20 dB and $\delta = 0.5$, the MSE increases with the number of iterations.

Let us focus on the effect of measurement noise on the performance of AMP+MUSIC and its prediction. For frequency-sparse signals under noiseless measurements, the convergence point of the state evolution sequence is (essentially) zero, while for noisy measurement the convergence point of state evolution is nonzero. In order to have

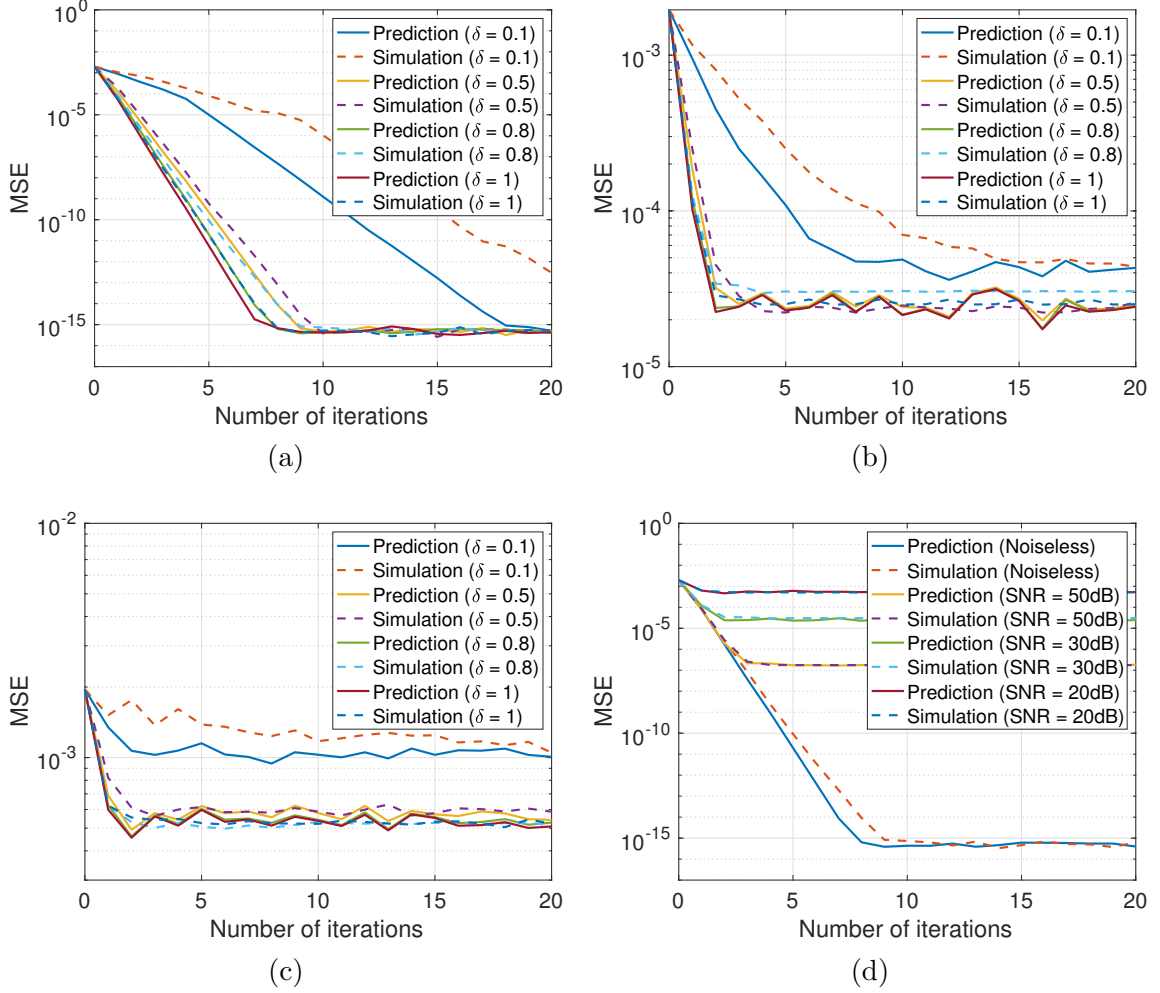


Figure 3.5: Evaluation of iteration-wise AMP+MUSIC recovery MSE prediction by the state evolution sequence with different measurement SNRs and undersampling ratios δ for frequency-sparse signals. (a) Noiseless measurements. (b) SNR = 30dB. (c) SNR = 20 dB. (d) $\delta = 0.8$.

a decreasing state evolution sequence (which is predictive of successful recovery with AMP+MUSIC), we need $\theta^{t+1}(\mathbf{x}, \delta, \sigma^2) < \theta^t(\mathbf{x}, \delta, \sigma^2)$ and the convergence point of the state evolution sequence is given by its fixed point $\theta^t(\mathbf{x}, \delta, \sigma^2) = \theta^{t+1}(\mathbf{x}, \delta, \sigma^2)$. Figures 3.7 and 3.8 illustrate the behavior of the sequences of state evolution predictions for frequency and bearing-sparse signals, respectively. For example, Figure 3.8 shows that for an SNR of 20 dB and $\delta = 0.5$, $\theta^{t+1}(\mathbf{x}, \delta, \sigma^2) > \theta^t(\mathbf{x}, \delta, \sigma^2)$, i.e., the MSE prediction grows as the number of iterations increases. Moreover, the fixed point of

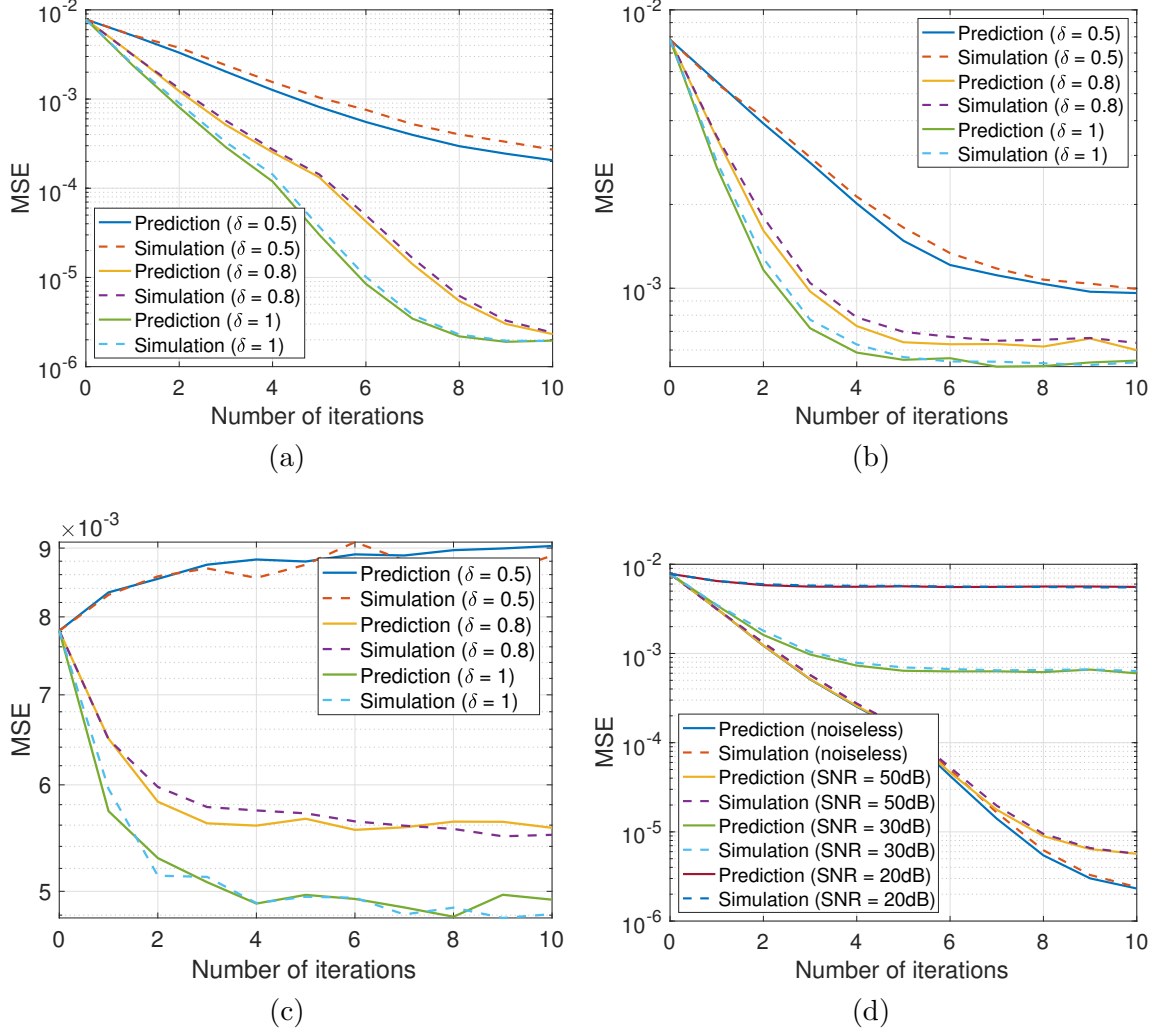


Figure 3.6: Evaluation of iteration-wise AMP+MUSIC recovery MSE prediction by the state evolution sequence with different measurement SNRs and undersampling ratios δ for bearing-sparse signals. (a) Noiseless measurements. (b) SNR = 30dB. (c) SNR = 20 dB. (d) $\delta = 0.8$.

the state evolution sequence is approximately 9×10^{-3} , which is in agreement with the results of Figure 3.6.

3.1.4 D-AMP Performance Analysis

We now focus on the performance analysis of the D-AMP algorithm in the noiseless and noisy measurement scenarios. Afterward, we will discuss the performance analysis of state evolution.

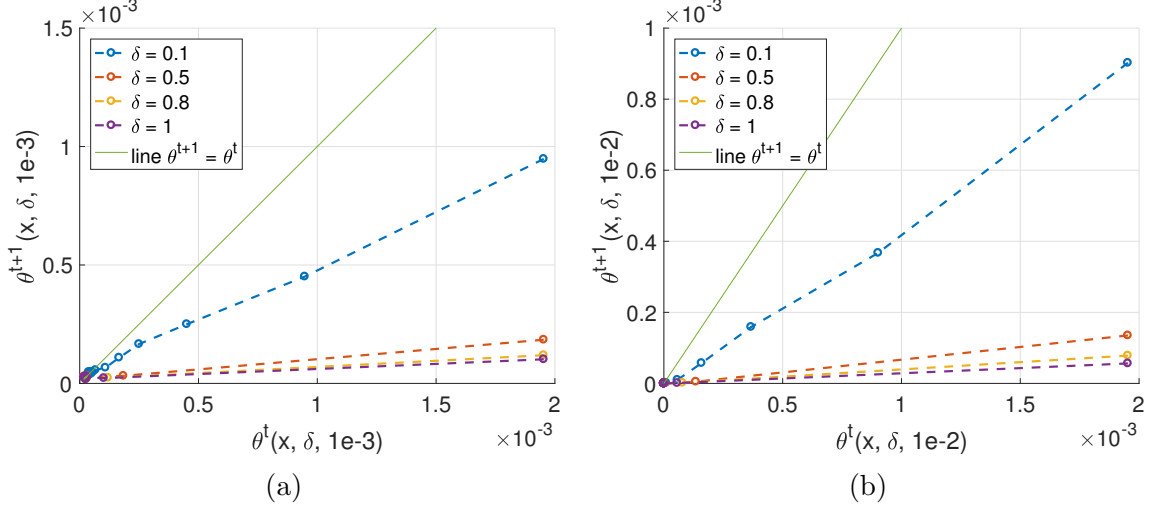


Figure 3.7: Convergence of state evolution sequence in frequency-sparse signals. (a) SNR = 30dB. (b) SNR = 20dB.

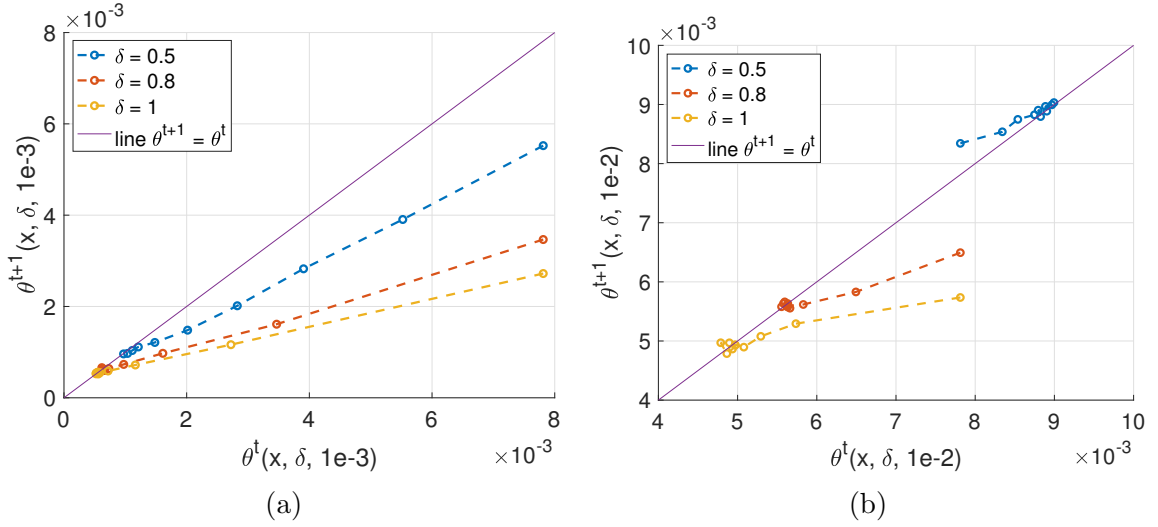


Figure 3.8: Convergence of state evolution sequence in bearing-sparse signals. (a) Noiseless measurements. (b) SNR = 20dB.

3.1.4.1 Noiseless Measurements

For monotone denoisers (cf. Section 3.1.2.2), the state evolution with $\sigma = 0$ may fall in one of two cases, depending on the value of δ : (i) $\theta^t(\mathbf{x}, \delta, 0) \rightarrow 0$ as $t \rightarrow \infty$, or (ii) $\theta^t(\mathbf{x}, \delta, 0) \not\rightarrow 0$ as $t \rightarrow \infty$. The former case implies successful D-AMP recovery,

while the latter implies failure. We therefore study success and failure regions under this prerogative.

Lemma 3.1.1. [25] *For monotone denoisers, if $\theta^t(\mathbf{x}, \delta_0, 0) \rightarrow 0$ as $t \rightarrow \infty$, then for any $\delta > \delta_0$, $\theta^t(\mathbf{x}, \delta, 0) \rightarrow 0$ as $t \rightarrow \infty$ as well.*

Note that for sufficiently small values of δ , our experimental results show that $\theta(\mathbf{x}, \delta, 0) \not\rightarrow 0$ as $t \rightarrow \infty$. If we combine this result with Lemma 3.1.1, we conclude that for small values of δ AMP+MUSIC fails in recovering \mathbf{s} . Nonetheless, as δ increases past a certain point, AMP+MUSIC will successfully recover \mathbf{s} from its CS measurements. Therefore, we define the minimum undersampling ratio required for the successful recovery of \mathbf{x} :

$$\delta^*(\mathbf{x}) = \inf_{\delta \in (0,1)} \{\delta : \theta^t(\mathbf{x}, \delta, 0) \rightarrow 0 \text{ as } t \rightarrow \infty\}. \quad (3.7)$$

Our goal is to characterize $\delta^*(\mathbf{x})$ in terms of the performance of the denoiser. However, since the undersampling ratio $\delta^*(\mathbf{x})$ depends on the signal \mathbf{x} , it is more practical to evaluate the number of measurements required by D-AMP to recover any signal $\mathbf{x} \in \mathcal{C}$. The following result addresses this question:

Proposition 3.1.1. [25] *Suppose that for the signal class \mathcal{C} the denoiser $D_{\sigma,\beta}$ is proper at level κ . Then the minimum undersampling ratio*

$$\delta_C \triangleq \sup_{\mathbf{x} \in \mathcal{C}} \delta^*(\mathbf{x}) \leq \kappa. \quad (3.8)$$

In order to check the correctness of this proposition, we compute the minimum undersampling ratio δ_C using a binary search algorithm on the range $\delta \in (0, 1)$ until the gap between a pair of δ values, one leading to successful recovery and the other leading to failure in recovery, reaches $\frac{1}{2N}$ for frequency-sparse signals or $\frac{1}{P}$ for

bearing-sparse signals. Success is defined as achieving normalized signal recovery error of 10^{-6} after 20 iterations of AMP+MUSIC in all cases. We then compare these results against the analog denoiser levels obtained in Section 3.1.2.1. As before, the values of non-varying parameters are set to be equal to their default values listed in Section 3.1.1.

Figures 3.9 and 3.10 illustrate the effect of different parameters of the signal models and analog denoisers on the terms of Proposition 3.1.1. It is observed that this proposition holds for both of our analog denoisers and for all tested parameters. Nonetheless, there is an often nontrivial gap between the minimum undersampling ratio δ_C and its upper bound, the level κ . This is likely due to the fact that (3.8) is a worst-case bound among all signals $\mathbf{x} \in \mathcal{C}$.

3.1.4.2 Noisy Measurements

Let us now consider the analysis of D-AMP in the presence of measurement noise. Let the measurement vector $\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}$, where \mathbf{A} is the Gaussian sensing matrix as before and $\mathbf{n} \sim \mathcal{N}(0, \sigma^2\mathbf{I})$ is AWGN. We define the noise sensitivity of D-AMP as the worst-case convergence point of the state evolution for a given noise variance σ^2 and undersampling ratio δ [25]:

$$NS(\sigma^2, \delta) = \sup_{\mathbf{x} \in \mathcal{C}} \theta^\infty(\mathbf{x}, \sigma^2, \delta), \quad (3.9)$$

where $\theta^\infty(\mathbf{x}, \delta, \sigma^2)$, defined in (3.6), denotes the convergence point of the state evolution equation. The following result shows that D-AMP is robust to the measurement noise.

Proposition 3.1.2. [25] *Let $D_{\sigma, \beta}$ denote a proper family of denoisers at level κ . Then, for $\delta > \kappa$, the noise sensitivity of D-AMP satisfies*

$$NS(\sigma^2, \delta) \leq \frac{\kappa\sigma^2}{1 - \frac{\kappa}{\delta}}. \quad (3.10)$$

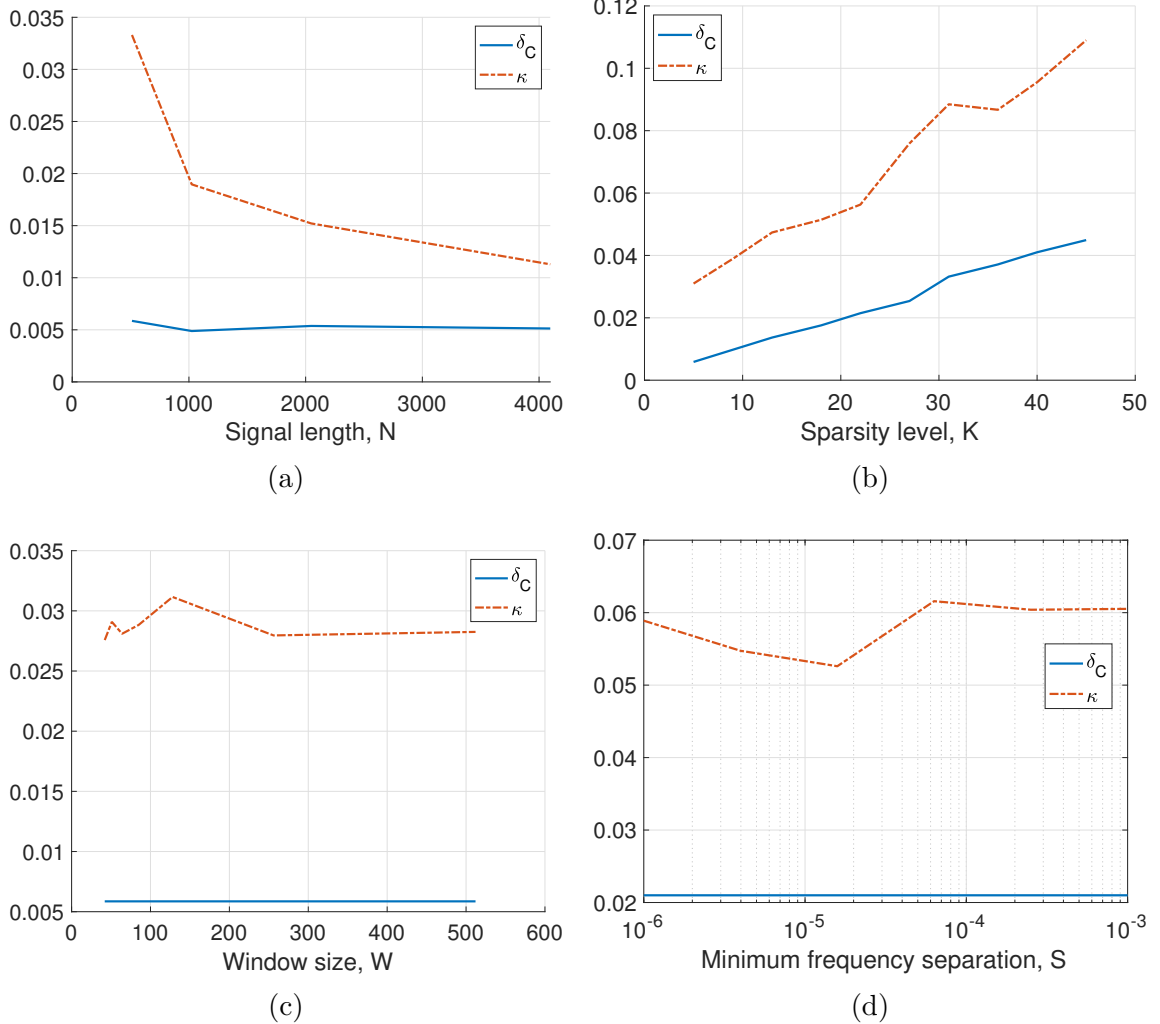


Figure 3.9: *Effect of different signal generation and denoiser parameters on the terms of Proposition 3.1.1 for frequency-sparse signals. (a) Signal length N . (b) Sparsity level K . (c) Window size W . (d) Minimum frequency separation S .*

We note that in the noisy scenario where $\sigma > 0$, the upper bound in (3.10) goes to infinity as the undersampling ratio approaches the minimum number of measurements required for exact recovery from noiseless measurements (i.e., $\delta \rightarrow \kappa$).

Proposition 3.1.2 connects the noise sensitivity and the variance of the measurement noise via the level of the denoiser. More precisely, for any signal \mathbf{x} chosen from a certain class of signals \mathcal{C} with denoiser level κ , the convergence value of the state evolution sequence is at worst proportional to the measurement noise σ^2 with a scaling

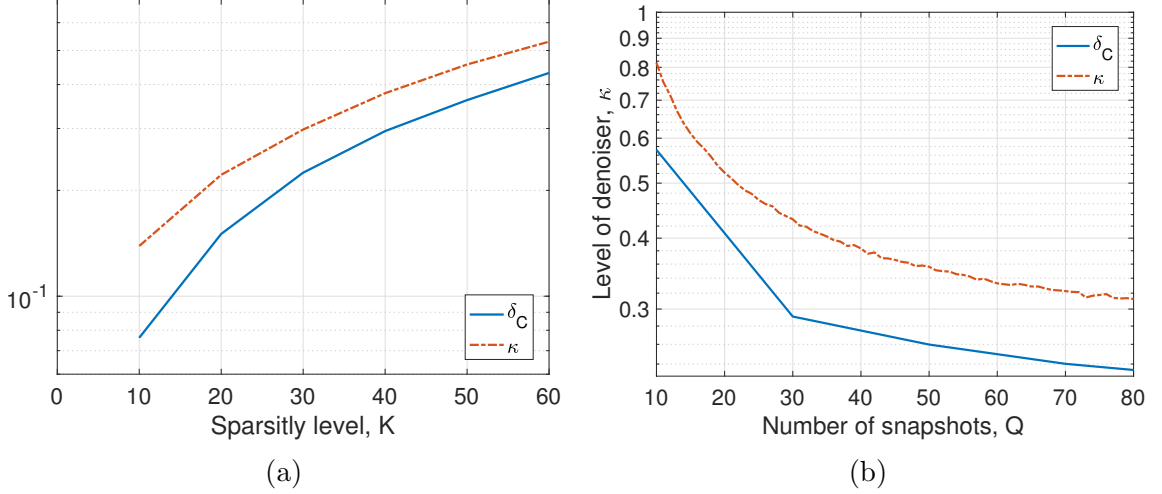


Figure 3.10: *Effect of different signal generation parameters on Proposition 3.1.1 for bearing-sparse signals. (a) Sparsity level K . (b) Number of snapshots Q .*

factor of $\frac{\delta\kappa}{\delta-\kappa}$. Increasing the measurement noise leads to a higher value of the state evolution sequence convergence, which is predictive of larger signal recovery MSE.

To evaluate Proposition 3.1.2 for our analog denoisers, we perform recovery on signals with varying parameter values and evaluate the relationship between the recovery error and the bound given by this proposition. In order to study the effect of these parameters on the aforementioned proposition, the noise sensitivity for a fixed value of δ and σ^2 is computed by taking the maximum value of the state evolution sequence over 100 randomly chosen signals after 20 iterations, when we usually have achieved convergence. We use the analog denoiser levels obtained in Section 3.1.2.1.

The results in Figures 3.11-3.12 show that although the noise sensitivity $NS(\sigma^2, \delta)$ and the bound in (3.10) have different behavior as different parameters change, Proposition 3.1.2 holds for our analog denoisers for all tested parameters and for both examples of parameter sparse signals. These results also provide some justification for parameter tuning. It is straightforward to see that increasing the value of κ while δ is fixed leads to increase in the upper bound for the noise sensitivity in (3.10). More precisely, decreasing κ is equivalent to reducing the coefficient of the effective mea-

surement noise variance at the output. Therefore, we desire parameter choices which lead to a smaller κ . As an example, Figure 3.1 shows that for a given value of N and K , we notice that the optimal window size in terms of the denoiser level is $W \approx \lceil \frac{N}{4} \rceil$.

3.2 Analysis of parameter estimation performance

In contrast to the focus on signal estimation performance in Section 3.1, we now attempt to leverage analysis on the performance of parameter estimation within analog denoisers to provide theoretical guarantees on compressive parameter estimation from AMP+MUSIC. To begin, we investigate the performance of the analog denoiser both in terms of parameter and signal estimation error from noisy signal observations. Our expectation is that, as in the recovery case, we can leverage the performance of the statistical estimator (instead of the statistical denoiser) within a state evolution loop for D-AMP to complete our analysis.

To start, we consider bearing-sparse signals and recall that the signal of interest \mathbf{x} has a parametrization consisting of the parameter values $\{\theta_k\}_{k=1}^K$ and their corresponding amplitudes $\{a_k\}_{k=1}^K$. We set up an experiment where we add AWGN to these signals with variance σ^2 and apply the Root MUSIC estimator to the noisy signal. We then estimate a recovered signal from the estimated parameters via the synthesis step of (2.13), i.e., the complete analog denoiser. We measure the average frequency estimation error by computing the cost of the Hungarian matching [24, 28] between the vectors containing the frequency values and their estimates, and normalizing by the vector length. One interesting issue observed in these experiments is that for some bearing values (close to $\pm \frac{\pi}{2}$) the average analog denoiser error might be unexpectedly high. We address this issue later in this section, and for now we choose the bearings in signal generation from the range $[-1.4, 1.4)$ instead. Furthermore, the sparsity level is set to $K = 10$. The experiment is repeated for 50 different randomly

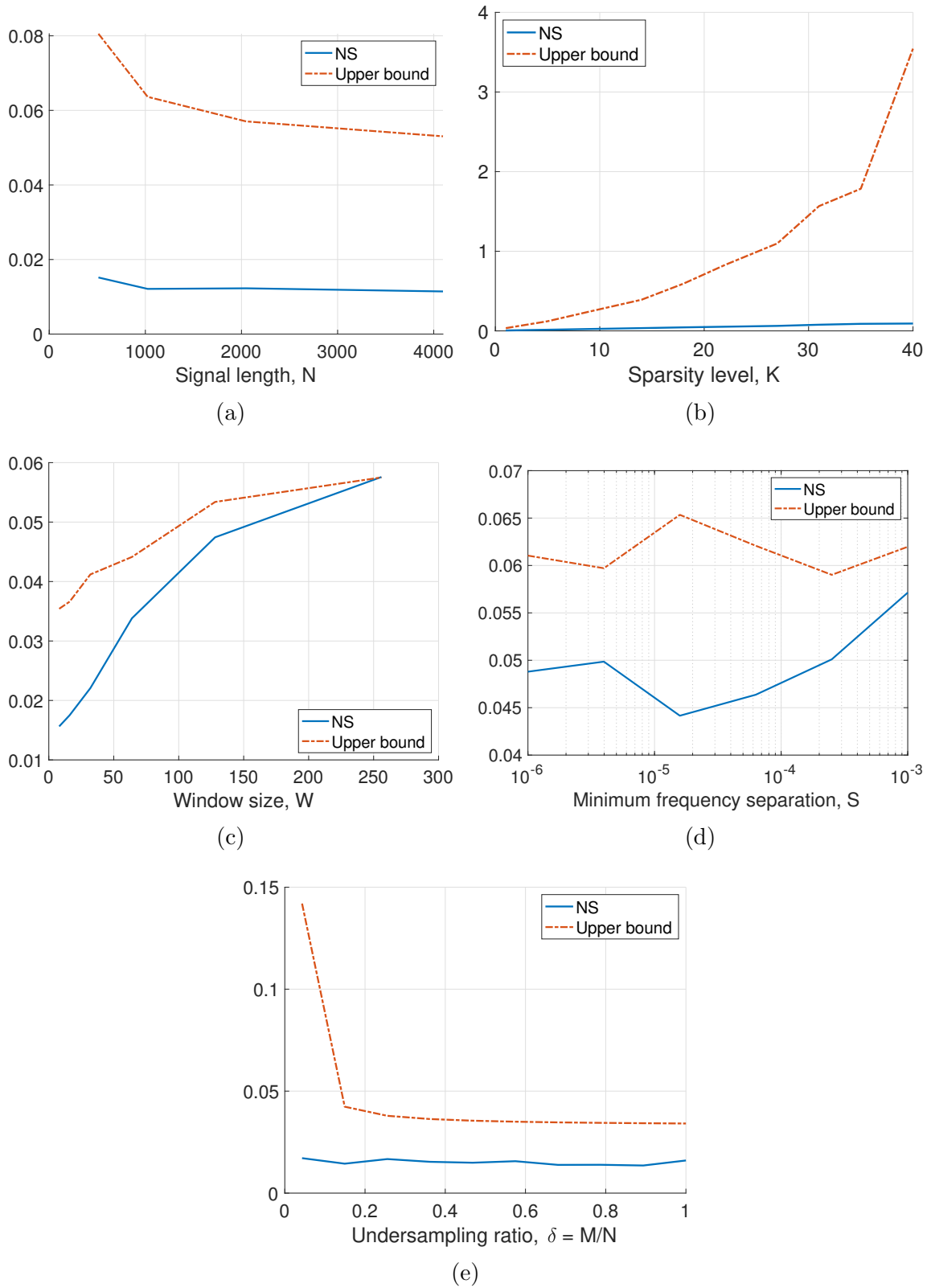


Figure 3.11: Effect of different signal generation, denoiser, and CS parameters on the terms of Proposition 3.1.2 for frequency-sparse signals. (a) Signal length N . (b) Sparsity level K . (c) Window size W . (d) Minimum frequency separation S . (e) Undersampling ratio δ .

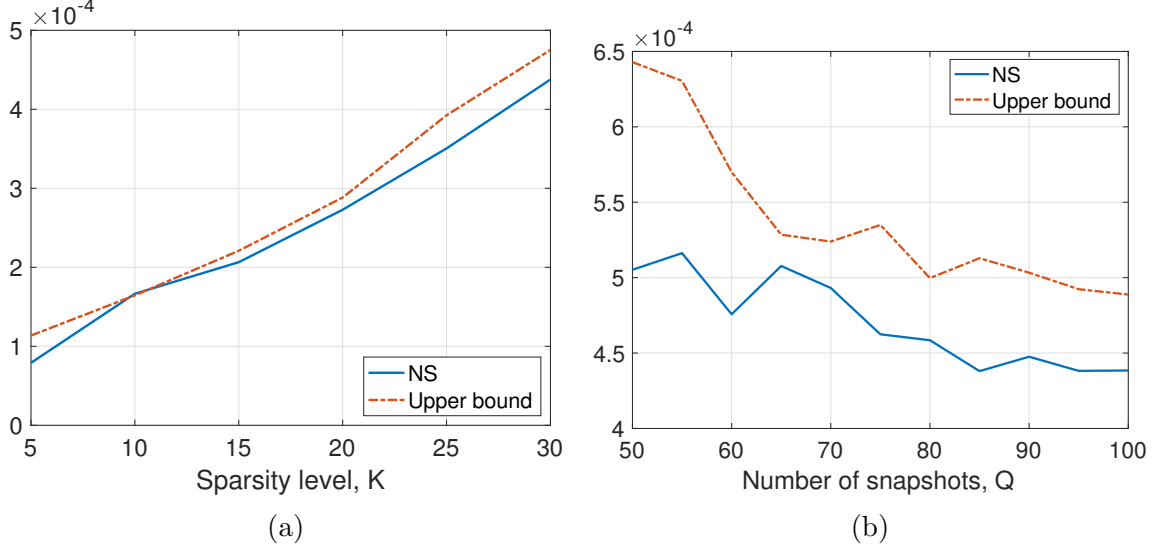


Figure 3.12: *Effect of different signal generation parameters on the terms of Proposition 3.1.2 for bearing-sparse signals. (a) Sparsity level K . (b) Number of snapshots Q .*

generated signals as described in Section 3.1.1, with each color corresponding to the average error for a specific signal over 50 instances of noise trials. Figure 3.13 illustrates the error in signal and bearing estimation as a function of the variance of the measurement noise. These results show the dependence of the error on the values of the bearings present in the signals, which we will analyze next.

We consider the analysis in [29] of the expected parameter estimation error from Root MUSIC, which focuses on bearing estimation. Recall from Section 2.11.2 that the bearings of interest are $\{\theta_k\}_{k=1}^K$ and $\{\omega_k\}_{k=1}^K = \{2\pi(d_x/\lambda) \sin \theta_k\}_{k=1}^K$ are the angular frequencies of the complex exponentials present in each snapshot. It has been shown in [29] that the Root MUSIC estimation errors $\Delta\theta_k$ obey

$$\mathbb{E}(|\Delta\theta_k|^2) \approx \left(\frac{\lambda}{2\pi d_x \cos \theta_k} \right)^2 \frac{S_{MU}(P-K)\sigma^2}{2P^2Q} \sum_{i=1}^K \left(\frac{\lambda_i}{(\tilde{\lambda}_i)^2} |\mathbf{v}^H(\omega_k)\mathbf{q}_i|^2 \right). \quad (3.11)$$

Here, $\{\lambda_i\}_{i=1}^P$ are the sorted eigenvalues of \mathbf{R}_{ss} , with $\lambda_1 > \lambda_2 > \dots > \lambda_K > \lambda_{K+1} = \dots = \lambda_P = \sigma^2$, and $\{\mathbf{q}_i\}_{i=1}^P$ are their corresponding eigenvectors. Recall that $\tilde{\lambda}_k =$

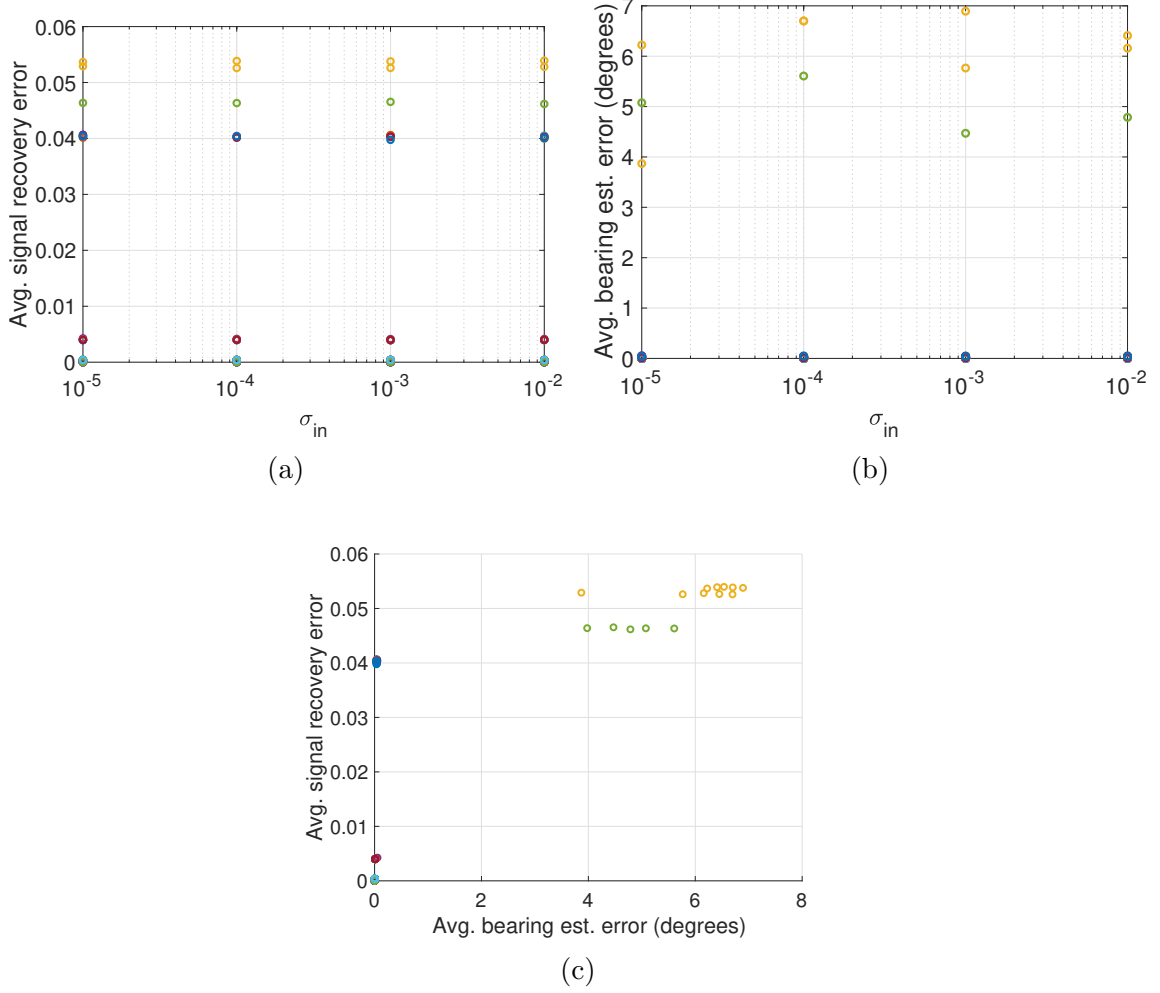


Figure 3.13: (a) Performance of the analog denoiser for bearing-sparse signals as a function of the measurement noise variance. (b) Performance of Root MUSIC frequency estimation as a function of the measurement noise variance. (c) Performance of Root MUSIC bearing estimation as a function of the analog denoiser error for bearing-sparse signals. Different colors correspond to different signals in the experiment.

$\lambda_k - \sigma^2$, and denote the length- P complex exponential vector by $\mathbf{v}(\omega)$ (similar to (2.14) but with angular frequency). We also denote $\mathbf{v}_1(\omega)$ to be the derivative of $\mathbf{v}(\omega)$, i.e.,

$$\mathbf{v}_1(\omega) = \frac{1}{\sqrt{P}} \left[0, j e^{j\omega}, 2j e^{j2\omega}, \dots, j(P-1) e^{j(P-1)\omega} \right]^T.$$

We can write

$$\mathbf{R}_{ss} = \mathbf{E} \mathbf{\Lambda} \mathbf{E}^H = \mathbf{E}_x \mathbf{\Lambda}_x \mathbf{E}_x^H + \sigma^2 \mathbf{I} = \sum_{i=1}^P \lambda_i \mathbf{q}_i \mathbf{q}_i^H,$$

where $\mathbf{E} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_P]$, $\mathbf{E}_x = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_K]$, and $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_P)$, $\mathbf{\Lambda}_x = \text{diag}(\tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_K)$. Finally, noting $\mathbf{P}_N = \mathbf{I} - \mathbf{E}_x \mathbf{E}_x^H$, S_{MU} in (3.11) is defined as

$$S_{MU} = \frac{P}{\mathbf{v}_1^H(\omega_k) \mathbf{P}_N \mathbf{v}_1(\omega_k)},$$

Recall that one can easily link the frequency estimation problem to bearing estimation and see that the frequency estimation error for the frequency ω_k is a function of ω_k itself; it can similarly be shown from [29] that for frequency estimation we have

$$\mathbb{E}(|\Delta\omega_k|^2) \approx \frac{S_{MU}(P-K)\sigma^2}{2P^2Q} \sum_{i=1}^K \left(\frac{\lambda_i}{(\tilde{\lambda}_i)^2} |\mathbf{v}^H(\omega_k) \mathbf{q}_i|^2 \right).$$

Note also that (3.11) implies a high expected value of estimation error for the cases where a bearing θ_k is close to $\pm\frac{\pi}{2}$, i.e., $\cos\theta_k \approx 0$. Our numerical experiments are consistent with this prediction; therefore, we limit the parameters to the range $[-1.4, 1.4)$, rather than $[-\frac{\pi}{2}, \frac{\pi}{2})$, to avoid such cases of inaccurate estimation.

In order to verify the accuracy of (3.11), we perform an experiment to compare the accuracy of the predicted expected error with empirical observations for the Root MUSIC estimator. We consider in this experiment bearing-sparse signals with different sparsity levels ($K = 10, 20$, or 30) and values of measurement SNR (20, 30, 50dB, or noiseless). In each scenario with a given sparsity level and measurement noise level, we consider 100 randomly generated signals drawn as described in Section 3.1.1. We then compute the predicted expectation of the parameter estimation error for each of the bearings using (3.11) and compare it to the actual average parameter estimation error from Root MUSIC. Note that the average parameter estimation error is calculated over Root MUSIC estimates for a given signal under 50 different realizations of the random noise and uses the Hungarian matching to align true bearings with their estimates.

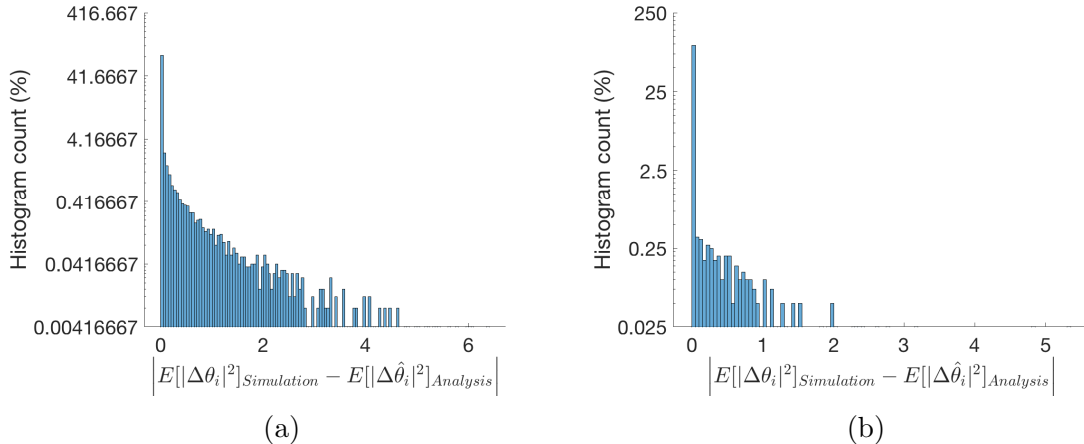


Figure 3.14: Histogram for the absolute difference between the empirical average bearing estimation error from Root MUSIC and the error predicted by (3.11) from [29] for two sets of numbers of transmitters K . (a) $K = 10$, $K = 20$, and $K = 30$. (b) $K = 10$.

Figure 3.14 shows the histogram of the absolute differences between the empirical and predicted parameter estimation errors. We observe that a significant number of predictions are accurate, but a nontrivial proportion also features significant prediction error. We then reduce the sparsity level to $K = 10$ and repeat our experiment; the results are shown in Figure 3.14, with almost 99% of the cases corresponding to highly accurate predictions (e.g., a discrepancy between predicted and actual error below 1°), while this number is significantly smaller in our earlier experiment. Our empirical results show that the performance prediction available in the literature for our chosen statistical parameter estimation algorithm (Root MUSIC) is not sufficiently accurate for us to provide analysis of the compressive parameter estimation performance of AMP+MUSIC. Nonetheless, our numerical results in the sequel show that our proposed approach to compressive parameter estimation outperforms those existing in the literature.

3.3 Numerical Evaluation of AMP+MUSIC

In this section, we show empirical results that evaluate the performance of AMP+MUSIC for both frequency and bearing-sparse signals. We evaluate both signal recovery and parameter estimation performance in each case. All experiments use CS matrices with i.i.d. Gaussian entries with variance $\sigma^2 = 1/M$.

3.3.1 Signal recovery

We study the signal recovery performance of AMP+MUSIC in terms of its MSE and its phase transition plot for both types of signals studied. The phase transition plot of a given recovery algorithm finds the maximum value of the normalized sparsity $\rho = K/N$ that allows for successful recovery with high probability as a function of the normalized measurement rate $\delta = M/N$ [11]. The phase transition plot is usually interpreted as showing the division between the (δ, ρ) regions for which the probability of successful signal recovery (i.e., sufficiently small MSE) goes to one as $N \rightarrow \infty$ (below the curve) from the (δ, ρ) region for which the probability of successful recovery goes to zero as $N \rightarrow \infty$ (above the curve). Thus, curves with higher values of ρ for a given value of δ are better.

3.3.1.1 Frequency-sparse Signals

In the first experiment we evaluate the performance of the AMP+MUSIC algorithm for signal recovery by studying the average normalized MSE (NMSE) of the recovered signal over 100 independent signal trials as a function of the number of measurements. The values of the signal length N , window size W , and minimum separation S are set to their default values from Section 3.1.1, while the sparsity level $K = 8$. Figure 3.15a shows that the NMSE converges to zero as the number of measurements increases, and $M > 50$ suffices for accurate signal recovery, e.g., $K/M \approx 1/6$.

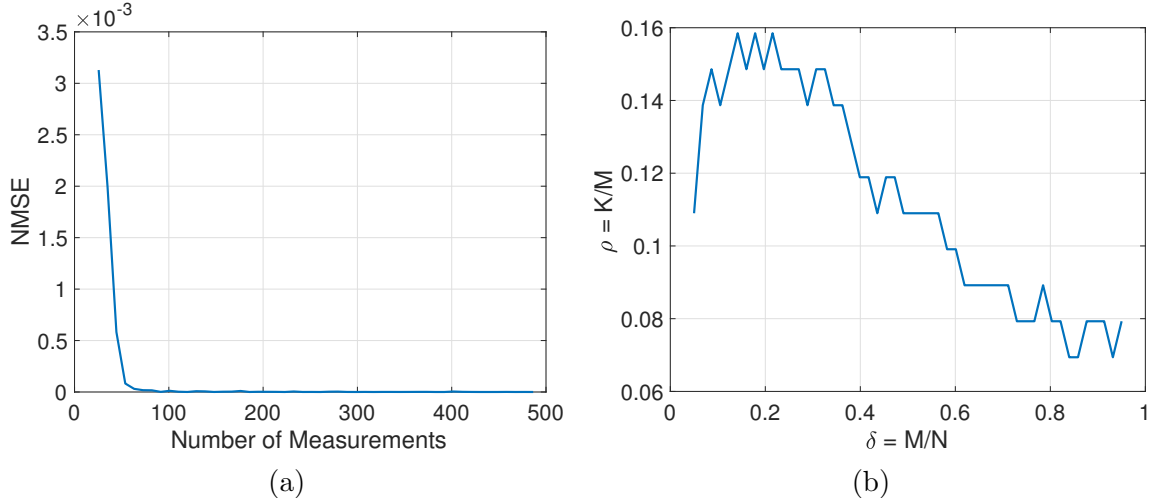


Figure 3.15: Performance of AMP+MUSIC for frequency-sparse signal recovery. (a) Average normalized mean squared error. (b) Phase transition plot.

In the second experiment we study the signal recovery performance of AMP+MUSIC in terms of the phase transition criterion. As before, the signal generation parameters are described in Section 3.1.1. We say recovery is successful if $\text{NMSE} < 10^{-6}$, following the behavior observed in Figure 3.15a. Figure 3.15b shows the curve that denotes the 50% chance of success for the algorithm, obtained via binary search, while considering 100 signal trials for each (δ, ρ) point evaluated. Note that AMP+MUSIC is run for 20 iterations for each trial. We observe that there is a certain value of δ that optimizes the value of ρ in the phase transition. Our conjecture is that there is a maximum value of K that can be resolved by the analog denoiser even as the noise appearing within the AMP+MUSIC iterations is reduced as M grows larger; the value is connected to the denoiser’s window size parameter W . Thus, the rate at which the break-even value of K increases is much slower than the rate of increase of M once we go past this “optimal” value of $\delta = M/N$.

3.3.1.2 Bearing-sparse Signals

We repeat our experiments for bearing-sparse signal with the corresponding analog denoiser in AMP+MUSIC. Figure 3.16a shows that the NMSE converges to zero as the number of measurements increases, albeit more slowly than for frequency-sparse signals. For the phase transition plot experiment, we need to choose a proper threshold for declaring success in signal recovery. As before, we observe the behavior in Figure 3.16a to determine a signal recovery success as $\text{NMSE} < 10^{-2}$. The signal generation parameters are also chosen as described in Section 3.1.1. Note that since in bearing estimation the number of bearings that can be identified is upper bounded by the number of antennas, we set $\rho = K/P$; furthermore, for experiments with bearing-sparse signals we denote by M the total number of measurements so that $\delta = M/(QP)$. Figure 3.16b shows the phase transition curve from 100 signal trials for each (δ, ρ) point evaluated. As before, AMP+MUSIC is run for 20 iterations for each trial. As observed in this figure, the normalized sparsity level ρ for which AMP+MUSIC succeeds in signal recovery increases with δ (i.e., with increasing the number of measurements) up to $\delta \approx 0.7$. After this point, AMP+MUSIC consistently succeeds in recovering the signal with acceptable accuracy even for $\rho = 1$, i.e., $K = P$.

3.3.2 Parameter Estimation

Next, we study the performance of AMP+MUSIC for compressive parameter estimation, including a comparison to existing approaches to this problem for the two signal classes of interest.

3.3.2.1 Frequency Estimation

In order to study the estimation performance of AMP+MUSIC, we again measure the frequency estimation error using Hungarian matching for the true frequencies and their estimates. In our experiments, we compare the performance of AMP+MUSIC to that of several alternative baselines: CS recovery followed by standard frequency

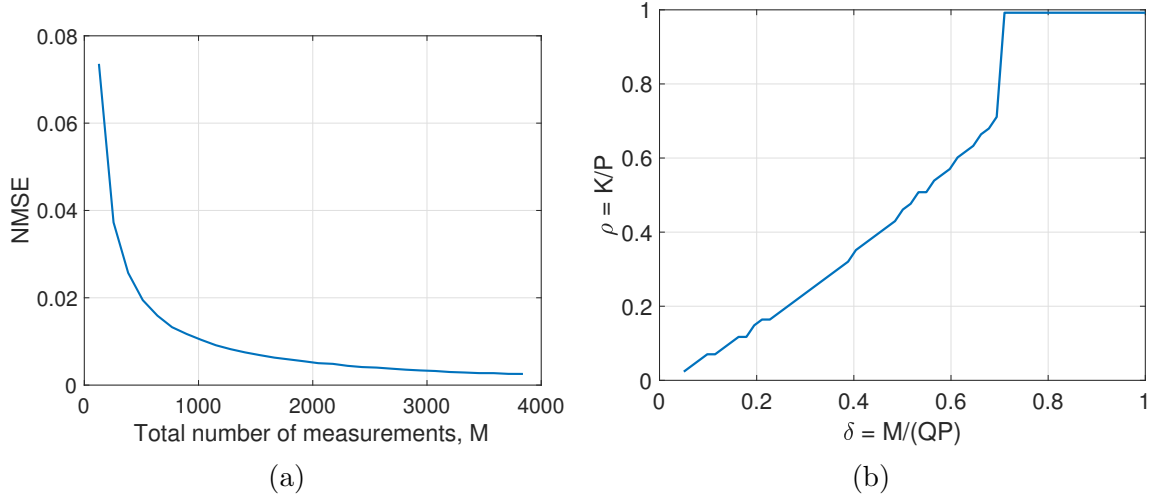


Figure 3.16: *Performance of AMP+MUSIC for bearing-sparse signal recovery. (a) Average normalized mean squared error. (b) Phase transition plot.*

estimation with Root MUSIC (AMP \rightarrow MUSIC and ℓ_1 -min. \rightarrow MUSIC); IHT + MUSIC, which is akin to AMP+MUSIC without the Onsager correction term, as proposed in [14]; and band-exclusion interpolated subspace pursuit (BISP) [18], a coherence-controlling sparsity-based algorithm. All iterative algorithms are run for 20 iterations. Figure 3.17a compares the average frequency estimation error over 100 signal trials as a function of the number of measurements. One can clearly see that once a sufficiently high number of measurements is taken, AMP+MUSIC again outperforms its competitors.

For the phase transition experiment, we define success as having an average frequency estimation error less than 1 Hz. Figure 3.17b compares the phase transition plots of these algorithms. Each figure shows the 50% success level for each algorithm, while considering 100 signal trials for each (δ, ρ) point. AMP+MUSIC not only outperforms in terms of estimation error but also in terms of the maximum number of recoverable frequencies. The numerical simulations indicate that the AMP+MUSIC algorithm for frequency estimation outperforms previously suggested algorithms both

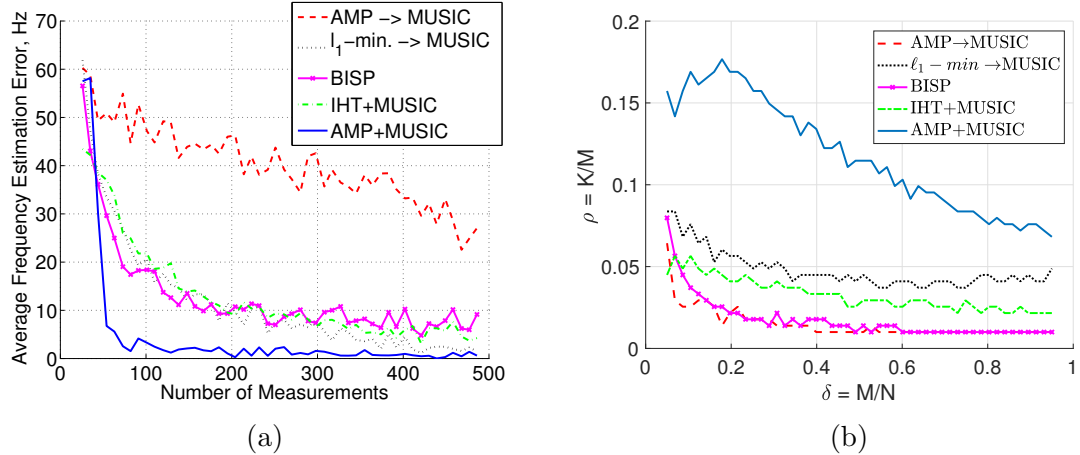


Figure 3.17: Numerical comparison of algorithms for compressive frequency estimation. (a) Average frequency estimation error as a function of the number of measurements M . (b) Phase transition plot for frequency estimation.

from the aspects of phase transition and average frequency estimation error; it also exhibits a similar trend as that observed in Figure 3.15 for signal recovery.

3.3.2.2 Bearing Estimation

In our first experiment, we compare the performance of the AMP+MUSIC algorithm among randomly drawn signals following the signal generation model of Section 3.1.1 and compare AMP+MUSIC with three alternative baselines: (i) $G\ell_1 \rightarrow$ MUSIC, an ℓ_1 -norm based method that exploits the common sparsity pattern observed across different snapshots [22]; (ii) subsampling, i.e., acquisition of $\frac{M}{P}$ antennas with $Q = 100$ snapshots followed by standard DOA estimation, without any CS; (iii) simultaneous recovery of all snapshots using ℓ_1 -norm minimization followed by standard DOA estimation ($\ell_1 \rightarrow$ MUSIC); and (iv) IHT+MUSIC, an analog of AMP+MUSIC based on Iterative Hard Thresholding that does not include the Onsager correction [3]. We evaluate the average bearing estimation error as a function of the number of measurements from the array MP over 100 trials for each value of M . Figure 3.18a illustrates the performance of AMP+MUSIC as a function of the

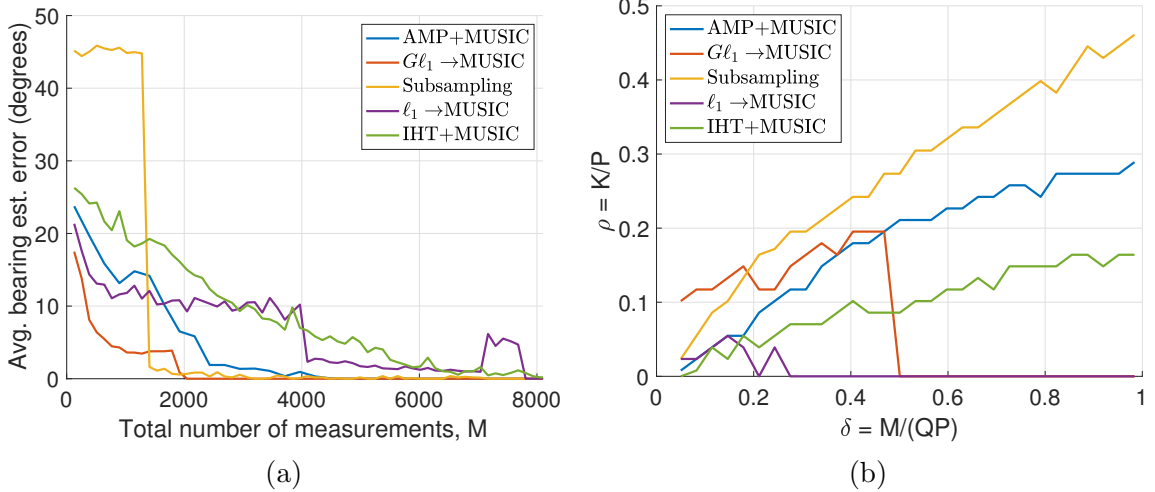


Figure 3.18: Numerical comparison of algorithms for compressive bearing estimation. (a) Average bearing estimation error as a function of the number of measurements for different sparsity levels. (b) Phase transition plot for bearing estimation.

total number of measurements. Comparing AMP+MUSIC with other baselines, we notice that AMP+MUSIC is outperformed only by the $G\ell_1 \rightarrow \text{MUSIC}$ approach for lower values of M ; nonetheless, that method is based on optimization and requires significantly more computation time than our proposed approach.¹

Our second experiment considers the compressive bearing estimation phase transition plots. We define success in this case as having an average bearing estimation error (over the K bearing angles) of up to 5° . For each value of the (δ, ρ) duplet, we execute 50 trials with bearing angles drawn uniformly at random from the range $[-\frac{\pi}{2}, \frac{\pi}{2})$, and amplitudes drawn uniformly at random from $(0, 1)$. Figure 3.18b shows the phase transition plot of AMP+MUSIC, where we see similar results as before: AMP+MUSIC is outperformed only by the $G\ell_1 \rightarrow \text{MUSIC}$ approach for lower values of δ , but requires significantly higher computation time.

¹Note that the computation time for the $G\ell_1 \rightarrow \text{MUSIC}$ algorithm for the distributed setting in [22] is significantly lower due to the fact that the equivalent CS matrix in the distributed setting can be expressed as a Kronecker product.

CHAPTER 4

CONCLUSION AND FUTURE WORK

In this thesis, we proposed *analog denoisers* as a new family of denoisers that allow the D-AMP algorithm to also solve the compressive parameter estimation problem for parameter-sparse signals. These denoisers leverage statistical models to link the signal observations obtained by D-AMP to the parametric representation of interest. As examples of parameter-sparse signals, we considered frequency and bearing-sparse signals, where the time-domain signal is generated through a transformation from a low-dimensional parameter space. We used the Root MUSIC algorithm for an analog denoiser and evaluated the performance requirements for this denoiser while considering signal recovery. Our numerical verifications show that our analog denoisers are appropriate for use within the D-AMP algorithm while considering signal recovery. We then considered the link between the analog denoiser's signal denoising error and its parameter estimation error, and observed that the latter is known to be dependent on the specific parameter values, and that the latter may not always be indicative of the former. Existing results we are aware of on statistical parameter estimation performance guarantees did not bear out in our experiments, and we are still pursuing a search for existing results that provide accurate performance guarantees for relevant statistical parameter estimation algorithms. Our belief is that finding such theoretical guarantees will enable a tractable analysis of the performance of D-AMP with analog denoisers, based on state evolution, for compressive parameter estimation. Nonetheless, our numerical experiment show that the AMP+MUSIC

algorithm proposed here matches or outperforms its existing competitors in accuracy and speed.

In the future, we intend to consider more parameter estimation examples and applications. Moreover, we will

BIBLIOGRAPHY

- [1] Barabell, A. Improving the resolution performance of eigenstructure-based direction-finding algorithms. In *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)* (Boston, MA, Apr. 1983), pp. 336–339.
- [2] Baraniuk, R. G. Compressive sensing. *IEEE Signal Proc. Mag.* 24, 4 (Jul 2007), 118–121.
- [3] Blumensath, T., and Davies, M. E. Iterative hard thresholding for compressed sensing. *Appl. Comput. Harmon. Anal.* 27, 3 (Nov. 2009), 265–274.
- [4] Candès, Emmanuel J. Compressive sampling. In *Int. Congress of Mathematicians* (Madrid, Spain, Aug. 2006), vol. 3, pp. 1433–1452.
- [5] Cevher, V., Duarte, M. F., and Baraniuk, R. G. Localization via spatial sparsity. In *European Signal Process. Conf. (EUSIPCO)* (Lausanne, Switzerland, 2008).
- [6] Chen, Scott Shaobing, David L. Donoho, and Saunders., Michael A. Atomic decomposition by basis pursuit. *SIAM review*, 43.1 (2001), 129–159.
- [7] Daubechies, I., Defrise, M., and Mol, C. De. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Comm. Pure Applied Math* 57, 11 (Nov. 2004), 1413–1457.
- [8] Donoho, D. Denoising by soft thresholding. *IEEE Trans. Inf. Theory* 41, 3 (May 1995), 613–627.
- [9] Donoho, D. L. Compressed sensing. *IEEE Trans. Info. Theory* 52, 4 (Apr. 2006), 1289–1306.

- [10] Donoho, D. L., Johnstone, I., and Montanari, A. Accurate prediction of phase transitions in compressed sensing via a connection to minimax denoising. *IEEE Trans. Inf. Theory* 59, 6 (June 2013), 3396–3433.
- [11] Donoho, D. L., and Tanner, J. Observed universality of phase transitions in high-dimensional geometry, with implications for modern data analysis and signal. *Philosophical Trans. of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 367, 1906 (2009), 4273–4293.
- [12] Donoho, D.L., Maleki, A., and Montanari, A. Message passing algorithms for compressed sensing. *Proc. Nat. Academy Sci.* 106, 45 (Nov. 2009), 18914–18919.
- [13] Duarte, M. F. Localization and bearing estimation via structured sparsity models. In *Proc. IEEE Stat. Signal Process. Workshop (SSP)* (Ann Arbor, MI, Aug. 2012), pp. 333–336.
- [14] Duarte, M. F., and Baraniuk, R. G. Spectral compressive sensing. *Appl. Comput. Harmonic Anal.* 35, 1 (July 2013), 111–129.
- [15] Fang, J., Li, J., Shen, Y., and Li, H. Super-resolution compressed sensing: An iterative reweighted algorithm for joint parameter learning and sparse signal recovery. *IEEE Signal Proc. Letters* 21, 6 (2014), 761–765.
- [16] Fang, J., Wang, F., Shen, Y., Li, H., and Blum, R. Super-resolution compressed sensing for line spectral estimation: An iterative reweighted approach. *IEEE Trans. Signal Proc.* 64, 18 (2016), 4649–4662.
- [17] Feng, C., Valaee, S., and Tan, Z. Multiple target localization using compressive sensing. In *Global Telecommunications Conf. (GLOBECOM)* (Honolulu, HI, Dec. 2009), pp. 1–6.

- [18] Fyhn, K., Dadkhahi, H., and Duarte, M. F. Spectral compressive sensing with polar interpolation. *IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)* (May 2013), 6225–6229.
- [19] Fyhn, K., Duarte, M. F., and Jensen, S. H. Compressive parameter estimation for sparse translation-invariant signals using polar interpolation. *IEEE Trans. Signal Process.* 63, 4 (Feb. 2015), 870–881.
- [20] Fyhn, Karsten, Dadkhahi, Hamid, and Duarte, Marco F. Spectral compressive sensing with polar interpolation. In *IEEE Int. Conf. Acoustics, Speech, and Signal Proc. (ICASSP)* (Vancouver, BC, May 2013), pp. 6225–6229.
- [21] Hamzehei, S., and Duarte, M. F. Compressive parameter estimation via approximate message passing. In *IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)* (Brisbane, Australia, Apr. 2015), pp. 3327–3331.
- [22] Hamzehei, S., and Duarte, M. F. Compressive direction-of-arrival estimation off the grid. in: *50th Asilomar Conf. Signals, Systems, and Computers* (Nov. 2016), 1081–1085.
- [23] Johnson, D., and Dudgeon, D. Array signal processing: Concepts and techniques. *Prentice-Hall Inc.* (1993).
- [24] Kuhn, H. W., and Yaw, B. The hungarian method for the assignment problem. *Naval Res. Logist. Quart.* (1995), 83–97.
- [25] Metzler, C. A., Maleki, A., and Baraniuk, R. G. From denoising to compressed sensing. *IEEE Trans. Info. Theory* 62, 9 (Sep. 2016), 5117–5144.
- [26] Mi, T., Li, S., and Liu, Y. Fast thresholding algorithms with feedbacks for sparse signal recovery. *Appl. Comput. Harmonic Anal.* 37, 1 (Jul 2014), 69–88.

- [27] Mo, D., and Duarte, M. F. Compressive parameter estimation with earth mover's distance via k-median clustering. In *Wavelets and Sparsity XV* (San Diego, CA, Aug. 2013), vol. 8858 of *Proc. SPIE*.
- [28] Munkres, J. Algorithms for the assignment and transportation problems. *J. Soc. Industrial and Applied Math.* 5, 1 (Mar. 1957), 32–38.
- [29] Rao, B. D., and Hari, K. V. S. Performance analysis of Root-MUSIC. *IEEE Trans. Acoust. Speech, and Signal Proc.* 37, 12 (Dec. 1989), 1939–1949.
- [30] Safavi, S., and Khan, U. A. Localization in mobile networks via virtual convex hulls. *IEEE Trans. on Signal and Info. Proc. over Networks* 4, 1 (2017), 188–201.
- [31] Stoica, P., and Moses, R. L. *Spectral Analysis of Signals*. Prentice Hall, 2005.
- [32] Tan, Z., Yang, P., and Nehorai, A. Joint sparse recovery method for compressed sensing with structured dictionary mismatches. *IEEE Trans. Signal Proc.* 62, 19 (2014), 4997–5008.
- [33] Yang, Z., Li, J., Stoica, P., and Xie, L. Sparse methods for direction-of-arrival estimation. *Academic Press Library in Signal Proc.* 7 (2018), 509–581.
- [34] Yang, Z., Xie, L., and Zhang, C. Off-grid direction of arrival estimation using sparse bayesian inference. *IEEE Trans. Signal Proc.* 16, 1 (2013), 38–43.
- [35] Zhu, H., Leus, G., and Giannakis, G. B. Sparsity-cognizant total least-squares for perturbed compressive sampling. *IEEE Trans. Signal Proc.* 59, 5 (2011), 2002–2016.