PERFORMANCE ANALYSIS OF CONVEX DATA DETECTION IN MIMO

Ehsan Abbasi, Fariborz Salehi, Babak Hassibi

Department of Electrical Engineering, Caltech, Pasadena, USA

ABSTRACT

We study the performance of a convex data detection method in large multiple-input multiple-output (MIMO) systems. The goal is to recover an n-dimensional complex signal whose entries are from an arbitrary constellation $\mathcal{D} \subset \mathbb{C}$, using m noisy linear measurements. Since the Maximum Likelihood (ML) estimation involves minimizing a loss function over the discrete set $\mathcal{D}^n$, it becomes computationally intractable for large n. One approach is to relax $\mathcal{D}$ to a convex set and to utilize convex programing to solve the problem and then to map the answer to the closest point in the set $\mathcal{D}$. We assume an i.i.d. complex Gaussian channel matrix and derive *precise* expressions for the symbol error probability of the proposed convex method in the limit of $m, n \to \infty$. Prior work was only able to do so for real valued constellations such as BPSK and PAM. The main contribution of this paper is to extend the results to complex valued constellations. In particular, we use our main theorem to calculate the performance of the complex algorithm for PSK and QAM constellations. In addition, we introduce a closed-form formula for the symbol error probability in the high-SNR regime and determine the minimum number of measurements m required for consistent signal recovery.

1. INTRODUCTION

We consider the problem of recovering a transmit signal, $\mathbf{x}_0 \in \mathcal{D}^n$, form m (noisy) linear observations of the form $\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \mathbf{z}$, where $\mathcal{D} \subset \mathbb{C}$ denotes the discrete transmit constellation, and $\mathbf{z} \in \mathbb{C}^m$ is the noise vector. This problem has a pivotal role in signal detection in multiple-input, multiple-output (MIMO) communication systems [1–3], where $\mathbf{A} \in \mathbb{C}^{m\times n}$, often referred to as the channel state information, is a known matrix. In such settings, m and n correspond to the number of transmit and receive antennas, respectively.

The Maximum Likelihood(ML) estimator is the desirable theoretical solution for this problem. There has been numerous studies to investigate algorithms that can generate exact or approximate solutions for this problem. Due to the combinatorial nature [4] of the problem, exact algorithms (e.g. sphere decoding [5]) are computationally prohibitive, especially in a very large system (e.g. massive MIMO) [6]. Therefore, various heuristics have been proposed and used in practice [7, 8] to approximate the ML solution. Despite tractable computational complexity, the precise performance analysis of such methods are often challenging.

Due to the practical advantages of convex algorithms, one conventional approach to solve this problem is to relax the discrete set $\mathcal{D}$ to a continuous convex set $\mathbf{S}$ and utilize convex programming to search over $\mathbf{S}$ instead of $\mathcal{D}$ [9–11]. The performance of this method for data recovery has been investigated in the works of [3, 12, 13] for the real valued constellations, specifically, BPSK and PAM when the channel matrix is Gaussian. To do so, Thrampoulidis et. al. [12] utilized a framework that they had developed, known as the CGMT framework [14, 15]. The CGMT framework has been successfully applied to analyze the performance in a number of other applications including analysis of regularized M- estimators [14], and PhaseMax in phase retrieval [16–18]. Unfortunately, The CGMT framework can not be readily extended to the complex settings (which indeed is the desirable case in many practical applications).

The major result of this paper is to introduce a new comparison lemma for complex Gaussian processes to study the convex detection problem for complex constellations. In particular, we *precisely* characterize the symbol error rate performance of the convex method, for a general constellation $\mathcal{D}$ and a convex relaxation $\mathbf{S}$. Our theorem also allows us to derive the necessary and sufficient number of antennas, m, required for data recovery in the high-SNR regime which enables us to precisely characterize the phase-transition regions. Through our analysis, we can further observe the relationship between the choice of the convex relaxation with its corresponding phase transition. As an example, we analyze the loss in performance when choosing a relaxation that is easier to implement in a convex program for the case of PSK modulation.

2. PROBLEM SETUP

Notations We gather here the basic notations that are used throughout this paper. We reserve the letter j for the complex unit. For a complex scalar $x \in \mathbb{C}$, x_{Re} and x_{Im} correspond to the real are imaginary parts of x, respectively and $|x| = \sqrt{x_{\text{Re}}^2 + x_{\text{Im}}^2}$. $\mathcal{N}(\mu, \sigma^2)$ denotes real Gaussian distribution with mean μ and variance σ^2. Similarly, $\mathcal{N}_{\mathbb{C}}(\mu, \sigma^2)$ refers to a *complex* Gaussian distribution with real and imaginary parts drawn independently from $\mathcal{N}(\mu_{\text{Re}}, \sigma^2/2)$ and $\mathcal{N}(\mu_{\text{Im}}, \sigma^2/2)$, respectively. $X \sim p_X$ implies that the random variable X has a density p_X. We reserve the letters G and H to denote (scalar) standard normal random variables. Similarly, $H_{\mathbb{C}}$ is reserved to denote a *complex* $\mathcal{N}_{\mathbb{C}}(0, 2)$ random variable. The bold lower letters are reserved for vectors and for a vector $\mathbf{v}$, $\mathbf{v}_i$ denotes its i^{th} entry. Finally, for a *convex* set $\mathbf{S} \subset \mathbb{C}$, the projection and distance functions with respect to $\mathbf{S}$ are defined as

$$\mathcal{P}_{\mathbf{S}}(\mathbf{x}) := \arg\min_{\mathbf{y}\in\mathbf{S}} \|\mathbf{x} - \mathbf{y}\|$$
$$\mathcal{D}_{\mathbf{S}}(\mathbf{x}) := \min_{\mathbf{y}\in\mathbf{S}} \|\mathbf{x} - \mathbf{y}\|. \tag{1}$$

Setup Our goal is to recover an n-dimensional vector $\mathbf{x}_0 \in \mathbb{C}^n$ where the entries of $\mathbf{x}_0$ are independenty drawn from the discrete set $\mathcal{D} \subset \mathbb{C}$ with distribution $\mathbf{x}_{0,i} \sim p_X$. The set $\mathcal{D}$ defines the modulation used for data transmission (e.g. QAM, PSK, etc.). For this purpose, we are given the noisy multiple-input multiple-output (MIMO) relation of the form

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{z} \in \mathbb{C}^m, \tag{2}$$

where $\mathbf{A} \in \mathbb{C}^{m\times n}$ is the known MIMO channel matrix with i.i.d. entries drawn from $\mathcal{N}_{\mathbb{C}}(0, \frac{1}{n})$ and $\mathbf{z} \in \mathbb{C}^m$ is the unknown noise vector with i.i.d. random complex Gaussian $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ entries.

978-1-5386-4658-8/18/$31.00 ©2019 IEEE

Estimator The ML estimator of $\mathbf{x}_0$ in this scenario is

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathcal{D}^n} \frac{1}{2m}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2. \quad (3)$$

Since solving (3) is computationally intractable, a variety of heuristic methods, such as zero-forcing, MMSE, decision-feedback, have been proposed. In this paper, we make use of convex programming to estimate $\mathbf{x}_0$. In the first step, we relax $\mathcal{D}$ to a convex set $\mathbf{S}$ and minimize the objective function of (3) over this relaxed convex set,

$$\tilde{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbf{S}^n} \frac{1}{2m}\|\mathbf{y}-\mathbf{A}\mathbf{x}\|^2. \quad (4)$$

Next, we map each entry of $\tilde{\mathbf{x}}$ to the closest point in $\mathcal{D}$ to build our final estimation of $\mathbf{x}_0$,

$$\hat{\mathbf{x}}_i = \arg\min_{x\in\mathcal{D}} |x-\tilde{\mathbf{x}}_i|, \quad i=1,\dots,n. \quad (5)$$

We refer to this method as the Convex Decoder Algorithm (CDA). In this paper, we will precisely analyze the performance of the CDA as a function of the problem parameters such as σ, m/n, $\mathcal{D}$ and $\mathbf{S}$. Note that the performance of CDA depends on the constellation $\mathcal{D}$ and the way we relax it to the convex set $\mathbf{S}$. Later in Section 3.1, we observe the impact of choosing two different relaxations on the performance of CDA and its phase-transition regions with the help of our main theorem.

Symbol error probability We characterize the performance of CDA in terms of the symbol error probability, defined as the expected value of the Symbol Error Rate (SER) where,

$$SER := \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}_{\hat{\mathbf{x}}_i\neq \mathbf{x}_{0,i}},$$

$$P_e := \mathbb{E}[SER] = \frac{1}{n}\sum_{i=1}^{n}\mathbb{P}\left(\hat{\mathbf{x}}_i \neq (\mathbf{x}_0)_i\right). \quad (6)$$

Here $\hat{\mathbf{x}}$ is the output of CDA in (5), $\mathbb{1}_{\mathcal{E}}$ is the indicator of the event $\mathcal{E}$ and the probability $\mathbb{P}(\cdot)$ is over the randomness of $\mathbf{A}$, $\mathbf{z}$ and $\mathbf{x}_0$. We introduce the notation $\mathbf{S}_x$ for $x\in\mathcal{D}$, as the set of all points in $\mathbf{S}$ that will be mapped to x in (5). Equivalently,

$$\mathbf{S}_x := \{x'\in\mathbf{S} : \forall \mathbf{y}\in\mathcal{D},\ |x'-x| < |x'-y|\}. \quad (7)$$

This notation helps us interpret our main theorem more clearly. Using this notation, we can rewrite the symbol error probability defined in (6) as

$$P_e = \frac{1}{n}\sum_{i=1}^{n}\mathbb{P}\left(\tilde{\mathbf{x}}_i \notin \mathbf{S}_{(\mathbf{x}_0)_i}\right), \quad (8)$$

where $\tilde{\mathbf{x}}$ is the minimizer of (4).

Assumptions We impose two mild assumptions on the problem. First, we assume that the entries of $\mathbf{x}_0$ are i.i.d. random variables with $\mathbf{x}_{0,i}\sim p_{\mathbf{x}}$ and also $\mathbb{P}_{\mathbf{x}}(r_1+jr_2)=\mathbb{P}_{\mathbf{x}}(r_2+jr_1),\ \forall r_1,r_2\in\mathbb{R}$. Second, we want the convex set to be *symmetric* in the sense that if $(r_1+jr_2)\in\mathbf{S}$, then also $(r_2+jr_1)\in\mathbf{S}$.

2.1. Modulations

Using our main theorem, we can precisely analyze the SER of the CDA in terms of the SER for a constellation $\mathcal{D}$ which we relax to an arbitrary convex set $\mathbf{S}$. For a better understanding of the theorem and to show how to apply it to different schemes, we will work with two conventional modulations; Phase-Shift Keying (PSK) and Quadrature Amplitude Modulation (QAM).

N-PSK Constellation: In the N-PSK constellation, each entry of $\mathbf{x}_0$ is randomly drawn from $\mathcal{D}=\left\{e^{\frac{j2\pi}{N}i} : i=0,\dots,N-1\right\}$. The entries of $\mathcal{D}$ are distributed over the unit circle in the complex space and therefore the Signal to Noise Ratio (SNR) will be $1/\sigma^2$. Next, we need an appropriate convex relaxation of $\mathcal{D}$ for CDA. We suggest two candidates for this purpose and compare their performances later in Section 3.1. In one which we will refer to as the Circular Relaxation (CR), we choose the set $\mathbf{S}^{(\text{CR})}=\{x\in\mathbb{C}:|x|\le 1\}$ as the convex set in (4). The simple structure of $\mathbf{S}^{(\text{CR})}$ makes its implementation easier in the convex program (4). In another scenario, we consider the convex hull of $\mathcal{D}$ as the relaxed set $\mathbf{S}$ and refer to it as Convex Hull Relaxation (CHR). Thus, $\mathbf{S}^{(\text{CHR})}=\text{Conv}(\mathcal{D})$ will be used in (4) which might be harder to implement compared to $\mathbf{S}^{(\text{CR})}$. But we will show that since (CHR) is a tighter relaxation, its corresponding CDA performs better in terms of SER.

N^2-QAM Constellation: We also briefly talk about the N^2-QAM modulation where

$$\mathcal{D}=\left\{(\alpha+j\beta)-\frac{N-1}{2}(1+j) : \alpha,\beta\in\{0,\dots,N-1\}\right\}.$$

Under this constellation the SNR will be $\frac{N(N^2-1)}{6\sigma^2}$. The relaxation that is often used for this modulation is known as the Box Relaxation (BR) [12] which is

$$\mathbf{S}=\left\{(x+jy)\in\mathbb{C} : |x|\le\frac{N-1}{2}, |y|\le\frac{N-1}{2}\right\} \quad (9)$$

Using our main theorem, we can calculate the SER of CDA under box relaxation and rederive the results of [3, 12].

3. MAIN RESULT

Our main result explicitly characterizes the limiting behavior of the symbol error rate of the convex decoder algorithm, under the high dimensional regime where $n,m\to\infty$ with a constant ratio $\delta := m/n$.

Theorem 3.1. *(SER analysis of CDA) Let SER denote the symbol error rate of the Convex Decoder Algorithm (CDA), for random signal $\mathbf{x}_0\in\mathcal{D}$ with entries drawn independently from the distribution p_X. Let $\mathbf{S}$ be a convex relaxation of $\mathcal{D}$ and $\mathbf{S}$ and p_X satisfy the assumptions in Section 2. Fix SNR and $\delta=m/n$ and consider the optimization*

$$\min_{\tau>0} \frac{\delta-1}{2\tau\delta}+\frac{\sigma^2\tau}{4}+\frac{\tau}{4}\mathbb{E}[\mathcal{D}^2_{\mathbf{S}}(X+\frac{H_{\mathbb{C}}}{\tau\sqrt{\delta}})]. \quad (10)$$

If (10) has a unique answer τ^*, then in the limit of $m,n\to\infty$

$$\lim_{m,n\to\infty} P_e = \mathbb{P}\left(\mathcal{P}_{\mathbf{S}}(X+\frac{H_{\mathbb{C}}}{\tau^*\sqrt{\delta}})\notin\mathbf{S}_X\right). \quad (11)$$

The expected value and probability in (10) and (11) are over $X\sim p_X$ and $H_{\mathbb{C}}\sim\mathcal{N}_{\mathbb{C}}(0,2)$, respectively.

Theorem 3.1 provides a formula to calculate the SER of the convex decoder, under a general constellation in the high dimensional regime.

Remark 1. (Computing τ^*) The objective function in (10) is convex and only involves one scalar variable. Thus, τ^* can, in principle, be efficiently numerically computed. It can be shown that τ^* is the minimizer of (10) if and only if it is the answer to the corresponding first-order optimality condition,

$$\frac{1}{\tau^{*2}}=\frac{1}{2}\left(\sigma^2+\mathbb{E}\left[\left|X-\mathcal{P}_{\mathbf{S}}(X+\frac{H_{\mathbb{C}}}{\tau^*\sqrt{\delta}})\right|^2\right]\right). \quad (12)$$

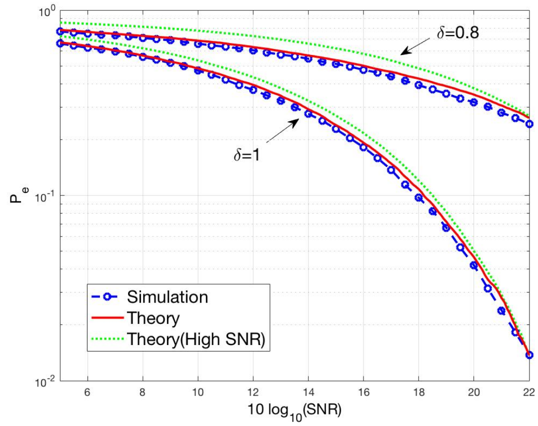

Fig. 1: SER Performance of the Circular Relaxation (CR) for 16-PSK: P_e as a function of SNR for the two cases where $\delta = .8$ and $\delta = 1$. The theoretical prediction follows from Theorem 3.1 and the high-SNR analysis comes from Section 3.1. For the simulation, we used signals of size $n = 128$ with each entry chosen randomly uniform from the set $\mathbf{S}_{PSK} = \left\{ e^{\frac{j\pi}{8} i} : i = 0, \dots, 15 \right\}$. The data are averages over 30 independent realizations of the channel matrix and the noise vector.

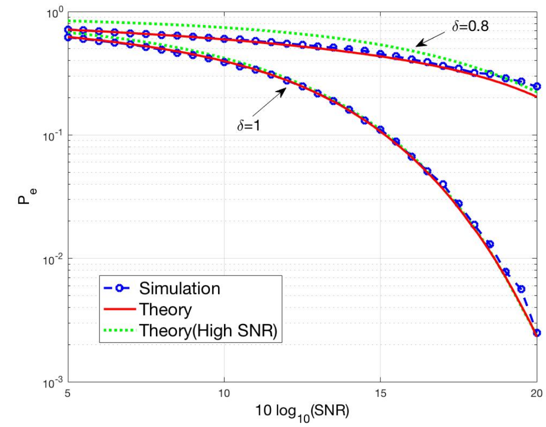

Fig. 2: SER Performance of the Box Relaxation for 16-QAM: P_e as a function of SNR for the two cases where $\delta = .8$ and $\delta = 1$. The theoretical prediction follows from Theorem 3.1 and the high-SNR analysis comes from Section 3.2. For the simulation, we used signals of size $n = 128$ with each entry chosen uniformly at random in the set $\mathbf{S}_{QAM} = \{\pm 1, \pm 3\}^2$. The data are averages over 30 independent realizations of the channel matrix and the noise vector.

Although, this does not provide us with a closed form formula to calculate τ^*, in all our simulations a fixed-point iterative method converges to τ^* over a handful of iterations. It can be also shown that (12) has a unique solution if $\delta > \delta^*$ for some $\delta^* \in (0, 1)$ which depends on $\mathbf{S}$.

Next, we apply Theorem 3.1 to the N-PSK and N^2-QAM modulations introduced in Section 2.1 to calculate their corresponding symbol error probabilities and phase-transition thresholds in the high-SNR regime. Figures 1 and 2 verify the accuracy of the prediction of Theorem 3.1 for 16-PSK and 16-QAM modulations, respectively. Note that although the theorem requires $n \to \infty$, the prediction is already accurate for $n = 128$. In these figures, we have also plotted the high-SNR expressions for SER that we derive in the next Section for both modulations. Interestingly, we observe that this high-SNR expression gives us a good enough approximation of the exact value of SER, even for small practical values of SNR.

3.1. N-PSK Constellation

Under the N-PSK setup, the set $\mathcal{D}$ is defined in Section 2.1. We investigate the error performance of the convex decoder algorithm for two different convex relaxations for the set $\mathcal{D}$; Circular Relaxation (CR) and Convex-Hull Relaxation (CHR). The effect of using different relaxations shows up in the projection function in equations (12) and (11). Define $\mathbf{S}^{(\mathrm{SR})} = \{c \in \mathbb{C} : |c| \leq 1\}$ as the circular relaxation of $\mathcal{D}$. The projection function on this set has the following form,

$$\mathcal{P}_{\mathbf{S}^{(\mathrm{CR})}}(\mathbf{x}) = \begin{cases} \mathbf{x} & \text{if } |\mathbf{x}| \leq 1 \\ \mathbf{x}/|\mathbf{x}| & \text{otherwise.} \end{cases} \tag{13}$$

Therefore, τ^* can be efficiently calculated using a fixed-point iterative method to solve (12). Furthermore, due to the symmetric nature of the N-PSK constellation, the probability of error for each symbol in $\mathcal{D}$ can be derived in the following closed form,

$$P_e = \mathbb{P}\left(|G| > \tan(\frac{\pi}{N})(H + \tau^* \sqrt{\delta})\right), \tag{14}$$

where G and H are i.i.d. $\mathcal{N}(0, 1)$.

High-SNR Analysis Let $\mathbf{S}^{(\mathrm{CR})}$ and $\mathbf{S}^{(\mathrm{CHR})}$ denote the circular relaxation and convex-hull relaxation of the set $\mathcal{D}$. It can be shown that for SNR $\gg 1$, τ^* grows large proportional to $\sqrt{\mathrm{SNR}}$. As a consequence, the last term in (10) can be approximated by $\frac{1}{8\tau\delta}$ and $\frac{N+4}{8N\tau\delta}$ for the cases of (CR) and (CHR), respectively. This results in $\tau^* = \sqrt{\frac{2\mathrm{SNR}(\delta - 3/4)}{\delta}}$ for (CR) and $\tau^* = \sqrt{\frac{2\mathrm{SNR}(\delta - 3/4 + 1/N)}{\delta}}$ for (CHR). Putting these values for τ^* in (14) yields their corresponding high-SNR symbol error probabilities,

$$P_e^{(\mathrm{CR})} = \mathbb{P}\left(|G| > \tan(\frac{\pi}{N})(H + \sqrt{2\mathrm{SNR} \cdot (\delta - 3/4)})\right), \tag{15}$$

$$P_e^{(\mathrm{CHR})} = \mathbb{P}\left(|G| > \tan(\frac{\pi}{N})(H + \sqrt{2\mathrm{SNR} \cdot (\delta - 3/4 + 1/N)})\right). \tag{16}$$

The difference between phase-transitions of these two cases can be observed from equations (15) and (16). While for (CR) we need $\delta > 3/4$ for consistent data recovery, this threshold changes to $\delta > (3/4 - 1/N)$ for (CHR). This essentially means that n/N additional MIMO receivers is required at the expense of having a simpler convex set. This verifies the fact that while the optimization over $\mathbf{S}^{(\mathrm{CR})}$ might be done faster over the Circular Relaxation due to its simple structure, we need more measurements (or higher SNR) to get the same performance for (CR) compared to (CHR). In other words, the performance of (CR) is $10\log_{10}(\frac{\delta - 3/4 + 1/n}{\delta - 3/4})$ off that of (CHR).

Comparison to the matched filter bound. The matched filter is the ideal impractical case where we assume to have the first $n-1$ entries of $\mathbf{x}_0$ and we want to recover the last entry. We compare the symbol error probability of this scenario, referred to as the Matched Filter Bound (MFB), with the P_e of the convex decoder that can be derived from Theorem 3.1. The matched filter bound corresponds to the probability of error in detecting $X \in \mathcal{D}$ from $\tilde{\mathbf{y}} = X\mathbf{a} + \mathbf{z}$, where $\mathbf{a} \in \mathbb{C}^n$ with Gaussian entries drawn from $\mathcal{N}_{\mathbb{C}}(0, \frac{1}{\sqrt{n}})$, and $\mathbf{z}$ is the noise vector with entries $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$. Then, the probability of error of the ML estimator of X in N-PSK will be

$$\mathbb{P}\left(|G| > \tan(\frac{\pi}{N})(H + \sqrt{2\text{SNR}\cdot\delta})\right) . \quad (17)$$

Comparison of (15) with (17) shows that in the high-SNR regime the performance of (CR) is $10\log_{10}(\frac{\delta}{\delta-3/4})$dB off from the (MFB). In particular, in the square case ($\delta = 1$), where the number of receive and transmit antennas are the same, the (CR) is 6dB off the (MFB). Besides, as $\delta \to \infty$ (meaning that the number of antennas grows largecompared to users), the performance of (CR) and (CHR) approaches (MFB).

3.2. N^2-QAM Constellation

In N^2-QAM, each entry of $\mathbf{x}_0$ is randomly chosen from the set $\mathcal{D}$ defined in Section 2.1 with distribution p_X. The conventional relaxation for this constellation is the Box Relaxation (BR) [9–11] defined in (9). Similar to the previous section, In order to use Theorem 3.1, we need to form the projection function to $\mathbf{S}$ in (1) which is straightforward for a box set. Once τ^* is obtained using equation (12) (or recruiting other methods to solve (10)), we shall use (11) to calculate P_e of N^2-QAM constellation. Here, unlike the N-PSK case, the probability of error in the recovery is not the same for different symbols in $\mathcal{D}$.

Using the same set of arguments in Section 3.1, it can be shown that in the high-SNR regime, the last term in the objective function of (10) approaches $\frac{1}{2\tau\delta N}$. Therefore the answer to the minimization problem will be $\tau^* = \sqrt{\frac{2\text{SNR}(\delta-(N-1)/N)}{\delta}}$. This implies that $\delta^\star = \frac{N-1}{N}$ is the recovery threshold for the Box Relaxation of the set $\mathcal{D}$. It can also be shown that for $\delta > \delta^\star$, the problem (10) is strictly convex and therefore has a unique solution. This is consistent with the result of [3] which proves the same phase-transition region for the Box Relaxation.

4. PROOF OUTLINE

In this part we introduce the main ideas used in the proof of Theorem 3.1. The goal is to analyze the performance of the following optimization problem:

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\mathbf{S}^n} \frac{1}{2m}||\mathbf{y} - \mathbf{A}\mathbf{x}||^2 \quad (18)$$

We rewrite (18) by changing variable to vector $\mathbf{w} = \mathbf{x} - \mathbf{x}_0$

$$\hat{\mathbf{w}} = \arg\min_{\mathbf{w}\in\mathbf{S}^n-\mathbf{x}_0} \frac{1}{2m}||\mathbf{z} - \mathbf{A}\mathbf{w}|| \quad (19)$$

Now let $\tilde{\mathbf{A}} = [\mathbf{A}_R, \mathbf{A}_I; -\mathbf{A}_I, \mathbf{A}_R] \in \mathbb{R}^{2m\times 2n}$ and $\mathbf{z}' = [\mathbf{z}_R; \mathbf{z}_I] \in \mathbb{R}^{2m}$, where $\mathbf{A}_R$ and $\mathbf{z}_R$ ($\mathbf{A}_I$ and $\mathbf{z}_I$) are the real (imaginary) parts of $\mathbf{A}$ and $\mathbf{z}$, respectively. Now (19) can be written as

$$\mathbf{w}^\star = \arg \min_{\substack{\mathbf{w}\in\mathbb{R}^{2n} \\ \mathbf{w}_i + j\mathbf{w}_{n+i}\in\mathbf{S}-(\mathbf{x}_0)_i}} \frac{1}{4m}||\mathbf{z}' - \frac{1}{\sqrt{2n}}\tilde{\mathbf{A}}\cdot\mathbf{w}||^2 . \quad (20)$$

This optimization is difficult to analyze and current methods for asymptotic analysis of such optimizations fail here, because of the dependence between the entries of $\tilde{\mathbf{A}}$. The main step of our proof is to show that in the asymptotic regime when $m, n =\to \infty$ with $m/n = \delta$, the SER in the optimization (19) converges to the one in the following.

$$\mathbf{w}^\star = \arg \min_{\substack{\mathbf{w}\in\mathbb{R}^{2n} \\ \mathbf{w}_i + j\mathbf{w}_{n+i}\in\mathbf{S}-(\mathbf{x}_0)_i}} \frac{1}{4m}||\mathbf{z}' - \frac{1}{\sqrt{2n}}\mathbf{B}\cdot\mathbf{w}||^2 . \quad (21)$$

Here, $\mathbf{z}' \in \mathbb{R}^{2m}$ is a vector with i.i.d. $\mathcal{N}(0, \sigma^2/2)$ entries and $\mathbf{B} \in \mathbb{R}^{(2m)\times(2n)}$ is a matrix whose entries are independently drawn from $\mathcal{N}(0, 1)$. To do so, we first show this in the case that both the objective functions have an extra strongly convex term $\epsilon\|\mathbf{w}\|^2/2$. Under this scenario, we can utilize the Lindeberg method as in [19]. The idea is to replace the rows of $\tilde{\mathbf{A}}$ to $\mathbf{B}$ in m steps. In each step, we replace the rows i and $n+i$ in the $\tilde{\mathbf{A}}$ (that are independent from the rest of $\tilde{\mathbf{A}}$) with the the rows i and $i+n$ in the $\mathbf{B}$. We can show that each step changes the SER on the order of $\mathcal{O}(n^{5/4})$. So as $n \to \infty$, the SER doesn't change. Next, we use the RIP condition for Gaussian matrices to show that removing the extra $\epsilon\|\mathbf{w}\|^2/2$ term in the optimization does not affect the SER for small enough ϵ (See Sections 3.1 and 3.3.2 in the appendix of [20] for more details regarding these two steps). Then, we just need to analyze performance of (21) instead of (19). For the rest of the proof, we apply the CGMT framework and the same tools as in [14](Section 5.3). The idea is to rewrite (21) as the following min-max problem,

$$\min_{\mathbf{w}_i + j\mathbf{w}_{n+i}\in\mathbf{S}-(\mathbf{x}_0)_i} \max_{\mathbf{u}\in\mathbb{R}^{2m}} \frac{1}{\sqrt{2m}}\mathbf{u}^t\mathbf{z}' - \frac{1}{2\sqrt{mn}}\mathbf{u}^t\mathbf{B}\mathbf{w} - \frac{1}{2}||\mathbf{u}||^2 , \quad (22)$$

This enables us to apply the CGMT which associates with (22), the following simplified optimization whose analysis provides us with the desired properties of the initial optimization.

$$\min_{\mathbf{w}_i + j\mathbf{w}_{n+i}\in\mathbf{S}-(\mathbf{x}_0)_i} \max_{\mathbf{u}} \frac{1}{\sqrt{2m}}\mathbf{u}^t\mathbf{z}' - \frac{1}{2}||\mathbf{u}||^2 + \frac{1}{2\sqrt{mn}}(\mathbf{g}^t\mathbf{u}||\mathbf{w}|| + \mathbf{h}^t\mathbf{w}||\mathbf{u}||) ,$$

where $\mathbf{g} \in \mathbb{R}^{2m}$ and $\mathbf{h} \in \mathbb{R}^{2n}$ have i.i.d. standard Gaussian entries. It can be shown that the optimization over $\mathbf{u}$ results in

$$\min_{\mathbf{w}_i + j\mathbf{w}_{n+i}\in\mathbf{S}-(\mathbf{x}_0)_i} \frac{1}{\sqrt{2m}}||\mathbf{z}' + \frac{||\mathbf{w}||}{\sqrt{2n}}\mathbf{g}|| + \frac{1}{2\sqrt{mn}}\mathbf{h}^t\mathbf{w} . \quad (23)$$

Using $\sqrt{x} = \min_{\tau>0} \frac{1}{2\tau} + \frac{\tau x}{2}$, optimization (23) can be written as

$$\min_{\tau>0} \frac{1}{2\tau} + \frac{\tau\|\mathbf{z}'\|^2}{4m} + \min_{\mathbf{w}_i + j\mathbf{w}_{n+i}\in\mathbf{S}-(\mathbf{x}_0)_i} \frac{\tau||\mathbf{w}||^2\|\mathbf{g}\|^2}{8nm} + \frac{1}{2\sqrt{mn}}\mathbf{h}^t\mathbf{w}. \quad (24)$$

Using dimension reduction techniques, we can show that from the following deterministic optimization, we can tightly infer the properties of (24).

$$\min_{\tau>0} \frac{1}{2\tau} + \frac{\tau\sigma^2}{4} + \min_{\mathbf{w}_i + j\mathbf{w}_{n+i}\in\mathbf{S}-(\mathbf{x}_0)_i} \frac{\tau||\mathbf{w}||^2}{4n} + \frac{1}{2\sqrt{mn}}\mathbf{h}^t\mathbf{w}. \quad (25)$$

A completion of squares in the minimization over $\mathbf{w}$, the weak law of large numbers and convex techniques (See Section A.3 in [21] to see how WLLN can be applied here) results in the final deterministic optimization

$$\min_{\tau>0} \frac{\delta-1}{2\tau\delta} + \frac{\sigma^2\tau}{4} + \frac{\tau}{4}\mathbb{E}[\mathcal{D}^2_{\mathbf{S}}(X + \frac{H_c}{\tau\sqrt{\delta}})]. \quad (26)$$

Besides, the optimal $\mathbf{w}$ can be obtained by putting the optimizer of (26) in the minimization over $\mathbf{w}$ in the last term of the (25). Similar to the proof of [12](section 3), SER of $\mathbf{w}^*$ derived here is equal to the one from (19) which is

$$P_e \to \mathbb{P}\left(\mathcal{P}_{\mathbf{S}}(X + \frac{H_{\mathbb{C}}}{\tau^*\sqrt{\delta}}) \notin \mathbf{S}_X\right) \quad (27)$$

5. REFERENCES

[1] Hien Quoc Ngo, Erik G Larsson, and Thomas L Marzetta, "Energy and spectral efficiency of very large multiuser mimo systems," *Communications, IEEE Transactions on*, vol. 61, no. 4, pp. 1436–1449, 2013.

[2] Charles Jeon, Ramina Ghods, Arian Maleki, and Christoph Studer, "Optimality of large mimo detection via approximate message passing," in *Information Theory (ISIT), 2015 IEEE International Symposium on*. IEEE, 2015, pp. 1227–1231.

[3] Charles Jeon, Arian Maleki, and Christoph Studer, "On the performance of mismatched data detection in large mimo systems," in *Information Theory (ISIT), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 180–184.

[4] Sergio Verdu, *Multiuser detection*, Cambridge university press, 1998.

[5] Babak Hassibi and Haris Vikalo, "On the sphere-decoding algorithm i. expected complexity," *Signal Processing, IEEE Transactions on*, vol. 53, no. 8, pp. 2806–2818, 2005.

[6] Fredrik Rusek, Daniel Persson, Buon Kiong Lau, Erik G Larsson, Thomas L Marzetta, Ove Edfors, and Fredrik Tufvesson, "Scaling up mimo: Opportunities and challenges with very large arrays," *IEEE signal processing magazine*, vol. 30, no. 1, pp. 40–60, 2013.

[7] Martin Grötschel, László Lovász, and Alexander Schrijver, *Geometric algorithms and combinatorial optimization*, vol. 2, Springer Science & Business Media, 2012.

[8] Gerard J Foschini, "Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas," *Bell labs technical journal*, vol. 1, no. 2, pp. 41–59, 1996.

[9] Wing-Kin Ma, Timothy N Davidson, Kon Max Wong, Zhi-Quan Luo, and Pak-Chung Ching, "Quasi-maximum-likelihood multiuser detection using semi-definite relaxation with application to synchronous cdma," *IEEE transactions on signal processing*, vol. 50, no. 4, pp. 912–922, 2002.

[10] Aylin Yener, Roy D Yates, and Sennur Ulukus, "Cdma multiuser detection: A nonlinear programming approach," *IEEE Transactions on Communications*, vol. 50, no. 6, pp. 1016–1024, 2002.

[11] Peng Hui Tan, Lars K Rasmussen, and Teng J Lim, "Constrained maximum-likelihood detection in cdma," *IEEE Transactions on Communications*, vol. 49, no. 1, pp. 142–153, 2001.

[12] Christos Thrampoulidis, Ehsan Abbasi, Weiyu Xu, and Babak Hassibi, "Ber analysis of the box relaxation for bpsk signal recovery," in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*. IEEE, 2016, pp. 3776–3780.

[13] Ismail Ben Atitallah, Christos Thrampoulidis, Abla Kammoun, Tareq Y Al-Naffouri, Babak Hassibi, and Mohamed-Slim Alouini, "Ber analysis of regularized least squares for bpsk recovery," in *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*. IEEE, 2017, pp. 4262–4266.

[14] Christos Thrampoulidis, Ehsan Abbasi, and Babak Hassibi, "Precise error analysis of regularized m-estimators in high-dimensions," *arXiv preprint arXiv:1601.06233*, 2016.

[15] Ehsan Abbasi, Christos Thrampoulidis, and Babak Hassibi, "General performance metrics for the lasso," in *Information Theory Workshop (ITW), 2016 IEEE*. IEEE, 2016, pp. 181–185.

[16] Fariborz Salehi, Ehsan Abbasi, and Babak Hassibi, "A precise analysis of phasemax in phase retrieval," in *2018 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2018, pp. 976–980.

[17] Oussama Dhifallah, Christos Thrampoulidis, and Yue M Lu, "Phase retrieval via linear programming: Fundamental limits and algorithmic improvements," in *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. IEEE, 2017, pp. 1071–1077.

[18] Fariborz Salehi, Ehsan Abbasi, and Babak Hassibi, "Learning without the phase: Regularized phasemax achieves optimal sample complexity," in *Advances in Neural Information Processing Systems*, 2018, pp. 8655–8666.

[19] Samet Oymak and Joel A Tropp, "Universality laws for randomized dimension reduction, with applications," *Information and Inference: A Journal of the IMA*, vol. 7, no. 3, pp. 337–446, 2017.

[20] Ashkan Panahi and Babak Hassibi, "A universal analysis of large-scale regularized least squares solutions," in *Advances in Neural Information Processing Systems*, 2017, pp. 3381–3390.

[21] Christos Thrampoulidis, Ehsan Abbasi, and Babak Hassibi, "Precise error analysis of regularized m-estimators in high-dimensions," *IEEE Transactions on Information Theory*, 2018.