

# Box-Relaxation for BPSK Recovery in Massive MIMO: A Precise Analysis under Correlated Channels

Ayed M. Alrashdi, Housseem Sifaou, Abla Kammoun, Mohamed-Slim Alouini and Tareq Y. Al-Naffouri  
 King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia  
 Emails: {ayed.alrashdi, housseem.sifaou, abla.kammoun, slim.alouini, tareq.alnaffouri}@kaust.edu.sa

**Abstract**—In this paper, we consider the problem of recovering a binary phase shift keying (BPSK) modulated signal in a massive multiple-input-multiple-output (MIMO) system. The recovery process is done using the box-relaxation method, in which the discrete set  $\{\pm 1\}^n$  is relaxed to the convex set  $[-1, +1]^n$  and solved by a convex optimization program followed by hard thresholding. We assume that the system has a Gaussian channel matrix with one sided left correlation. The entries of the noise vector are assumed to be independent and identically distributed (iid) zero-mean Gaussian. In this work, we precisely characterize the mean squared error (MSE) and the bit error rate (BER) of the box-relaxation decoder in the asymptotic regime where both dimensions grow simultaneously large at a fixed ratio. Numerical simulations validate the theoretical expressions derived in this paper.

**Index Terms**—least squares, BER, MSE, box relaxation, asymptotic performance, Gaussian processes

## I. INTRODUCTION

The problem of recovering an unknown vector of symbols from a noisy linear measurements has many applications in wireless communications and signal processing. Of special interest is the setting where the symbols belong to a finite discrete constellation such as binary phase shift keying (BPSK) signals. This problem arises in many applications such as Massive multiple-input-multiple-output (MIMO) decoding in wireless communication systems [1]–[5], global navigation satellite system (GNSS) positioning [6] and many other problems.

In modern Massive MIMO systems, exact algorithms such as maximum-likelihood (ML) and sphere decoding, become computationally prohibitive as the problem dimensions grow. To obtain a reasonable computational complexity, one popular heuristic that is used is the so called box-relaxation optimization (BRO) decoder [7], [8], which is a natural convex relaxation of the optimum ML decoder. In this decoder, the discrete set  $\{\pm 1\}^n$  is relaxed to the convex set  $[-1, +1]^n$ , and now the signal can be recovered via efficient convex optimization followed by hard thresholding. The BRO decoder has been shown to outperform conventional decoders such as zero-forcing (ZF) decoder as we will see later on.

This work was supported by the King Abdullah University of Science and Technology (KAUST)’s Office of Sponsored Research under Award OSR-2016-KKI-2899.

The performance of the box-relaxation decoder in the case of independent and identically distributed (iid) channel matrices has been recently analyzed in [9]–[11]. However, in many practical situations in wireless communications, it is important to take into account the effect of correlation [12], [13]. In this work, we characterize the asymptotic performance of the BRO decoder in the presence of channel correlation using BPSK signals for simplicity. In particular, we characterize the asymptotic mean squared error (MSE) and bit error rate (BER) of the BRO using the recently developed convex Gaussian min-max theorem (CGMT). The CGMT framework has been used to characterize the error performance of many algorithms under the iid assumption on the entries of the channel matrix [11], [14]–[18]. To the best of our knowledge, this work is the first one to analyze the asymptotic performance of the BRO under correlated channel matrices.

### A. Problem Formulation

We consider a massive MIMO system with  $n$  transmitters and  $m$  receivers. The received signal  $\mathbf{y} \in \mathbb{R}^m$  is modeled as

$$\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \mathbf{z},$$

where  $\mathbf{x}_0 \in \{\pm 1\}^n$  is the unknown transmitted BPSK signal (to be recovered),  $\mathbf{A} \in \mathbb{R}^{m \times n}$  is the MIMO channel, and  $\mathbf{z} \in \mathbb{R}^m$  is the noise vector that has iid entries  $\mathcal{N}(0, \sigma^2)$ . The MIMO channel matrix can be modeled as [4], [19]

$$\mathbf{A} = \mathbf{\Sigma}^{\frac{1}{2}} \mathbf{H},$$

where  $\mathbf{\Sigma} \in \mathbb{R}^{m \times m}$  is known Hermitian nonnegative left correlation matrix, satisfying  $\frac{1}{m} \text{tr}(\mathbf{\Sigma}) = \mathcal{O}(1)$ , while  $\mathbf{H} \in \mathbb{R}^{m \times n}$  is a Gaussian matrix with iid entries  $\mathcal{N}(0, \frac{1}{n})$ . In this case, the signal-to-noise ratio (SNR) becomes  $\text{SNR} = \frac{1}{\sigma^2}$ .

The heuristic we consider in this paper is the box-relaxation optimization (BRO) decoder [7], [8]. It consists of two steps. The first one involves solving a convex quadratic program with linear constraints. The output of the optimization is hard-thresholded in the second step to produce the final binary estimate. Formally, the algorithm outputs an estimate  $\mathbf{x}^*$  of  $\mathbf{x}_0$  given as

$$\hat{\mathbf{x}} := \arg \min_{-1 \leq x \leq 1} \frac{1}{n} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|^2, \quad (1a)$$

$$\mathbf{x}^* := \text{sign}(\hat{\mathbf{x}}), \quad (1b)$$

where  $\|\cdot\|$  denotes the  $\ell_2$ -norm of a vector.

### B. Performance Metrics

We consider the following two performance metrics:

**Mean squared error:** The reconstruction *mean squared error* (MSE) measures the deviation of  $\hat{\mathbf{x}}$  from the true signal  $\mathbf{x}_0$ . This assesses the performance of the *first* step of the decoding algorithms. Formally, the MSE is defined as

$$\text{MSE} := \frac{1}{n} \|\hat{\mathbf{x}} - \mathbf{x}_0\|^2. \quad (2)$$

**Bit Error Rate:** We also use the bit error rate (BER) to characterize the performance of the detection process, which is defined as

$$\text{BER} := \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{x_i^* \neq x_{0,i}\}}, \quad (3)$$

where  $i$  indicates the  $i$ -th element of the corresponding vector, and  $\mathbb{1}_{\{\cdot\}}$  is the indicator function.

The rest of the paper is organized as follows. In Section II, we present and discuss the main results of the paper. Numerical results are included in Section III, while a proof outline is given in Section IV. Finally, the paper is concluded in Section V.

## II. MAIN RESULTS

In this section, we summarize the asymptotic analysis of the MSE and BER of the BRO decoder in (1). The analysis is performed when the system dimensions  $m$  and  $n$  grow simultaneously to infinity ( $m, n \rightarrow +\infty$ ) at a fixed rate  $\frac{m}{n} \rightarrow \delta \in (0, \infty)$ . The SNR is assumed to be constant.

We use the standard notation  $\text{plim}_{n \rightarrow \infty} X_n = X$  to denote that a sequence of random variables  $X_n$  converges in probability towards a constant  $X$ . Define the spectral decomposition of  $\Sigma^{\frac{1}{2}}$  as  $\Sigma^{\frac{1}{2}} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$ . Finally, let  $Q(\cdot)$  denote the  $Q$ -function associated with the standard normal probability density function (pdf).

**Theorem 1** (MSE of the BRO). Let MSE denote the mean squared error of the BRO decoder in (1) for some fixed but unknown BPSK signal  $\mathbf{x}_0$ , then in the limit of  $m, n \rightarrow \infty, m/n = \delta$ , it holds

$$\text{plim}_{n \rightarrow \infty} \text{MSE} = \varphi(\tau_*), \quad (4)$$

where  $\tau_*$  is the unique solution to the following scalar optimization problem:

$$\min_{\tau > 0} \sup_{\beta > 0} D(\tau, \beta) := \frac{1}{2n} \sum_{j=1}^m \frac{\lambda_j^2 \varphi(\tau) + \sigma^2}{\frac{1}{2} + \frac{\lambda_j^2}{\beta}} - S^2(\tau) \frac{\beta}{2}, \quad (5)$$

and  $\varphi(\tau) := 4(Q(\tau) + \frac{f(\tau)}{\tau^2})$ ,  $S := \frac{1}{\tau}(1 - 2Q(\tau))$ ,  $\lambda_j$  is the  $j$ -th eigenvalue of the matrix  $\Sigma^{\frac{1}{2}}$ , and  $f(\tau) := \frac{1}{2} - Q(\tau) - \frac{\tau}{\sqrt{2\pi}} e^{-\frac{\tau^2}{2}}$ .

*Proof.* A proof outline of this theorem is given in Section IV.  $\square$

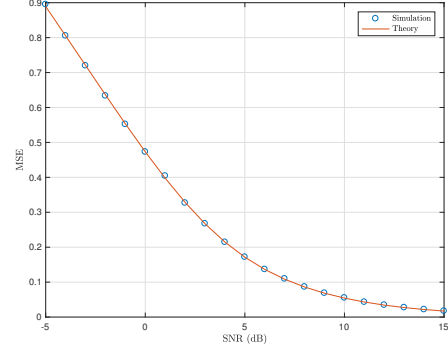


Fig. 1. MSE performance of the BRO decoder for the correlation model in (7).  $\delta = 1.5$ ,  $n = 500$ .

**Remark 1.** The objective function in (5) is strictly convex-concave. Hence, the optimal solution  $\tau_*$  can be computed numerically by writing the first order optimality conditions, i.e., by solving  $\nabla_{(\tau, \beta)} D(\tau, \beta) = 0$ .

In the next theorem we present the asymptotic BER analysis of the BRO decoder.

**Theorem 2** (BER of the BRO). Let BER denote the bit-error-rate of the BRO decoder in (1). Then, under the same setting of Theorem 1, it holds that

$$\text{plim}_{n \rightarrow \infty} \text{BER} = Q\left(\frac{\tau_*}{2}\right). \quad (6)$$

*Proof.* The proof is given in Section IV.  $\square$

**Remark 2.** Note that although the theorem requires  $n \rightarrow \infty$ , the prediction is already accurate for  $n$  as low as 64, for example see Fig. 2.

## III. NUMERICAL RESULTS

To validate our theoretical predictions of the MSE as given by Theorem 1 and BER as stated in Theorem 2, we consider the following examples for the correlation matrix  $\Sigma$ :

a) We generate  $\Sigma$  as

$$[\Sigma]_{i,j} = J_0(\pi|i-j|), i, j = 1, 2, \dots, m, \quad (7)$$

where  $J_0(\cdot)$  is the zero-order Bessel function of the first kind. Such matrix is used to model the correlation between transmit antennas in a dense scattering environment [12]. Fig. 1 shows the MSE performance of the BRO decoder for different values of SNR. Monte Carlo Simulations are used to validate the theoretical prediction of Theorem 1. Comparing the simulation results to the asymptotic MSE prediction of Theorem 1 shows the close match between the two. We used  $n = 500$ ,  $\delta = 1.5$ , and the data are averaged over 5000 realizations of the channel matrix and the noise vector.

Fig. 2 also shows the close match between simulation results and the asymptotic BER prediction of Theorem 2. We notice the match even for small values of  $n$  as low

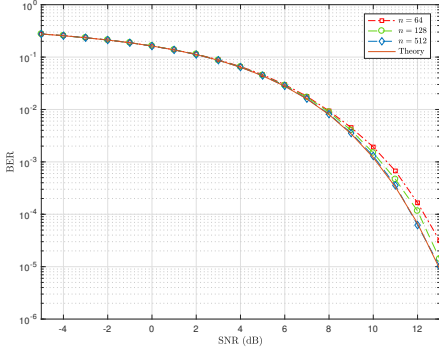


Fig. 2. BER performance of the BRO decoder for the correlation model in (7), with  $\delta = 1.5$ .

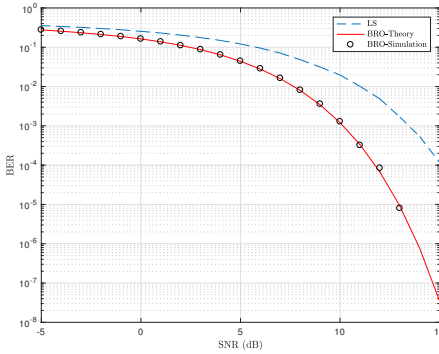


Fig. 3. BER performance of the BRO decoder and LS for the correlation model in (7), with  $\delta = 1.5$ ,  $n = 500$ .

as 64. In Fig. 3, we compare the BER performance of the considered BRO decoder to the standard least squares (LS) decoder (which is also known as zero-forcing (ZF) decoder in the communications literature). This figure shows that the BRO decoder significantly outperforms the LS decoder.

**b)** We consider the exponential correlation model for  $\Sigma$  which is defined as [13]

$$\Sigma(\rho) = \left[ \rho^{|i-j|} \right]_{i,j=1,2,\dots,m}, \rho \in [0, 1]. \quad (8)$$

Again, for this correlation model, we compare Monte Carlo simulations to the asymptotic predictions of Theorems 1 and 2. The results are shown in Fig. 4 and Fig. 5 respectively, for different values of the correlation coefficient  $\rho$ .

#### IV. APPROACH AND PROOF OVERVIEW

In this section, we provide a proof outline of Theorems 1 and 2. The proof idea is mainly based on the framework of the CGMT which is summarized next.

##### A. Convex Gaussian Min-max Theorem (CGMT)

The key ingredient of the analysis is the CGMT. Here, we recall the statement of the theorem, and we refer the

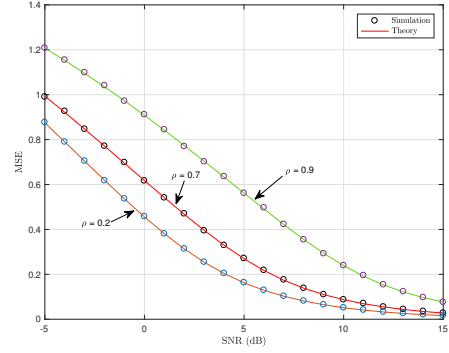


Fig. 4. MSE performance of the BRO decoder for the exponential correlation model in (8),  $\delta = 1.5$ ,  $n = 500$ .

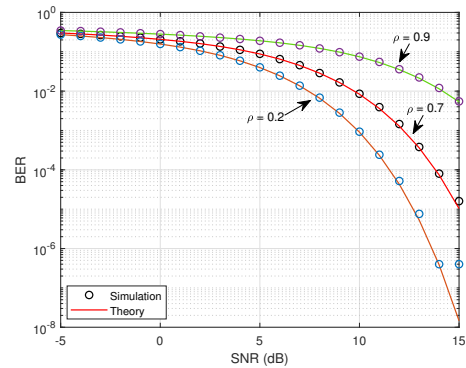


Fig. 5. BER performance of the BRO decoder for the exponential correlation model in (8),  $\delta = 1.5$ ,  $n = 500$ .

reader to [11], [14] for the complete technical details. Consider the following two min-max problems, which we refer to, respectively, as the Primary Optimization (PO) and Auxiliary Optimization (AO):

$$\Phi(\mathbf{C}) := \min_{\mathbf{w} \in \mathcal{S}_w} \max_{\mathbf{u} \in \mathcal{S}_u} \mathbf{u}^T \mathbf{C} \mathbf{w} + \psi(\mathbf{w}, \mathbf{u}), \quad (9a)$$

$$\phi(\mathbf{g}, \mathbf{h}) := \min_{\mathbf{w} \in \mathcal{S}_w} \max_{\mathbf{u} \in \mathcal{S}_u} \|\mathbf{w}\| \mathbf{g}^T \mathbf{u} - \|\mathbf{u}\| \mathbf{h}^T \mathbf{w} + \psi(\mathbf{w}, \mathbf{u}), \quad (9b)$$

where  $\mathbf{C} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{g} \in \mathbb{R}^m$ ,  $\mathbf{h} \in \mathbb{R}^n$ ,  $\mathcal{S}_w \subset \mathbb{R}^m$ ,  $\mathcal{S}_u \subset \mathbb{R}^n$  and  $\psi : \mathbb{R}^m \times \mathbb{R}^n \mapsto \mathbb{R}$ . Denote by  $\mathbf{w}_\Phi := \mathbf{w}_\Phi(\mathbf{C})$  and  $\mathbf{w}_\phi := \mathbf{w}_\phi(\mathbf{g}, \mathbf{h})$  any optimal minimizers of (9a) and (9b), respectively. Further let  $\mathcal{S}_w, \mathcal{S}_u$  be convex and compact sets,  $\psi(\cdot, \cdot)$  is convex-concave continuous on  $\mathcal{S}_w \times \mathcal{S}_u$  and,  $\mathbf{C}, \mathbf{g}$  and  $\mathbf{h}$  all have iid standard normal entries.

Let  $\mathcal{S}$  be any arbitrary open subset of  $\mathcal{S}_w$ , and  $\mathcal{S}^c = \mathcal{S}_w \setminus \mathcal{S}$ . Denote  $\phi_{\mathcal{S}^c}(\mathbf{g}, \mathbf{h})$  the optimal cost of the optimization in (9b), when the minimization over  $\mathbf{w}$  is constrained over  $\mathbf{w} \in \mathcal{S}^c$ . Suppose that there exist constants  $\bar{\phi}$  and  $\eta > 0$  such that in the limit as  $n \rightarrow +\infty$ , it holds with probability approaching one: (i)  $\phi(\mathbf{g}, \mathbf{h}) \leq \bar{\phi} + \eta$ , and, (ii)  $\phi_{\mathcal{S}^c}(\mathbf{g}, \mathbf{h}) \geq \bar{\phi} + 2\eta$ . Then,  $\lim_{n \rightarrow \infty} \mathbb{P}(\mathbf{w}_\phi \in \mathcal{S}) = 1$ , and  $\lim_{n \rightarrow \infty} \mathbb{P}(\mathbf{w}_\Phi \in \mathcal{S}) = 1$ .

### B. Identifying the PO and the AO

For notational convenience, we consider the error vector  $\mathbf{w} := \mathbf{x} - \mathbf{x}_0$ , then the problem in (1) can be reformulated as

$$\hat{\mathbf{w}} := \arg \min_{-2 \leq \mathbf{w} \leq 0} \frac{1}{n} \|\mathbf{A}\mathbf{w} - \mathbf{z}\|^2. \quad (10)$$

Without loss of generality, we assume for the analysis that  $\mathbf{x}_0 = \mathbf{1}_n = [1, 1, \dots, 1]^T$ . Therefore, the BER can be written as,  $\text{BER} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\hat{w}_i \leq -1\}}$ . Using the invariance of the Gaussian distribution under orthogonal transformations, we have

$$\hat{\mathbf{w}} = \arg \min_{-2 \leq \mathbf{w} \leq 0} \frac{1}{n} \|\mathbf{A}\mathbf{G}\mathbf{w} - \mathbf{z}\|^2, \quad (11)$$

where  $\mathbf{G}$  has iid Gaussian entries with mean 0 and variance  $\frac{1}{n}$ . The loss function can be expressed in its dual form through the Fenchel conjugate as

$$\|\mathbf{A}\mathbf{G}\mathbf{w} - \mathbf{z}\|^2 = \max_{\mathbf{u}} \sqrt{n} \mathbf{u}^T (\mathbf{A}\mathbf{G}\mathbf{w} - \mathbf{z}) - \frac{n \|\mathbf{u}\|^2}{4}.$$

Then, our PO can be written as

$$\Phi^{(n)} = \frac{1}{n} \min_{-2 \leq \mathbf{w} \leq 0} \max_{\mathbf{u}} \sqrt{n} \mathbf{u}^T \mathbf{A}\mathbf{G}\mathbf{w} - \sqrt{n} \mathbf{u}^T \mathbf{z} - \frac{n \|\mathbf{u}\|^2}{4}. \quad (12)$$

Redefining  $\mathbf{u}$  as  $\mathbf{u} = \mathbf{A}\mathbf{u}$  yields

$$\Phi^{(n)} = \frac{1}{n} \min_{-2 \leq \mathbf{w} \leq 0} \max_{\mathbf{u}} \sqrt{n} \mathbf{u}^T \mathbf{G}\mathbf{w} - \sqrt{n} \mathbf{u}^T \mathbf{A}^{-1} \mathbf{z} - \frac{n}{4} \mathbf{u}^T \mathbf{A}^{-2} \mathbf{u}. \quad (13)$$

The above optimization is in a PO form, and its corresponding AO is

$$\phi^{(n)} = \frac{1}{n} \min_{-2 \leq \mathbf{w} \leq 0} \max_{\mathbf{u}} \|\mathbf{w}\| \mathbf{g}^T \mathbf{u} - \|\mathbf{u}\| \mathbf{h}^T \mathbf{w} - \sqrt{n} \mathbf{u}^T \mathbf{A}^{-1} \mathbf{z} - \frac{n}{4} \mathbf{u}^T \mathbf{A}^{-2} \mathbf{u}. \quad (14)$$

Fixing the norm of  $\mathbf{w}$  to  $\alpha = \frac{\|\mathbf{w}\|}{\sqrt{n}}$ . We assume that <sup>1</sup>  $\alpha^2 \leq 2$ , the AO can be expressed as

$$\phi^{(n)} = \min_{0 < \alpha \leq \sqrt{2}} \max_{\mathbf{u}} \frac{\alpha}{\sqrt{n}} \mathbf{g}^T \mathbf{u} - \frac{1}{\sqrt{n}} \mathbf{u}^T \mathbf{A}^{-1} \mathbf{z} - \frac{1}{4} \mathbf{u}^T \mathbf{A}^{-2} \mathbf{u} + \|\mathbf{u}\| \min_{\substack{\|\mathbf{w}\| = \sqrt{n}\alpha \\ -2 \leq \mathbf{w} \leq 0}} -\frac{1}{n} \mathbf{h}^T \mathbf{w}. \quad (15)$$

Now, applying Lemma 1, we get under the condition,  $\alpha^2 \leq \frac{4}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i < 0\}}$ ,

$$\hat{S}(\alpha) := \min_{\substack{\|\mathbf{w}\| = \sqrt{n}\alpha \\ -2 \leq \mathbf{w} \leq 0}} -\frac{1}{n} \mathbf{h}^T \mathbf{w} \quad (16)$$

$$= \frac{2}{n} \left[ \sum_{i=1}^n |h_i| \mathbb{1}_{\{h_i \leq -\hat{\tau}(\alpha)\}} + \sum_{i=1}^n \frac{|h_i|^2}{\hat{\tau}(\alpha)} \mathbb{1}_{\{-\hat{\tau}(\alpha) < h_i < 0\}} \right], \quad (17)$$

<sup>1</sup>From our extensive simulations, we found that MSE of BRO is always less than 2. This result is expected since a random estimation of the signal vector  $\mathbf{x}$  by -1 and 1 with equal probabilities yields an MSE equal to 2. It is expected that the performance of BRO is better than a random estimation. Thus, we assume without loss of generality that  $\text{MSE}_{\text{BRO}} = \alpha^2 \leq 2$ .

where  $\hat{\tau}(\alpha)$  verifies

$$\frac{4}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i \leq -\hat{\tau}(\alpha)\}} + \frac{4}{\hat{\tau}(\alpha)^2 n} \sum_{i=1}^n |h_i|^2 \mathbb{1}_{\{-\hat{\tau}(\alpha) < h_i < 0\}} = \alpha^2. \quad (18)$$

Taking the limit as  $n \rightarrow \infty$ , equation (18) becomes

$$4Q(\tau(\alpha)) + 4 \frac{f(\tau(\alpha))}{\tau^2(\alpha)} = \alpha^2, \quad (19)$$

with  $f(x) = 1/2 - Q(x) - \frac{x}{\sqrt{2\pi}} e^{-x^2/2}$ , and the following convergence results hold<sup>2</sup>

$$\hat{S}(\alpha) - S(\tau(\alpha)) \xrightarrow{P} 0, \quad (20)$$

$$\hat{\tau}(\alpha) - \tau(\alpha) \xrightarrow{P} 0, \quad (21)$$

with  $S(x) = \frac{1}{x} [1 - 2Q(x)]$  and  $\tau(\alpha)$  is the unique positive solution of the following fixed-point equation

$$\tau(\alpha) = 2 \sqrt{\frac{f(\tau(\alpha))}{\alpha^2 - 4Q(\tau(\alpha))}}, \quad (22)$$

where (22) is obtained from (19). In the sequel  $\tau(\alpha)$  will be just denoted by  $\tau$  for notational simplicity. Now, noting that  $\frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i < 0\}} \xrightarrow{P} \frac{1}{2}$ , then the condition  $\alpha^2 \leq \frac{4}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i < 0\}}$  becomes  $\alpha^2 \leq 2$  asymptotically. Thus, applying Lemma 10 in [14], we have  $\phi^{(n)} - \tilde{\phi}^{(n)} \xrightarrow{P} 0$ , where

$$\tilde{\phi}^{(n)} = \min_{0 < \alpha \leq \sqrt{2}} \max_{\mathbf{u}} \frac{\alpha}{\sqrt{n}} \mathbf{g}^T \mathbf{u} - \frac{1}{\sqrt{n}} \mathbf{u}^T \mathbf{A}^{-1} \mathbf{z} - \frac{1}{4} \mathbf{u}^T \mathbf{A}^{-2} \mathbf{u} - \|\mathbf{u}\| S(\tau). \quad (23)$$

Using the variational form,

$$\|\mathbf{u}\| = \inf_{\chi > 0} \frac{\chi}{2} + \frac{\|\mathbf{u}\|^2}{2\chi},$$

we can write

$$\tilde{\phi}^{(n)} = \min_{0 < \alpha \leq \sqrt{2}} \max_{\mathbf{u}} \sup_{\chi > 0} \left( \frac{\alpha}{\sqrt{n}} \mathbf{g} - \frac{1}{\sqrt{n}} \mathbf{A}^{-1} \mathbf{z} \right)^T \mathbf{u} - \frac{\chi}{2} S(\tau) - \frac{S(\alpha) \|\mathbf{u}\|^2}{2\chi} - \frac{1}{4} \mathbf{u}^T \mathbf{A}^{-2} \mathbf{u}. \quad (24)$$

Let  $\tilde{\mathbf{g}} = \alpha \mathbf{g} - \mathbf{A}^{-1} \mathbf{z}$  and  $\mathbf{\Gamma} = \frac{1}{2} \mathbf{A}^{-2} + \frac{S(\tau)}{\chi} \mathbf{I}_m$ , then

$$\tilde{\phi}^{(n)} = \min_{0 < \alpha \leq \sqrt{2}} \sup_{\chi > 0} \max_{\mathbf{u}} \frac{1}{\sqrt{n}} \tilde{\mathbf{g}}^T \mathbf{u} - \frac{1}{2} \mathbf{u}^T \mathbf{\Gamma} \mathbf{u} - \frac{\chi}{2} S(\tau). \quad (25)$$

From the first order optimality condition, the optimal  $\mathbf{u}^*$  is given by  $\mathbf{u}^* = \frac{1}{\sqrt{n}} \mathbf{\Gamma}^{-1} \tilde{\mathbf{g}}$ . Thus, the AO can be written as

$$\tilde{\phi}^{(n)} = \min_{0 < \alpha \leq \sqrt{2}} \sup_{\chi > 0} \frac{1}{2n} \tilde{\mathbf{g}}^T \mathbf{\Gamma}^{-1} \tilde{\mathbf{g}} - \frac{S(\tau) \chi}{2}. \quad (26)$$

<sup>2</sup>We write " $\xrightarrow{P}$ " to denote convergence in probability as  $n \rightarrow \infty$ .

### C. Probabilistic Asymptotic Analysis of the AO

Note that  $\tilde{\mathbf{g}} \sim \mathcal{N}(\mathbf{0}, \mathbf{C}_{\tilde{\mathbf{g}}})$ , where  $\mathbf{C}_{\tilde{\mathbf{g}}}$  is given by

$$\mathbf{C}_{\tilde{\mathbf{g}}} = \alpha^2 \mathbf{I}_m + \sigma^2 \mathbf{\Lambda}^{-2}.$$

Thus, applying the trace lemma [20], we have

$$\frac{1}{n} \tilde{\mathbf{g}}^T \mathbf{\Gamma}^{-1} \tilde{\mathbf{g}}^T - \frac{1}{n} \text{tr}(\mathbf{C}_{\tilde{\mathbf{g}}} \mathbf{\Gamma}^{-1}) \xrightarrow{P} 0.$$

Therefore, again using Lemma 10 of [14],  $\tilde{\phi}^{(n)} - \bar{\phi}^{(n)} \xrightarrow{P} 0$ , where

$$\bar{\phi}^{(n)} = \min_{0 < \alpha \leq \sqrt{2}} \sup_{\chi > 0} \frac{1}{2n} \sum_{j=1}^m \frac{\lambda_j^2 \alpha^2 + \sigma^2}{\frac{1}{2} + \frac{S(\tau) \lambda_j^2}{\chi}} - \frac{S(\tau) \chi}{2}. \quad (27)$$

Using the change of variable  $\beta := \frac{\chi}{S(\alpha)}$ , we get

$$\bar{\phi}^{(n)} = \min_{0 < \alpha \leq \sqrt{2}} \sup_{\beta > 0} \frac{1}{2n} \sum_{j=1}^m \frac{\lambda_j^2 \alpha^2 + \sigma^2}{\frac{1}{2} + \frac{\lambda_j^2}{\beta}} - S^2(\tau) \frac{\beta}{2}. \quad (28)$$

Using equation (19), we can write

$$\alpha^2 = \varphi(\tau) := 4 \left( Q(\tau) + \frac{f(\tau)}{\tau^2} \right). \quad (29)$$

Noting that  $\varphi(\cdot)$  is a strictly decreasing function on  $(0, \infty)$  and according to the change of variables rule on page 130 of [21], we can use the change of variable  $\alpha^2 = \varphi(\tau)$  to obtain

$$\bar{\phi}^{(n)} = \min_{\tau > 0} \sup_{\beta > 0} \frac{1}{2n} \sum_{j=1}^m \frac{\lambda_j^2 \varphi(\tau) + \sigma^2}{\frac{1}{2} + \frac{\lambda_j^2}{\beta}} - S^2(\tau) \frac{\beta}{2}. \quad (30)$$

Investigating the second derivative of the objective function in (30) with respect to  $\tau$ , one can show that it is strictly positive for all  $\beta > 0$ . Thus, (30) has a unique minimizer  $\tau_*$ . This automatically implies that (28) has a unique minimizer  $\alpha_*$ .

### D. Applying the CGMT

We are now in a position to study the convergence limit of the MSE of the BRO. Let  $\tilde{\mathbf{w}}$  be the optimal solution to the AO defined as the solution to (15). Let  $\tau_*$  be the optimal solution to (30). For any  $\epsilon > 0$ , define the set:

$$\mathcal{S}_\epsilon = \left\{ \mathbf{v} : \left| \frac{1}{n} \|\mathbf{v}\|^2 - \varphi(\tau_*) \right| < \epsilon \right\},$$

where  $\varphi(\tau)$  is as defined in (29). Define  $\hat{\alpha}$  as the minimizer of (15). By definition,  $\hat{\alpha}^2 = \frac{\|\tilde{\mathbf{w}}\|^2}{n}$ . We have shown in the previous section that  $\tilde{\phi}^{(n)} - \bar{\phi}^{(n)} \xrightarrow{P} 0$ , and since  $\bar{\phi}^{(n)}$  in (28) has a unique minimizer  $\alpha_*$ , then  $\hat{\alpha} - \alpha_* \xrightarrow{P} 0$  which implies that

$$\left| \frac{1}{n} \|\tilde{\mathbf{w}}\|^2 - \varphi(\tau_*) \right| \xrightarrow{P} 0.$$

This proves that  $\tilde{\mathbf{w}} \in \mathcal{S}_\epsilon$  with probability approaching 1. Then, applying the CGMT yields that  $\hat{\mathbf{w}} \in \mathcal{S}_\epsilon$  with probability approaching 1 as well. This completes the proof of Theorem 1.

For the BER analysis, we need to change the set to the following:

$$\mathcal{S}_\epsilon = \left\{ \mathbf{v} : \left| \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{v_i \leq -1\}} - Q\left(\frac{\tau_*}{2}\right) \right| < \epsilon \right\}.$$

Using the expression of  $\tilde{\mathbf{w}}$  in (32), we have

$$\begin{aligned} & \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\tilde{w}_i \leq -1\}} \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i \leq -\hat{\tau}\}} + \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\frac{2h_i}{\hat{\tau}} \leq -1\}} \mathbb{1}_{\{-\hat{\tau} < h_i < 0\}} \\ &= \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i \leq -\hat{\tau}\}} + \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{-\hat{\tau} < h_i \leq -\frac{\hat{\tau}}{2}\}}. \end{aligned}$$

It can be easily seen that

$$\left| \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{\tilde{w}_i \leq -1\}} - Q\left(\frac{\tau_*}{2}\right) \right| \xrightarrow{P} 0.$$

Therefore,  $\tilde{\mathbf{w}} \in \mathcal{S}_\epsilon$  with probability approaching 1. Note that the indicator function  $\mathbb{1}_{\{\tilde{w}_i \leq -1\}}$  is not Lipschitz, so we cannot directly apply the CGMT. However, as discussed in [11, Lemma A.4], this function can be appropriately approximated with Lipschitz functions. Therefore, we can conclude by applying the CGMT that  $\hat{\mathbf{w}} \in \mathcal{S}_\epsilon$  with probability approaching 1, which proves the result of Theorem 2.

## V. CONCLUSION

In this paper, we derived precise asymptotic performance analysis of the BRO decoder in terms of its MSE and BER in the presence of one left-sided correlation in the measurement matrix. For simplicity, we used BPSK signals in the analysis but our results can be extended to other modulation schemes such as  $M$ -PAM and is left for future work. Simulation results show precise matching between the numerical simulations and the derived asymptotic expressions even for small dimensions of the problem. Possible future extensions include studying the double-sided correlation model (Kronecker model), imperfect channel models and extensions to other decoders such as the regularized-BRO decoder.

## APPENDIX

**Lemma 1.** Let  $\mathbf{h} \in \mathbb{R}^n$  and let  $a$  and  $\alpha$  be strictly positive constants such as

$$\alpha^2 \leq a^2 \left( \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i < 0\}} \right). \quad (31)$$

Then,

$$\begin{aligned} & \min_{\substack{\|\mathbf{w}\| = \sqrt{n}\alpha \\ -a \leq \mathbf{w} \leq 0}} \frac{1}{n} \mathbf{h}^T \mathbf{w} = \\ & - \frac{a}{n} \sum_{i=1}^n |h_i| \mathbb{1}_{\{h_i \leq -\tau\}} - \frac{a}{\tau n} \sum_{i=1}^n |h_i|^2 \mathbb{1}_{\{-\tau < h_i < 0\}}, \end{aligned}$$

where  $\tau$  verifies

$$\frac{a^2}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i \leq -\tau\}} + \frac{a^2}{\tau^2 n} \sum_{i=1}^n |h_i|^2 \mathbb{1}_{\{-\tau < h_i < 0\}} = \alpha^2$$

The corresponding optimal  $\mathbf{w}^*$  is given by

$$w_i^* = \begin{cases} 0, & \text{if } h_i \geq 0 \\ \frac{a}{\tau} h_i, & \text{if } -\tau < h_i < 0 \\ -a, & \text{if } h_i \leq -\tau. \end{cases} \quad (32)$$

*Proof.* Let us consider the relaxed version of this optimization

$$\min_{\substack{\|\mathbf{w}\| \leq \sqrt{n}\alpha \\ -a \leq \mathbf{w} \leq 0}} -\frac{1}{n} \mathbf{h}^T \mathbf{w}.$$

We will show later that the optimal solution of the relaxed problem verifies the constraints of the original problem and thus the two problem are equivalent under condition (31). Note that the relaxed problem is convex, its corresponding dual problem is,  $\max_{\gamma \geq 0} g(\gamma)$ , where  $g(\gamma) = \min_{-a \leq \mathbf{w} \leq 0} -\frac{1}{n} \mathbf{h}^T \mathbf{w} + \gamma \left( \frac{\|\mathbf{w}\|^2}{n} - \alpha^2 \right)$  and the corresponding optimal  $\mathbf{w}^*$  is

$$w_i^* = \begin{cases} 0, & \text{if } \frac{h_i}{2\gamma} \geq 0 \\ \frac{h_i}{2\gamma}, & \text{if } -a < \frac{h_i}{2\gamma} < 0 \\ -a, & \text{if } \frac{h_i}{2\gamma} \leq -a. \end{cases} \quad (33)$$

The optimal  $\gamma$  that maximizes  $g(\gamma)$  is such as

$$\gamma^* \left( \frac{\|\mathbf{w}^*\|^2}{n} - \alpha^2 \right) = 0.$$

Assume that  $\frac{\|\mathbf{w}^*\|^2}{n} < \alpha^2$ , then  $\gamma^* = 0$  with

$$w_i^* = \begin{cases} 0, & \text{if } h_i \geq 0 \\ -a, & \text{if } h_i < 0. \end{cases} \quad (34)$$

Thus,  $\frac{a^2}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i < 0\}} < \alpha^2$ , which contradicts condition (31), thus  $\gamma^* > 0$  and  $\frac{\|\mathbf{w}^*\|^2}{n} = \alpha^2$ . Replacing  $w_i^*$  by its expression in (33), we get

$$\frac{a^2}{n} \sum_{i=1}^n \mathbb{1}_{\{h_i \leq -2a\gamma\}} + \frac{1}{4\gamma^2 n} \sum_{i=1}^n h_i^2 \mathbb{1}_{\{-2a\gamma < h_i < 0\}} = \alpha^2. \quad (35)$$

Note that under condition (31), we can always find  $\gamma > 0$  verifying equation (35). Since the constraint  $\|\mathbf{w}\| \leq \sqrt{n}\alpha$  in the relaxed problem is verified with equality under condition (31),  $\mathbf{w}^*$  is also optimal for the original problem. Finally, Letting  $\tau = 2a\gamma$ , yields the result of Lemma 1.  $\square$

## REFERENCES

- [1] Erik G Larsson, Ove Edfors, Fredrik Tufvesson, and Thomas L Marzetta, "Massive mimo for next generation wireless systems," *IEEE communications magazine*, vol. 52, no. 2, pp. 186–195, 2014.
- [2] Lu Lu, Geoffrey Ye Li, A Lee Swindlehurst, Alexei Ashikhmin, and Rui Zhang, "An overview of massive mimo: Benefits and challenges," *IEEE journal of selected topics in signal processing*, vol. 8, no. 5, pp. 742–758, 2014.
- [3] Abba Kammoun, Axel Müller, Emil Björnson, and Mérouane Debbah, "Linear precoding based on polynomial expansion: Large-scale multi-cell mimo systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 861–875, 2014.
- [4] Axel Mueller, Abba Kammoun, Emil Björnson, and Mérouane Debbah, "Linear precoding based on polynomial expansion: Reducing complexity in massive mimo," *EURASIP journal on wireless communications and networking*, vol. 2016, no. 1, pp. 63, 2016.
- [5] H. Sifaou, A. Kammoun, L. Sanguinetti, M. Debbah, and M. Alouini, "Maxmin sinr in large-scale single-cell mu-mimo: Asymptotic analysis and low-complexity transceivers," *IEEE Transactions on Signal Processing*, vol. 65, no. 7, pp. 1841–1854, April 2017.
- [6] Peter JG Teunissen, "Integer least-squares theory for the gnss compass," *Journal of Geodesy*, vol. 84, no. 7, pp. 433–447, 2010.
- [7] 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.
- [8] 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.
- [9] Christos Thrampoulidis, Ehsan Abbasi, Weiyu Xu, and Babak Hassibi, "Ber analysis of the box relaxation for bpsk signal recovery," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2016, pp. 3776–3780.
- [10] Ismail Ben Atitallah, Christos Thrampoulidis, Abba Kammoun, Tareq Y Al-Naffouri, Babak Hassibi, and Mohamed-Slim Alouini, "Ber analysis of regularized least squares for bpsk recovery," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2017, pp. 4262–4266.
- [11] Christos Thrampoulidis, Weiyu Xu, and Babak Hassibi, "Symbol error rate performance of box-relaxation decoders in massive mimo," *IEEE Transactions on Signal Processing*, vol. 66, no. 13, pp. 3377–3392, 2018.
- [12] Giuseppa Alfano, Antonia M Tulino, Angel Lozano, and Sergio Verdú, "Capacity of mimo channels with one-sided correlation," in *Eighth IEEE International Symposium on Spread Spectrum Techniques and Applications-Programme and Book of Abstracts (IEEE Cat. No. 04TH8738)*. IEEE, 2004, pp. 515–519.
- [13] Hyundong Shin, Moe Z Win, Jae Hong Lee, and Marco Chiani, "On the capacity of doubly correlated mimo channels," *IEEE Transactions on Wireless Communications*, vol. 5, no. 8, pp. 2253–2265, 2006.
- [14] Christos Thrampoulidis, Ehsan Abbasi, and Babak Hassibi, "Precise error analysis of regularized  $m$ -estimators in high dimensions," *IEEE Transactions on Information Theory*, vol. 64, no. 8, pp. 5592–5628, 2018.
- [15] Ismail Ben Atitallah, Christos Thrampoulidis, Abba Kammoun, Tareq Y Al-Naffouri, Mohamed-Slim Alouini, and Babak Hassibi, "The box-lasso with application to gssk modulation in massive mimo systems," in *2017 IEEE International Symposium on Information Theory (ISIT)*. IEEE, 2017, pp. 1082–1086.
- [16] Ayed M Alrashdi, Ismail Ben Atitallah, and Tareq Y Al-Naffouri, "Precise performance analysis of the box-elastic net under matrix uncertainties," *IEEE Signal Processing Letters*, vol. 26, no. 5, pp. 655–659, 2019.
- [17] Ayed M Alrashdi, Ismail Ben Atitallah, Tarig Ballal, Christos Thrampoulidis, Anas Chaaban, and Tareq Y Al-Naffouri, "Optimum training for mimo bpsk transmission," in *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2018, pp. 1–5.
- [18] 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.
- [19] Ansuman Adhikary, Junyoung Nam, Jae-Young Ahn, and Giuseppe Caire, "Joint spatial division and multiplexing the large-scale array regime," *IEEE transactions on information theory*, vol. 59, no. 10, pp. 6441–6463, 2013.
- [20] T. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*, U.K., Cambridge: Cambridge Univ. Press, 2011.
- [21] S. Boyd and L. Vandenberghe, *Convex Optimization*, Cambridge University Press, 2003.