A SIMPLE BOUND ON THE BER OF THE MAP DECODER FOR MASSIVE MIMO SYSTEMS

Christos Thrampoulidis[⋆], *Ilias Zadik*[†], *Yury Polyanskiy*[†]

[⋆] University of California, Santa Barbara, USA.
[†] Massachusetts Institute of Technology, USA.

ABSTRACT

The deployment of massive MIMO systems has revived much of the interest in the study of the large-system performance of multiuser detection systems. In this paper, we prove a non-trivial upper bound on the bit-error rate (BER) of the MAP detector for BPSK signal transmission and equal-power condition. In particular, our bound is approximately tight at high-SNR. The proof is simple and relies on Gordon's comparison inequality. Interestingly, we show that under the assumption that Gordon's inequality is tight, the resulting BER prediction matches that of the replica method under the replica symmetry (RS) ansatz. Also, we prove that, when the ratio of receive to transmit antennas exceeds 0.9251, the replica prediction matches the matched filter lower bound (MFB) at high-SNR. We corroborate our results by numerical evidence.

***Index Terms*—** massive mimo, large-system analysis, JO detector, Gaussian process inequalities, replica method.

1. INTRODUCTION

Massive multiple-input multiple-output (MIMO) systems, where the base station is equipped with hundreds of thousands of antennas, promise improved spectral efficiency, coverage and range compared to small-scale systems. As such, they are widely believed to play an important role in 5G wireless communication systems [1]. Their deployment has revived much of the recent interest for the study of multiuser detection schemes in high-dimensions, e.g., [2, 3, 4, 5].

A large host of exact and heuristic detection schemes have been proposed over the years. Decoders such as zero-forcing (ZF) and linear minimum mean square error (LMMSE) have inferior performances [6], and others such as local neighborhood search-based methods [7] and lattice reduction-aided (LRA) decoders [8, 9] are often difficult to precisely characterize. Recently, [10] studied in detail the performance of the box-relaxation optimization (BRO), which is a natural convex relaxation of the maximum a posteriori (MAP) decoder, and which allows one to recover the signal via efficient convex optimization followed by hard thresholding. In particular, [10] precisely quantifies the performance gain of the BRO compared to the ZF and the LMMSE. Despite such gains, it remains unclear the degree of sub-optimality of the convex relaxation compared to the combinatorial MAP detector. The challenge lies in the complexity of analyzing the latter. In particular, known predictions of the performance of the MAP detector are known only via the (non-rigorous) replica method from statistical physics [11].

In this paper, we derive a simple, yet non-trivial, upper bound on the bit error rate (BER) of the MAP detector. We show (in a precise manner) that our bound is approximately tight at high-SNR, since it is close to the matched filter lower bound (MFB). Our numerical simulations verify our claims and further include comparisons to the replica prediction and to the BER of the BRO. Our proof relies on Gordon's Gaussian comparison inequality [12]. While Gordon's inequality is not guaranteed to be tight, we make the following possibly interesting and useful observation. If Gordon's inequality was asymptotically tight, then its BER prediction would match the prediction of the replica method (under replica-symmetry).

2. SETTING

We assume a real Gaussian wireless channel, additive Gaussian noise and and uncoded modulation scheme. For concreteness, we focus on the binary-phase-shift-keying (BPSK) transmission; but, the techniques naturally extend to other constellations. Formally, we seek to recover an n-dimensional BPSK vector $\mathbf{x}_0 \in \{\pm 1\}^n$ from the noisy MIMO relation $\mathbf{y} = \mathbf{A}\mathbf{x}_0 + \sigma \mathbf{z} \in \mathbb{R}^m$, where $\mathbf{A} \in \mathbb{R}^{m\times n}$ is the channel matrix (assumed to be known) with entries iid $\mathcal{N}(0, 1/n)$. and $\mathbf{z} \in \mathbb{R}^m$ the noise vector with entries iid $\mathcal{N}(0, 1)$. The normalization is such that the reciprocal of the noise variance σ^2 is equal to the SNR, i.e., $\mathrm{SNR} = 1/\sigma^2$. The performance metric of interest is the bit-error rate (BER). The BER of a detector which outputs $\hat{\mathbf{x}}$ as an estimate to $\mathbf{x}_0$ is formally defined as $\mathrm{BER} := \frac{1}{n}\sum_{i=1}^n \mathbb{1}_{\{\hat{\mathbf{x}}_i \neq \mathbf{x}_{0,i}\}}$.

In this paper, we study the BER of the MAP (also commonly referred to in this context as the jointly-optimal (JO) multiuser) detector, which is defined by

$$\hat{\mathbf{x}} = \arg\min_{\mathbf{x}\in\{\pm 1\}^n} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2. \tag{1}$$

We state our results in the large-system limit where $m, n \to \infty$, while the ratio of receive to transmit antennas is maintained fix to $\delta = m/n$[1] It is well known that in the worst case, solving (1) is an NP-hard combinatorial optimization problem in the number of users [13]. The BRO is a relaxation of (1) to an efficient convex quadratic program, namely $\hat{\mathbf{x}} = \mathrm{sign}\big(\arg\min_{\mathbf{x}\in[-1,1]^n} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2\big)$. Its performance in the large-system limit has been recently analyzed in [10]. Regarding the performance of (1), Tse and Verdu [14] have shown that the BER approaches zero at high-SNR. Beyond that, there is a now long literature that studied (1) using the replica method, developed in the field of spin-glasses. The use of the method in the context of multiuser detection was pioneered by Tanaka [11] and several extensions have followed up since then [15, 16]. The replica method has the remarkable ability to yield highly nontrivial predictions, which in certain problem instances they can been formally shown to be correct (e.g., [17, 18, 19]). However, it is still lacking a complete rigorous mathematical justification.

This material is based upon work supported by the National Science Foundation under Grant No CCF-17-17842.

[1] The proof of our main result Theorem 3.1 reveals that a non-asymptotic bound is also possible with only slight more effort.

3. RESULTS

3.1. Upper bound

This section contains our main result: a simple upper bound on the BER of (1). First, we introduce some useful notation. We say that an event $\mathcal{E}(n)$ holds with probability approaching 1 (wpa 1) if $\lim_{n\rightarrow} \Pr(\mathcal{E}(n)) = 1$. Let X_n a sequence of random variables indexed by n and X some constant. We write $X_n \overset{P}{=} X$ and $X_n \overset{P}{\leq} X$, if for all $\epsilon > 0$ the events $\{|X_n - X| \leq \epsilon\}$ and $\{X_n \leq X + \epsilon\}$ hold wpa 1. Finally, let $\phi(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, $Q(x) = \int_x^\infty \phi(\tau)\mathrm{d}\tau$ the Gaussian tail function and Q^{-1} its inverse.

Theorem 3.1. *Fix constant noise variance $\sigma^2 > 0$ and $\delta > 0$. Let* BER *denote the bit-error-rate of the MAP detector in* (1) *for fixed but unknown $\mathbf{x}_0 \in \{\pm 1\}^n$. Define the function $\ell(\theta) : (0,1) \rightarrow \mathbb{R}$:*

$$\ell(\theta) := \sqrt{\delta}\,\sqrt{4\theta + \sigma^2} - \sqrt{\frac{2}{\pi}}e^{-\frac{(Q^{-1}(\theta))^2}{2}}, \qquad (2)$$

and let $\theta_0 \in (0,1)$ be the largest solution to the equation $\ell(\theta) = \sigma\sqrt{\delta}$. Then, in the limit of $m, n \rightarrow \infty$, $\frac{m}{n} = \delta$, it holds $\mathrm{BER} \overset{P}{\leq} \theta_0$.

Propositions A.1 and A.2 in the Appendix gathers several useful properties of the function ℓ. Notice that $\ell(1^-) > \ell(0^+) = \sqrt{\delta}\sigma$. Also, ℓ is continuous and $\ell'(0^+) < 0$. Thus, θ_0 in Theorem 3.1 is well-defined. Moreover, we show in Proposition A.2(i) that if $\delta > 0.9251$ or $\sigma^2 > 0.1419$, then θ_0 is the *unique* solution of the equation $\ell(\theta) = \sigma\sqrt{\delta}$ in $(0,1)$.

Remark 1 (On the function $\ell(\theta)$). Let us elaborate on the operational role of the function ℓ. We partition the feasible vectors $\mathbf{x} \in \{\pm 1\}^n$ according to their Hamming distance from the true vector $\mathbf{x}_0$. Specifically, for $\theta \in [0,1]$ let $\mathcal{S}_\theta := \{\mathbf{x} \in \{\pm 1\}^n \;:\; \|\mathbf{x} - \mathbf{x}_0\|_0 = \theta n\}$ and consider the optimal cost of (1) for each partition, i.e.,

$$c_\star(\theta) := \min_{\mathbf{x} \in \mathcal{S}_\theta} \frac{1}{\sqrt{n}}\|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2. \qquad (3)$$

Evaluating the BER of (1) is of course closely related to understanding the typical behavior of $c_\star(\theta)$ in the large system limit. The proof of the theorem in Section 3.3 shows that $\ell(\theta)$ is a *high-probability lower bound* on $c_\star(\theta)$. Hence, we get an estimate on the BER via studying $\ell(\theta)$ instead. In this direction, note that the value $\sigma\sqrt{\delta}$, to which $\ell(\theta)$ is compared to, is nothing but the typical value of $c_\star(0) = \frac{\|\mathbf{y} - \mathbf{A}\mathbf{x}_0\|_2}{\sqrt{n}} = \frac{\|\mathbf{z}\|_2}{\sqrt{n}}$. Finally, we make the following note for later reference: the value $\inf_{\theta \in (0,1)} \ell(\theta)$ is a high-probability lower bound to the optimal cost of (1), i.e., to $c_\star = \inf_{\theta \in (0,1)} c_\star(\theta)$. An illustration of these is included in Figure 2.

Remark 2 (A genie lower bound). A lower bound on the BER of (1) can be obtained easily via comparison to the idealistic matched filter bound (MFB), where one assumes that all $n-1$, but 1, bits of $\mathbf{x}_0$ are known. In particular, the MFB corresponds to the probability of error in detecting (say) $\mathbf{x}_{0,1} \in \{\pm 1\}$ from $\widetilde{\mathbf{y}} = \mathbf{x}_{0,1}\mathbf{a}_1 + \mathbf{z}$, where $\widetilde{\mathbf{y}} = \mathbf{y} - \sum_{i=2}^n \mathbf{x}_{0,i}\mathbf{a}_i$ is assumed known, and $\mathbf{a}_i$ is the i^{th} column of $\mathbf{A}$ (eqv., the MFB is the error probability of an isolated transmission of only the first bit over the channel). It can be shown (e.g., [10]) that the MFB is given by $Q(\sqrt{\delta\,\mathrm{SNR}})$. Combining this with (a straightforward re-parametrization of) Theorem 3.1 it follows that the BER of (1) satisfies

$$Q(\sqrt{\delta\,\mathrm{SNR}}) \leq \mathrm{BER} \leq Q(\tau_0), \qquad (4)$$

where $\tau_0 \in \mathbb{R}$ is the *smallest* solution to the equation $\sqrt{\delta \mathrm{SNR}} + 2\phi(\tau) = \sqrt{\delta \mathrm{SNR}}\,\sqrt{1 + 4\mathrm{SNR}\,Q(\tau)}$.

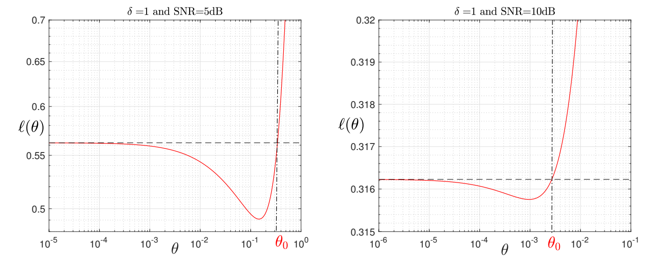


Fig. 1: Plots of the function $\ell(\theta)$ defined in (2) for two problem instances: $(\delta = 1, \mathrm{SNR} = 5\mathrm{dB})$, $(\delta = 1, \mathrm{SNR} = 10\mathrm{dB})$. Also depicted the value of θ_0 for each instance (see Theorem 3.1).

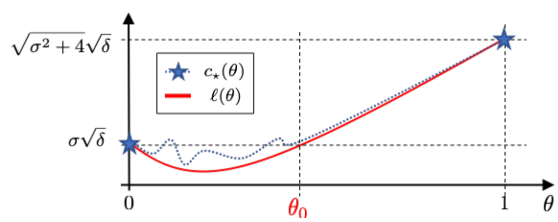
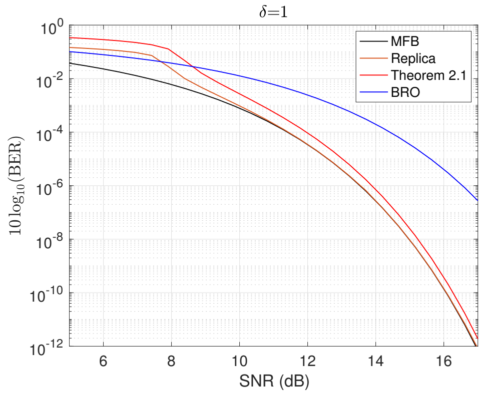


Fig. 2: The function $\ell(\theta)$ (in red) is a high-probability lower bound on the typical value of $c_\star(\theta)$ (in dashed blue) defined in (3). See Remark 1.

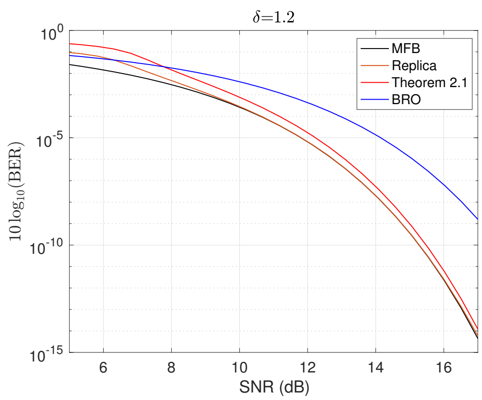
Remark 3 (Behavior at high-SNR). In Proposition A.2(iv) we prove that at high values of $\mathrm{SNR} \gg 1$: $\theta_0 \rightarrow 0$. Thus, from Theorem 3.1 we have that BER approaches zero (thus, providing an alternative proof to the corresponding result in [14]). This thinking confirms already that our upper bound is non-trivial. In fact, an even stronger statement can be shown, namely, at $\mathrm{SNR} \gg 1$: $\theta_0 \approx Q(\sqrt{\delta\,\mathrm{SNR}} - \eta)$ for an arbitrarily small $\eta > 0$ (see Proposition A.2(iv) for exact statement). This, when combined with the MFB in (4) shows that our upper bound is approximately tight at high-SNR.

Remark 4 (Gordon's comparison inequality). The proof of Theorem 3.1 uses Gordon's comparison inequality for Gaussian processes (also known as the Gaussian min-max Theorem (GMT)). In essence, the GMT provides a simple lower bound on the typical value of $c_\star(\theta)$ in (3) in the large-system limit. Gordon's inequality is classically used to establish (non)-asymptotic probabilistic lower bounds on the minimum singular value of Gaussian matrices [20], and has a number of other applications in high-dimensional convex geometry [21]. In general, the inequality is not tight. Recently, Stojnic [22] proved that the inequality is tight when applied to convex problems. The result was refined in [23] and has been successfully exploited to precisely analyze the BER of the BRO [10]. Unfortunately, the minimization in (3) is not convex, thus there are no immediate tightness guarantees regarding the lower bound $\ell(\theta)$. Interestingly, in Section 3.4 we show that if GMT was (asymptotically) tight then it would result in a prediction that matches the replica prediction in [24].

Remark 5 (Replica prediction). The replica prediction on the BER of (1) is given by [11] (based on the ansatz of replica-symmetry (RS)) as the solution to a system of nonlinear equations. It is reported in [24, Eqn. (15)] that as long as δ is not too small, the saddle-point equations reduce to the solution of the following fixed-point equa-



(a) $\delta = 1$



(b) $\delta = 1.2$

Fig. 3: BER curve as a function of the SNR (in dB) for the following: matched-filter lower bound (MFB) (cf. Remark 2); replica prediction corresponding to (5)); upper bound of Theorem 3.1 for (1); box-relaxation optimization (BRO) [10].

tion[2]:

$$\theta = Q\Big(\sqrt{\frac{\delta}{\sigma^2 + 4\theta}}\Big). \tag{5}$$

Onwards, we refer to (5) as the replica-symmetry prediction. In Proposition A.1, we prove that equation (5) has either one or three solutions. In the later case, the BER formula can exhibit complicated behavior, such as anomalous, non-monotonic dependence on the SNR [11]. On the other hand, the solution is unique if either $\delta > 0.9251$ or $\sigma^2 \geq 0.1419$. This proves the numerical observations reported in [25, Fig. 3]. Finally, Proposition A.2(iii) shows that when $\delta > \underline{0.9251}$ and $\mathrm{SNR} \gg 1$, the unique solution of (5) satisfies $\theta_\star \approx Q(\sqrt{\delta\,\mathrm{SNR}})$. This suggests that at high-SNR, the BER of the (1) decreases at an optimal rate.

3.2. Numerical Evaluations

Figure 3 includes numerical illustrations that help visualize the prediction of Theorem 3.1 and several of the remarks that followed. For two values of δ, we plot BER as a function of $\mathrm{SNR} = 1/\sigma^2$. Each plot includes four curves: (i) the MFB; (ii) the solution to (5) corresponding to the replica prediction; (iii) the upper bound θ_0 of Theorem 3.1; (iv) the BER of the BRO according to [10, Thm. II.I]. We make several observations. First, it is interesting to note that our upper bound follows the same trend as the replica prediction. For example, note the kink at values of $\mathrm{SNR} \sim 7$dB in both the curves in Figure 3a. Second, note that the upper bound of Theorem 3.1 approaches the MFB at high-SNR confirming our theoretical findings in Remark 3. Also, as predicted in Remark 5, the solution $\theta_\star$ to (5) goes to zero exactly at the rate of the MFB. Finally, let us compare the upper bound θ_0 of Theorem 3.1 to the BER of the BRO. At low SNR, θ_0 takes values larger than the latter. We remark that Theorem 3.1 is not entirely to be blamed for this behavior, since the replica prediction experiences the very same one. There is no contradiction here: the MAP detector is *not* optimal for minimizing the BER (e.g., [11, Sec. 2]), thus it is likely that its convex approximation (aka, the BRO) shows better BER performance at low-SNR. On the other hand, for high-SNR our upper bound takes values significantly smaller than the BER of the BRO. This proves that at high-SNR the latter is still quite far from that of the combinatorial optimization it tries to approximate.

3.3. Proof Theorem 3.1

Let $\hat{\mathbf{x}}$ be the solution to (1). First, observe that $\|\hat{\mathbf{x}} - \mathbf{x}_0\|_2^2 = 2n - 2(n - \sum_{i=1}^n \mathbb{1}_{\{\hat{\mathbf{x}}_i \neq \mathbf{x}_{0,i}\}}) = 4n\,\mathrm{BER}$. Hence, we will prove that

$$\frac{\|\hat{\mathbf{x}} - \mathbf{x}_0\|_2}{\sqrt{n}} \overset{P}{\leq} \alpha_0 =: 2\sqrt{\theta_0} \in (0,1). \tag{6}$$

Second, due to rotational invariance of the Gaussian measure we can assume without loss of generality that $\mathbf{x}_0 = +\mathbf{1}$. For convenience, define the (normalized) error vector $\mathbf{w} := n^{-1/2}(\mathbf{x} - \mathbf{1})$ and consider the set of feasible such vectors that do not satisfy (6), i.e.,

$$\mathcal{S}(\alpha_0) := \big\{\mathbf{w} \in \{-2/\sqrt{n}, 0\}^n \ : \ \|\mathbf{w}\|_2 \geq \alpha_0 + \epsilon\big\},$$

for some fixed (but arbitrary) $\epsilon > 0$. Also, denote the (normalized) objective function of (1) as $F(\mathbf{w}) = F(\mathbf{w}; \mathbf{z}, \mathbf{G}) := n^{-1/2}\|\mathbf{z} - \mathbf{G}\mathbf{w}\|_2$, where $\mathbf{G} = \sqrt{n}\mathbf{A}$ has entries iid standard normal. With this notation, our goal towards establishing (6) is proving that there exists constant $\eta := \eta(\epsilon) > 0$ such that the following holds wpa 1,

$$\min_{\mathbf{w} \in \mathcal{S}(\alpha_0)} F(\mathbf{w}) \geq \min_{\mathbf{w} \in \{-2/\sqrt{n}, 0\}^n} F(\mathbf{w}) + \eta. \tag{7}$$

Our strategy in showing the above is as follows.

First, we use Gordon's inequality to obtain a high-probability lower bound on the left-hand side (LHS) of (7). In particular, it can be shown (see for example [10, Sec. D.3]) that the primary optimization (PO) in the (LHS) of (7) can be lower bounded with high-probability by an auxiliary optimization (AO) problem, which is defined as follows:

$$\min_{\mathbf{w} \in \mathcal{S}(\alpha_0)} G(\mathbf{w}; \mathbf{g}, \mathbf{h}) := \sqrt{\|\mathbf{w}\|_2^2 + \sigma^2}\,\|\mathbf{g}\|_2 - \mathbf{h}^T\mathbf{w}, \tag{8}$$

where $\mathbf{g} \in \mathbb{R}^m$ and $\mathbf{h} \in \mathbb{R}^n$ have entries iid Gaussian $\mathcal{N}(0, 1/n)$. Specifically, the following statement holds for all $c \in \mathbb{R}$:

$$\Pr\Big(\min_{\mathbf{w} \in \mathcal{S}(\alpha_0)} F(\mathbf{w}; \mathbf{z}, \mathbf{G}) \leq c\Big) \leq 2\Pr\Big(\min_{\mathbf{w} \in \mathcal{S}(\alpha_0)} G(\mathbf{w}; \mathbf{g}, \mathbf{h}) \leq c\Big). \tag{9}$$

The AO can be easily simplified as follows

$$\min_{1 \geq \alpha \geq \alpha_0 + \epsilon} \sqrt{\alpha^2 + \sigma^2}\,\|\mathbf{g}\|_2 - \frac{2}{\sqrt{n}} \sum_{i=1}^{(\alpha^2/4)n} \mathbf{h}_i^{\downarrow}, \tag{10}$$

[2]For the reader's convenience we note the following mapping between notation here and [24]: $\delta \leftrightarrow \alpha$, $\sigma^2 \leftrightarrow \beta_s^{-1}$ and $\mathrm{BER} \leftrightarrow (1-m)/2$.

where, $\mathbf{h}_1^{\downarrow} \geq \mathbf{h}_2^{\downarrow} \geq \ldots \geq \mathbf{h}_n^{\downarrow}$ denotes the ordered statistics of the entries of $\mathbf{h}$ and we have used the fact that for $\mathbf{w} \in \{-2/\sqrt{n}, 0\}^n$ it holds $\|\mathbf{w}\|_2 = \alpha \Leftrightarrow \|\mathbf{w}\|_0 = \alpha^2/4$. Furthermore, note that $\|\mathbf{g}\|_2 \overset{P}{=} \sqrt{\delta}$ and [3] for any fixed $\theta \in (0,1)$: $\frac{1}{\sqrt{n}} \sum_{i=1}^{\theta n} \mathbf{h}_i^{\downarrow} \overset{P}{=} \phi\big(Q^{-1}(\theta)\big)$. Thus, the objective function in (10) converges in probability, pointwise on α, to $\ell(\alpha^2/4)$ (cf. (2)). In fact, since the minimization in (10) is over a compact set, uniform convergence holds and the minimum value converges to $\min_{1 \geq \alpha \geq \alpha_0 + \epsilon} \ell(\alpha^2/4)$. Combining the above, shows that for all $\eta > 0$ the following event holds wpa 1:

$$\min_{\mathbf{w} \in \mathcal{S}(\alpha_0)} G(\mathbf{w}; \mathbf{g}, \mathbf{h}) \geq \min_{1 \geq \alpha \geq \alpha_0 + \epsilon} \ell\big(\alpha^2/4\big) - \eta. \qquad (11)$$

Hence, from (9) the above statement holds with $G(\mathbf{w}; \mathbf{g}, \mathbf{h})$ replaced by $F(\mathbf{w}; \mathbf{z}, \mathbf{G})$.

Next, we obtain a simple upper bound on the RHS in (7):

$$\min_{\mathbf{w} \in \{-2/\sqrt{n}, 0\}^n} F(\mathbf{w}) \leq F(\mathbf{0}) = \frac{\|\mathbf{z}\|_2}{\sqrt{n}}, \qquad (12)$$

which we combine with the fact that wpa 1 it holds $\|\mathbf{z}\|_2/\sqrt{n} \leq \sqrt{\delta}\,\sigma + \eta$.

Combining the two displays in (11) and (12), we have shown that (7) holds as long as there exists $\eta > 0$ such that

$$\min_{1 \geq \alpha \geq \alpha_0 + \epsilon} \ell\big(\alpha^2/4\big) \geq \sqrt{\delta}\,\sigma + 3\eta. \qquad (13)$$

At this point, recall that $\alpha_0^2/4 = \theta_0$ and the definition of θ_0 as the largest solution to the equation $\ell(\theta) = \sqrt{\delta}\,\sigma$. By this definition and the fact that $\ell(\theta)$ is continuous and satisfies $\ell(1^-) > \sqrt{\delta}\,\sigma$ we have that $\ell(\theta) > \sqrt{\delta}$ for all $\theta > \theta_0$. Thus, there always exist $\eta(\epsilon)$ satisfying (13) and the proof is complete.

3.4. Gordon's prediction meets Tanaka

Inspecting the proof of Theorem 3.1 reveals two possible explanations for why the resulting upper bound might be loose. First, recall that we obtain a lower bound in the LHS of (7) via Gordon's inequality. As mentioned, in Remark 4 the inequality is not guaranteed to be tight in this instance. Second, recall that in upper bounding the RHS of (7) we use the crude bound (12). Specifically, we upper bound the optimal cost $c_\star$ of the MAP in (1) simply by the value of the objective function at a known feasible solution, namely $\mathbf{x} = \mathbf{x}_0$.

In this section, we make the following leap of faith. We assume that $\inf_{\theta \in (0,1)} \ell(\theta)$ is an asymptotically *tight* high-probability lower bound of $c_\star$, i.e., for all $\eta > 0$ wpa 1:

$$\min_{\mathbf{x} \in \{\pm 1\}^n} \frac{1}{\sqrt{n}} \|\mathbf{y} - \mathbf{A}\mathbf{x}\|_2 \overset{?}{\leq} \inf_{\theta \in (0,1)} \ell(\theta) + \eta. \qquad (14)$$

Assuming (14) is true and repeating the arguments of Section 3.3 leads to the following conclusion: the BER of the MAP detector is upper bounded by $\theta_\star = \arg\min_{\theta \in (0,1)} \ell(\theta)$. This can be also be expressed as the solution to the fixed-point equation $\ell'(\theta_\star) = 0$. Interestingly, this is shown in Proposition A.1(i) to be equivalent to (5). In other words, under the assumption above, Gordon's prediction on the BER of the MAP detector coincides with the replica prediction under the RS ansatz. While it is known that the MAP detector exhibits replica symmetry breaking (RSB) behavior [25], we believe that our observation on a possible connection between Gordon's inequality and the replica symmetric prediction is worth exploring further.

[3]Let $\gamma_i \overset{\text{iid}}{\sim} \mathcal{N}(0,1)$ and $\theta \in (0,1)$. Then, for large n: $\frac{1}{n} \sum_{i=1}^{\theta n} \gamma_i^{\downarrow} \approx \frac{1}{n} \sum_{\{i : \gamma_i \geq Q^{-1}(\theta)\}} h_i \approx \mathbb{E}[\gamma \,|\, \gamma \geq Q^{-1}(\theta)] = \phi\big((Q^{-1}(\theta)\big)$.

4. CONCLUSION

In this paper, we prove a simple yet highly non-trivial upper bound on the BER of the MAP detector in the case of BPSK signals and of equal-power condition. Theorem 3.1 naturally extends to allow for other constellation types (such as M-PAM) and power control and it also enjoys a non-asymptotic version. Perhaps more challenging, but certainly of interest, is the extension of our results to complex Gaussian channels. Also, we wish to develop a deeper understanding of the connection between Gordon's inequality and the replica-symmetric prediction.

Acknowledgement

The authors would like to thank Dr. Tom Richardson for sharing with us his proof of Proposition A.1. In particular, Proposition A.1 as it appears here is a refined version of what appeared in our initial submission.

A. PROPERTIES OF $\ell(\theta)$

Let $\ell : (0,1) \to \mathbb{R}$ and θ_0 be defined as in Theorem 3.1.

Proposition A.1. *The following statements are true:*

(i) $\theta \in (0,1)$ *is a critical point of* ℓ *if and only if it solves* (5). *All critical points belong in* $(0, \frac{1}{2})$.

(ii) ℓ *has either one or three critical points.*

(iii) ℓ *has a unique critical point if either one of the following two holds:* $\delta > 0.9251$ *or* $\sigma^2 > 0.1419$.

Proposition A.2. *If* ℓ *has a unique critical point (see Prop. A.1(iii)), then the following are true:*

(i) θ_0 *is the* unique *solution of the equation* $\ell(\theta) = \sigma\sqrt{\delta}$ *in* $(0,1)$.

(ii) The unique solution of (5) *is the unique* $\theta_\star = \arg\min_\theta \ell(\theta)$.

Moreover, if $\delta > 0.9251$ *it holds that:*

(iii) The unique solution $\theta_* = \theta_*(\sigma)$ *of* (5) *satisfies* $\frac{\theta_*}{Q\left(\sqrt{\delta}/\sigma\right)} \to 1$, *in the limit of* $\sigma^2 \to 0$.

(iv) For $\eta > 0$, $\theta_0 = \theta_0(\sigma)$ *satisfies* $\limsup_{\sigma \to 0} \frac{\theta_0}{Q\left(\frac{\sqrt{\delta}}{\sigma} - \eta\right)} \leq 1$.

Due to space limitations the detailed proof can be found in [26]. Below, we include the proof of statement (i) of the proposition.

Proof of Proposition A.1(i). Setting $u := Q^{-1}(\theta)$, we will equivalently study the critical points of the function $\tilde{\ell}(u) := \ell(Q^{-1}(\theta)) = \sqrt{\delta}\sqrt{4Q(u) + \sigma^2} - 2\phi(u)$. By simple algebra,

$$\tilde{\ell}'(u) = -2Q'(u)\big(u - \sqrt{\delta}(4Q(u) + \sigma^2)^{-1/2}\big),$$

where $Q'(u) = -\phi(u)$. Clearly, $\tilde{\ell}'(u) < 0$ for all $u \leq 0$. Thus, all critical points of $\tilde{\ell}$ are in $(0, +\infty)$. Now, let us define $F(u) := u^2(4Q(u) + \sigma^2)$. Note that for $u > 0$: $\tilde{\ell}'(u) = 0 \Leftrightarrow F(u) = \delta$. Using the transformation $\theta = Q(u)$ and simple algebra shows that the equation on the RHS of (16) is identical to (5). This proves statement (i) of the proposition. $\square$

B. REFERENCES

[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive mimo for next generation wireless systems," *IEEE communications magazine*, vol. 52, no. 2, pp. 186–195, 2014.

[2] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *Communications, IEEE Transactions on*, vol. 61, no. 4, pp. 1436–1449, 2013.

[3] C.-K. Wen, J.-C. Chen, K.-K. Wong, and P. Ting, "Message passing algorithm for distributed downlink regularized zero-forcing beamforming with cooperative base stations," *Wireless Communications, IEEE Transactions on*, vol. 13, no. 5, pp. 2920–2930, 2014.

[4] T. L. Narasimhan and A. Chockalingam, "Channel hardening-exploiting message passing (chemp) receiver in large-scale mimo systems," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 8, no. 5, pp. 847–860, 2014.

[5] J. Charles, G. Ramina, M. Arian, and S. Christoph, "Optimality of large mimo detection via approximate message passing," in *Information Theory (ISIT), 2015 IEEE International Symposium on*. IEEE, 2015.

[6] S. Verdu, *Multiuser detection*. Cambridge university press, 1998.

[7] K. V. Vardhan, S. K. Mohammed, A. Chockalingam, and B. S. Rajan, "A low-complexity detector for large mimo systems and multicarrier cdma systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 3, pp. 473–485, 2008.

[8] C. Windpassinger and R. F. Fischer, "Low-complexity near-maximum-likelihood detection and precoding for mimo systems using lattice reduction," in *Information Theory Workshop, 2003. Proceedings. 2003 IEEE*. IEEE, pp. 345–348.

[9] Q. Zhou and X. Ma, "Element-based lattice reduction algorithms for large mimo detection," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 274–286, 2013.

[10] C. Thrampoulidis, W. Xu, and B. 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.

[11] T. Tanaka, "A statistical-mechanics approach to large-system analysis of cdma multiuser detectors," *IEEE Transactions on Information theory*, vol. 48, no. 11, pp. 2888–2910, 2002.

[12] Y. Gordon, *On Milman's inequality and random subspaces which escape through a mesh in* $\mathbb{R}^n$. Springer, 1988.

[13] S. Verdú, "Computational complexity of optimum multiuser detection," *Algorithmica*, vol. 4, no. 1-4, pp. 303–312, 1989.

[14] D. N. C. Tse and S. Verdú, "Optimum asymptotic multiuser efficiency of randomly spread cdma," *IEEE Transactions on Information Theory*, vol. 46, no. 7, pp. 2718–2722, 2000.

[15] D. Guo and S. Verdú, "Randomly spread cdma: Asymptotics via statistical physics," *IEEE Transactions on Information Theory*, vol. 51, no. 6, pp. 1983–2010, 2005.

[16] G. Caire, R. R. Muller, and T. Tanaka, "Iterative multiuser joint decoding: Optimal power allocation and low-complexity implementation," *IEEE Transactions on Information Theory*, vol. 50, no. 9, pp. 1950–1973, 2004.

[17] M. Talagrand, *Spin glasses: a challenge for mathematicians: cavity and mean field models*. Springer Science & Business Media, 2003, vol. 46.

[18] G. Reeves and H. D. Pfister, "The replica-symmetric prediction for compressed sensing with gaussian matrices is exact," in *Information Theory (ISIT), 2016 IEEE International Symposium on*. IEEE, 2016, pp. 665–669.

[19] J. Barbier, M. Dia, N. Macris, and F. Krzakala, "The mutual information in random linear estimation," in *Communication, Control, and Computing (Allerton), 2016 54th Annual Allerton Conference on*. IEEE, 2016, pp. 625–632.

[20] K. R. Davidson and S. J. Szarek, "Local operator theory, random matrices and banach spaces," *Handbook of the geometry of Banach spaces*, vol. 1, no. 317-366, p. 131, 2001.

[21] M. Ledoux and M. Talagrand, *Probability in Banach Spaces: isoperimetry and processes*. Springer, 1991, vol. 23.

[22] M. Stojnic, "A framework to characterize performance of lasso algorithms," *arXiv preprint arXiv:1303.7291*, 2013.

[23] C. Thrampoulidis, S. Oymak, and B. Hassibi, "Regularized linear regression: A precise analysis of the estimation error," in *Proceedings of The 28th Conference on Learning Theory, 2015*.

[24] T. Tanaka, "Analysis of bit error probability of direct-sequence cdma multiuser demodulators," in *Advances in Neural Information Processing Systems*, 2001, pp. 315–321.

[25] ——, "Statistical mechanics of cdma multiuser demodulation," *EPL (Europhysics Letters)*, vol. 54, no. 4, p. 540, 2001.

[26] C. Thrampoulidis, I. Zadik, and Y. Polyanskiy, "A simple bound on the BER of the MAP decoder for massive MIMO systems," https://web.ece.ucsb.edu/~cthrampo/MAP, 2018.