Non–Asymptotic Performance Analysis of the Semidefinite Relaxation Detector in Digital Communications

Anthony Man–Cho So*

April 16, 2010

Abstract

We consider the problem of detecting a vector of symbols that is being transmitted over a fading multiple–input multiple–output (MIMO) channel, where each symbol is an M–th root of unity for some fixed $M \geq 2$. Although the symbol vector that minimizes the error probability can be found by the so–called *maximum–likelihood (ML) detector*, its computation is intractable in general. In this paper we analyze a popular polynomial–time heuristic for the problem, called the *semidefinite relaxation (SDR) detector*, and establish its first non–asymptotic performance guarantee. Specifically, our contribution is twofold. First, in the low signal–to–noise ratio (SNR) region, it can be shown that there exists a universal constant $c > 0$ such that for any $M \geq 2$, *every* detector will yield a c–approximation to the optimal log–likelihood value with probability increasing *exponentially fast* to 1 as the channel size increases. We improve upon this result by showing that the SDR detector can achieve a much tighter approximation bound than the universal bound c. Secondly, in the high SNR region, it is known that for $M = 2$, the SDR detector will yield an *exact* solution to the ML detection problem with probability converging to 1. We sharpen this result by relaxing its assumptions and establishing the rate of convergence. Our work can be viewed as an average–case analysis of a certain SDP relaxation, and the input distribution we use is motivated by physical considerations. Our results also refine and extend those in previous work, which are all asymptotic in nature and apply only to the problem of detecting *binary* (i.e., when $M = 2$) vectors. In particular, our results can give better insight into the performance of the SDR detector in practical settings.

*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N. T., Hong Kong. E–mail: manchoso@se.cuhk.edu.hk

1 Introduction

A fundamental problem in modern digital communication is that of the joint detection of several information carrying symbols that are being transmitted over a multiple–input multiple–output (MIMO) communication channel [21, 19, 5]. Such a problem arises in many contexts. For instance, consider a wireless communication setting, where there are multiple antennae at both ends of the channel. While it is known that there could be significant gain in capacity and reliability in such a setting (see, e.g., [2, 19]), there is also much interference among the different transmitter–receiver pairs. Thus, in order to capitalize on the gains in capacity and reliability, one has to deal with the problem of detecting multiple signals across different transmitter–receiver pairs. For other applications of the detection problem, we refer the reader to [21, 19].

Before we formulate the detection problem, let us fix some notation and specify the channel model. Let $\mathbb{F}$ be either the real or complex scalar field. Let $\mathcal{S}$ be a finite set representing the signal constellation (e.g., $\mathcal{S} = \{-1, +1\}$), and let $x \in \mathcal{S}^n$ be a vector of transmitted symbols. The input–output relationship of the MIMO channel can be modeled as

$$y = \sqrt{\frac{\rho}{n}} Hx + v, \tag{1}$$

where $H \in \mathbb{F}^{m \times n}$ is the *channel matrix* for n inputs and m $(\geq n)$ outputs; $v \in \mathbb{F}^m$ is an *additive white Gaussian noise (AWGN)* with unit variance (i.e., v is a standard Gaussian random vector that is independent of H); $y \in \mathbb{F}^m$ is the vector of received signals; and $\rho > 0$ is the *signal–to–noise ratio (SNR)* of the channel. Such a model captures a wide variety of communication channels, including the one mentioned in the preceding paragraph. We refer the reader to [21, 19] for further details. Now, the goal of the detection problem is to recover the vector of transmitted symbols x, assuming that we have full knowledge of both the vector of received signals y and the channel matrix H. Specifically, we would like to design a *detector* $\varphi : \mathbb{F}^m \times \mathbb{F}^{m \times n} \to \mathcal{S}^n$ that takes the vector of received signals $y \in \mathbb{F}^m$ and the channel matrix $H \in \mathbb{F}^{m \times n}$ as inputs and produces an estimate $\hat{x} = \varphi(y, H) \in \mathcal{S}^n$ of the transmitted vector $x \in \mathcal{S}^n$ as output. Of course, such a detector should not be arbitrary, and a natural property it should possess is that it has a small *error probability*, i.e., the quantity $p_e \equiv \Pr(\hat{x} \neq x)$ should be small. It turns out that if every vector in $\mathcal{S}^n$ is equally likely to be transmitted, then the *maximum–likelihood* (ML) detector, which is given by

$$\hat{x} = \arg \min_{x \in \mathcal{S}^n} \|y - Hx\|_2^2, \tag{2}$$

minimizes the error probability (see, e.g., [21, Chapter 3]). Unfortunately, whenever $|\mathcal{S}| > 1$, the problem of computing $\hat{x}$ via (2) is NP–hard in general [20]. In fact, even when the entries of the channel matrix are standard Gaussian random variables (as is usually assumed in the communications context), it is still not known whether there exists an efficient algorithm for solving such instances. Thus, much of the recent research has focused on developing detection heuristics that not only are efficient but also achieve near–ML performance. One such heuristic (or more precisely, a family of heuristics) is the *semidefinite relaxation (SDR) detector*, which solves an SDP relaxation of (2) and produces, via some rounding procedure, an approximate solution to the detection problem in polynomial time. The SDR detector was first proposed by Tan and Rasmussen [18] and Ma et al. [12] to handle the case where $\mathcal{S} = \{-1, +1\}$ (also known as the *binary phase–shift keying (BPSK) constellation*), and simulation results (see, e.g., [12, 13]) indicate that it often achieves a performance that is comparable to that of the ML detector. In an attempt to understand this phenomenon, Kisialiou and Luo [9] considered the case where $\mathbb{F} = \mathbb{C}$, $\mathcal{S} = \{-1, +1\}$ and analyzed the *asymptotic* performance of a version of the SDR detector under the additional assumption that the entries of H are suitably normalized i.i.d. random variables. They showed that in the low SNR region (i.e., when ρ is sufficiently small), the probability of the SDR detector yielding a constant factor approximation to problem (2) will tend to 1 as $n \to \infty$. Here, the probability is computed over all possible realizations of (H, v) and the randomness used in the rounding procedure. Furthermore, they showed that in the high SNR region, the probability (over all possible realizations of (H, v)) of a natural SDP relaxation of (2) being *exact* (i.e., solving problem (2) is equivalent to solving the SDP relaxation) will also tend to 1 as $n \to \infty$. These results should be contrasted with those that can be obtained from a worst–case analysis. In particular, the ratio between the optimal value of (2) and that of its natural SDP relaxation can be unbounded in the worst case (see Section 3.1). This is perhaps not very surprising, as the difficulty of analyzing SDP relaxations of quadratic minimization problems is well known. Nevertheless, the results of Kisialiou and Luo have some limitations. First, in the context of wireless communications, signal constellations other than the BPSK constellation are often used in practice to increase the data rate of the channel. However, the analyses of Kisialiou and Luo do not extend to cover these settings. Secondly, all the aforementioned results hold only asymptotically. Consequently, they offer limited insight into the performance of the SDR detector in practical settings, where the channel size parameters m, n are finite. Our research is motivated in part by the desire to remedy this situation.

Our Contribution. In this paper, we establish the first non–asymptotic performance guarantee of the SDR detector under the scenario where $\mathbb{F} = \mathbb{C}$, $H \in \mathbb{C}^{m \times n}$ is the so–called *i.i.d. Rayleigh fading channel* with $m \geq n$ (i.e., the entries of H are i.i.d. *complex* standard Gaussian random variables), and $\mathcal{S}$ is an *M–ary phase–shift keying (MPSK) constellation* (for some fixed $M \geq 2$), i.e.,

$$\mathcal{S} = \mathcal{S}_M \equiv \{\exp(2\pi l j/M) : l = 0, 1, \ldots, M-1\}, \tag{3}$$

where $j \equiv \sqrt{-1}$ (in other words, $\mathcal{S}_M$ is the set of M–th roots of unity; see, e.g., [10, 11] for its use in the communications context). Specifically, in the low SNR region (i.e., when ρ is sufficiently small), we show that for any $M \geq 2$, a version of the SDR detector will yield a constant factor approximation to problem (2) with probability approaching 1 *exponentially fast*. Again, the probability here is computed over all possible realizations of (H, v) and the randomness used in the rounding procedure. Note that since the BPSK constellation is simply a 2–PSK constellation, our results refine and extend those in [9]. A key step in our proof is to show that the optimal value of the SDP is large with high probability. This is achieved using SDP duality theory, as well as results from random matrix theory. We remark that when the SNR is sufficiently small, there exists a universal constant $c > 0$ such that for any $M \geq 2$, *every* detector will yield a c–approximation to problem (2) with exponentially high probability (communicated to us by a referee; see Section 3.2.3). However, as we shall see, the approximation bound we derive for the SDR detector is much tighter than the universal bound c.

Next, we complement the above result by considering the high SNR region and establishing the rate at which the probability (over all possible realizations of (H, v)) of having an exact SDP relaxation tends to 1 for the case where $M = 2$. The proof involves analyzing a sufficient condition for having an exact SDP relaxation, and results from random matrix theory again play an important role. Our work can be viewed as an average–case analysis of a certain SDP relaxation, and the input distribution we use is motivated by physical considerations. We believe that the non–asymptotic nature of our results can give better insight into the performance of the SDR detector in practical settings. Furthermore, our techniques seem to be more general than those in [9] and can be used to analyze the performance of the SDR detector for other signal constellations as well (see, e.g., [15] for the case of quadrature amplitude modulation (QAM) constellations).

Outline of the Paper. The rest of the paper is organized as follows. In Section 2 we specify the channel model and give a formal description of the version of the SDR detector that we are going to analyze. In Section 3 we present the main results of this paper. Specifically, we analyze the performance of the SDR detector, both in the worst case setting and in the probabilistic setting defined by the channel model. Finally, we close with some concluding remarks and future directions in Section 4.

2 The Semidefinite Relaxation Detector

We begin with some notation and definitions. For a complex number $z \in \mathbb{C}$, let $\bar{z}$, $|z|$, $\Re(z)$ and $\Im(z)$ denote the conjugate, modulus, real and imaginary part of z, respectively. For a complex matrix $H \in \mathbb{C}^{m \times n}$, let $H^* \in \mathbb{C}^{n \times m}$ denote the conjugate transpose of H. A complex matrix $H \in \mathbb{C}^{n \times n}$ is said to be *Hermitian* if $H = H^*$. The inner product of two complex vectors $u, v \in \mathbb{C}^n$ is defined as $u^*v = \sum_{i=1}^n \bar{u}_i v_i$. In particular, we have $u^*v = \overline{v^*u}$ and $\|u\|_2^2 = u^*u = \sum_{i=1}^n |u_i|^2$. We say that a complex random variable $Y = Y_R + jY_I$ is distributed as $\mathcal{CN}(0, \sigma^2)$ (denoted by $Y \sim \mathcal{CN}(0, \sigma^2)$) if Y_R and Y_I are i.i.d. real Gaussian random variables with mean 0 and variance $\sigma^2/2$. If $Y \sim \mathcal{CN}(0, 1)$, then we say that Y is a *complex standard Gaussian random variable*. Note that if $Y \sim \mathcal{CN}(0, \sigma^2)$, then $e^{j\theta}Y \sim \mathcal{CN}(0, \sigma^2)$ for any $\theta \in \mathbb{R}$. In other words, the random variable Y is *circular symmetric*.

Now, consider the channel model (1), where the entries of H and v are i.i.d. complex standard Gaussian random variables, with H and v being independent. Let $x \in \mathcal{S}_M^n$ be the vector of transmitted symbols, where $\mathcal{S}_M$ is given by (3) and $M \geq 2$ is fixed. As mentioned in the Introduction, the ML detector attempts to recover the transmitted symbol vector $x \in \mathcal{S}_M^n$ from both the received signal vector $y \in \mathbb{C}^m$ and the realized channel $H \in \mathbb{C}^{m \times n}$ by solving the following discrete least squares problem:

$$v_{ml} = \min_{x \in \mathcal{S}_M^n} \left\| y - \sqrt{\rho/n}\, Hx \right\|_2^2 = \min_{(x,t) \in \mathcal{S}_M^{n+1}} \left\| yt - \sqrt{\rho/n}\, Hx \right\|_2^2. \tag{4}$$

Since problem (4) is intractable in general, many heuristics have been proposed for solving it. One such heuristic is based on solving an SDP relaxation of (4). To derive the SDP relaxation, observe that $v_{ml} = \min_{z \in \mathcal{S}_M^{n+1}} \mathrm{tr}(Qzz^*)$, where

$$Q = \begin{bmatrix} (\rho/n)H^*H & -\sqrt{\rho/n}\, H^*y \\ -\sqrt{\rho/n}\, y^*H & \|y\|_2^2 \end{bmatrix} \in \mathbb{C}^{(n+1)\times(n+1)}. \tag{5}$$

Thus, we may relax problem (4) to the following *complex* SDP (cf. [10, 11]):

$$\begin{array}{rcll} v_{sdp} & = & \text{minimize} & \text{tr}(QZ) \\ & & \text{subject to} & \text{diag}(Z) = \mathbf{e}, \\ & & & Z \succeq \mathbf{0}. \end{array} \tag{6}$$

Here, $\mathbf{e} \in \mathbb{R}^{n+1}$ is the vector of all ones and $Z \in \mathbb{C}^{(n+1)\times(n+1)}$ is a Hermitian positive semidefinite matrix. Note that since $Q \succeq \mathbf{0}$ and problem (6) is a relaxation of problem (4), we clearly have $0 \le v_{sdp} \le v_{ml}$. We emphasize that both v_{ml} and v_{sdp} depend on the particular realizations of H *and* v, since y is related to H and v via (1).

Now, the complex SDP (6) can be solved to any desired accuracy in polynomial time (see, e.g., [1] and the discussion in [3]). However, we still need a procedure that can convert any feasible solution $\hat{Z} \in \mathbb{C}^{(n+1)\times(n+1)}$ to (6) into a feasible solution $\hat{x} \in \mathcal{S}_M^n$ to (4). Below is one such procedure.

Randomized Rounding Procedure

1. Let $\hat{Z} \in \mathbb{C}^{(n+1)\times(n+1)}$ be a feasible solution to (6). Partition it as

$$\hat{Z} = \begin{bmatrix} U & u \\ u^* & 1 \end{bmatrix}, \tag{7}$$

where $u \in \mathbb{C}^n$ and $U \in \mathbb{C}^{n\times n}$. Note that since $\hat{Z} \succeq \mathbf{0}$ and $\text{diag}(\hat{Z}) = \mathbf{e}$, we must have $|u_k| \le 1$ for $k = 1, \ldots, n$.

2. Let $z = (z_1, \ldots, z_{n+1}) \in \mathbb{C}^{n+1}$ be a random vector, whose entries are independently distributed according to the following distribution:

$$\begin{aligned} \Pr\left(z_k = e^{2\pi jl/M}\right) &= \frac{1 + \Re(u_k e^{-2\pi jl/M})}{M} \quad \text{for } k = 1, \ldots, n;\ l = 0, 1, \ldots, M-1, \\ \Pr\left(z_{n+1} = e^{2\pi jl/M}\right) &= \frac{1 + \Re(e^{-2\pi jl/M})}{M} \quad \text{for } \ l = 0, 1, \ldots, M-1. \end{aligned} \tag{8}$$

Note that (8) defines a valid probability distribution on $\mathcal{S}_M$. Indeed, let $u = (u_1, \ldots, u_n)$, and set $u_{n+1} = 1$. Since $|u_k| \le 1$, we have $(1 + \Re(u_k e^{-2\pi jl/M}))/M \ge 0$ for $k = 1, \ldots, n+1$ and $l = 0, 1, \ldots, M-1$. Moreover, we have

$$\frac{1}{M} \sum_{l=0}^{M-1} \left(1 + \Re(u_k e^{-2\pi jl/M})\right) = 1 + \frac{1}{M} \Re\left(u_k \sum_{l=0}^{M-1} e^{-2\pi jl/M}\right) = 1$$

for $k = 1, \ldots, n+1$, as required. Consequently, we have $z \in \mathcal{S}_M^{n+1}$.

3. Now, let $z^1, \ldots, z^m \in \mathcal{S}_M^{n+1}$ be m independent copies of the random vector z. These can be obtained, for instance, by repeating Step 2 above m times. Let $i' = \arg\min_{1\le i\le m} \left(z^i\right)^* Q z^i$ and define $\hat{z} = z^{i'}$. Set $v_{sdr} = \hat{z}^* Q \hat{z}$ and return $\hat{x} = \overline{\hat{z}_{n+1}}(\hat{z}_1, \ldots, \hat{z}_n) \in \mathcal{S}_M^n$ as our candidate solution to (4). In other words, we have $\hat{x}_k = \overline{\hat{z}_{n+1}}\hat{z}_k \in \mathcal{S}_M$ for $k = 1, \ldots, n$. Note that $v_{sdr} = \|y - \sqrt{\rho/n}\, H\hat{x}\|_2^2$, and hence $v_{ml} \le v_{sdr}$ for any realization of $\hat{x} \in \mathcal{S}_M^n$.

We remark that a rounding procedure similar to the one given in Step 2 has been used before in the context of complex quadratic *maximization* [4, 17]. Now, we are interested in the quality of the solution $\hat{x}$. Specifically, we would like to bound the approximation ratio v_{sdr}/v_{ml}. Intuitively, if the ratio is close to 1, then we may conclude that the solution generated by the rounding procedure is close (in terms of the log–likelihood value) to the optimal ML solution. In the next section, we will analyze the approximation ratio v_{sdr}/v_{ml}, both in the worst case setting and in the probabilistic setting defined by the channel model.

3 Analysis of the SDP Relaxation

3.1 Worst–Case Analysis

A standard approach to bounding the ratio v_{sdr}/v_{ml} is to first establish a relationship between v_{sdr} and v_{sdp}, and then use the fact that $v_{sdp} \le v_{ml}$ to obtain a bound on v_{sdr}/v_{ml}. However, such an approach may not always yield useful results. To further motivate our consideration of a probabilistic model for the detection problem and to put our results in subsequent sections in perspective, we now show that for *any* $M \ge 2$, the ratio v_{sdp}/v_{ml} can be zero *in the worst case*, even when $m = n = 2$.

Proposition 1 *Let $M \ge 2$ be fixed. Then, for any given $\rho^{(M)} > 0$, there exists an instance $(H^{(M)}, y^{(M)})$ of problem (4) such that $v_{ml} > 0$ and $v_{sdp} = 0$.*

Proof Suppose that $M \ge 3$ is odd. Let $J \in \mathbb{R}^{2\times 2}$ be the matrix of all ones, and set $H^{(M)} = \sqrt{2/\rho^{(M)}}\, J$. Furthermore, let $y^M = \mathbf{0} \in \mathbb{R}^2$. We claim that $v_{ml} > 0$. Indeed, for any $x = (x_1, x_2) \in \mathcal{S}_M^2$, we have

$$\left\| y^{(M)} - \sqrt{\frac{\rho^{(M)}}{2}}\, H^{(M)} x \right\|_2^2 = 2|x_1 + x_2|^2 > 0,$$

since $x_1 + x_2 \ne 0$ whenever M is odd. Now, to show that $v_{sdp} = 0$, we observe that the objective matrix Q in (6) has the form

$$Q = \begin{bmatrix} 2J & \mathbf{0} \\ \mathbf{0}^T & 0 \end{bmatrix}.$$

Let $Z' \in \mathbb{R}^{3\times 3}$ be the matrix

$$Z' = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Note that $Z' \succeq \mathbf{0}$, since it is diagonally dominant. Thus, we see that Z' is feasible for (6). Moreover, we have $\mathrm{tr}(QZ') = 0$, and hence $v_{sdp} = 0$ as desired.

Next, consider the case where $M \ge 2$ is even. Let $H^{(M)} \in \mathbb{R}^{2\times 2}$ be the matrix

$$H^{(M)} = \sqrt{\frac{2}{\rho^{(M)}}} \begin{bmatrix} -1 & -1 \\ 0 & 0 \end{bmatrix},$$

and let $y^{(M)} = (a,0) \in \mathbb{R}^2$, where $a \in [1,2]$ is chosen so that $\cos(2\pi l/M) \neq -a/2$ for all $l = 0,1,\ldots,M-1$. Again, we claim that $v_{ml} > 0$. To prove this, observe that for any $x = (x_1, x_2) \in \mathcal{S}_M^2$, we have

$$\left\| y^{(M)} - \sqrt{\frac{\rho^{(M)}}{2}} H^{(M)} x \right\|_2^2 = |a + x_1 + x_2|^2.$$

Suppose to the contrary that $a + x_1 + x_2 = 0$ for some $x_1, x_2 \in \mathcal{S}_M$. Then, upon writing $x_1 = \exp(j\theta_1)$ and $x_2 = \exp(j\theta_2)$ with $\theta_1, \theta_2 \in [0, 2\pi)$, we see that

$$\cos\theta_1 + \cos\theta_2 = -a, \tag{9}$$

$$\sin\theta_1 + \sin\theta_2 = 0. \tag{10}$$

Now, equation (10) implies that $\theta_1 \equiv -\theta_2$ or $\pi + \theta_2 \pmod{2\pi}$. In the former case, equation (9) becomes $\cos\theta_1 = -a/2$. However, since $\theta_1 \in \{2\pi l/M : l = 0,1,\ldots,M-1\}$, we obtain a contradiction. In the latter case, equation (9) becomes $a = 0$, which again is a contradiction. It follows that $v_{ml} > 0$ as claimed.

On the other hand, observe that the objective matrix Q in (6) has the form

$$Q = \begin{bmatrix} 1 & 1 & a \\ 1 & 1 & a \\ a & a & a^2 \end{bmatrix}.$$

Let $Z'' \in \mathbb{R}^{3\times3}$ be the matrix

$$Z'' = \begin{bmatrix} 1 & \beta & -(1+\beta) \\ \beta & 1 & -(1+\beta) \\ -(1+\beta) & -(1+\beta) & 1 \end{bmatrix},$$

where

$$\beta = \frac{-a^2 + 4a - 2}{2(1-2a)}.$$

Using the fact that $a \in [1,2]$, it is straightforward to verify that $-1/2 \le \beta < 0$. It follows that Z'' is diagonally dominant, whence $Z'' \succeq \mathbf{0}$. In particular, Z'' is feasible for (6). Furthermore, we have

$$\mathrm{tr}(QZ'') = 2 + a^2 + 2\beta - 4a(1+\beta) = 0,$$

and hence $v_{sdp} = 0$ as desired. □

3.2 Probabilistic Analysis: The Low SNR Region

3.2.1 The $M \ge 3$ Case

We now return to the problem of analyzing the approximation quality of the SDP relaxation (6) under the probabilistic model described in Section 2. For reasons that would become clear, the case where $M = 2$ requires a slightly different treatment than the $M \ge 3$ case. In order to illustrate the main ideas of our approach and to simplify the exposition, we shall first consider the case where $M \ge 3$. The $M = 2$ case will be dealt with afterwards.

Our goal in this section is to prove the following theorem:

Theorem 1 *Let $\gamma \equiv m/n \geq 1$ and $M \geq 3$ be fixed. Define*

$$\rho_0 \equiv \frac{\gamma}{28(1+\gamma)-\gamma} \qquad \text{and} \qquad \Lambda \equiv \frac{1}{2}(\rho+1) - \sqrt{\frac{7\rho(\rho+1)(\gamma+1)}{\gamma}}.$$

Suppose that the SNR ρ satisfies $\rho \in (0, \rho_0)$. Then, we have $\Lambda > 0$ and

$$\begin{aligned} & \Pr_{(H,v,\hat{x})}\left[v_{sdr} \leq 2\left(\frac{1}{4} + \frac{7\rho+3}{2\Lambda} \right) v_{ml} \right] \\ \geq \quad & 1 - \exp(-m/6) - \exp(-(m+n)/5) - \exp(-mn/4) - \exp(-m/4) - 2^{-m} \qquad (11) \\ \geq \quad & 1 - 5\exp(-m/6), \end{aligned}$$

where $\Pr_{(H,v,\hat{x})}(\cdot)$ means that the probability is computed over all possible realizations of (H,v) and the random vector $\hat{x}$ obtained in Step 3 of the randomized rounding procedure.

Theorem 1 implies that in the low SNR region (i.e., when $\rho \in (0, \rho_0)$), the SDR detector will produce a constant factor approximate solution to the detection problem (4) with exponentially high probability as the channel size increases. To the best of our knowledge, this is the first performance guarantee of the SDR detector for the case of an MPSK constellation, where $M \geq 3$. We remark that the constants in our proofs are chosen to simplify the exposition and have not been optimized. With a more refined analysis, those constants can certainly be improved.

The proof of Theorem 1 consists of two steps. The first step is to show that, conditioned on a particular realization of (H,v), the value v_{sdr} is, with high probability, at most $O(v_{sdp} + \Delta)$, where Δ depends only on (H,v). Here, the probability is computed over all possible realizations of $\hat{x}$. Then, in the second step, we analyze what effect does the distribution of (H,v) have on the value v_{sdr}. In particular, we will show that v_{sdp} and Δ are of the same order with high probability. This will then imply Theorem 1.

To begin, consider a particular realization of (H,v) (and hence of Q). Let $\hat{Z} \in \mathbb{C}^{(n+1)\times(n+1)}$ be a feasible solution to (6) with objective value v_{sdp}, and let $z \in \mathcal{S}_M^{n+1}$ be the random vector generated in Step 2 of the randomized rounding procedure. Set $\Gamma \equiv \mathbb{E}_z\left[z^* Q z\right]$, where $\mathbb{E}_z$ denotes the mathematical expectation w.r.t. the distribution defined in (8). Then, by Markov's inequality and the fact that the random vectors $z^1, \ldots, z^m$ are i.i.d. as z, we have

$$\Pr_{\hat{x}}\left(v_{sdr} \geq 2\Gamma\right) = \left[\Pr_z\left(z^* Q z \geq 2\Gamma\right)\right]^m \leq 2^{-m}. \qquad (12)$$

To get a hold on the value of Γ, we need the following result:

Fact 1 *(Huang and Zhang [4, Lemma 2.1]) Let $u = (u_1, \ldots, u_n) \in \mathbb{C}^n$ be given by (7), and set $u_{n+1} = 1$. Let $z \in \mathcal{S}_M^{n+1}$ be generated according to the distribution defined in (8). Then, for $M \geq 3$ and $1 \leq k \neq k' \leq n+1$, we have*

$$\mathbb{E}_z\left[z_k \overline{z_{k'}}\right] = \frac{1}{4} u_k \overline{u_{k'}}.$$

Proposition 2 *For $M \geq 3$, we have*

$$\Gamma \leq \frac{1}{4} v_{sdp} + \frac{\rho}{n} tr(H^* H) + \frac{3}{4}\|y\|_2^2. \qquad (13)$$

Proof Let $u = (u_1, \ldots, u_n) \in \mathbb{C}^n$ be given by (7). Set $\hat{u} = (u, 1) \in \mathbb{C}^{n+1}$ and write $\hat{u} = \hat{u}_R + j\hat{u}_I$, where $\hat{u}_R, \hat{u}_I \in \mathbb{R}^{n+1}$. By Fact 1, we have

$$\begin{aligned}
\mathbb{E}_z\left[\sum_{k,k'=1}^{n+1} Q_{kk'} z_k \overline{z_{k'}}\right] &= \sum_{k=1}^{n+1} Q_{kk} + 2\Re\left(\frac{1}{4}\sum_{1\le k<k'\le n} Q_{kk'} u_k \overline{u_{k'}}\right) + 2\Re\left(\frac{1}{4}\sum_{k=1}^{n} Q_{k,n+1} u_k\right) \\
&= \frac{1}{4}\hat{u}^* Q \hat{u} + \sum_{k=1}^{n} Q_{kk}\left(1 - \frac{1}{4}|u_k|^2\right) + \frac{3}{4} Q_{n+1,n+1}.
\end{aligned}$$

Now, we claim that $\hat{Z} \succeq \hat{u}\hat{u}^*$. Indeed, observe that since $\hat{Z} \succeq \mathbf{0}$, we have $U \succeq uu^*$ in (7) by the Schur complement. This in turn implies that $\hat{Z} \succeq \hat{u}\hat{u}^*$, as desired. In particular, since $Q \succeq \mathbf{0}$, we have $\hat{u}^* Q \hat{u} = \mathrm{tr}(Q\hat{u}\hat{u}^*) \le \mathrm{tr}(Q\hat{Z}) = v_{sdp}$, whence

$$\Gamma = \mathbb{E}_z\left[\sum_{k,k'=1}^{n+1} Q_{kk'} z_k \overline{z_{k'}}\right] \le \frac{1}{4} v_{sdp} + \frac{\rho}{n}\mathrm{tr}(H^*H) + \frac{3}{4} Q_{n+1,n+1} = \frac{1}{4} v_{sdp} + \frac{\rho}{n}\mathrm{tr}(H^*H) + \frac{3}{4}\|y\|_2^2,$$

as desired. □

Now, if we could show that v_{sdp}, $(\rho/n)\cdot\mathrm{tr}(H^*H)$ and $\|y\|_2^2$ are all within a constant factor of each other with high probability (w.r.t. the realizations of (H, v)), then (12) and (13) would imply that v_{sdr} and v_{sdp} are within a constant factor of each other with high probability (w.r.t. the realizations of (H, v) *and* $\hat{x}$). To carry out this idea, we first need estimates of $\mathrm{tr}(H^*H)$ and $\|y\|_2^2$. These are given below, and the proofs can be found in the Appendix:

Proposition 3 *The following hold:*

- **Estimate of tr(H^*H).** *Let $H \in \mathbb{C}^{m\times n}$ (where $m \ge n \ge 2$) be a random matrix whose entries are i.i.d. complex standard Gaussian random variables. Then, we have*

$$\Pr_H\left[tr(H^*H) \ge 2mn\right] \le \exp(-mn/4) \tag{14}$$

  (see Appendix A.1 for the proof).

- **Estimates of $\|y\|_2^2$.** *Consider the channel model (1), where the entries of H and v are i.i.d. complex standard Gaussian random variables, with H and v being independent. Then, we have*

$$\Pr_{(H,v)}\left[\|y\|_2^2 \le \frac{1}{2}(\rho+1)m\right] \le \exp(-m/6), \tag{15}$$

$$\Pr_{(H,v)}\left[\|y\|_2^2 \ge 2(\rho+1)m\right] \le \exp(-m/4) \tag{16}$$

  (see Appendix A.2 for the proof).

Next, we need to show that v_{sdp} is large with high probability (w.r.t. the realizations of (H, v)). By the SDP weak duality theorem, it suffices to consider the dual of (6) and exhibit a dual feasible solution with large objective value. Such an idea has been used in the work of Kisialiou and Luo [8, 9]. However, our approach differs from that of [8, 9] in that we are able to obtain a non–asymptotic result.

To begin, let us write down the dual of (6):

$$\begin{array}{ll} \text{maximize} & \mathrm{tr}(W) \\ \text{subject to} & Q - W \succeq \mathbf{0}, \\ & W \in \mathbb{R}^{(n+1)\times(n+1)} \text{ diagonal.} \end{array} \tag{17}$$

Let $\alpha > 0$ and $\beta \in \mathbb{R}$ be parameters to be chosen, and define

$$\hat{W} = \begin{bmatrix} -\alpha I & \mathbf{0} \\ \mathbf{0}^T & \beta \end{bmatrix}.$$

In order for $\hat{W}$ to be feasible for (17), we must have $Q - \hat{W} \succeq \mathbf{0}$. By the Schur complement, this is equivalent to

$$y^* \left[I - \frac{\rho}{n} H \left(\frac{\rho}{n} H^* H + \alpha I \right)^{-1} H^* \right] y \ge \beta \tag{18}$$

(note that $(\rho/n)H^*H + \alpha I$ is invertible for any $\alpha > 0$). Now, we are interested in choices of $\alpha > 0$ and $\beta \in \mathbb{R}$ that would make (18) a valid inequality with high probability (w.r.t. the realizations of (H, v)). Towards that end, observe that

$$\begin{aligned} y^* \left[I - \frac{\rho}{n} H \left(\frac{\rho}{n} H^* H + \alpha I \right)^{-1} H^* \right] y &\ge \left[1 - \frac{\rho}{n} \lambda_{\max} \left(H \left(\frac{\rho}{n} H^* H + \alpha I \right)^{-1} H^* \right) \right] \cdot \|y\|_2^2 \\ &\ge \left(1 - \frac{\rho \lambda_{max}(H^*H)}{n\alpha} \right) \cdot \|y\|_2^2, \end{aligned}$$

where the last inequality follows from the fact that

$$\begin{aligned} \lambda_{max} \left(H \left(\frac{\rho}{n} H^* H + \alpha I \right)^{-1} H^* \right) &\le \lambda_{max}(HH^*) \cdot \lambda_{max} \left(\left(\frac{\rho}{n} H^* H + \alpha I \right)^{-1} \right) \\ &\le \frac{\lambda_{max}(H^*H)}{\alpha}. \end{aligned}$$

The following proposition provides an estimate of $\lambda_{max}(H^*H)$. Its proof can be found in the Appendix:

Proposition 4 *The following holds:*

- **Estimate of $\lambda_{max}(H^*H)$.** *Let $H \in \mathbb{C}^{m\times n}$ (where $m \ge n \ge 2$) be a random matrix whose entries are i.i.d. complex standard Gaussian random variables. Then, we have*

$$\Pr_H \left[\lambda_{max}(H^*H) \ge \frac{7}{2}(m+n) \right] \le \exp(-(m+n)/5) \tag{19}$$

(see Appendix A.3 for the proof).

Now, by setting $\beta = \beta_0$ in (18), where

$$\beta_0 \equiv \frac{1}{2}\left(1 - \frac{7\rho(\gamma+1)}{2\alpha}\right)(\rho+1)m \tag{20}$$

with $\gamma \equiv m/n \geq 1$, we conclude from (15) and (19) that the matrix $\hat{W}$ will be feasible for (17) with probability at least $1 - \exp(-m/6) - \exp(-(m+n)/5) \geq 1 - 2\exp(-m/6)$. In that event we have $v_{sdp} \geq \operatorname{tr}(\hat{W}) = \beta_0 - n\alpha$ by the SDP weak duality theorem, and upon optimizing over $\alpha > 0$, we obtain the following:

Proposition 5 *Let $\gamma \equiv m/n \geq 1$, and let β_0 be as in (20). Suppose that the SNR ρ satisfies $\rho \in (0, \rho_0)$, where*

$$\rho_0 \equiv \frac{\gamma}{28(1+\gamma) - \gamma}.$$

Then, with probability (over all possible realizations of (H, v)) at least $1 - \exp(-m/6) - \exp(-(m+n)/5) \geq 1 - 2\exp(-m/6)$, we will have $v_{sdp} \geq \Lambda m$, where

$$\Lambda \equiv \frac{1}{2}(\rho+1) - \sqrt{\frac{7\rho(\rho+1)(\gamma+1)}{\gamma}} > 0.$$

We are now ready to finish the proof of Theorem 1.

Proof of Theorem 1 By Propositions 2, 5 and inequalities (14), (16), we will have

$$\Gamma \leq \frac{1}{4}v_{sdp} + 2\rho m + \frac{3}{2}(\rho+1)m \leq \left(\frac{1}{4} + \frac{7\rho+3}{2\Lambda}\right)v_{sdp}$$

with probability (over all possible realizations of (H, v)) at least $1 - \exp(-m/6) - \exp(-(m+n)/5) - \exp(-mn/4) - \exp(-m/4) \geq 1 - 4\exp(-m/6)$. This, together with (12), implies the result claimed in Theorem 1. $\square$

3.2.2 The $M = 2$ Case

It is a bit unfortunate that the above argument does not readily extend to cover the $M = 2$ case. The main difficulty is the following. Fact 1 now gives

$$\mathbb{E}_z[z_k \overline{z_{k'}}] = \Re(u_k)\Re(u_{k'})$$

for $1 \leq k \neq k' \leq n+1$, which implies the bound

$$\Gamma \leq \Re(\hat{u})^T Q \Re(\hat{u}) + \frac{\rho}{n}\operatorname{tr}(H^* H),$$

where $\Re(\hat{u}) = (\Re(\hat{u}_1), \ldots, \Re(\hat{u}_{n+1})) \in \mathbb{R}^{n+1}$ (cf. Proposition 2). However, there is no clear relationship between the quantities $\Re(\hat{u})^T Q \Re(\hat{u})$ and $v_{sdp} = \hat{u}^* Q \hat{u}$. To circumvent this difficulty, we may proceed as follows. Observe that the discrete least squares problem (4) can be written as

$$v_{ml} = \min_{(x,t)\in\{-1,+1\}^{n+1}} \|\tilde{y}t - \sqrt{\rho/n}\,\tilde{H}x\|_2^2 = \min_{z\in\{-1,+1\}^{n+1}} \operatorname{tr}(\tilde{Q}zz^T), \tag{21}$$

where

$$\tilde{y} = \left[\begin{array}{c} \Re(y) \\ \Im(y) \end{array} \right] \in \mathbb{R}^{2m}, \quad \tilde{H} = \left[\begin{array}{c} \Re(H) \\ \Im(H) \end{array} \right] \mathbb{R}^{2m \times n}, \quad \tilde{v} = \left[\begin{array}{c} \Re(v) \\ \Im(v) \end{array} \right] \in \mathbb{R}^{2m}, \tag{22}$$

and $\tilde{Q} \in \mathbb{R}^{(n+1)\times(n+1)}$ is obtained from Q in (5) by replacing H and y by $\tilde{H}$ and $\tilde{y}$, respectively. Thus, problem (21) can be relaxed to a *real* SDP of the form (6), where Q is replaced by $\tilde{Q}$. Now, let $\hat{Z}$ be a feasible solution to the SDP with objective value v_{sdp} (note that $\hat{Z}$ is now an $(n+1)\times(n+1)$ *real* matrix). Clearly, we can still apply the randomized rounding procedure in Section 2 on $\hat{Z}$. Since $\Re(u_k) = u_k$ for $k = 1, \ldots, n+1$, the candidate solution $\hat{x}$ returned by the rounding procedure will be feasible (i.e., $\hat{x} \in \{-1, +1\}^n$) and satisfy

$$\Pr_{\hat{x}} \left(v_{sdr} \geq 2\mathbb{E}_z[z^T \tilde{Q} z] \right) \leq 2^{-m}.$$

Moreover, it can be readily verified that $\mathbb{E}_z[z_k z_{k'}] = u_k u_{k'}$ for $1 \leq k \neq k' \leq n+1$, whence

$$\mathbb{E}_z \left[z^T \tilde{Q} z \right] \leq v_{sdp} + \frac{\rho}{n} \mathrm{tr}(\tilde{H}^T \tilde{H})$$

(cf. the argument in the proof of Proposition 2). Now, observe that $\tilde{H}$ is a random matrix whose entries are i.i.d. real Gaussian random variables with mean 0 and variance 1/2. Thus, the random variable $2 \cdot \mathrm{tr}(\tilde{H}^T \tilde{H})$ follows a chi–square distribution with $2mn$ degrees of freedom (cf. Appendix A.1). In particular, by following the same argument as in the previous section and using the estimate of $\lambda_{max}(\tilde{H}^T \tilde{H})$ derived in Appendix A.4, one can show that v_{sdp} and $\mathrm{tr}(\tilde{H}^T \tilde{H})$ are of the same order with high probability (w.r.t. the realizations of (H, v)). This in turn leads to the following theorem:

Theorem 2 *Consider the case where $M = 2$, and let $\gamma \equiv m/n \geq 1$ be fixed. Define*

$$\rho_0 \equiv \frac{\gamma}{16(1+2\gamma) - \gamma} \qquad \text{and} \qquad \Lambda \equiv \frac{1}{2}(\rho+1) - 2\sqrt{\frac{\rho(\rho+1)(2\gamma+1)}{\gamma}}.$$

Suppose that the SNR ρ satisfies $\rho \in (0, \rho_0)$. Then, we have $\Lambda > 0$ and

$$\begin{aligned} & \Pr_{(H,v,\hat{x})} \left[v_{sdr} \leq 2\left(1 + \frac{2\rho}{\Lambda}\right) v_{ml} \right] \\ \geq \;& 1 - \exp(-m/6) - \exp(-(m+n/2)/4) - \exp(-mn/4) - 2^{-m} \qquad (23) \\ \geq \;& 1 - 4\exp(-m/6). \end{aligned}$$

Theorem 2 refines a result of Kisialiou and Luo [9] by establishing the rate at which the probability on the left–hand side of (23) tends to 1.

3.2.3 Quality of the Approximation Bounds

Given the approximation results in Sections 3.2.1 and 3.2.2, it is natural to ask about their quality. Curiously, as a referee has pointed out, when the SNR is sufficiently small, *every* detector of MPSK signals (where $M \geq 2$ is fixed) will yield a constant factor approximation to the optimal log–likelihood value with exponentially high probability. Specifically, using the appropriate probabilistic estimates, one can prove the following result:

Theorem 3 *Let $M \geq 2$ and $\gamma \equiv m/n \geq 1$ be fixed, and let $\varphi : \mathbb{C}^m \times \mathbb{C}^{m\times n} \to \mathcal{S}_M^n$ be an arbitrary detector of MPSK signals. Define*

$$\rho_0 \equiv \frac{\gamma}{28(1+\gamma)}.$$

Suppose that the SNR ρ satisfies $\rho \in (0, \rho_0)$. Then, with probability at least $1 - 2\exp(-m/6) - \exp(-(m+n)/5)$, the detector φ will yield a c–approximation to the optimal log–likelihood value, where

$$c \equiv \left(\frac{\sqrt{7\gamma} + \sqrt{56\rho(1+\gamma)}}{\sqrt{2\gamma} - \sqrt{56\rho(1+\gamma)}} \right)^2.$$

Proof Let $x \in \mathcal{S}_M^n$ be the vector of transmitted symbols, so that $y = \sqrt{\rho/n}\, Hx + v$ according to (1). Let $x' = \varphi(y, H)$ be the vector of symbols returned by the detector φ. Upon noting that

$$\left\| y - \sqrt{\frac{\rho}{n}}\, Hx' \right\|_2 = \left\| v + \sqrt{\frac{\rho}{n}}\, H(x - x') \right\|_2$$

and using the triangle inequality, we obtain

$$\|v\|_2 - \left\| \sqrt{\frac{\rho}{n}}\, H(x-x') \right\|_2 \leq \left\| y - \sqrt{\frac{\rho}{n}}\, Hx' \right\|_2 \leq \|v\|_2 + \left\| \sqrt{\frac{\rho}{n}}\, H(x-x') \right\|_2. \tag{24}$$

Now, observe that $\|x - x'\|_2 \leq 2\sqrt{n}$ for any $x, x' \in \mathcal{S}_M^n$. Moreover, the random variable $2\|v\|_2^2$ follows a chi–square distribution with $2m$ degrees of freedom. Thus, using standard concentration results for the chi–square random variable (see Proposition 8 in the Appendix), we have

$$\Pr_v \left(\frac{1}{2} m \leq \|v\|_2^2 \leq \frac{7}{4} m \right) \geq 1 - 2\exp(-m/6).$$

In addition, by Proposition 4, we have

$$\Pr_H \left[\lambda_{max}(H^* H) \leq \frac{7}{2}(m+n) \right] \geq 1 - \exp(-(m+n)/5).$$

Hence, we conclude from (24) that with probability (over all possible realizations of (H, v)) at least $1 - 2\exp(-m/6) - \exp(-(m+n)/5)$, the inequalities

$$\sqrt{\frac{m}{2}} - \sqrt{14\rho(m+n)} \leq \left\| y - \sqrt{\frac{\rho}{n}}\, Hx' \right\|_2 \leq \sqrt{\frac{7m}{4}} + \sqrt{14\rho(m+n)} \tag{25}$$

will hold. Since φ is arbitrary, it follows from (25) that whenever $\sqrt{m/2} - \sqrt{14\rho(m+n)} > 0$, or equivalently, $\rho < \rho_0$, we have

$$\Pr_{(H,v)} \left[v_\varphi \leq \left(\frac{\sqrt{7m} + \sqrt{56\rho(m+n)}}{\sqrt{2m} - \sqrt{56\rho(m+n)}} \right)^2 v_{ml} \right] \geq 1 - 2\exp(-m/6) - \exp(-(m+n)/5), \tag{26}$$

where $v_\varphi \equiv \|y - \sqrt{\rho/n}\, Hx'\|_2^2$ is the log–likelihood value of the estimate x' returned by φ. This completes the proof. $\square$

Theorem 3 naturally calls for a justification of the analyses done in Sections 3.2.1 and 3.2.2. First, observe that the results in Theorems 1 and 2 apply to a larger range of SNRs than that in Theorem 3. Now, let us investigate the tightness of the approximation bounds. For simplicity, consider the case where $\gamma = 1$. To ensure fairness in our comparison, the probabilities (11), (23) and (26) are set to be comparable. Figure 1 below shows a plot of the approximation bounds established in Theorems 1 (labeled "SDP, $M \geq 3$"), 2 (labeled "SDP, $M = 2$") and 3 (labeled "Universal"), respectively. As can be verified and seen from the figure, the bound in Theorem 2 is tighter than that in Theorem 3 for all $\rho > 0$ in the low SNR region. Moreover, there exists a $\rho' > 0$ such that the bound in Theorem 1 is tighter than that in Theorem 3 for all $\rho > \rho'$ in the low SNR region. A similar phenomenon occurs when $\gamma > 1$, as can be easily verified.

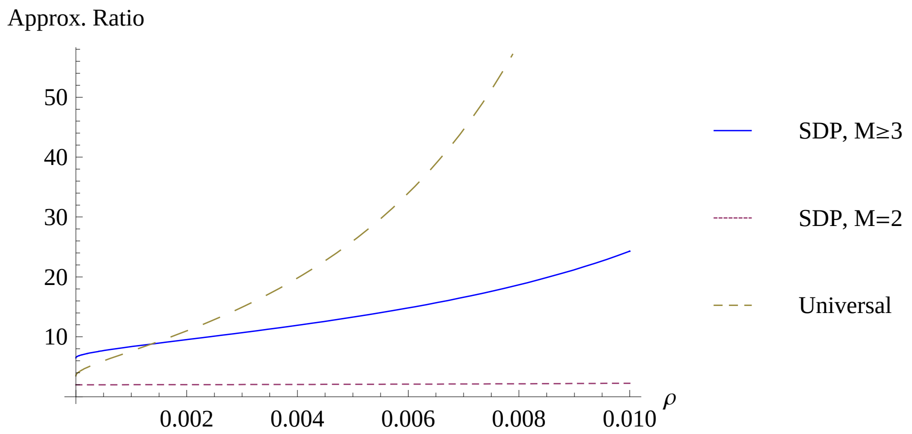

Figure 1: Comparison of the Approximation Bounds

To further demonstrate the value of our analyses, let us compare the theoretical performance of the SDR detector with that of a random detector. Specifically, consider a detector φ that generates m independent random vectors $\hat{x}^1, \ldots, \hat{x}^m$, each of which is chosen from $\mathcal{S}_M^n$ uniformly at random and is independent of (H, v), and returns $\hat{x} \equiv \hat{x}^{i'} \in \mathcal{S}_M^n$, where

$$i' = \arg \min_{1 \leq i \leq m} \left\| y - \sqrt{\frac{\rho}{n}} H \hat{x}^i \right\|_2^2,$$

as the output. Let $x \in \mathcal{S}_M^n$ be the vector of transmitted symbols, so that $y = \sqrt{\rho/n}\, Hx + v$ according to (1). Furthermore, let

$$v_{rnd}^1 = \left\| y - \sqrt{\frac{\rho}{n}} H \hat{x}^1 \right\|_2^2 = \left\| \sqrt{\frac{\rho}{n}} H(x - \hat{x}^1) + v \right\|_2^2$$

be the log–likelihood value of $\hat{x}^1 \in \mathcal{S}_M^n$. Since $\hat{x}^1$ is chosen uniformly at random from $\mathcal{S}_M^n$, we have $\mathbb{E}_{\hat{x}^1}\left[\hat{x}^1\right] = \mathbf{0}$ and $\mathbb{E}_{\hat{x}^1}\left[(\hat{x}^1)(\hat{x}^1)^*\right] = I$. It then follows that

$$\mathbb{E}_{\hat{x}^1}\left[v_{rnd}^1\right] = \frac{\rho}{n}\mathrm{tr}(H^*H) + \frac{\rho}{n} x^* H^* H x + 2\sqrt{\frac{\rho}{n}} \cdot \Re\left(v^* H x\right) + v^* v = \frac{\rho}{n}\mathrm{tr}(H^*H) + \|y\|_2^2.$$

Now, let $v_{rnd} = \|y - \sqrt{\rho/n}\, H\hat{x}\|_2^2$ be the log–likelihood value of $\hat{x}$. Using the arguments in Sections 3.2.1 and 3.2.2 and the fact that $v_{sdp} \le v_{ml}$, it can be shown that

$$\begin{aligned}
& \Pr_{(H,v,\hat{x})}\left(v_{rnd} \le \frac{4(2\rho+1)}{\Lambda} v_{ml} \right) \qquad (27) \\
\ge \; & 1 - \exp(-m/6) - \exp(-(m+n)/5) - \exp(-mn/4) - \exp(-m/4) - 2^{-m} \\
\ge \; & 1 - 5\exp(-m/6)
\end{aligned}$$

for all $\rho \in (0, \rho_0)$, where

$$\rho_0 = \begin{cases} \dfrac{\gamma}{28(1+\gamma)-\gamma} & \text{for } M \ge 3, \\ \dfrac{\gamma}{16(1+2\gamma)-\gamma} & \text{for } M = 2, \end{cases}$$

and

$$\Lambda = \begin{cases} \dfrac{1}{2}(\rho+1) - \sqrt{\dfrac{7\rho(\rho+1)(\gamma+1)}{\gamma}} & \text{for } M \ge 3, \\ \dfrac{1}{2}(\rho+1) - 2\sqrt{\dfrac{\rho(\rho+1)(2\gamma+1)}{\gamma}} & \text{for } M = 2. \end{cases}$$

However, as can be readily verified, the approximation bound in (27) is worse than those in Theorems 1 and 2.

3.3 Probabilistic Analysis: The High SNR Region

In the previous section, we investigated the performance of the SDR detector when the SNR ρ is small. Let us now consider the performance of the SDR detector when the SNR is large. Intuitively, when the SNR is large, the additive noise will be drowned out by the signal, and hence the SDR detector is more likely to detect the vector of transmitted symbols. Such an intuition can indeed be made precise when $\mathcal{S} = \{-1, +1\}$ (i.e., $M = 2$). Recall that when $M = 2$, the detection problem (21) can be relaxed to the following real SDP:

$$\begin{array}{rl} \text{minimize} & \operatorname{tr}(\tilde{Q}Z) \\ \text{subject to} & \operatorname{diag}(Z) = \mathbf{e}, \\ & Z \succeq \mathbf{0}, \end{array} \qquad (28)$$

where $\tilde{Q} \in \mathbb{R}^{(n+1)\times(n+1)}$ is defined in Section 3.2.2. The following proposition gives a sufficient condition under which the SDP relaxation (28) is exact for problem (21).

Proposition 6 *Consider a realization of $(\tilde{H}, \tilde{v})$, where $\tilde{H}$ and $\tilde{v}$ are given by (22). Suppose that the SNR ρ satisfies*

$$\sqrt{\frac{\rho}{n}}\, \lambda_{min}(\tilde{H}^T\tilde{H}) > \|\tilde{H}^T\tilde{v}\|_\infty. \qquad (29)$$

Then, the SDP relaxation (28) is exact *for problem (21), i.e., solving problem (21) is equivalent to solving problem (28).*

Proof Let $x \in \{-1,+1\}^n$ be the vector of transmitted symbols, so that $y = \sqrt{\rho/n}\, Hx + v$ according to (1). Clearly, the matrix $Z' = (x,1)(x,1)^T \in \mathbb{R}^{(n+1)\times(n+1)}$ is feasible for the SDP (28), and we have

$$v_{sdp} = \min\{\mathrm{tr}(\tilde{Q}Z) : \mathrm{diag}(Z) = \mathbf{e},\ Z \succeq \mathbf{0}\} \leq \mathrm{tr}(\tilde{Q}Z') = \|\tilde{v}\|_2^2,$$

where

$$\tilde{Q} = \left[\begin{array}{cc} (\rho/n)\tilde{H}^T\tilde{H} & -\sqrt{\rho/n}\,\tilde{H}^T\tilde{y} \\ -\sqrt{\rho/n}\,\tilde{y}^T\tilde{H} & \|\tilde{y}\|_2^2 \end{array} \right] \in \mathbb{R}^{(n+1)\times(n+1)},$$

and $\tilde{H}, \tilde{y}$ are given by (22). Now, let $\tilde{Z} \in \mathbb{R}^{(n+1)\times(n+1)}$ be an optimal solution to (28). We partition $\tilde{Z}$ as

$$\tilde{Z} = \left[\begin{array}{cc} U & u \\ u^T & 1 \end{array} \right],$$

where $u \in \mathbb{R}^n$ and $U \in \mathbb{R}^{n\times n}$. Note that by the Schur complement, we have $\tilde{Z} \succeq \mathbf{0}$ iff $U - uu^T \succeq \mathbf{0}$. In particular, if $u = x$, then $\mathrm{diag}(U - uu^T) = \mathbf{0}$, which implies that $U = xx^T \in \mathbb{R}^{n\times n}$ and $\tilde{Z} = Z'$. Thus, our goal is to show that $\Delta u \equiv x - u = \mathbf{0}$ whenever the condition in the proposition statement holds. Towards that end, we first use the definition of $\tilde{Q}$ and compute

$$\begin{aligned} v_{sdp} &= \frac{\rho}{n}\mathrm{tr}\left(\tilde{H}^T\tilde{H}(U - uu^T)\right) + \left\|\sqrt{\frac{\rho}{n}}\,\tilde{H}\Delta u + \tilde{v}\right\|_2^2 \\ &\geq \frac{\rho}{n}\cdot\lambda_{min}(\tilde{H}^T\tilde{H})\cdot(n - \|u\|_2^2) + \frac{\rho}{n}(\Delta u)^T\tilde{H}^T\tilde{H}(\Delta u) + 2\sqrt{\frac{\rho}{n}}(\Delta u)^T\tilde{H}^T\tilde{v} + \|\tilde{v}\|_2^2 \\ &\geq \frac{2\rho}{n}\cdot\lambda_{min}(\tilde{H}^T\tilde{H})\cdot x^T\Delta u + 2\sqrt{\frac{\rho}{n}}(\Delta u)^T\tilde{H}^T\tilde{v} + \|\tilde{v}\|_2^2, \end{aligned} \tag{30}$$

where the last inequality follows from the facts that

$$\begin{aligned} \|x\|_2^2 &= n, \\ n - \|u\|_2^2 &= 2x^T\Delta u - \|\Delta u\|_2^2, \\ (\Delta u)^T\tilde{H}^T\tilde{H}\Delta u &\geq \lambda_{min}(\tilde{H}^T\tilde{H})\cdot\|\Delta u\|_2^2. \end{aligned}$$

Since $v_{sdp} \leq \|\tilde{v}\|_2^2$, we conclude from (30) that

$$\begin{aligned} 0 &\geq \sqrt{\frac{\rho}{n}}\cdot\lambda_{min}(\tilde{H}^T\tilde{H})\cdot x^T\Delta u + (\Delta u)^T\tilde{H}^T\tilde{v} \\ &\geq \sqrt{\frac{\rho}{n}}\cdot\lambda_{min}(\tilde{H}^T\tilde{H})\cdot x^T\Delta u - \|\Delta u\|_1\cdot\|\tilde{H}^T\tilde{v}\|_\infty. \end{aligned} \tag{31}$$

Now, recall that $|u_i| \leq 1$ for $i = 1, \ldots, n$. Thus, if $x_i = 1$, then $(\Delta u)_i = x_i - u_i = 1 - u_i \geq 0$. On the other hand, if $x_i = -1$, then $(\Delta u)_i = -1 - u_i \leq 0$. In particular, we have $x^T\Delta u = \|\Delta u\|_1$, whence we conclude by (31) that

$$0 \geq \left(\sqrt{\frac{\rho}{n}}\cdot\lambda_{min}(\tilde{H}^T\tilde{H}) - \|\tilde{H}^T\tilde{v}\|_\infty\right)\|\Delta u\|_1. \tag{32}$$

If $\Delta u \neq \mathbf{0}$, or equivalently, $\|\Delta u\|_1 > 0$, then (32) implies that

$$\sqrt{\frac{\rho}{n}} \cdot \lambda_{min}(\tilde{H}^T \tilde{H}) \leq \|\tilde{H}^T \tilde{v}\|_\infty,$$

which is a contradiction. Hence, we must have $\Delta u = \mathbf{0}$, and the proof is completed. □

REMARKS.

1. It should be noted that Proposition 6 can also be derived directly from Theorem 7.1 of [5]. However, the proof presented here is more elementary, in the sense that it does not involve the use of SDP duality theory.

2. In Theorem 3.4 of [9] a condition similar to (29) is proposed, except that the term $\|\tilde{H}^T \tilde{v}\|_\infty$ is replaced by $\|\tilde{H}^T \tilde{v}\|_2$. Since $\|\tilde{H}^T \tilde{v}\|_2 \geq \|\tilde{H}^T \tilde{v}\|_\infty$, we see that the result in Theorem 3.4 of [9] is weaker than that in Proposition 6. In fact, it can be shown that under the condition stipulated in Theorem 3.4 of [9], the zero–forcing (ZF) detector (see, e.g., Section 1.4 of [5] for a description), which is much simpler and has much lower complexity than the SDR detector, will also be able to solve the discrete least squares problem (21) exactly. In other words, the condition stipulated in Theorem 3.4 of [9] does not help in differentiating the performance of the SDR detector from that of the ZF detector in the high SNR region.

3. Currently, we do not know whether the straightforward modification of condition (29) (i.e., with $\tilde{H}$ and $\tilde{v}$ replaced by H and v, respectively) is sufficient for the complex SDP (6) to be exact for all $M \geq 2$.

Now, given Proposition 6, it is natural to ask whether there is a threshold value ρ_0 so that condition (29) will hold with positive probability whenever $\rho \geq \rho_0$. Our goal in this section is to prove the following theorem:

Theorem 4 *Suppose that the SNR ρ satisfies $\rho \geq e^{16} n$. Then, we have*

$$\Pr_{(\tilde{H},\tilde{v})} \left(\sqrt{\frac{\rho}{n}}\, \lambda_{min}(\tilde{H}^T \tilde{H}) > \|\tilde{H}^T \tilde{v}\|_\infty \right) \geq 1 - \exp(-\Omega(m)).$$

In particular, the SDP relaxation (28) is exact for problem (21) with exponentially high probability.

To the best of our knowledge, Theorem 4 is the first non–asymptotic guarantee on the equivalence of the detection problem (21) and its SDP relaxation (28).

The proof of Theorem 4 relies on the following probabilistic estimates whose proofs can be found in the Appendix:

Proposition 7 *The following hold:*

- **Estimate of $\lambda_{min}(\tilde{H}^T \tilde{H})$.** *Let $H \in \mathbb{C}^{m \times n}$ (where $m \geq n \geq 2$) be a random matrix whose entries are i.i.d. complex standard Gaussian random variables, and let $\tilde{H}$ be given by (22). Then, we have*

  $$\Pr_{\tilde{H}} \left(\lambda_{min}(\tilde{H}^T \tilde{H}) \leq \frac{m+1}{e^8} \right) < 2\exp(-m/4)$$

  (see Appendix A.4 for the proof).

- **Estimate of $\|\tilde{H}^T\tilde{v}\|_\infty$.** *Consider the channel model (1), where the entries of H and v are i.i.d. complex standard Gaussian random variables, with H and v being independent. Let $\tilde{H}$ and $\tilde{v}$ be given by (22). Then, for any $\gamma > 1/2$, we have*

$$\Pr_{(\tilde{H},\tilde{v})}\left(\|\tilde{H}^T\tilde{v}\|_\infty \geq m^\gamma\right) \leq \sqrt{\frac{2}{\pi}} \cdot n \cdot \exp\left(-m^{2\gamma-1}/2\right) + 4\exp(-m/8)$$

(see Appendix A.5 for the proof).

Proof of Theorem 4 By Proposition 7, the event

$$\mathcal{E} = \left\{\lambda_{min}(\tilde{H}^T\tilde{H}) \geq \frac{(m+1)}{e^8}\right\} \bigcap \left\{\|\tilde{H}^T\tilde{v}\|_\infty \leq m\right\}$$

will occur with probability at least

$$1 - 2\exp(-m/4) - \sqrt{\frac{2}{\pi}} \cdot n \cdot \exp\left(-m/2\right) - 4\exp(-m/8) \geq 1 - \exp(-\Omega(m)).$$

It follows that whenever the SNR ρ satisfies $\rho \geq e^{16}n$, we have

$$\Pr_{(\tilde{H},\tilde{v})}\left(\sqrt{\frac{\rho}{n}} \cdot \lambda_{min}(\tilde{H}^T\tilde{H}) > \|\tilde{H}^T\tilde{v}\|_\infty\right) \geq \Pr_{(\tilde{H},\tilde{v})}(\mathcal{E}) > 1 - \exp(-\Omega(m)),$$

as desired. $\square$

4 Conclusion

In this paper we gave the first non–asymptotic performance analysis of the SDR detector, which is a widely used heuristic in the communications community for detecting symbol vectors that are transmitted over an MIMO channel. We considered the scenario where symbols from an MPSK constellation are transmitted over an i.i.d. Rayleigh fading channel, and showed that in both low and high SNR regions, the SDR detector will achieve a performance that is close to that of the optimal but computationally intractable ML detector with high probability. Our results were established by means of SDP duality theory, as well as results from random matrix theory. We believe that these tools will be valuable for analyzing SDPs in some other probabilistic (or even non–probabilistic) settings.

Our work also opens up several directions for future research. Perhaps the most immediate one is to derive, for any fixed $M \geq 2$, a sufficient condition under which the complex SDP (6) is exact for the detection problem (4). On another front, recall that we have analyzed the approximation guarantee of the SDR detector. However, there is another interesting measure of the quality of the SDR detector, namely its error probability, which is defined as the probability that the vector $\hat{x}$ returned by the SDR detector differs from the transmitted vector x. In [6] the authors analyzed the error probability of a version of the SDR detector under the assumptions that $H \in \mathbb{R}^{m\times n}$ (with $m \geq n$) is a real Gaussian random matrix and $\mathcal{S} = \{-1, +1\}$ is the BPSK constellation. They showed that the error probability is asymptotically (as the SNR ρ tends to infinity) on the order of $\rho^{-m/2}$. It would be interesting to derive a non–asymptotic version of this result and/or extend it to the complex channel model and other signal constellations. Finally, note that our approximation results (Theorems 1 and 2) apply only when the SNR is extremely low. Indeed, when $\gamma = 1$, the threshold SNR value ρ_0 in Theorem 1 (resp. Theorem 2) is roughly -17.40 dB (resp. -16.72 dB). In practice, it is almost impossible to communicate reliably over an MIMO channel at such low SNRs. Thus, a natural problem would be to determine the approximation guarantee of the SDR detector for the range of SNRs encountered in practice.

Acknowledgements

This research is supported by CUHK Direct Grant No. 2050401. The author would like to express his gratitude to Professor Ken Ma for many insightful discussions, and to the referee who suggested Theorem 3, which motivates the discussion in Section 3.2.3.

References

[1] F. Alizadeh and S. Schmieta. Symmetric Cones, Potential Reduction Methods and Word–by–Word Extensions. In H. Wolkowicz, R. Saigal, and L. Vandenberghe, editors, *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, volume 27 of *International Series in Operations Research and Management Science*, pages 195–233. Kluwer Academic Publishers, Boston, Massachusetts, 2000.

[2] S. N. Diggavi, N. Al–Dhahir, A. Stamoulis, and A. R. Calderbank. Great Expectations: The Value of Spatial Diversity in Wireless Networks. *Proceedings of the IEEE*, 92(2):219–270, 2004.

[3] M. X. Goemans and D. P. Williamson. Approximation Algorithms for Max–3–Cut and Other Problems via Complex Semidefinite Programming. *Journal of Computer and System Sciences*, 68(2):442–470, 2004.

[4] Y. Huang and S. Zhang. Approximation Algorithms for Indefinite Complex Quadratic Maximization Problems. Technical Report SEEM2005–03, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N. T., Hong Kong, 2005.

[5] J. Jaldén. *Detection for Multiple Input Multiple Output Channels: Analysis of Sphere Decoding and Semidefinite Relaxation.* PhD thesis, School of Electrical Engineering, Royal Institute of Technology (KTH), Stockholm, Sweden, 2006.

[6] J. Jaldén and B. Ottersten. The Diversity Order of the Semidefinite Relaxation Detector. *IEEE Transactions on Information Theory*, 54(4):1406–1422, 2008.

[7] A. T. James. Distributions of Matrix Variates and Latent Roots Derived from Normal Samples. *The Annals of Mathematical Statistics*, 35(2):475–501, 1964.

[8] M. Kisialiou and Z.-Q. Luo. Performance Analysis of Quasi–Maximum–Likelihood Detector Based on Semidefinite Programming. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005 (ICASSP 2005)*, volume 3, pages III–433—III–436, 2005.

[9] M. Kisialiou and Z.-Q. Luo. Probabilistic Analysis of Semidefinite Relaxation for Binary Quadratic Minimization. *SIAM Journal on Optimization*, 20(4):1906–1922, 2010.

[10] Z.-Q. Luo, X. Luo, and M. Kisialiou. An Efficient Quasi–Maximum Likelihood Decoder for PSK Signals. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing, 2003 (ICASSP 2003)*, volume 6, pages VI–561—IV–564, 2003.

[11] W.-K. Ma, P.-C. Ching, and Z. Ding. Semidefinite Relaxation Based Multiuser Detection for M–Ary PSK Multiuser Systems. *IEEE Transactions on Signal Processing*, 52(10):2862–2872, 2004.

[12] W.-K. Ma, T. N. Davidson, K. M. Wong, Z.-Q. Luo, and P.-C. Ching. Quasi–Maximum–Likelihood Multiuser Detection Using Semi–Definite Relaxation with Application to Synchronous CDMA. *IEEE Transactions on Signal Processing*, 50(4):912–922, 2002.

[13] A. Mobasher, M. Taherzadeh, R. Sotirov, and A. K. Khandani. A Near–Maximum–Likelihood Decoding Algorithm for MIMO Systems Based on Semi–Definite Programming. *IEEE Transactions on Information Theory*, 53(11):3869–3886, 2007.

[14] M. P. Quine. A Calculus–Based Proof of a Stirling Formula for the Gamma Function. *International Journal of Mathematical Education in Science and Technology*, 28(6):914–917, 1997.

[15] A. M.-C. So. On the Performance of Semidefinite Relaxation MIMO Detectors for QAM Constellations. In *Proceedings of the 2009 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2009)*, pages 2449–2452, 2009.

[16] A. M.-C. So, Y. Ye, and J. Zhang. A Unified Theorem on SDP Rank Reduction. *Mathematics of Operations Research*, 33(4):910–920, 2008.

[17] A. M.-C. So, J. Zhang, and Y. Ye. On Approximating Complex Quadratic Optimization Problems via Semidefinite Programming Relaxations. *Mathematical Programming, Series B*, 110(1):93–110, 2007.

[18] P. H. Tan and L. K. Rasmussen. The Application of Semidefinite Programming for Detection in CDMA. *IEEE Journal on Selected Areas in Communications*, 19(8):1442–1449, 2001.

[19] D. Tse and P. Viswanath. *Fundamentals of Wireless Communication.* Cambridge University Press, New York, 2005.

[20] S. Verdú. Computational Complexity of Optimum Multiuser Detection. *Algorithmica*, 4:303–312, 1989.

[21] S. Verdú. *Multiuser Detection.* Cambridge University Press, Cambridge, 1998.

A Appendix: Some Probabilistic Estimates

We first state the following concentration result for a chi–square random variable, which will be used repeatedly in the sequel.

Proposition 8 *(cf. [16, Propositions 2.1 and 2.2]) Let $\xi_1, \ldots, \xi_h$ be i.i.d. real standard Gaussian random variables. Let $\alpha \in (1, \infty)$ and $\beta \in (0, 1)$ be constants, and set $U_h = \sum_{i=1}^{h} \xi_i^2$. Note that U_h is a chi–square random variable with h degrees of freedom. Then, the following hold:*

$$\Pr(U_h \geq \alpha h) \leq \exp\left[\frac{h}{2}(1 - \alpha + \ln \alpha)\right],$$

$$\Pr(U_h \leq \beta h) \leq \exp\left[\frac{h}{2}(1 - \beta + \ln \beta)\right].$$

A.1 Trace of a Complex Wishart Matrix

Let $H \in \mathbb{C}^{m \times n}$ (where $m \geq n \geq 2$) be a random matrix whose entries are i.i.d. complex standard Gaussian random variables. Let $h_k \in \mathbb{C}^m$, where $k = 1, \ldots, n$, be the k–th column of the matrix H. Then, we have $\mathrm{tr}(H^*H) = \sum_{k=1}^{n} \|h_k\|_2^2$. Since $h_1, \ldots, h_n$ are independent random vectors with independent entries, we see that $2 \cdot \mathrm{tr}(H^*H)$ is a chi–square random variable with $2mn$ degrees of freedom. In particular, by Proposition 8, we have

$$\Pr_H[\mathrm{tr}(H^*H) \geq 2mn] \leq \exp(-mn/4).$$

A.2 Norm of the Vector of Received Signals

Consider the channel model (1), where the entries of H and v are i.i.d. complex standard Gaussian random variables, with H and v being independent. Note that

$$y_k = \sqrt{\frac{\rho}{n}} \sum_{l=1}^{n} x_l H_{kl} + v_k \qquad \text{for } k = 1, \ldots, m.$$

Since $H_{kl} \sim \mathcal{CN}(0, 1)$ is circular symmetric and $x_l \in \mathcal{S}_M$, we have $x_l H_{kl} \sim \mathcal{CN}(0, 1)$ for $k = 1, \ldots, m$ and $l = 1, \ldots, n$. Moreover, for $k = 1, \ldots, m$, the random variables $x_1 H_{k1}, \ldots, x_n H_{kn}$ are independent. It follows that $y_k \sim \mathcal{CN}(0, \rho + 1)$ for $k = 1, \ldots, m$. In particular, this implies that $(2/(\rho + 1))\|y\|_2^2$ is a chi–square random variable with $2m$ degrees of freedom. Hence, by Proposition 8, we have

$$\begin{aligned} \Pr_{(H,v)}\left[\|y\|_2^2 \leq \frac{1}{2}(\rho + 1)m\right] &\leq \exp(-m/6), \\ \Pr_{(H,v)}\left[\|y\|_2^2 \geq 2(\rho + 1)m\right] &\leq \exp(-m/4). \end{aligned}$$

A.3 The Largest Eigenvalue of a Complex Gaussian Random Matrix

Let $H \in \mathbb{C}^{m \times n}$ (where $m \geq n \geq 2$) be a random matrix whose entries are i.i.d. complex standard Gaussian random variables. Let $\lambda_{max} = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n = \lambda_{min} \geq 0$ be the n eigenvalues of H^*H. The exact joint density function for the n ordered eigenvalues of H^*H is given by (see, e.g., [7, Section 8])

$$C_{m,n} \exp\left(-\sum_{k=1}^{n} \lambda_k\right) \prod_{1 \leq k < k' \leq n} (\lambda_k - \lambda_{k'})^2 \prod_{k=1}^{n} \lambda_k^{m-n} d\lambda_k,$$

where

$$C_{m,n}^{-1} = \prod_{k=1}^{n} \Gamma(m-k+1)\Gamma(n-k+1).$$

Let $f_{\lambda_{max}}$ be the density function for the largest eigenvalue of H^*H, and define

$$R_\lambda = \left\{(\lambda_2, \ldots, \lambda_n) \in \mathbb{R}^{n-1} : \lambda \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0\right\}.$$

Then, we have

$$\begin{aligned}
f_{\lambda_{max}}(\lambda) &= C_{m,n} e^{-\lambda} \lambda^{m-n} \cdot \\
&\quad \left(\int_{R_\lambda} \exp\left(-\sum_{k=2}^{n} \lambda_k\right) \prod_{k=2}^{n} \lambda_k^{m-n} \prod_{2 \leq k < k' \leq n} (\lambda_k - \lambda_{k'})^2 \prod_{k=2}^{n} (\lambda - \lambda_k)^2 \prod_{k=2}^{n} d\lambda_k \right) d\lambda \\
&\leq C_{m,n} e^{-\lambda} \lambda^{m+n-2} \left(\int_{R_\lambda} \exp\left(-\sum_{k=2}^{n} \lambda_k\right) \prod_{k=2}^{n} \lambda_k^{m-n} \prod_{2 \leq k < k' \leq n} (\lambda_k - \lambda_{k'})^2 \prod_{k=2}^{n} d\lambda_k \right) d\lambda \\
&\leq \frac{C_{m,n}}{C_{m-1,n-1}} e^{-\lambda} \lambda^{m+n-2} d\lambda \\
&= \frac{e^{-\lambda} \lambda^{m+n-2}}{\Gamma(m)\Gamma(n)} d\lambda.
\end{aligned}$$

It follows that

$$\begin{aligned}
\Pr_H(\lambda_{max}(H^*H) \geq t) &\leq \frac{1}{\Gamma(m)\Gamma(n)} \int_t^\infty e^{-\lambda} \lambda^{m+n-2} d\lambda \\
&= \frac{e^{-t} t^{m+n-2}}{m!\, n!} \left[1 + \sum_{i=3}^{m+n} \left(\prod_{k=2}^{i-1} \frac{m+n-k}{t} \right) \right].
\end{aligned}$$

By putting $t = 7(m+n)/2$ and using Stirling's formula, we obtain

$$\begin{aligned}
\Pr_H\left(\lambda_{max}(H^*H) \geq \frac{7}{2}(m+n) \right) &< \frac{e^{-7(m+n)/2}(m+n)^{m+n-2}}{2\pi e^{-(m+n)} m^{m+1/2} n^{n+1/2}} \cdot \left(\frac{7}{2}\right)^{m+n-2} \cdot \sum_{i=0}^{m+n-2} \left(\frac{2}{7}\right)^i \\
&< \left(\frac{7e^{-3/2}}{2} \right)^{m+n} \\
&< \exp(-(m+n)/5).
\end{aligned}$$

A.4 The Largest and Smallest Eigenvalues of a Real Gaussian Random Matrix

Let $\tilde{H} \in \mathbb{R}^{2m \times n}$ (where $m \geq n \geq 2$) be a random matrix whose entries are i.i.d. real Gaussian variables with mean 0 and variance 1/2. Let $\lambda_{max} = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n = \lambda_{min} \geq 0$ be the n

eigenvalues of $\tilde{H}^T\tilde{H}$. The exact joint density function for the n ordered eigenvalues of $\tilde{H}^T\tilde{H}$ is given by (see, e.g., [7, Section 7])

$$K_{2m,n}\exp\left(-\sum_{k=1}^{n}\lambda_k\right)\cdot\prod_{1\le k<k'\le n}(\lambda_k-\lambda_{k'})\prod_{k=1}^{n}\lambda_k^{(2m-n-1)/2}d\lambda_k,$$

where

$$K_{2m,n}^{-1}=\pi^{-n/2}\prod_{k=1}^{n}\Gamma\left(\frac{2m-k+1}{2}\right)\Gamma\left(\frac{n-k+1}{2}\right).$$

Let $f_{\lambda_{max}}$ be the density function for the largest eigenvalue of $\tilde{H}^T\tilde{H}$, and define

$$R_\lambda=\left\{(\lambda_2,\ldots,\lambda_n)\in\mathbb{R}^{n-1}:\lambda\ge\lambda_2\ge\cdots\ge\lambda_n\ge 0\right\}.$$

Then, we have

$$\begin{aligned}
f_{\lambda_{max}}(\lambda) &= K_{2m,n}e^{-\lambda}\lambda^{(2m-n-1)/2}\cdot \\
&\quad\left(\int_{R_\lambda}\exp\left(-\sum_{k=2}^{n}\lambda_k\right)\prod_{k=2}^{n}\lambda_k^{(2m-n-1)/2}\prod_{2\le k<k'\le n}(\lambda_k-\lambda_{k'})\prod_{k=2}^{n}(\lambda-\lambda_k)d\lambda_k\right)d\lambda \\
&\le K_{2m,n}e^{-\lambda}\lambda^{(2m+n-3)/2}\cdot \\
&\quad\left(\int_{R_\lambda}\exp\left(-\sum_{k=2}^{n}\lambda_k\right)\prod_{k=2}^{n}\lambda_k^{(2m-n-1)/2}\prod_{2\le k<k'\le n}(\lambda_k-\lambda_{k'})\prod_{k=2}^{n}d\lambda_k\right)d\lambda \\
&\le \frac{K_{2m,n}}{K_{2m-1,n-1}}e^{-\lambda}\lambda^{(2m+n-3)/2}d\lambda \\
&= \frac{\sqrt{\pi}\cdot e^{-\lambda}\lambda^{(2m+n-3)/2}}{\Gamma(m)\Gamma(n/2)}d\lambda.
\end{aligned}$$

It follows that

$$\begin{aligned}
\Pr_{\tilde{H}}(\lambda_{max}(\tilde{H}^T\tilde{H})\ge t) &\le \frac{\sqrt{\pi}}{\Gamma(m)\Gamma(n/2)}\int_t^\infty e^{-\lambda}\lambda^{(2m+n-3)/2}d\lambda \\
&\le \frac{\sqrt{\pi}\cdot e^{-t}t^{(2m+n-3)/2}}{\Gamma(m)\Gamma(n/2)}\cdot\left[1+\sum_{i=2}^{\lceil\frac{2m+n-1}{2}\rceil}\left(\prod_{k=1}^{i-1}\frac{2m+n-(2k+1)}{2t}\right)\right].
\end{aligned}$$

Now, set $t=2(2m+n)$. To bound the quantity on the right–hand side, we use the Stirling formula for the Gamma function, i.e.,

$$\sqrt{2\pi}\cdot x^{x-1/2}e^{-x}<\Gamma(x)<\frac{6}{5}\sqrt{2\pi}\cdot x^{x-1/2}e^{-x} \tag{33}$$

which is valid for all $x \geq 1/2$ (see, e.g., [14]). In particular, we obtain

$$\begin{aligned}
\Pr_{\tilde{H}}(\lambda_{max}(\tilde{H}^T\tilde{H}) \geq 2(2m+n)) &< \frac{e^{-2(2m+n)}(2m+n)^{m+n/2-3/2}2^{m+n/2-3/2}}{2\sqrt{\pi}e^{-(m+n/2)}m^{m+1/2}(n/2)^{(n-1)/2}} \cdot \sum_{i=0}^{\infty}\frac{1}{4^i} \\
&< \exp(-(m+n/2)) \cdot 2^{m+n/2} \\
&< \exp(-(m+n/2)/4).
\end{aligned}$$

Similarly, let $f_{\lambda_{min}}$ be the density function for the smallest eigenvalue of $\tilde{H}^T\tilde{H}$, and define

$$S_\lambda = \left\{(\lambda_1,\ldots,\lambda_{n-1}) \in \mathbb{R}^{n-1} : \lambda_1 \geq \cdots \geq \lambda_{n-1} \geq \lambda \geq 0\right\}.$$

Then, we have

$$\begin{aligned}
f_{\lambda_{min}}(\lambda) &= K_{2m,n}e^{-\lambda}\lambda^{(2m-n-1)/2} \cdot \\
&\quad \left(\int_{S_\lambda} \exp\left(-\sum_{k=1}^{n-1}\lambda_k\right)\prod_{k=1}^{n-1}\lambda_k^{(2m-n-1)/2}\prod_{1\leq k<k'\leq n-1}(\lambda_k-\lambda_{k'})\prod_{k=1}^{n-1}(\lambda_k-\lambda)d\lambda_k\right)d\lambda \\
&\leq K_{2m,n}e^{-\lambda}\lambda^{(2m-n-1)/2} \cdot \\
&\quad \left(\int_{S_\lambda} \exp\left(-\sum_{k=1}^{n-1}\lambda_k\right)\prod_{k=1}^{n-1}\lambda_k^{(2m-n+1)/2}\prod_{1\leq k<k'\leq n-1}(\lambda_k-\lambda_{k'})\prod_{k=1}^{n-1}d\lambda_k\right)d\lambda \\
&\leq \frac{K_{2m,n}}{K_{2m+1,n-1}}e^{-\lambda}\lambda^{(2m-n-1)/2}d\lambda \\
&= \frac{\sqrt{\pi}\cdot\Gamma(m+1/2)}{\Gamma(m-(n-1)/2)\Gamma(m-(n-2)/2)\Gamma(n/2)}e^{-\lambda}\lambda^{(2m-n-1)/2}d\lambda. \qquad (34)
\end{aligned}$$

Now, we use the Stirling formula for the Gamma function (see (33)) to bound the quantity on the right–hand side. First, we compute

$$\begin{aligned}
\frac{\sqrt{\pi}\cdot\Gamma(m+1/2)}{\Gamma(m-(n-1)/2)\Gamma(m-(n-2)/2)\Gamma(n/2)} &< \frac{3e}{5\sqrt{\pi}}\cdot(2e)^{m-n/2}\cdot\frac{(2m+1)^m(2m-n+1)^{n/2}}{(2m-n+1)^m}\cdot \\
&\quad (2m-n+2)^{-(m-n/2+1/2)}\cdot n^{(1-n)/2}.
\end{aligned}$$

Observe that

$$\begin{aligned}
\frac{(2m+1)^m(2m-n+1)^{n/2}}{(2m-n+1)^m} &= \left(1+\frac{n}{2m-n+1}\right)^{2m-n+1}\cdot\frac{(2m-n+1)^{m-n/2+1}}{(2m+1)^{m-n+1}} \\
&< e^n(2m+1)^{n/2}.
\end{aligned}$$

Upon putting the pieces together, we obtain

$$\begin{aligned}
\frac{\sqrt{\pi}\cdot\Gamma(m+1/2)}{\Gamma(m-(n-1)/2)\Gamma(m-(n-2)/2)\Gamma(n/2)} &< \frac{2^{m-n/2}e^{m+n/2}(2m+1)^{n/2}n^{(1-n)/2}}{(2m-n+2)^{m-n/2+1/2}} \\
&< \frac{\sqrt{n}\cdot 2^{m-n/2}e^{2m+n/2}}{(2m-n+2)^{m-n/2+1/2}}. \qquad (35)
\end{aligned}$$

Now, set $t = e^{-8}(m+1)$ in (34). We bound

$$\int_0^{e^{-8}(m+1)} e^{-\lambda}\lambda^{(2m-n-1)/2}d\lambda \le \frac{2}{2m-n+1}\left(\frac{m+1}{e^8}\right)^{(2m-n+1)/2} \le \left(\frac{2m+2}{2e^8}\right)^{(2m-n+1)/2}. \tag{36}$$

Upon substituting (35) and (36) into (34), we obtain

$$\begin{aligned}
\Pr_{\tilde{H}}\left(\lambda_{min}(\tilde{H}^T\tilde{H}) \le \frac{2m+2}{2e^8}\right) &< \frac{\sqrt{n}\cdot 2^{m-n/2}e^{2m+n/2}}{2^{m-n/2+1/2}e^{4(2m-n+1)}}\cdot\frac{(2m+2)^{m-n/2+1/2}}{(2m-n+2)^{m-n/2+1/2}} \\
&< \frac{\sqrt{n}\cdot 2^{m-n/2}e^{2m+n/2}\cdot e^n}{2^{m-n/2+1/2}e^{4(2m-n+1)}} \\
&< \sqrt{m}\cdot\exp(-m/2) \\
&< 2\exp(-m/4).
\end{aligned}$$

A.5 Norm of the Vector $\tilde{H}^T\tilde{v}$

Consider the channel model (1), where the entries of H and v are i.i.d. complex standard Gaussian random variables, with H and v being independent. Let $\tilde{H}$ and $\tilde{v}$ be given by (22). Observe that

$$\left(\tilde{H}^T\tilde{v}\right)_k = \sum_{i=1}^{2m}\tilde{H}_{ik}\tilde{v}_i \qquad \text{for } k=1,\ldots,n.$$

Thus, given a realization of $\tilde{H}$, we see that $(\tilde{H}^T\tilde{v})_k$ is a Gaussian random variable with mean 0 and variance $\sigma_k^2 \equiv (1/2)\sum_{i=1}^{2m}\tilde{H}_{ik}^2$, for $k=1,\ldots,n$. It follows that

$$\Pr_{\tilde{v}}\left(\|\tilde{H}^T\tilde{v}\|_\infty \ge t\right) \le \sum_{k=1}^n \Pr_{\tilde{v}}\left(|(\tilde{H}^T\tilde{v})_k| \ge t\right) \le \sqrt{\frac{2}{\pi}}\sum_{k=1}^n\frac{\sigma_k}{t}\exp\left(-t^2/2\sigma_k^2\right).$$

Now, note that $4\sigma_k^2$ is a chi–square random variable with $2m$ degrees of freedom. Thus, by standard concentration results for the chi–square random variable (see Proposition 8), we have

$$\Pr_{\tilde{H}}\left(\max_{1\le k\le n}\sigma_k^2 \ge m\right) \le n\cdot\exp(-m/4) \le 4\exp(-m/8)$$

(recall that $m \ge n$). In particular, we conclude that

$$\begin{aligned}
\Pr_{(\tilde{H},\tilde{v})}\left(\|\tilde{H}^T\tilde{v}\|_\infty \ge t\right) &= \Pr_{(\tilde{H},\tilde{v})}\left(\|\tilde{H}^T\tilde{v}\|_\infty \ge t,\ \max_{1\le k\le n}\sigma_k^2 < m\right) \\
&+ \Pr_{(\tilde{H},\tilde{v})}\left(\|\tilde{H}^T\tilde{v}\|_\infty \ge t,\ \max_{1\le k\le n}\sigma_k^2 \ge m\right) \\
&\le \sqrt{\frac{2m}{\pi}}\cdot\frac{n}{t}\cdot\exp(-t^2/2m)\cdot(1-4\exp(-m/8)) + 4\exp(-m/8).
\end{aligned}$$

Upon setting $t = m^\gamma$ for some $\gamma > 1/2$, we conclude that

$$\Pr_{(\tilde{H},\tilde{v})}\left(\|\tilde{H}^T\tilde{v}\|_\infty \ge m^\gamma\right) \le \sqrt{\frac{2}{\pi}}\cdot n\cdot\exp\left(-m^{2\gamma-1}/2\right) + 4\exp(-m/8),$$

as desired.