PERFORMANCE ANALYSIS OF QUASI-MAXIMUM-LIKELIHOOD DETECTOR BASED ON SEMI-DEFINITE PROGRAMMING

Mikalai Kisialiou and Zhi-Quan Luo

Department of Electrical and Computer Engineering
University of Minnesota, Minneapolis, MN, 55455, U.S.A.
Email:{rainbow,luozq}@ece.umn.edu

ABSTRACT

Despite its optimal bit-error-rate (BER) performance, the maximum-likelihood (ML) detection is known to be NP-hard and suffers from high computational complexity. The currently popular suboptimal detectors either achieve a polynomial time complexity at the expense of BER performance degradation (e.g., MMSE Detector), or offer a near ML performance with a complexity that is exponential in the worst case. This paper considers a highly efficient (polynomial worst case complexity) quasi-ML detection method based on Semi-Definite (SDP) relaxation. It is shown that, for a standard vector Rayleigh fading channel, this SDP-based quasi-ML detector achieves, in the high signal-to-noise ratio (SNR) region, a BER which is identical to that of the exact ML detector. In the low SNR region we use the random matrix theory to show that the SDP-based detector serves as a constant factor approximation to the ML detector for large systems.

1. INTRODUCTION

Consider a Rayleigh fading vector communication channel with n transmit and m receive antennas:

$$\mathbf{y} = \sqrt{\rho/n}\,\mathbf{Hs} + \mathbf{v}, \tag{1}$$

where ρ is expected SNR at each receive antenna, $\mathbf{s} \in \mathcal{C}^n$ is the vector of modulated transmitted signals and $\mathcal{C}^n$ denotes a constellation of dimension n, $\mathbf{y} \in \mathbb{R}^m$ is the real-valued vector of received signals, $\mathbf{H} \in \mathbb{R}^{m\times n}$ is the matrix of fading coefficients, $H_{ik} \sim \mathcal{N}(0,1),\ \forall\ i,k$, and $\mathbf{v} \in \mathbb{R}^m$ is additive white Gaussian noise (AWGN), $v_i \sim \mathcal{N}(0,1),\ \forall\ i$. Notice that the above channel model (1) is quite generic and can be used to describe other communication systems, such as a synchronous CDMA multi-access channel, with n representing the number of users.

For a memoryless channel with equiprobable input signals, the maximum-likelihood detection achieves the minimum error probability of detection. In general, ML Detector solves the following optimization problem:

$$\mathbf{s}_{ML} = \arg\max_{\mathbf{s}\in\mathcal{C}^n} p(\mathbf{y}|\mathbf{s},\mathbf{H}),$$

where $p(\mathbf{y}|\mathbf{s},\mathbf{H})$ is the conditional probability density function of $\mathbf{y}$ and $\mathbf{s}_{ML}$ is the ML estimate of the transmitted signals. For the Gaussian noise ML detection can be written in the form:

$$\mathbf{s}_{ML} = \arg\min_{\mathbf{s}\in\mathcal{C}^n} \|\mathbf{y} - \sqrt{\rho/n}\,\mathbf{Hs}\|^2. \tag{2}$$

This research is supported in part by the National Science Foundation, Grant No. DMS-0312416.

In a practical communication system, constellation $\mathcal{C}^n$ consists of discrete n-dimensional vectors $\mathbf{s}$. This discrete structure generally causes the problem (2) to be NP-hard. For systems with small n, exhaustive search can be applied to determine the ML solution. However, with relatively large n, m, such approach becomes prohibitively expensive. For example, in decoding the so called Linear Dispersion Codes [1] for a wireless channel, one must solve (2) for an equivalent channel model with dimensions up to 100. In that situation, the exhaustive search is no longer practically feasible. Instead, efficient suboptimal detection algorithms must be used. In other words, we need to design computationally efficient detectors (i.e., with polynomial time complexity) which can achieve BER close to the one provided by ML Detector.

One of the popular detectors, the so called Sphere Decoder [2], originated from the work in [3], runs exhaustive search in the region around zero-forcing solution to localize the ML solution. Although Sphere Decoder offers excellent practical performance, its expected complexity is exponential in the problem size [4]. As an alternative, quasi-ML detector [5] based on SDP relaxation and PSK Decoder [6] based on low rank relaxation provide BER close to that of ML detector in the high SNR region. The SDP-based detector (called SDP Detector hereafter) solves an SDP problem of polynomial complexity by means of recently developed Interior Point methods. The BER performance of SDP detector compares favorably to that of MMSE Detector since the latter relaxes ML detection to a linear problem. This paper presents a probabilistic analysis of the optimality conditions for SDP Detector. Our analysis in the high SNR region leads us to a sufficient condition for SDP Detector to output the ML solution. In the low SNR region, we use the classical results on the limiting eigenvalue distribution of large random matrices [7] to establish a constant factor optimality for the SDP Detector.

We adopt the following notations: bold letters denote vectors while bold capital letters will signify matrices.

2. SDP DETECTOR

We will focus on the BPSK case when $\mathcal{C}^n = \{-1,+1\}^n$. First of all, we would like to linearize the objective function, so we proceed in the following way:

$$\|\mathbf{y} - \sqrt{\rho/n}\,\mathbf{Hs}\|^2 = \mathrm{Trace}(\mathbf{Qxx}^T),$$

where matrix $\mathbf{Q} \in \mathcal{R}^{(n+1)\times(n+1)}$ and vector $\mathbf{x} \in \mathcal{R}^{n+1}$ are defined as

$$\mathbf{Q} = \begin{bmatrix} (\rho/n)\,\mathbf{H}^T\mathbf{H} & -\sqrt{\rho/n}\,\mathbf{H}^T\mathbf{y} \\ -\sqrt{\rho/n}\,\mathbf{y}^T\mathbf{H} & \|\mathbf{y}\|^2 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} \mathbf{s} \\ 1 \end{bmatrix} \tag{3}$$

Thus, problem (2) can be reformulated as

$$\begin{array}{rl} f_{ML} := \min & \text{Trace}(\mathbf{QX}) \\ \text{s.t.} & \mathbf{X} = \mathbf{xx}^T \\ & X_{i,i} = 1,\ i = 1, \ldots, n+1. \end{array} \quad (4)$$

The new problem formulation (4) is insensitive to the sign of vector $\mathbf{x}$, hence, if we obtain solution $\hat{\mathbf{x}}$ with the last entry equal to -1, then we simply declare $-\hat{\mathbf{x}}$ as the solution of (4) for consistency with (3). The constraints in (4) can be written without variable $\mathbf{x}$ if we impose rank-1 constraint on matrix $\mathbf{X}$ and make sure that it is positive semi-definite (PSD). This observation leads to the problem with linear objective, linear and PSD constraints:

$$\begin{array}{rl} f_{ML} := \min & \text{Trace}(\mathbf{QX}) \\ \text{s.t.} & \mathbf{X} \succeq 0 \\ & \mathbf{X} \text{ is rank-1} \\ & X_{i,i} = 1,\ i = 1, \ldots, n+1. \end{array} \quad (5)$$

This problem is equivalent to the ML detection problem in (2) over BPSK constellation, therefore, it is still NP-hard in general, although there is no explicit integer constraint. This simple reformulation allows us to spot the constraint making the problem difficult. Rank-1 constraint is the only non-convex constraint in (5), so we drop it and solve the convex optimization problem:

$$\begin{array}{rl} f_{SDP} := \min & \text{Trace}(\mathbf{QX}) \\ \text{s.t.} & \mathbf{X} \succeq 0, \\ & X_{i,i} = 1,\ i = 1, \ldots, n+1. \end{array} \quad (6)$$

The problem that naturally arises after solving (6) is that the solution $\mathbf{X}_{opt}$ is no longer restricted to be rank-1, so we still have to project optimal matrix $\mathbf{X}_{opt}$ to the set of rank-1 matrices. It turns out that solution matrix $\mathbf{X}_{opt}$ with high probability comprises enough information about ML solution $\mathbf{x}_{ML}$. There may be several ways to generate an estimate $\hat{\mathbf{x}}_{SDR}$ of transmitted signals from matrix $\mathbf{X}_{opt}$. In this paper we consider the randomized procedure that can be described by the following algorithm:

1. Take a spectral decomposition $\mathbf{X}_{opt} = \sum_{i=1}^{n+1} \lambda_i \mathbf{u}_i \mathbf{u}_i^T$ and let $\mathbf{v}_i = \sqrt{\lambda_i}\, \mathbf{u}_i,\ i = 1, \ldots, n+1$.
2. Pick k such that vector $\mathbf{v}_k$ corresponds to the largest eigenvalue: $\mathbf{v}_k = \arg\max_{1 \le i \le n+1}\{\|\mathbf{v}_i\|\}$.
3. Define distribution $P_{\mathbf{x}}$:

$$\begin{array}{l} \Pr\{x_i = +1\} = (1 + v_{ki})/2, \\ \Pr\{x_i = -1\} = (1 - v_{ki})/2. \end{array} \quad (7)$$

4. Generate L i.i.d. vector samples $\bar{\mathbf{x}}_l,\ l = 1, \ldots, L$, such that each entry $\bar{x}_{li},\ i = 1, \ldots, n+1$ is drawn from the distribution defined in (7).
5. For all L samples set $\bar{\mathbf{x}}_l := -\bar{\mathbf{x}}_l$ if (n+1)-st entry of $\bar{\mathbf{x}}_l$ is equal to -1;
6. Pick $\hat{\mathbf{x}} := \arg\min_l \bar{\mathbf{x}}_l^T \mathbf{Q} \bar{\mathbf{x}}_l$ and assign $f_{SDR} := \hat{\mathbf{x}}^T \mathbf{Q} \hat{\mathbf{x}}$.
7. Output quasi-ML estimate $\hat{\mathbf{x}}_{SDR}$ which is vector $\hat{\mathbf{x}}$ with the last bit discarded.

3. MAIN RESULTS

In the sequel we assume that $m = n$. Since SDP Detector (SDP solver of (6) followed by a randomized rounding procedure) represents an approximation algorithm for the original ML detection problem, we are interested in its optimality criteria.

Theorem 1 *For a given ρ and n, the solution $\boldsymbol{X}_{opt}$ of relaxed problem (6) is rank-1, if channel matrix $\boldsymbol{H}$ and noise $\boldsymbol{v}$ satisfy the following inequality:*

$$\lambda_{min}(\boldsymbol{H}^T\boldsymbol{H}) > \sqrt{n/\rho}\, \|\boldsymbol{H}^T\boldsymbol{v}\|_1, \quad (8)$$

where $\lambda_{min}(\boldsymbol{H}^T\boldsymbol{H})$ denotes the minimum eigenvalue of $\boldsymbol{H}^T\boldsymbol{H}$ and 1-norm of any vector $\boldsymbol{a}$ is defined as $\|\boldsymbol{a}\|_1 = \sum_i |a_i|$.

Since a random channel matrix is full rank with probability 1, we can also interpret this claim as follows: for most channel matrices $\mathbf{H}$ and noise vectors $\mathbf{v}$ there exists a sufficiently high SNR such that the solution of SDP solver (6) is rank-1 and sufficiency is determined by (8).

We can see that the solution of SDP problem (6) falls in the feasible set of the original ML detection problem (5) when condition (8) is satisfied. Since SDP problem is the relaxed version of ML detection problem, then solution $\mathbf{X}_{opt}$ of SDP problem must also be the solution of the ML detection problem, that is, SDP Detector solves the ML detection problem $\mathbf{X}_{opt} = \mathbf{x}_{ML}\mathbf{x}_{ML}^T$ (in polynomial time) when (8) holds true. For large systems SDP Detector can provide BER arbitrary close to that of ML Detector by choosing a sufficiently high SNR. This result is summarized in the theorem which is stated here without a proof.

Theorem 2 *Suppose* $\lim_{n\to\infty} \rho^{-1} n^6 = 0$. *Then*

$$\Pr\left\{\lambda_{\min}(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}) > \sqrt{\mathrm{n}/\rho}\, \|\boldsymbol{H}^{\mathrm{T}}\boldsymbol{v}\|_1\right\} \to 1, \quad \textit{as } n \to \infty.$$

The results in Theorems 1 and 2 are important for communication systems operating in the high SNR region. The next theorem characterizes the asymptotic behavior of SDP Detector in the low SNR region for large systems. We will use notation $\le^P (=^P)$ to imply that inequality(equality) is satisfied in probability as $n \to \infty$, that is, for two random variables a and b we write $a \le^P b$ if $\Pr\{a \le b\} \to 1$ with n going to infinity.

Theorem 3 *The objective value of SDP Detector f_{SDR} is within a constant factor $c(\rho)$ away from the ML objective value in probability:*

$$\frac{f_{SDR}}{n} \le^P c(\rho)\frac{f_{ML}}{n},\ \ c(\rho) = 1 + \frac{9\rho(2\rho+1)}{\sqrt{12\rho+1}-1}. \quad (9)$$

We can see from (9) that $\lim_{\rho\to 0} c(\rho) = 5/2$, which means that in probability SDP Detector provides a $5/2$-approximation algorithm for the ML detection problem in low SNR region. Next section will be devoted to proving these claims.

4. PERFORMANCE ANALYSIS OF SDP DETECTOR

In our notations $\mathbf{e}$ ($\mathbf{E}$) will denote the vector (matrix) with all entries equal to 1. Since $\mathbf{H}$ is a Gaussian zero-mean matrix, we can assume that $\mathbf{s} = \mathbf{e}$ has been transmitted due to symmetry.

Lemma 1 *For any full rank matrix $\boldsymbol{H}$, we have*

$$\lim_{\rho\to\infty} f_{SDP}(\boldsymbol{X}_{opt})/\rho = 0, \textit{ and } \lim_{\rho\to\infty} \boldsymbol{X}_{opt} = \boldsymbol{E}.$$

Proof of Lemma 1: Note that, although $\mathbf{X} = \mathbf{E}$ may not be the optimal solution of SDP problem (6), it is always feasible and, therefore, the objective value at the optimum is always less than or equal to $f_{SDP}(\mathbf{E})$. After simple derivations we can see that $f_{SDP}(\mathbf{E}) = \text{Trace}(\mathbf{QE}) = \|\mathbf{v}\|^2$. Since both $\mathbf{Q}$ and any feasible $\mathbf{X}$ are PSD matrices, then $f_{SDP}(\mathbf{X}_{opt}) \geq 0$. That is, we always have $0 \leq f_{SDP}(\mathbf{X}_{opt}) \leq \|\mathbf{v}\|^2$. Divide it by ρ and obtain $f_{SDP}(\mathbf{X}_{opt})/\rho \to 0$ as $\rho \to \infty$.

For convenience we would also like to represent the optimal matrix $\mathbf{X}_{opt}$ in the block form:

$$\mathbf{X}_{opt} = \begin{bmatrix} \mathbf{Z} & \mathbf{z} \\ \mathbf{z}^T & 1 \end{bmatrix} \succeq 0, \tag{10}$$

where block $\mathbf{Z} \in \mathcal{R}^{n\times n}$ is also a PSD matrix with ones on the diagonal. Evaluation of ratio $f_{SDP}(\mathbf{X}_{opt})/\rho$ in terms of blocks $\mathbf{Z}$, $\mathbf{z}$ and ρ leads to the following expression:

$$\begin{aligned} f_{SDP}(\mathbf{X}_{opt})/\rho &= \text{Trace}(\mathbf{Q}\mathbf{X}_{opt})/\rho \\ &= (1/n)\ \text{Trace}\left(\mathbf{H}^T\mathbf{H}(\mathbf{Z} - \mathbf{z}\mathbf{z}^T)\right) \\ &\quad + \|\sqrt{1/n}\,\mathbf{H}(\mathbf{e} - \mathbf{z}) + \sqrt{1/\rho}\,\mathbf{v}\|^2. \end{aligned} \tag{11}$$

Using Schur complement for PSD matrix $\mathbf{X}_{opt}$ we can see that $\mathbf{Z} - \mathbf{z}\mathbf{z}^T$ is PSD matrix and, hence, both terms in (11) are nonnegative and must vanish with ρ going to infinity. If $\mathbf{H}$ is full rank, we have $\mathbf{z} \to \mathbf{e}$ as $\rho \to \infty$ to ensure that the second term in (11) goes to 0. For the first term with full rank matrix $\mathbf{H}^T\mathbf{H}$ we must have $\mathbf{Z} \to \mathbf{z}\mathbf{z}^T$ to ensure that trace goes to 0. Therefore, $\mathbf{X}_{opt} \to \mathbf{E}$ as $\rho \to \infty$. □

Proof of Theorem 1: Define a dual variable $\mathbf{G}$. One of KKT optimality conditions for problem (6) is

$$(\mathbf{Q} - \mathbf{G})\mathbf{X}_{opt} = \mathbf{0}. \tag{12}$$

Now, we represent $\mathbf{G}$ in the same block form as $\mathbf{X}_{opt}$ in (10):

$$\mathbf{G} = \begin{bmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & d \end{bmatrix}.$$

Then, the last column of the first equation (12) will be described by the system of equations:

$$\begin{aligned} &\tfrac{\rho}{n}\mathbf{H}^T\mathbf{H}\mathbf{z} - \tfrac{\rho}{n}\mathbf{H}^T\mathbf{H}\mathbf{e} - \mathbf{D}\mathbf{z} - \sqrt{\tfrac{\rho}{n}}\mathbf{H}^T\mathbf{v} = \mathbf{0} \\ &\tfrac{\rho}{n}\mathbf{e}^T\mathbf{H}^T\mathbf{H}\mathbf{e} - \tfrac{\rho}{n}\mathbf{e}^T\mathbf{H}^T\mathbf{H}\mathbf{z} - \sqrt{\tfrac{\rho}{n}}\mathbf{v}^T\mathbf{H}\mathbf{z} \\ &\qquad + \|\mathbf{v}\|^2 + 2\sqrt{\tfrac{\rho}{n}}\mathbf{v}^T\mathbf{H}\mathbf{e} - d = 0, \end{aligned}$$

Define $\Delta\mathbf{z} = \mathbf{e} - \mathbf{z} \geq 0$, where the componentwise inequality follows from feasibility constraints of (6). The diagonal entries of dual variables $\mathbf{D}$ and d can be written as functions of difference $\Delta\mathbf{z}$:

$$\frac{D_{i,i}}{\rho} = -\frac{1}{z_i}\left[\frac{1}{n}\,\mathbf{H}^T\mathbf{H}\Delta\mathbf{z} + \frac{1}{\sqrt{n\rho}}\mathbf{H}^T\mathbf{v}\right]_i,\ i = 1,\ldots,n,$$

$$\frac{d}{\rho} = \frac{1}{n}\mathbf{e}^T\mathbf{H}^T\mathbf{H}\Delta\mathbf{z} + \frac{1}{\sqrt{n\rho}}\mathbf{v}^T\mathbf{H}\Delta\mathbf{z} + \frac{1}{\rho}\|\mathbf{v}\|^2 + \frac{1}{\sqrt{n\rho}}\mathbf{v}^T\mathbf{H}\mathbf{e}. \tag{13}$$

If inequality $\lambda_{min}(\mathbf{H}^T\mathbf{H}) > \sqrt{n/\rho}\,\|\mathbf{H}^T\mathbf{v}\|_1$ holds, then matrix $\mathbf{H}$ must be full rank, and according to Lemma 1 we have $\Delta\mathbf{z} \to \mathbf{0}$ as $\rho \to \infty$. For sufficiently big ρ it leads to:

$$\frac{1}{z_i} = \frac{1}{1 - \Delta z_i} = 1 + \Delta z_i + O\left((\Delta z_i)^2\right).$$

So, the asymptotic behavior of the sum $\frac{1}{\rho}\sum_{i=1}^n D_{i,i}$ can be expressed as

$$\begin{aligned} \tfrac{1}{\rho}\textstyle\sum_{i=1}^n D_{i,i} = &- \tfrac{1}{n}\mathbf{e}^T\mathbf{H}^T\mathbf{H}\Delta\mathbf{z} - \tfrac{1}{\sqrt{n\rho}}\mathbf{v}^T\mathbf{H}\mathbf{e} \\ &- \tfrac{1}{n}(\Delta\mathbf{z})^T\mathbf{H}^T\mathbf{H}\Delta\mathbf{z} - \tfrac{1}{\sqrt{n\rho}}\mathbf{v}^T\mathbf{H}\Delta\mathbf{z} \\ &+ O\left(\tfrac{1}{\sqrt{n\rho}}\|\Delta\mathbf{z}\|^2 + \|\Delta\mathbf{z}\|^3\right). \end{aligned} \tag{14}$$

From the strong duality we have $f_{SDP}/\rho = (d + \sum_{i=1}^n D_{i,i})/\rho$, hence, summing (13) and (14) we obtain the asymptotic of the objective function:

$$\begin{aligned} \tfrac{1}{\rho}f_{SDP} = &\tfrac{1}{\rho}\|\mathbf{v}\|^2 - \tfrac{1}{n}(\Delta\mathbf{z})^T\mathbf{H}^T\mathbf{H}\Delta\mathbf{z} \\ &+ O\left(\tfrac{1}{\sqrt{n\rho}}\|\Delta\mathbf{z}\|^2 + \|\Delta\mathbf{z}\|^3\right). \end{aligned} \tag{15}$$

At the same time, from the exact expression of the objective function in (11) we can derive the lower bound

$$\begin{aligned} \tfrac{f_{SDP}}{\rho} &= \tfrac{1}{n}\text{Trace}(\mathbf{H}^T\mathbf{H}(\mathbf{Z} - \mathbf{z}\mathbf{z}^T)) + \left\|\tfrac{1}{\sqrt{n}}\mathbf{H}\Delta\mathbf{z} + \tfrac{1}{\sqrt{\rho}}\mathbf{v}\right\|^2 \\ &\geq \tfrac{\lambda_{min}(\mathbf{H}^T\mathbf{H})}{n}(n - \|\mathbf{z}\|^2) + \left\|\tfrac{1}{\sqrt{n}}\mathbf{H}\Delta\mathbf{z} + \tfrac{1}{\sqrt{\rho}}\mathbf{v}\right\|^2 \\ &\geq \tfrac{\lambda_{min}(\mathbf{H}^T\mathbf{H})}{n}(2\mathbf{e}^T\Delta\mathbf{z} - \|\Delta\mathbf{z}\|^2) + \tfrac{1}{\rho}\|\mathbf{v}\|^2 \\ &\quad + \tfrac{1}{n}(\Delta\mathbf{z})^T\mathbf{H}^T\mathbf{H}\Delta\mathbf{z} - \tfrac{2}{\sqrt{n\rho}}\|\mathbf{H}^T\mathbf{v}\|_1\|\Delta\mathbf{z}\|_1. \end{aligned}$$

Now, we notice that terms $\|\Delta\mathbf{z}\|^2$ and $\Delta\mathbf{z}^T\mathbf{H}^T\mathbf{H}\Delta\mathbf{z}$ are definitely high order term compared to $\mathbf{e}^T\Delta\mathbf{z}$ because we know that all entries of $\Delta\mathbf{z}$ are nonnegative. Taking into account the result in (15) we can write:

$$\frac{1}{\sqrt{n\rho}}\ \|\mathbf{H}^T\mathbf{v}\|_1\|\Delta\mathbf{z}\|_1 \geq \frac{1}{n}\lambda_{min}(\mathbf{H}^T\mathbf{H})\mathbf{e}^T\Delta\mathbf{z} + O\left(\|\Delta\mathbf{z}\|^2\right).$$

Note that $\mathbf{e}^T\Delta\mathbf{z} = \|\Delta\mathbf{z}\|_1$, which leads to

$$\sqrt{\frac{n}{\rho}}\ \|\mathbf{H}^T\mathbf{v}\|_1\|\Delta\mathbf{z}\|_1 \geq \lambda_{min}(\mathbf{H}^T\mathbf{H})\|\Delta\mathbf{z}\|_1 + O\left(\|\Delta\mathbf{z}\|^2\right). \tag{16}$$

According to Lemma 1 we have $\Delta\mathbf{z} \to 0$ as $\rho \to \infty$ and the higher order terms denoted as $O\left(\|\Delta\mathbf{z}\|^2\right)$ become negligible. For sufficiently big ρ we must have

$$\sqrt{\frac{n}{\rho}}\ \|\mathbf{H}^T\mathbf{v}\|_1\ <\ \lambda_{min}(\mathbf{H}^T\mathbf{H}), \tag{17}$$

since $\mathbf{H}$ is a full rank matrix. Then, for such ρ, $\|\Delta\mathbf{z}\|_1$ should be equal to 0 for optimality condition (16) to hold. This implies that $\mathbf{z} = \mathbf{e}$ whenever (17) holds. Now, recall that $\mathbf{Z} - \mathbf{z}\mathbf{z}^T$ must be PSD matrix, and feasibility constraint requires $\mathbf{Z}$ to have ones on the diagonal. Therefore, PSD matrix $\mathbf{Z} - \mathbf{e}\mathbf{e}^T$ will have zeros on the diagonal which can be true only if $\mathbf{Z} - \mathbf{e}\mathbf{e}^T = \mathbf{0}$. Thus, all entries of $\mathbf{X}_{opt}$ are equal to 1 which is rank-1 matrix. □

Lemma 2 *For any fixed $\rho \geq 0$, the optimal objective value f_{SDP} of SDP (6) satisfies:*

$$\frac{1}{n}f_{SDP} \geq^P \frac{\sqrt{12\rho + 1} - 1}{9\rho}$$

in probability as $n \to \infty$.

Proof of Lemma 2: The dual problem for (6) has the form:

$$\begin{aligned} f_{SDP} := \max\ &\text{Trace}(\mathbf{G}) \\ \text{s.t.}\ &\mathbf{Q} - \mathbf{G} \succeq 0, \\ &\mathbf{G} \text{ is diagonal.} \end{aligned}$$

Consider a sequence of dual feasible matrices $\mathbf{G}_{n+1}$ of dimension $n+1$ parameterized by α and β which will be chosen later to ensure dual feasibility:

$$\mathbf{G}_{n+1} = \begin{bmatrix} -\alpha\mathbf{I}_n & \mathbf{0} \\ \mathbf{0}^T & \beta \end{bmatrix}.$$

The objective value at $\mathbf{G}_{n+1}$ is given by $f(\mathbf{G}_{n+1}) = \beta - n\alpha$ and serves as a lower bound on f_{SDP}. To ensure that

$$\mathbf{Q} - \mathbf{G}_{n+1} = \begin{bmatrix} (\rho/n)\,\mathbf{H}^T\mathbf{H} + \alpha\mathbf{I} & -\sqrt{\rho/n}\,\mathbf{H}^T\mathbf{y} \\ -\sqrt{\rho/n}\,\mathbf{y}^T\mathbf{H} & \|\mathbf{y}\|^2 - \beta \end{bmatrix}$$

is a PSD matrix in probability as $n \to \infty$, we choose α to be nonnegative, and pick β so that the Schur complement is nonnegative:

$$(1/n)\,\beta \ \leq^P \ (1/n)\mathbf{y}^T\mathbf{M}\mathbf{y},$$

where $\mathbf{M} = \left(\mathbf{I} - \frac{\rho}{n}\mathbf{H}\left(\frac{\rho}{n}\mathbf{H}^T\mathbf{H} + \alpha\mathbf{I}\right)^{-1}\mathbf{H}^T\right)$. Consider the asymptotic behavior of the right hand side as $n \to \infty$. Substitute $\mathbf{y} = \sqrt{\frac{\rho}{n}}\mathbf{H}\mathbf{x} + \mathbf{v}$ we get three terms:

$$\frac{1}{n}\,\beta \ \leq^P \ \frac{\rho}{n^2}\mathbf{x}^T\mathbf{H}^T\mathbf{M}\mathbf{H}\mathbf{x} + \frac{1}{n}\mathbf{v}^T\mathbf{M}\mathbf{v} + \frac{1}{n}\sqrt{\frac{\rho}{n}}\mathbf{v}^T\mathbf{M}\mathbf{H}\mathbf{x}.$$

Using the results on the limiting eigenvalue distribution of large matrices from [7] we can calculate the limits for the first two terms:

$$\lim_{n\to\infty} \frac{\rho}{n^2}\mathbf{x}^T\mathbf{H}^T\mathbf{M}\mathbf{H}\mathbf{x} = \alpha - \frac{\alpha^2}{\rho}\int_0^{+\infty}\frac{dG(\lambda)}{\lambda + \alpha/\rho} \quad (18)$$

$$\lim_{n\to\infty} \frac{1}{n}\mathbf{v}^T\mathbf{M}\mathbf{v} = \frac{\alpha}{\rho}\int_0^{+\infty}\frac{dG(\lambda)}{\lambda + \alpha/\rho}, \quad (19)$$

where $G(\lambda)$ stands for the limiting empirical distribution of eigenvalues of $\frac{1}{n}\mathbf{H}^T\mathbf{H}$ as $n \to \infty$. Let $\beta_v = \frac{1}{n}\sqrt{\frac{\rho}{n}}\mathbf{v}^T\mathbf{M}\mathbf{H}\mathbf{x}$. We note that $E\{\beta_v\} = 0$, and $E\{\beta_v^2\}$ can be calculated using the results from [7]:

$$\lim_{n\to\infty} nE\{\beta_v^2\} = \alpha^2\int_0^{+\infty}\frac{\rho\lambda}{(\rho\lambda + \alpha)^2}dG(\lambda) = const.$$

Therefore, based on Chebyshev inequality we can conclude that β_v goes to 0 in probability as $n \to \infty$. Collecting the results from (18) and (19) and evaluating the integrals using the Stieltjes transform [7] we obtain

$$\frac{\beta}{n} \leq^P \alpha + \frac{1}{2\rho}\left[(1-\alpha)\left(\sqrt{\alpha^2 + 4\alpha\rho} - \alpha\right)\right].$$

This condition ensures dual feasibility of $\mathbf{G}_{n+1}$. The dual objective value evaluated at $\mathbf{G}_{n+1}$, which serves as a lower bound on the primal optimal value, is given by:

$$\lim_{n\to\infty}\frac{1}{n}f_{SDP} \geq^P \frac{1}{2\rho}\left[(1-\alpha)\left(\sqrt{\alpha^2 + 4\alpha\rho} - \alpha\right)\right].$$

Picking $\alpha = 1/3$ we arrive at the claim of the lemma. □

Proof of Theorem 3: We can bound $E\{f_{SDR}\}$ (where the expectation is taken with respect to the randomization procedure in Section 2) as follows:

$$\begin{aligned} E\{f_{SDR}\} &= E\{\bar{\mathbf{x}}^T\mathbf{Q}\bar{\mathbf{x}}\} \\ &= E\left\{\textstyle\sum_{i,j=1,i\neq j}^{n+1} Q_{ij}\bar{x}_i\bar{x}_j + \sum_{i=1}^{n+1} Q_{ii}\bar{x}_i^2\right\} \\ &= \textstyle\sum_{i,j=1,i\neq j}^{n+1} Q_{ij}v_{ki}v_{kj} + \sum_{i=1}^{n+1} Q_{ii} \\ &\leq \mathbf{v}_k^T\mathbf{Q}\mathbf{v}_k + \mathrm{Trace}(\mathbf{Q}) \\ &\leq f_{SDP} + \|\mathbf{y}\|^2 + \tfrac{\rho}{n}\mathrm{Trace}(\mathbf{H}^T\mathbf{H}). \end{aligned}$$

Using the fact that $\frac{1}{n}\|\mathbf{y}\|^2$ converges to $\rho + 1$ and $\frac{\rho}{n^2}\mathrm{Trace}(\mathbf{H}^T\mathbf{H})$ converges to ρ in probability when $n \to \infty$ we can write

$$\frac{1}{n}E\{f_{SDR}\} \leq^P \frac{1}{n}f_{SDP} + (2\rho + 1). \quad (20)$$

Now, we have the following chain of inequalities:

$$\begin{aligned} \frac{1}{n}f_{ML} \leq \frac{1}{n}f_{SDR} &\leq^P \frac{1}{n}E\{f_{SDR}\} \\ &\leq^P \frac{1}{n}f_{SDP} + (2\rho+1) \\ &\leq^P \frac{1}{n}\left(1 + \frac{9\rho(2\rho+1)}{\sqrt{12\rho+1}-1}\right)f_{SDP} \\ &\leq \frac{1}{n}\left(1 + \frac{9\rho(2\rho+1)}{\sqrt{12\rho+1}-1}\right)f_{ML}. \end{aligned}$$

The first inequality is due to the fact that the output of SDP Detector is a feasible point for the ML detection problem. The second inequality is satisfied in probability for sufficiently large number of samples L in randomized rounding procedure. The third inequality has been obtained in (20). The fourth one follows from Lemma 2, and the last one is true because SDP problem (6) is relaxation of the original ML detection problem (5). The claim of Theorem 3 follows from this chain of inequalities. □

5. CONCLUSIONS

The paper presents a performance analysis for the SDP detection algorithm [5] for a communication link operating at both high and low SNRs. It has been shown that for any fixed fading $\mathbf{H}$ and noise $\mathbf{v}$ there exists an SNR threshold such that the output of SDP Detector coincides with the solution of the ML detection problem for all SNR levels above the threshold. This result implies that for sufficiently high SNR, the ML detection problem allows a polynomial time algorithm (e.g., SDP Detector) if condition (8) is satisfied. In low SNR region, the SDP Detector is guaranteed to generate a log-likelihood value that is within a factor of $5/2$ to the optimal ML log-likelihood value for large systems.

6. REFERENCES

[1] B. Hassibi, B.M. Hochwald, “High-Rate Codes That Are Linear in Space and Time,” *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1804 – 1824, 2002.

[2] E. Viterbo and J. Boutros, “A Universal Lattice Code Decoder for Fading Channels,” *IEEE Trans. Inform. Theory*, pp. 1639 – 1642, 1999.

[3] U. Fincke, M. Pohst, “Improved Methods for Calculating Vectors of Short Length in a Lattice, Including a Complexity Analysis,” *Math. Comput.*, vol. 44, pp. 463 – 471, 1985.

[4] J. Jalden, B. Ottersten, “An Exponential Lower Bound on the Expected Complexity of Sphere Decoding,” *Proc. ICASSP ’04*, vol. 4, pp. 393 – 396, 2004.

[5] 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,” *IEEE Trans. Signal Processing*, vol. 50, no. 4, pp. 912 – 922, 2002.

[6] Z.-Q. Luo, X. Luo and M. Kisialiou, “An Efficient Quasi-Maximum-Likelihood Decoder for PSK Signals,” in *Proc. ICASSP ’03*, vol. 6, pp. VI 561 – VI 564, 2003.

[7] J.W. Silverstein, Z.D. Bai, “On the Empirical Distribution of Eigenvalues of a Class of Large Dimensional Random Matrices,” *J. Multivariate Analysis*, vol. 54, no. 2, pp. 175 – 192, 1995.