Applied Convex Optimization

MIMO detection : A Semidefinite Relaxation approach

Name: Ricardo García Gutiérrez , Genís Floriach Pigem

Universitat Politècnica de Catalunya
(Spring, 2017)

10th June 2017

Contents

1 Introduction

During the past decades, the demand in transferring large amounts of data rapidly and reliably has been increasing drastically. This demand is not likely to reduce but rather increase. One way of increasing data speed/throughput is to increase the spectral bandwidth. However, in the frequency bands where communication is possible today, bandwidth is a very limited resource. One could keep the spectral bandwidth fixed and increase the transmission power to get better signal strengths and thus to be able to send more bits per transmission. However, this would cause interference to fellow "communicators" that would also like to send information with the same data speed. Another way, and an even more promising one, is to increase the spatial bandwidth (number of elements along the spatial dimension), which is not utilized fully today and therefore offers many degrees of freedom yet to be exploited. One way of doing that is by placing multiple antennas at both the transmitting and receiving side of the communication link. This technology is called multiple-input multiple output (MIMO). It has many advantages over the single-antenna technologies. For instance, it increases the information throughput and link reliability by exploiting link diversity in rich scattering environments. Nonetheless, this technology arises many problems to deal with.

A major implementation difficulty of the MIMO technology is the signal separation (detection) problem at the receiving side of the MIMO link. This is due to the fact that the transmitted signals superimpose and the information from the different antennas interfere with each other. Now, if the MIMO channel matrix is not well-conditioned and is close to singular, solving the detection problem becomes very difficult. Well known optimal methods are available but their complexity is exponential with number of transmitting antennas, and therefore they become unpractical when the number of antennas is large.

To overcome this difficulty , smart approaches employ approximations to lower the computational complexity burden at the expense of accuracy. There is an inherent trade-off between accuracy and complexity that decides what total performance a detector will have and how good the algorithm is. The goal is to achieve the accuracy of the optimal methods with as low complexity as possible.

2 Problem Statement

Let's consider a MIMO system with M_T transmitting antennas and M_R receiving antennas. The model is the following:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n} \tag{1}$$

where $\mathbf{y} \in \mathbb{C}^{M_R}$ is the received signal, $\mathbf{x} \in \boldsymbol{\chi}^{M_T}$ is the transmitted vector of symbols and $\mathbf{H} \in \mathbb{C}^{M_R \times M_T}$ is the channel matrix. We will assume a frequency flat fading Rayleigh channel, so the entries of the channel matrix will be IDD standard random Gaussian variables. Here, the channel is subject to AWGN noise, where $\mathbf{n}$ is modeled as $\mathcal{CN}(0, \sigma^2\mathbf{I})$. At the receiver side, the problem consists on recovering $\mathbf{x}$ from the observation, based on the knowledge of $\mathbf{H}$ (usually estimated beforehand), $\boldsymbol{\chi}$ and $\mathbf{y}$. The elements of $\mathbf{x}$, say x_i, belong to a finite alphabet $\boldsymbol{\chi}$ of size $\mid \boldsymbol{\chi} \mid$. We will also assume the system of linear equations in (1) is not undetermined, that is $M_T \leq M_R$. The model in (1) can

be applied to other digital communications scheme such as multiuser CDMA, space-time coding and frequency-space coding among others.

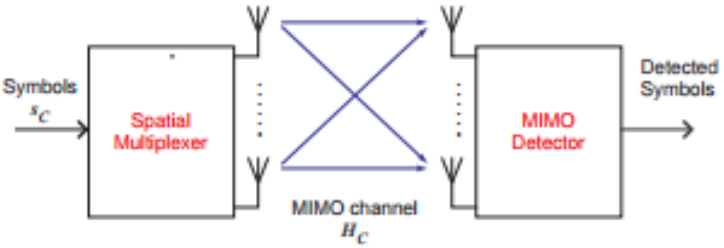

Figure 1: Spatial Multiplexing MIMO system

2.1 Detection

We will review briefly the main categories of MIMO detection algorithms and some of the key ideas on how they work. First of all, there are two main categories of detectors: hard decision detectors which decide whether a bit is zero or one and soft decision detectors which decide how likely it is that a bit is zero or one. The latter category of detectors in comparison with the previous one produces more information at the output and in a receiving chain with appropriate soft decoders will yield much better performance. In this report, we will focus on hard decision detectors.

- **Maximum Likelihood** This is the optimal detector from the point of view of minimizing the probability of error (assuming equiprobable $\mathbf{x}$).The maximum likelihood problem is denoted by:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \\ \text{subject to} \quad & \mathbf{x} \in \boldsymbol{\chi}^{M_T} \end{aligned} \tag{2}$$

  Solving this problem involves computing the objective function for all the potential values of $\mathbf{x}$, i.e. a brute-force search. Hence the ML detector has a very large computational complexity, in fact it is regarded as a NP-hard problem in general, meaning that no polynomial-time algorithm exists to solve it. For example, for binary constellation such as BPSK, the complexity of the brute-force algorithm would be $\mathcal{O}(2^{M_T})$, that is exponential with the number of transmitting antennas. For such simple modulation scheme it is still relatively low for small sized MIMO systems, however when high order constellations are used, the complexity becomes quickly significant. It is also worth saying, that such binary constraint sets are usually non-convex. The high complexity that arises in this problem statement is the main issue in MIMO detection. It has required serious attention and still does. Thus, many approximate methods have been proposed.

- **Linear Receivers**

  1. **Zero-forcing (ZF)** One of the crudest and the simplest approximations of the problem in (2) is the zero forcing (ZF) solution. It

consists of two steps: decoupling of the interfering symbols using the pseudo-inverse $\mathbf{H^H H^{-1} H^H}$, i.e., the least squares filter, and then quantizing the decoupled symbols independently per dimension.

2. **L-MMSE** The Linear Minimum Mean Square Error approach tries to find a coefficient $\mathbf{W}$ which minimizes the criterion,

$$\mathbb{E}\{||\mathbf{Wy} - \mathbf{x}||^2\} \tag{3}$$

Assuming AWGN, the solution is $\mathbf{W} = (\mathbf{H^T H} + \frac{N_0}{Es} I)^{-1} \mathbf{H^T}$. Then the estimated symbols are:

$$\hat{\mathbf{x}} = sign[(\mathbf{H^T H} + \frac{N_0}{Es} I)^{-1} \mathbf{H^T y}] \tag{4}$$

- **Box Relaxation**

  As its name indicates, this relaxation relaxes the non-affine equality constraint and allows the solution to live inside the hypercube centered in the origin of coordinates. Basically, we are changing the equality constraint to an inequality constraint, the problem is then,

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{y} - \mathbf{Hx}\|^2 \\ \text{subject to} \quad & \mathbf{x}_i^2 \leq 1 \quad i = 1, \ldots, M_T \end{aligned} \tag{5}$$

  This is a convex problem that falls into the category of QCQP.

- **Sphere decoders (SD)**

  This algorithm solves the ML problem in (2) exactly, however, as we have said before the problem in (2) is NP-Hard therefore the complexity of Sphere decoder grows exponentially with the number of antennas. Nonetheless, this algorithm is useful for moderate problem sizes ($M_T \leq 20$). It is outside the scope of this project to study the Sphere decoder.

3 Semidefinite Relaxation (SDR) Detector

3.1 SDP Relaxation

The SDR detector tries to relax the problem (2) associated to the ML detector in order to end up with a convex problem (Semidefinite Programming) that can be solved in polynomial time. For this case, we assume that the transmitted symbols follow a quaternary phase-shift-keying (QPSK) constellation, then we have the following problem:

$$\begin{aligned} \min_{\mathbf{x_c}} \quad & \|\mathbf{y}_c - \mathbf{Hx}_c\|^2 \\ \text{subject to} \quad & \mathbf{x}_c \in \{\pm 1 \pm j\}^{M_T} \end{aligned} \tag{6}$$

where the entries of $\mathbf{y}$ and $\mathbf{H}$ are complex values. The first step to relax this problem is convert it into a real valued homogeneous QCQP problem. To do so, we first define the new variables:

$$\mathbf{y} = \begin{bmatrix} \mathbb{R}\{\mathbf{y}_c\} \\ \mathbb{I}\{\mathbf{y}_c\} \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} \mathbb{R}\{\mathbf{x}_c\} \\ \mathbb{I}\{\mathbf{x}_c\} \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} \mathbb{R}\{\mathbf{H}_c\} & -\mathbb{I}\{\mathbf{H}_c\} \\ \mathbb{I}\{\mathbf{H}_c\} & \mathbb{R}\{\mathbf{H}_c\} \end{bmatrix} \tag{7}$$

Now the problem (6) can be rewritten as,

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \\ \text{subject to} \quad & \mathbf{x} \in \{\pm 1\}^{2M_T} \end{aligned} \tag{8}$$

However, it is still not a homogeneous QCQP as the objective function is not a pure quadratic form. In order to homogenize it we need to introduce the variable t:

$$\begin{aligned} \min_{\mathbf{x},t} \quad & \|t\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \\ \text{subject to} \quad & t^2 = 1,\ x_i^2 = 1,\ i = 1, \dots, 2M_T \end{aligned} \tag{9}$$

Problem (9) is equivalent to (8). To see that we have to realize that if (x^*, t^*) is an optimal solution to (9), then x^* is an optimal solution to (8) when $t^* = 1$. If $t^* = -1$ then $-x^*$ is an optimal solution to (8). Thanks to the introduction of the variable t the problem (9) can be expressed as a homogeneous QCQP:

$$\begin{aligned} \min_{\mathbf{x},t} \quad & \begin{bmatrix} \mathbf{x}^T & t \end{bmatrix} \begin{bmatrix} \mathbf{H}^T\mathbf{H} & -\mathbf{H}^T\mathbf{y} \\ -\mathbf{y}^T\mathbf{H} & \|\mathbf{y}\|^{\mathbf{2}} \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ t \end{bmatrix} \\ \text{subject to} \quad & t^2 = 1,\ x_i^2 = 1,\ i = 1, \dots, 2M_T \end{aligned} \tag{10}$$

Grouping the terms together yields:

$$\begin{aligned} \min_{\mathbf{z}} \quad & \mathbf{z}^T\mathbf{C}\mathbf{z} \\ \text{subject to} \quad & z_i^2 - 1 = 0, \quad i = 1, \dots, 2M_T + 1 \end{aligned} \tag{11}$$

Problem (11) is the well-known Boolean quadratic program (BQP). BQP is is not a convex problem because, regardless of the convexity of the objective function (given by the characteristics of the matrix $\mathbf{C}$), the equality constraint is not affine but quadratic.

Once we have expressed the initial problem (6) as a homogeneous QCQP, we express it as an SDP. Taking advantage of the cyclic property of the trace operator, we can rewrite the objective function:

$$\mathbf{z}^T\mathbf{C}\mathbf{z} = \mathbf{trace}\,(\mathbf{z}^T\mathbf{C}\mathbf{z}) = \mathbf{trace}\,(\mathbf{C}\mathbf{z}\mathbf{z}^T) = \mathbf{trace}\,(\mathbf{C}\mathbf{Z}) \tag{12}$$

where $\mathbf{Z} = \mathbf{z}\mathbf{z}^T$. Therefore, $\mathbf{Z} \succeq 0$, has rank one, and $\mathbf{Z}_{ii} = z_i^2 = 1$. Then, the problem (11) can be exactly rewritten as:

$$\begin{aligned} \min_{\mathbf{Z}} \quad & \mathbf{trace}\,(\mathbf{C}\mathbf{Z}) \\ \text{subject to} \quad & \mathbf{Z} \succeq 0 \\ & \mathbf{Z}_{ii} = 1 \\ & \mathbf{rank}\,(\mathbf{Z}) = 1 \end{aligned} \tag{13}$$

The objective function is linear with the matrix $\mathbf{Z}$, the inequality constraint is convex as it is the convex cone of positive semidefinite symmetric matrices, the first equality constraint is affine, however, the second equality constraint is not affine. Finally, the key idea in SDP relaxation is to relax the problem (13) by removing the rank constraint, then it becomes a convex problem that can be

solved in polynomial time using an interior-point algorithm with a worst-case complexity $\mathcal{O}\{M_T^{3.5}\}$. The SDR of problem (10) can now be express as:

$$\begin{aligned} \min_{\mathbf{Z}} \quad & \mathbf{trace}\,(\mathbf{CZ}) \\ \text{subject to} \quad & \mathbf{Z} \succeq 0 \\ & \mathbf{Z}_{ii} = 1 \end{aligned} \tag{14}$$

In this relaxation we are lifting the problem to a higher dimensional space, from $\mathbb{R}^{2M_T+1}$ to $\mathbb{S}^{2M_T+1}$. Using the convex optimization toolbox CVX for MATLAB, we can solve (13) with the code in algorithm 1.

Algorithm 1 CVX CODE FOR SDR

```
define matrix C
n = 2*Mt + 1
cvx_begin
    variable Z(n,n) symmetric
    minimize (trace(C*Z));
    subject to
        diag(Z) == 1;
    Z == semidefinite(n);
cvx_end
```

This SDP relaxation can also be derived through other methods such as the Schur complement or the bidual problem, it will be covered in the next section. In summary, the SDR is computationally efficient approximation approach to QCQP, in the sense that is complexity is polynomial in the problem size.

3.2 Extracting feasible points

There is still one fundamental issue that one must address when using the SDR. That is how to convert a globally optimal solution $\mathbf{Z}^*$ to (13) into a feasible solution $\hat{\mathbf{z}}$ to problem (11). Now, if $\mathbf{Z}^*$ is rank one, then there is nothing to do, we can write $\mathbf{Z} = \hat{\mathbf{z}}\hat{\mathbf{z}}^T$ and $\hat{\mathbf{z}}$ will be feasible. On the other hand, if the rank of $\mathbf{Z}^*$ is larger than one, then we must somehow extract from it, in a efficient manner, a feasible solution $\hat{\mathbf{z}}$. There are many ways to do this, we will mention a couple of approaches that are based on some intuitive heuristics.

3.2.1 SDR with rank-one approximation

The idea of applying rank-one approximation is very appealing to this problem. If we let $r = \mathbf{rank}\,(\mathbf{Z}^*)$, and then we express the Eigen-Decomposition of $\mathbf{Z}^*$ as:

$$\mathbf{Z}^* = \sum_{i=1}^{r} \lambda_i \mathbf{q_i} \mathbf{q_i}^T \tag{15}$$

where $\lambda_1 \geq \lambda_2 ... \geq \lambda_r > 0$ are the eigenvalues and $\mathbf{q_1}, ..., \mathbf{q_r} \in \mathbb{R}^{2M_T+1}$ are the associated eigenvectors, in terms of our problem size. Since the best rank-one approximation $\mathbf{Z_1}^*$ to $\mathbf{Z}^*$ (in the least 2-norm sense) is given by:

$$\mathbf{Z_1}^* = \lambda_1 \mathbf{q_1} \mathbf{q_1}^T \tag{16}$$

We can define a candidate solution as $\hat{\mathbf{z}} = \sqrt{\lambda_1}\mathbf{q_1}$ to problem (11), only if the candidate is feasible. Otherwise, there is one remaining step that consists on trying to map $\hat{\mathbf{z}}$ to a "nearby" feasible solution of problem (11), such mapping is usually problem dependent, in our BQP, we can obtain a feasible solution $\hat{\mathbf{z}}$ from $\hat{\mathbf{z}} = sign(\hat{\mathbf{z}})$.

3.2.2 SDR with randomization

Another way to extract an approximate QCQP from an SDR solution $\mathbf{Z}^*$ is called Gaussian randomization. To start off, we will consider again the real-valued homogeneous QCQP that arose in (10). Let $\mathbf{X} \in \mathbb{S}^n$ be an arbitrary symmetric positive semidefinite matrix and consider the random vector $\xi \sim \mathcal{N}(0, \mathbf{X})$. Now we are able to build the following standard form stochastic QCQP:

$$\begin{aligned} \min_{\mathbf{X}\in\mathbb{S}^n, \mathbf{X}\succeq 0} \quad & E[\xi^T \mathbf{C} \xi] \\ \text{subject to} \quad & E[\xi^T \mathbf{A_i} \xi] \trianglerighteq_i b_i, \quad i = 1, \ldots, m \end{aligned} \tag{17}$$

where we manipulate the covariance matrix of ξ so that the expected value of the objective function is minimized and the constraints are satisfied in expectation. By making the simple observation of the following relation, $\mathbf{X} = E[\xi\xi^T]$, it is appreciable that the stochastic QCQP is equivalent to the SDR. Thus, we can design an alternative method to generate approximate solutions to the QCQP (11). The procedure is detailed below:

Algorithm 2 Gaussian Randomization for BQP

given an SDR solution $\mathbf{Z}^*$, and a number of randomizations L.
for l=1 ...,L
    generate $\xi \sim \mathcal{N}(0, \mathbf{Z}^*)$
    construct a QCQP feasible point by : $\hat{\mathbf{z}}_l = sign(\xi_\mathbf{l})$
end
determine $l^* = argmin_{l=1,\ldots,L} \hat{\mathbf{z}}_l^T \mathbf{C} \hat{\mathbf{z}}_l$
output $\hat{\mathbf{z}} = \hat{\mathbf{z}}_{l^*}$

In 2, the rounding used to generate feasible points for the BQP from the random samples makes the procedure problem dependent again. Moreover, the experiment must be carried out L times in order to select the best objective value. It has been observed that not an excessive number of samples are needed to achieve a good enough performance.

One question that should be asked: How far is the optimal SDR objective value from the optimal QCQP? Although this analysis belongs to a particular field in mathematics and it is beyond the scope of this work to go into further details, the randomization approach was the one that opened the gateway to a host of theoretically provable worst-case approximation bounds for SDR. To give just a bit of flavor of the approximation accuracy results, the following summary is given.

Take γ as the so-called approximation ratio, $\nu(\hat{\mathbf{z}})$ is the objective value of the SDR with a randomization approach and ν_{QP} is the objective value of the real-valued homogeneous QCQP. Then, it has been stated that $\nu(\hat{\mathbf{z}})$ will satisfy with high probability:

$$\nu_{QP} \leq \nu(\hat{\mathbf{z}}) \leq \gamma \nu_{QP} \tag{18}$$

For our particular application of MIMO detection for BQPSK or QPSK constellations, it has been shown the following:

Theorem. *For $\sigma^2 \geq 60n$ (which corresponds to the Low SNR region), with probability at least $1 - 3\exp(-n/6)$: $\gamma \leq \frac{11}{2}$. For $\sigma^2 = \mathcal{O}(1)$ (which corresponds to the high SNR region), with probability at least $1 - exp(\mathcal{O}(1))$: $\gamma = 1$, i.e. the SDR is tight, it yields the optimal value of the original QCQP formulation, or in other words, it is not a relaxation anymore.*

This approximation accuracy framework can be generalized to other quadratic minimizations programs.

3.3 Constraint set: Higher order constellations

There is one assumption being made when solving (5), it is formulated for QPSK type constellations. The SDR is, in fact, a symbol-constellation dependent technique, therefore the problem structure might change. It was developed initially for BPSK and QPSK, given that the same ideas apply, there the SDR has been found to be capable of providing near-optimal performance. Back then, it stimulated a number of research endeavors that aimed to apply SDR to, for example, high-order QAM or PSK schemes. These independently developed SDRs are different in concept and structure.

For the M-QAM signaling schemes, several work have proposed different SDR approaches that based on some kind of algebraic extension to the problem. For example,in the work [1] they proposed a 16-QAM SDR detector, the key observation is that any finite alphabet constraint can be replaced by a polynomial constraint, e.g., $x \in (a, b, c)$ then $(x-a)(x-b)(x-c) = 0$. Then, by introducing slack variable the polynomials can be translated into multiple quadratic constraints. One of the drawbacks of the method was that it could not be extended to a higher order modulations scheme. There has also been attempts to propose a universal SDR detector at the expense of higher computational costs.

In the other hand, for M-PSK signaling, a generalized SDR approach has been proposed with fare more ease. The idea lies on considering a more general complex valued homogeneous QCQP. Consider the following k-ary quadratic program:

$$\begin{aligned} \min_{\mathbf{z}\in\mathbb{C}^{M_T}} \quad & \mathbf{z}^H\mathbf{C}\mathbf{z} \\ \text{subject to} \quad & z_i \in \{1, e^{\frac{j2\pi}{k}}, \ldots, e^{\frac{j2\pi(k-1)}{k}}\}, \\ & i = 1, \ldots, M_T \end{aligned} \tag{19}$$

Using the same SDR idea as in the real case, we can derive the following SDP:

$$\begin{aligned} \min_{\mathbf{Z}\in\mathbb{H}_T^M} \quad & \mathbf{trace}\,(\mathbf{C}\mathbf{Z}) \\ \text{subject to} \quad & \mathbf{Z} \succeq 0 \\ & \mathbf{Z}_{ii} = 1 \end{aligned} \tag{20}$$

Practically the only difference with the real value case is that our problem domain is now on $\mathbb{H}^N$. Notice that the constellation size k is no longer used in the formulation. In fact, SDP duality theory and results in non-asymptotic random

matrix theory can be used to establish approximation bounds for the SDR detector in the case of MPSK constellations. Nonetheless, the same guarantees of approximation accuracy with respect to the case of QPSK and BPSK ($k > 2$) are no longer satisfied

3.4 Duality

We are now interested in finding the Dual problem of the original problem, which is non-convex. Recalling the original problem in homogeneous form:

$$\begin{aligned} \min_{\mathbf{z}} \quad & \mathbf{z}^T \mathbf{C} \mathbf{z} \\ \text{subject to} \quad & z_i^2 - 1 = 0, \;\; i = 1, \ldots, 2M_T + 1 \end{aligned} \tag{21}$$

The Lagrangian function:

$$L(\mathbf{z}, \lambda) = \mathbf{z}^T \mathbf{C} \mathbf{z} - \sum_{i=1}^{2M_T+1} \lambda_i (z_i^2 - 1) = \mathbf{z}^T (\mathbf{C} - \mathbf{\Lambda}) \mathbf{z} + \mathbf{trace}\, \mathbf{\Lambda} \tag{22}$$

where $\mathbf{\Lambda} = diag(\lambda)$. Now the Dual function:

$$g(\Lambda) = \inf_{\mathbf{z}} L(\mathbf{z}, \mathbf{\Lambda}) = \begin{cases} \mathbf{trace}\, \mathbf{\Lambda} & \text{if} \quad \mathbf{C} - \mathbf{\Lambda} \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \tag{23}$$

Finally the Dual problem:

$$\begin{aligned} \max_{\mathbf{\Lambda}} \quad & \mathbf{trace}\, \mathbf{\Lambda} \\ \text{subject to} \quad & \mathbf{C} - \mathbf{\Lambda} \succeq 0 \\ & \mathbf{\Lambda} \quad \text{diagonal} \end{aligned} \tag{24}$$

As we can see the Dual problem is also an SDP. As the problem is non-convex strong duality does not hold, however, the Dual problem is still useful as it provides a lower bound for the real solution, it can be easily checked:

$$\mathbf{z}^T \mathbf{C} \mathbf{z} \geq \mathbf{z}^T \mathbf{\Lambda} \mathbf{z} = \sum_{i=1}^{2M_T+1} \lambda_i z_i^2 = \mathbf{trace}\, \Lambda \tag{25}$$

The first inequality comes from the constraint in the Dual problem (24), the first equality from $\mathbf{\Lambda}$ being diagonal and the last equality from $z_i \in \{+1, -1\}$.

Regarding the solution of the relaxed problem (14), most solvers use Primal-Dual Interior Point method in order to find the solution. This method solves both the Primal and the Dual and uses the Duality gap, that is the difference between the primal objective and the Dual objective, as a stopping criteria. Therefore, it may be also interesting to find the Dual Problem of the relaxed SDP. According to [2], the Primal and its Dual problem of an SDP are,

$$\begin{aligned} \text{P:} \quad \min_{\mathbf{X}} \quad & \mathbf{trace}\, \mathbf{C} \mathbf{X} \\ \text{subject to} \quad & \mathbf{trace}\, \mathbf{A_i} \mathbf{X} = b_i \quad i = 1 \ldots n \\ & \mathbf{X} \succeq 0 \end{aligned} \tag{26}$$

$$\begin{aligned} \text{D:} \quad & \max_{\mathbf{y},\mathbf{S}} && \mathbf{b^T y} \\ & \text{subject to} && \sum_{i=1}^{n} \mathbf{A_i y_i} + \mathbf{S} = \mathbf{C} \\ & && \mathbf{S} \succeq 0 \end{aligned} \tag{27}$$

Expressing the relaxed SDP (14) in the form of (26) yields,

$$\begin{aligned} \text{P:} \quad & \min_{\mathbf{Z}} && \mathbf{trace}\, \mathbf{CZ} \\ & \text{subject to} && \mathbf{trace}\, \mathbf{A_i Z} = 1 \quad i = 1 \dots n \\ & && \mathbf{Z} \succeq 0 \end{aligned} \tag{28}$$

where $\mathbf{A_i}$ is matrix with all entries equal to 0 except for the ith element of the diagonal which is equal to 1. Then, the Dual problem is,

$$\begin{aligned} \text{D:} \quad & \max_{\lambda,\mathbf{S}} && \mathbf{1}^T \lambda \\ & \text{subject to} && \sum_{i=1}^{n} \mathbf{A_i} \lambda_\mathbf{i} + \mathbf{S} = \mathbf{C} \\ & && \mathbf{S} \succeq 0 \end{aligned} \tag{29}$$

We can rewrite the equality constraint as

$$\mathbf{C} - \sum_{i=1}^{n} \mathbf{A_i} \lambda_\mathbf{i} = \mathbf{C} - diag(\lambda) = \mathbf{C} - \mathbf{\Lambda} = \mathbf{S} \tag{30}$$

and the objective function as

$$\mathbf{1}^T \lambda = \mathbf{trace}\, \mathbf{\Lambda} \tag{31}$$

Finally, rexpressing the Dual problem,

$$\begin{aligned} & \max_{\mathbf{\Lambda}} && \mathbf{trace}\, \mathbf{\Lambda} \\ & \text{subject to} && \mathbf{C} - \mathbf{\Lambda} \succeq 0 \\ & && \mathbf{\Lambda} \quad \text{diagonal} \end{aligned} \tag{32}$$

which is exactly the Dual problem of the original problem (24). However, as the SDP relaxation is a convex problem, strong duality holds between the primal SDP relaxation (14) and its Dual (32). Besides, as strong duality holds the SDP relaxation is also considered as the Dual of the Dual (bidual) of the original problem.

3.5 SDP Relaxation vs other relaxations

We have presented several relaxations techniques, however, we have not given any insight about which one is better. In this section we will show why the SDP relaxation outperforms all the other relaxations.

The reason why SDP outperforms other methods is that it is a tighter relaxation, that means, that it relaxes "less" the original problem. To see that we resort to the Dual problem. The optimal value of the Dual problem is always

a lower bound for the primal problem, therefore, the higher the optimal value of the Dual problem is, the tighter is the bound. Now, we have to see which relaxation gives the highest optimal value of the Dual problem. The optimal value of the Dual function of the Zero-forcing relaxation is the following,

$$p^*_{ZF} = \min_{\mathbf{x}} \ \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 \tag{33}$$

For the box relaxation,

$$p^*_{Box} = \max_{\lambda \geq 0} \ \min_{\mathbf{x}} \ \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \sum_{i=1}^{n} \lambda_i (x_i^2 - 1) \tag{34}$$

It is important to recall that the dual variables for the inequality constraint are constrained to be positive. Finally, for the SDP relaxation we recall that it was the same as the Dual for the orignal problem,

$$p^*_{SDR} = \max_{\lambda} \ \min_{\mathbf{x}} \ \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2 + \sum_{i=1}^{n} \lambda_i (x_i^2 - 1) \tag{35}$$

We can see that the feasible set in the SDP relaxation contains that in the Box relaxation and that in turn contains that in the Zero forcing, therefore $p^*_{ZF} \leq p^*_{Box} \leq p^*_{SDR} \leq p^*_{BQP}$.

In section 4 we confirm that SDP is a better relaxation by means of simulation.

3.6 Interior point method for SDP

There are several Interior point methods in the literature such as Primal Path, Dual Path or Primal-Dual Path. However, we will only focus on the main ideas of the Primal Path method.

The main idea behind Interior Point methods is to remove the inequality constraints by adding a term to the objective function that acts as a barrier, forcing the solution to live in the feasible set. In other words the typical optimization problem is rewritten as,

$$\begin{aligned} \min_{\mathbf{x}} \quad & f_0(\mathbf{x}) + I_{feas}(\mathbf{x}) \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned} \tag{36}$$

where $I_{feas}(\mathbf{x}) = 0$ for $\mathbf{x} \in$ feasible set. However, the barrier I_{feas} is not differentiable, then a good differentiable candidate that can act as a barrier is the *log* function. Therefore, the barrier I_{feas} can be approximated by $\hat{I}(x) = -\sum_{i=1}^{m} log(-f_i(\mathbf{x}))$. Now we can define an approximation to the original problem,

$$\begin{aligned} \hat{\text{P}}: \quad \min_{\mathbf{x}} \quad & f_0(\mathbf{x}) + \sum_{i=1}^{m} (\frac{-1}{t}) log(-f_i(\mathbf{x})) \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \end{aligned} \tag{37}$$

where t defines how bad or accurate we are approximating the problem, if $t \to \infty$ the approximation is very accurate whereas if $t \to 1$ the approximation is very coarse. Having that in mind, we can introduce the Central Path algorithm,

Algorithm 3 Central path

Given an initial strictly feasible point $\mathbf{x}$ and t

Repeat

1. Centering step: Starting at $\mathbf{x}$ solve $\hat{\mathrm{P}}(t)$ and obtain $\mathbf{x}^*(t)$
2. Update $\mathbf{x^{n+1}} := \mathbf{x^{n*}}(t)$
3. Quit if $\frac{m}{t} < \epsilon$
4. Increase $t := \mu t$

The collection of different steps $\mathbf{x}^n$ is what is called the central path. The proper way to set the variable t is the following: start with a small t (otherwise the problem at step 1 of the algorithm may be very difficult to minize), and iteratively increase t until a desired solution accuracy is reached. Now, the only point missing is how to solve the optimization problem in step 1. To do that, we resort to the Newton's Method, for which we need the gradient and hessian of the objective function.

Once we have presented the interior point method, it is time to apply it to our problem. The first thing that we need to do is define a barrier for the cone of positive semidefinite symmetric matrices. If we recall, we can characterize the interior of the positive semidefinite cone as,

$$\mathbf{S}^n_{++} = \{\mathbf{X} \mid \lambda_1(\mathbf{X}) \geq 0, \ldots, \lambda_m(\mathbf{X}) \geq 0\} \tag{38}$$

Therefore, a natural barrier function to repel $\mathbf{X}$ from the boundary of $\mathbf{S}^n_{++}$ is,

$$-\sum_{i=1}^{m} log(\lambda_i(\mathbf{X})) = -log(\prod_{i=1}^{m} \lambda_i(\mathbf{X})) = -log(\mathbf{det}(\mathbf{X})) \tag{39}$$

Then, for our particular problem (in the same canonical form of (28)), the approximated problem that has to be solved at each iteration is,

$$\begin{array}{lll} \hat{\mathrm{P}}(t): & \min\limits_{\mathbf{X}} & \mathbf{trace}\,\mathbf{CX} - (\frac{-1}{t})log(\mathbf{det}(\mathbf{X})) \\ & \text{subject to} & \mathbf{trace}\,\mathbf{A_i X} = b_i \;\; i = 1 \ldots m \end{array} \tag{40}$$

we recall that the matrix $\mathbf{A_i}$ is a matrix with all entries equal to 0 except for the ith element of the diagonal which is equal to 1. Now, to solve this problem we resort to the KKT and Newton's method. According to the KKT, the gradient (with respect to the optimization variable) of the lagrangian at the optimal solution has to be 0 (optimality condition) and the equality constraints have to be fullfiled. In our case, the lagrangian is,

$$L(\mathbf{X}, \lambda) = \mathbf{trace}\,\mathbf{CX} - (\frac{1}{t})log(\mathbf{det}(\mathbf{X})) - \sum_{i=1}^{m} \lambda_i\, \mathbf{trace}\,\mathbf{A_i X} \tag{41}$$

Taking the gradient with respect to $\mathbf{X}$,

$$\nabla_{\mathbf{X}} L(\mathbf{X}, \lambda) = \mathbf{C} - (\frac{1}{t})\mathbf{X}^{-1} - \sum_{i=1}^{m} \lambda_i \mathbf{A_i} = \mathbf{C} - (\frac{1}{t})\mathbf{X}^{-1} - \mathbf{\Lambda} \tag{42}$$

Therefore, the KKT conditions for the approximated problem $\hat{\mathrm{P}}(t)$ are,

$$\begin{cases} \mathbf{trace}\,\mathbf{A_i X} = 1 \quad i = 1, \ldots, m \\ \mathbf{C} - (\frac{1}{t})\mathbf{X}^{-1} = \mathbf{\Lambda} \end{cases} \tag{43}$$

To solve this equation system we resort to the Newton's method. The Newton's method solves at each step the linearized optimality condition, however, it requires the hessian of a matrix function and it belongs to the set $\mathbb{R}^{n\times n\times n\times n}$, therefore the size of this hessian is huge. To overcome this problem, there are several approximations of the quadratic form of the function $log(det(\mathbf{X}))$. However, this is outside the scope of this project.

4 Numerical Results

In the following simulation examples, we compare the bit error rate (BER) performance and computational complexity of the analyzed SDR detector and some of the conventional linear receivers and the Sphere Decoding (SD) detector. We have performed a *Montercarlo* simulation with up to 10000 trials. The simulation scenario follows the standard MIMO setting described at the beginning of the report. All simulations were done using a QPSK simulation. The metric of performance is the Bit Error Rate (BER). The results that were obtained are plotted below:

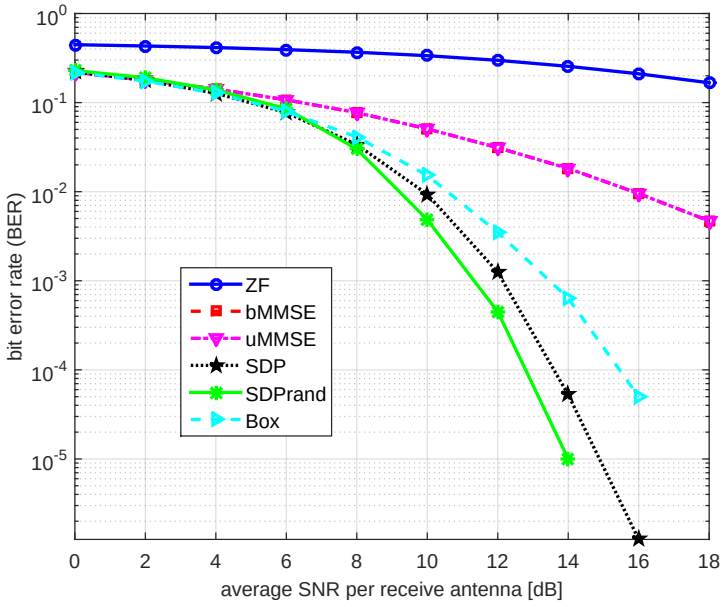

Figure 2: BER analysis of various MIMO detectors in a 40x40 QSPK MIMO system

The curve with the label 'SDP' corresponds to the SDR using a rank-one approximation with the dominant eigenvector method. It shows that it is

competitive in performance, but the alternative 'SDPrand', which is the SDR with the randomization procedure, is even slightly better. As expected, the Box relaxation lies behind these two SDR procedures. The other benchmark methods have a poorer performance in high SNR regime.

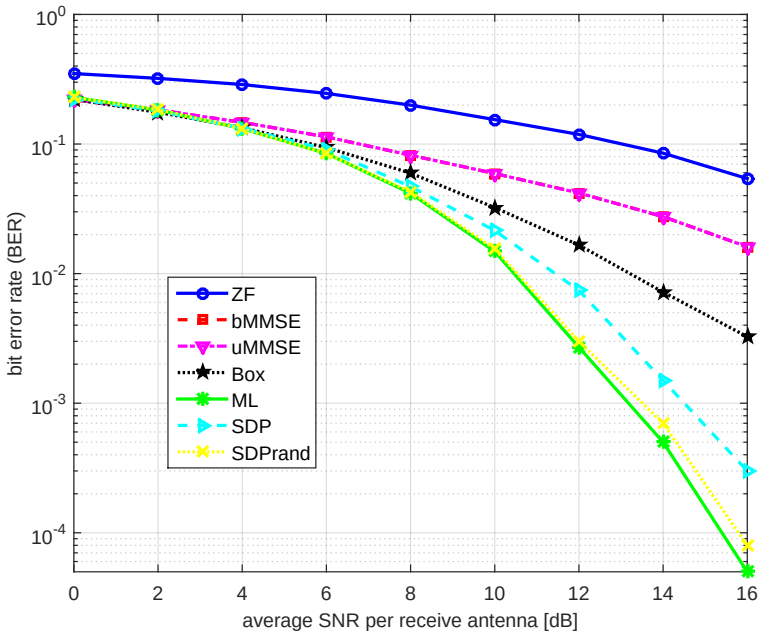

Figure 3: BER analysis of various MIMO detectors in a 5x5 QSPK MIMO system

In 5, we can see the SDR provides near-optimal BER performance, and gives notably better performance than other linear MIMO detectors. The ML curve corresponds to the sphere decoding method, which gives ML performance in small sized MIMO system.

Next, we evaluate the computational complexity of the SDR detector and the sphere decoding detector because of the particular interest in this comparison. The goal is to see that the SDR has a polynomial time complexity with respect to M_T, the number of TX antennas. For the case of SD, which is known to give ML performance, the complexity is small for $M_T \leq 16$, but it increases exponentially otherwise.

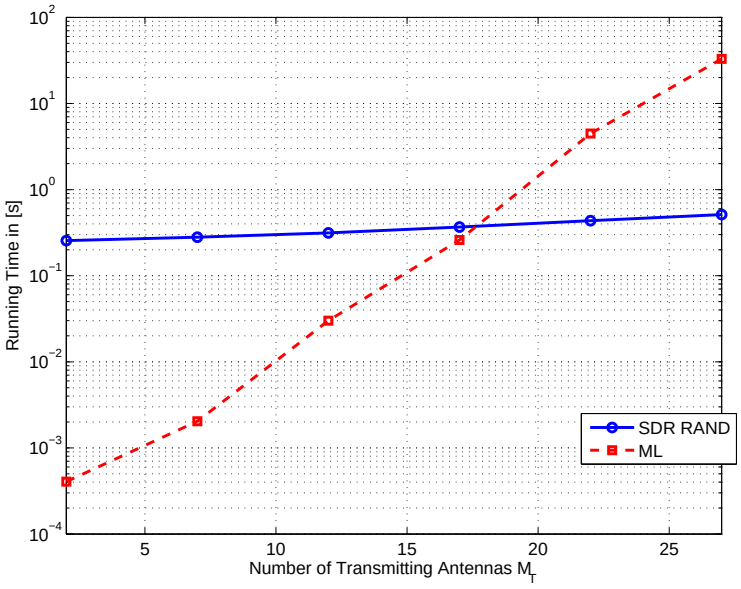

Figure 4: Complexity comparison between ML Sphere Decoding and SDR with randomization

The analysis of complexity has been done in a rather informal way, i.e. we have not considered the real number of operations of the IP method.

Lastly, we analyze if the SDR with the randomization procedure attached can be computed in a efficient manner. By that, we mean to explore what is the effect of the number of random experiments used to generate samples from the optimal SDR solution distribution.

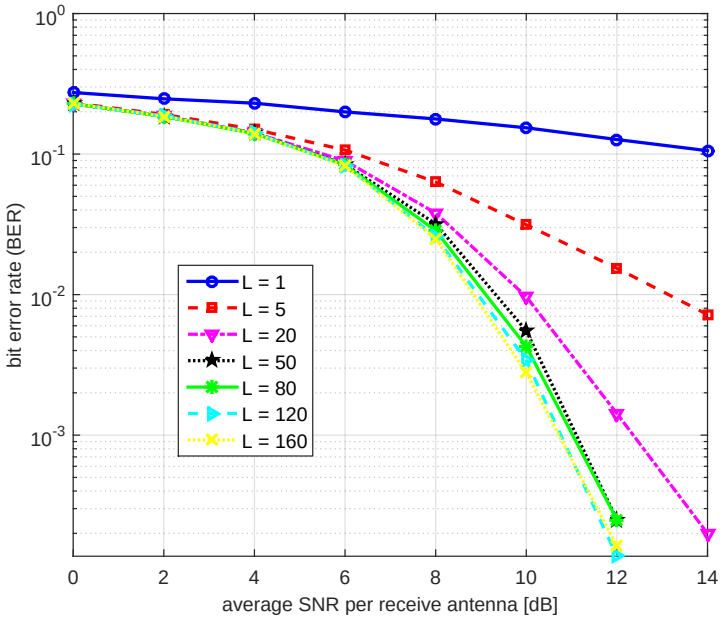

Figure 5: BER analysis of the SDR randomized method for different values of generated samples

Briefly, the plot confirms that with a not very excessive number of random experiments we are able to achieve a really good performance.

5 Conclusions

In summary, the Semidefinite Relaxation approach in MIMO detection is based on a convex relaxation of the ML problem. This relaxed problem is convex, and can be efficiently solved using so-called interior point methods. Most of the times, the error incurred by relaxation is small. Hence, SDR is a balanced method, offering high performance , even near optimal ML with BPSK or QPSK, with a computational complexity lower than other similar performing methods.

But we must ask ourselves, in the communications receiver context, which method is the best in practice? This depends much on the purpose of solving the : what error rate can be tolerated, what is the ultimate measure of performance (e.g., bit-error-rate, worst-case complexity, or average complexity) and what computational platform is used. Also, it is important to state where is the bottleneck of the system, is it in the decoding stage or in the detection of symbols ?. Additionally, the bits in s may be part of a larger code word and different s vectors in that code word may either see the same H (slow fading) or many different realizations of H (fast fading). This complicates the picture, because notions that are important in slow fading (such as spatial diversity) are less important in fast fading, where diversity is provided anyway by time variations.

Anyway, in state of the art systems, the detection module might not be the most critical design part, therefore relaxing the requirements of high performance of detection and focusing on other areas of the digital communications paradigm such as coding might not be practical. Despite this fact, the SDR stands out as a very competitive detector using convex optimization tools.

References

[1] A. Wiesel, Y. C. Eldar, and S. Shamai, “Semidefinite relaxation for detection of 16-QAM signaling in MIMO channels,” IEEE Signal Process. Lett., vol. 12, no. 9, pp. 653-656, 2005

[2] Freund, Robert M. "Introduction to semidefinite programming (SDP)." Massachusetts Institute of Technology (2004).

[3] Vandenberghe, Lieven, and Stephen Boyd. "Semidefinite programming." SIAM review 38.1 (1996): 49-95.

[4] Petersen, Kaare Brandt, and Michael Syskind Pedersen. "The matrix cookbook." Technical University of Denmark 7 (2008): 15.

[5] P. Parrilo and S. Lall. "SDP Relaxations for Quadratic Programming". Stanford University (2003).

[6] Junxiao Song and Daniel P. Palomar. "ML Detection via SDP Relaxation". The Hong Kong University of Science and Technology (HKUST) (2016).

[7] W. K. Ma, P. C. Ching, and Z. Ding, “Semidefinite relaxation based multiuser detection for M-ary PSK multiuser systems” IEEE Signal Process, Oct. 2004.

[8] Z.-Q. Luo, W.-K. Ma, A.-C. So, Y. Ye, and S. Zhang, “Semidefinite relaxation of quadratic optimization problems,” IEEE Signal Process. Mag., vol. 27, no. 3, pp. 20–34, May 2010.