TECHNICAL UNIVERSITY OF CRETE
ELECTRICAL AND COMPUTER ENGINEERING DEPARTMENT
TELECOMMUNICATIONS DIVISION

Antenna selection techniques for large-scale MIMO

by

Dimitrios Papaioannou

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DIPLOMA DEGREE OF

ELECTRICAL AND COMPUTER ENGINEERING

December 2018

THESIS COMMITTEE

Associate Professor George N. Karystinos, *Thesis Supervisor*
Professor Athanasios P. Liavas
Associate Professor Aggelos Bletsas

Acknowledgements

I would like to express my gratitude to every single person that helped me with this thesis.

I am extremely grateful towards my thesis supervisor, Associate Professor George N. Karystinos, for giving me the chance to work on this thesis, for his patient guidance, and for his useful critiques on this work.

I would also like to thank all my friends who shared many beautiful moments with me and supported me through all of these years. Special thanks to a girl, for her love and support.

Finally, and most importantly, i can't describe my appreciation towards my family. This work would have been impossible without their endless support and encouragement throughout my studies.

Abstract

The recent concept of large-scale multiple-input multiple-output (MIMO) systems has attracted substantial research attention and been regarded as a promising technique for next-generation wireless communications networks. Combined with the cost of analog radio-frequency chains it necessitates to work on the use of efficient antenna selection (AS) schemes. Signal-to-noise ratio (SNR) optimal transmit antenna selection and beamforming for a MIMO system that consists of a large number of transmit antennas has been considered in this work. We examine an algorithm that has polynomial complexity and solves the transmit AS problem for maximum-SNR joint beamforming with two receive antennas and under a total power constraint. Furthermore, this thesis works on the beamforming vector approximation under a per-antenna-element power constraint on the transmitter, a problem which has not been efficiently solved so far and many suboptimal algorithms have been produced to tackle this problem. Some of them are examined and compared in this thesis.

Table of Contents

Appendices

List of Abbreviations

SNR	Signal to Noise Ratio
UQP	Unimodular Quadratic Programming
MIMO	Multi-Input and Multi-Output
ULS	Unit-modulus Least Squares
MMSE	Minimum Mean Square Error
GP	Gradient Projection
BER	Bit Error Rate
AS	Antenna Selection

Chapter 1

Introduction

1.1 Motivation and related work

Considering the cost and complexity of the analog RF chains connected to the antenna elements at both sides of a multiple-input multiple-output system its straight forward to point out the necessity of antenna selection schemes as this cost is a limiting factor of the antennas that may operate in practice[1]-[3]. The need for AS algorithms that efficiently select the antennas has has been growing since the increas in the investigation of large-scale multiple antenna wireless systems [4],[5], often called massive MIMO with hundreds of low-power antennas that may be collocated at the base-station site or distributed geographically.

With the antenna selection criteria being either capacity or signal-to-noise ratio, previous works [6]-[21] to tackle the problem proposed suboptimal algorithms as the evaluation of the metric of interest demands an exhaustive search over all $\binom{N}{K}$ possible combinations of antenna selection sets, where N is the number of available transmit antennas and K is the number of the of selected antennas to be occupied for transmission.

This thesis examines and focuses in the antenna selection algorithm, developed and proposed in [22], for the case of $M = 2$ receive antennas and an arbitrary number of K selected transmit antennas. This algorithm identifies a polynomial number of candidate antenna selection sets for the maximum-SNR joint beamforming problem which interestingly enough, gives the same candidate solution subset for unit-norm beamforming (total power constraint scenario) and per-antenna-element power constraint scenario. For the latter, no efficient solver exists regarding the beamforming vector approximation, a problem that has been tackled by various suboptimal iterative algorithms.

Applying those algorithms to approach the per-antenna-element power constraint case, we test their performacnce towards the bit error rate metric, as well as, the complexity of each one of them. Such iterative algorithms that were examined in this thesis include cyclic maximization [23],[24] and gradient projection [25].

1.2 Thesis Outline

The thesis is organized as follows :

- Chapter 2 determines the properties of the signal model and introduces the reader to the maximum-SNR joint beamforming transmit antenna selection problem.
- Chapter 3 presents the antenna selection methods for the cases where we have one or two receive antennas occupied and computes the optimal beamforming vector for each case, under both per-antenna-element and total power constraints.
- Chapter 4 describes and tests algorithms that are produced to tackle the suboptimal beamforming approximation problem considering the per-antenna-element power constraint case.
- In Chapter 5, we test the performance of the suggested algorithms towards the bit error rate versus the total number of available transmit antennas N. We also compare the complexity of all the algorithms discussed in this thesis that are used for the beamforming approximation regarding the per-antenna-element power constraint scenario.
- Finally, Chapter 6 has a conclusion, and some suggestions for future work.

Chapter 2

Signal Model and Problem Statement

2.1 Signal Model

Considering a point-to- point MIMO downlink system from a base station with N transmit antennas to a mobile terminal with M receive antennas and assuming the channel between to be flat fading we denote the $M \times N$ complex baseband matrix:

$$\mathbf{G} = [\mathbf{g}_1 \;\; \mathbf{g}_2 \;\; \dots \;\; \mathbf{g}_M]^H \tag{2.1}$$

where $\mathbf{g}_m^H$ is a complex row vector with the channel coefficients between the N transmit antennas and the m_{th} $(m = 1, 2, \dots, M)$ receive antenna.

Based on receiver (who has knowledge of $\mathbf{G}$) feedback the transmitter selects K out of N antennas and applies either a per-antenna-element or total power constraint to beamform and transmit the symbol $x \in \mathbb{C}$. In other words we define the beamforming vector $\mathbf{w} \in W_N$, subject to the constraint $\|\mathbf{w}\|_0 = K$ where

$$\begin{aligned} W_N &= \left\{ \mathbf{w} \in \mathbb{C}^N : |w_n| \leq 1, n = 1, 2, \dots, N \right\} \\ &\quad or \\ W_N &= \left\{ \mathbf{w} \in \mathbb{C}^N : \|\mathbf{w}\| \leq 1 \right\} \end{aligned} \tag{2.2}$$

for per-antenna element or total, respectively, power constrainted beamforming.

The downconverted and pulse-matched filter received vector
of size $M \times 1$ is

$$\mathbf{y} \triangleq \mathbf{G}\mathbf{w}x + \mathbf{n}, \tag{2.3}$$

where $\mathbf{n} \in \mathbb{C}^M$ is a zero-mean additive colored complex noise vector with autocorrelation matrix $\mathbf{R}_{M \times M}$. We consider the minimum-mean-square-error (MMSE) filter being the maximum-SNR filter

$$\mathbf{f} \triangleq \mathbf{R}^{-1}\mathbf{G}\mathbf{w}. \tag{2.4}$$

Lemma 1. *Since* $\mathbf{y}$ *represents an unknown vector signal in colored vector noise, the maximum-SNR filter is the minimum-mean-square-error (MMSE) filter.*

The proof of Lemma 1 can be found in appendix A.1. Considering the output of the filter $\mathbf{f}$ given above we get

$$\begin{aligned}\mathbf{f}^H\mathbf{y} &= \left(\mathbf{R}^{-1}\mathbf{G}\mathbf{w}\right)^H(\mathbf{G}\mathbf{w}x+\mathbf{n})\\ &= \mathbf{R}^{-\frac{1}{2}}\left(\mathbf{R}^{-\frac{1}{2}}\right)^H(\mathbf{G}\mathbf{w})^H(\mathbf{G}\mathbf{w})\,x+\mathbf{w}^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{n}\\ &= \left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^2 x+\mathbf{w}^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{n}\end{aligned} \tag{2.5}$$

whose output SNR is

$$\frac{E\left\{\left|\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^2 x\right|^2\right\}}{E\left\{\left|\mathbf{w}^H\mathbf{G}^H\mathbf{R}^{-1}\mathbf{n}\right|^2\right\}} = \frac{\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^4 E\left\{|x|^2\right\}}{\left\|(\mathbf{G}\mathbf{w})^H\mathbf{R}^{-1}\right\|^2 E\left\{|\mathbf{n}|^2\right\}} = \frac{\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^4 E\left\{|x|^2\right\}}{\left\|(\mathbf{G}\mathbf{w})^H\mathbf{R}^{-1}\right\|^2 \mathbf{R}}$$

$$= \frac{\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^4 E\left\{|x|^2\right\}}{\left\|(\mathbf{G}\mathbf{w})^H\mathbf{R}^{-\frac{1}{2}}\right\|^2} = \frac{\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^4 E\left\{|x|^2\right\}}{\left\|(\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w})^H\right\|^2} = \frac{\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^4 E\left\{|x|^2\right\}}{\left\|(\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^2}$$

$$= E\left\{|x|^2\right\}\|\mathbf{H}\mathbf{w}\|^2,$$

where $\mathbf{H} = \mathbf{R}^{-\frac{1}{2}}\mathbf{G}$ is the $M \times N$ transformed channel matrix. It is easy to observe how the predetection SNR is proportional to the beamforming vector $\mathbf{w}$. We can jump to equivalent conclusions considering the importance of the beamforming vector $\mathbf{w}$, working on the detection of x at (2.5).

Let $x = \pm\sqrt{P_o}$ with P_o being the total transmit SNR

$$P_o = \frac{E\{|x|^2\}\|\mathbf{w}\|^2}{\sigma_n^2}.$$

Setting $Y = \mathbb{R}e\{\mathbf{f^H y}\} = \left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^2 x + \mathbb{R}e\{\mathbf{f^H n}\}$,

and $z = \mathbb{R}e\{\mathbf{f^H n}\}$, we decide $x = \sqrt{P_o}$ when $Y \geq 0$.

We take a moment, to present a statistical analysis of z in order to proceed on the Error Probability.

$n = n_R + jn_I \sim CN\left(0, N_0 I_M\right)$ and $f = f_R + f_I \in C^M$.

$$\begin{aligned}V = f^H n &= \sum_{i=1}^{M} f_i^* n_i = \sum_{i=1}^{M}\left(f_{R,i} - jf_{I,i}\right)\left(n_{R,i} - jn_{I,i}\right)\\ &= \sum_{i=1}^{M}\left(f_{R,i}n_{R,i} + f_{I,i}n_{I,i}\right) + j\sum_{i=1}^{M}\left(f_{R,i}n_{I,i} - f_{I,i}n_{R,i}\right)\\ &= V_R + jV_I.\end{aligned}$$

It is easily shown that V_R, V_I are real random Gaussian variables with $E\left\{V_R\right\} = E\left\{V_I\right\} = 0$.

But $z = \mathbb{R}e\left\{\mathbf{f^H n}\right\} = V_R \longrightarrow E\left\{z\right\} = E\left\{V_R\right\} = 0.$

Variance of z is derived from

$$\begin{aligned}
E\{V_R^2\} &= E\left\{\sum_{i_1=1}^{M}(f_{R,i_1}n_{R,i_1} + f_{I,i_1}n_{I,i_1})\sum_{i_2=1}^{M}(f_{R,i_2}n_{R,i_2} + f_{I,i_2}n_{I,i_2})\right\} \\
&= \sum_{i_1=1}^{M}\left(f_{R,i}^2 E\left\{n_{R,i}^2\right\} + f_{I,i}^2 E\left\{n_{I,i}^2\right\}\right) \\
&= \frac{N_0}{2}\sum_{i_1=1}^{M}\left(f_{R,i}^2 + f_{I,i}^2\right) = \frac{N_0}{2}\left\|f\right\|^2.
\end{aligned}$$

Given that $\mathbf{f} = \mathbf{R^{-1}Gw}$, $\mathbf{R} = N_0\mathbf{I}_M$ we get:

$$Var\{z\} = \frac{\mathbf{R}\left\|\mathbf{R^{-1}Gw}\right\|^2}{2} = \frac{\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{Gw}\right\|^2}{2}.$$

To summarize, $z = \mathbb{R}e\{\mathbf{f^H n}\} \sim N\left(0\,,\ \frac{\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{Gw}\right\|^2}{2}\right).$

For convenience we set $a = \left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{Gw}\right\|$ and continue to determine the Error probability

$$\begin{aligned}
P\left(e|x = \sqrt{P_o}\right) &= P\left(\mathbf{Y} < 0|x = \sqrt{P_o}\right) = P\left(\left\|a\right\|^2 x + z < 0|x = \sqrt{P_o}\right) \\
&= P\left(z < -\left\|a\right\|^2\sqrt{P_o}\right) = 1 - P\left(z \geq -\left\|a\right\|^2\sqrt{P_o}\right) \\
&= 1 - P\left(\frac{z}{\left\|a\right\|\big/\sqrt{2}} \ \geq\ -\frac{\left\|a\right\|^2\sqrt{P_o}}{\left\|a\right\|\big/\sqrt{2}}\right) \\
&= 1 - Q\left(-\sqrt{2}\left\|a\right\|\ \sqrt{P_o}\right) = Q\left(\left\|a\right\|\sqrt{2P_o}\right) \\
&= Q\left(\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{Gw}\right\|\sqrt{2P_o}\right) \\
&= Q\left(\left\|\mathbf{Hw}\right\|\sqrt{2P_o}\right).
\end{aligned}$$

Equivalently, for $x = -\sqrt{P_o}$ it is trivial to show the probability error has the same result and if we assume that $P(x = -\sqrt{P_o}) = P(x = \sqrt{P_o}) = \frac{1}{2}$, then easily we conclude that the total error probability is

$$P(e) = \frac{1}{2}Q\left(\|\mathbf{Hw}\|\sqrt{2P_o}\right) + \frac{1}{2}Q\left(\|\mathbf{Hw}\|\sqrt{2P_o}\right) = Q\left(\|\mathbf{Hw}\|\sqrt{2P_o}\right).$$

It is straightforward to see that the Error Propability is disproportionate to the beamforming vector $\mathbf{w}$.

2.2 Problem Statement

We now proceed to select K out of the N Base-station Antennas and maximize the beamforming vector $\mathbf{w}$ for this selection. If we name $\mathcal{I}$ the subset of {1,2,...,N} with the indices of the K selected antennas and

$$\mathcal{S} \triangleq \left\{\mathcal{I} \subset \{1,2,\ldots,N\} : |\mathcal{I}| = K\right\} = \left\{\mathcal{I}_1, \mathcal{I}_2, \ldots, \mathcal{I}_{\binom{N}{K}}\right\}$$

the feasible set with all possible realizations of $\mathcal{I}$, then we face the problem:

$$\max_{\substack{\mathbf{w}\in W_N \\ \|\mathbf{w}\|_0 = K}} \|\mathbf{Hw}\| = \max_{\mathcal{I}\in S} \max_{\mathbf{w}\in W_K} \left\|\mathbf{H}_{:,\mathcal{I}}\,\mathbf{w}\right\|, \tag{2.6}$$

dividing our problem into two nested ones. The antenna selection problem (outer maximization) and the beamforming vector selection (inner maximization) for the particular antenna selection.

Given a fixed antenna selection $\mathcal{I} \in \mathcal{S}$ the problem becomes:

$$\mathbf{w}(\mathcal{I}) \triangleq \arg\max_{\mathbf{w}\in W_K} \left\|\mathbf{H}_{:,\mathcal{I}}\,\mathbf{w}\right\|. \tag{2.7}$$

If we were able to compute (2.7) for every possible $\mathcal{I} \in \mathcal{S}$,then the optimal beamformer would be

$$\mathbf{w}_{opt} = \mathbf{w}\left(\mathcal{I}_{opt}\right) \tag{2.8}$$

with $\mathcal{I}_{opt}$ being the optimal Antenna Selection subset, i.e.,

$$\mathcal{I}_{opt} = \operatorname*{argmax}_{\mathcal{I}\in\mathcal{S}} \left\{\max_{\mathbf{w}\in W_K} \left\|\mathbf{H}_{:,\mathcal{I}}\,\mathbf{w}\right\|\right\}. \tag{2.9}$$

Under a per-antenna-element power constraint on the beamformer, i.e., when $W_K = \{\mathbf{w} \in \mathbb{C}^k : |w_k| \leq 1, k = 1,2,\ldots,K\}$, even for a fixed antenna selection $\mathcal{I} \in \mathcal{S}$, the problem of identyfying the optimal beamforming vector(i.e., the solution of (2.7)) is a UQP problem that has not been efficiently solved so far. Suboptimal approaches to tackle this problem include semidefinite relaxation and cyclic iterative algorithms.

Lemma 2. *Under a total power constraint on the beamforming vector* $\mathbf{w}$ ($W_K = \{\mathbf{w} \in \mathbb{C}^K : \|\mathbf{w}\| \leq 1\}$), *and for a fixed antenna selection* $\mathcal{I}$, *the optimal beamforming vector* $\mathbf{w}(\mathcal{I})$ *is given by the right singular vector of* $\mathbf{H}_{:,\mathcal{I}}$ *that corresponds to its principal singular value* $\sigma_{max}\left(\mathbf{H}_{:,\mathcal{I}}\right)$ *simply called the "principal right singular vector ".*

Proof of Lemma 2 is relegated to Appendix A.2.

To optimally select the K transmit antennas, according to Lemma 2, we need to find K columns of $\mathbf{H}$ that form the $M \times K$ submatrix with the maximum principal singular value. The elements of the optimal set $\mathcal{I}$ are the indices of the optimally selected antennas. Then, for these optimal indices, the optimal unit-norm pruned beamforming vector is given by the principal right singular vector of $\mathbf{H}_{:,\mathcal{I}}$.

Assuming that a solver exists for (2.7) , a straightforward approach to solve (2.6) and identify the optimal set would lead us to an exhaustive search among all $\binom{N}{k}$ cadinality-K subsets of 1,2,...N. However, if the number of selected Transmit antennas K is a linear function of the total number of transmit antennas N, then such a solver would be impractical even for moderate values of N, since its complexity grows exponentially with N. Even when K is not a function of N, when N is large enough (e.g. N=100) and K is greater than 5 then the exhaustive search would be still too large to consider for practical implementation.

In the next chapter we present an efficient algorithm that, for any fixed number of receive Antennas $M \leq 2$ and any unrestricted number of selected transmit antennas K (even if K is linear to N) identifies in time polynomial to N , a polynomial-size collection of antenna selection subsets $\mathcal{I}$ that includes the optimal one.For each antenna candidate subset we compute the corresponding beamforming vector $\mathbf{w}$ via a UQP algorithm for the per-antenna-element power constraint case or as the principal right singular vector for the total power one.The selected pair (Antenna subset-beamforming vector) is determined through a polynomial exhaustive search among the candidate pairs.

Chapter 3

Optimal Transmit Antenna Selection

3.1 One Receive Antenna

We consider now the straight forward case of the problem with $M = 1$ receive antenna. Since $M = 1$ then the transformed channel matrix becomes $\mathbf{H} = \mathbf{h}^H$. For the case of a particular antenna selection $\mathcal{I}$ with the K indices of the selected transmitted antennas it becomes $\mathbf{H}_{:,\mathcal{I}} = \mathbf{h}_{\mathcal{I}}^H$.

3.1.1 Per Antenna Element Power Constraint

In this power constraint scenario where $W_K = \left\{\mathbf{w} \in \mathbb{C}^N : |w_n| \leq 1, n = 1, 2, \ldots, N\right\}$, 2.6 becomes

$$\mathbf{w}(\mathcal{I}) = \underset{\mathbf{w} \in W_K}{\operatorname{argmax}} \left|\mathbf{h}_{\mathcal{I}}^H \mathbf{w}\right| = \underset{\mathbf{w} \in W_K}{\operatorname{argmax}} \left|\sum_{k=1}^{K} h_{\mathcal{I}}^*(k) w(k)\right|.$$

However,

$$\left|\sum_{k=1}^{K} h_{\mathcal{I}}^*(k) w(k)\right| = \left|\sum_{k=1}^{K} |h_{\mathcal{I}}^*(k)| e^{-j\theta_k} \ |w(k)| e^{j\phi_k}\right| =$$

$$\left|\sum_{k=1}^{K} |h_{\mathcal{I}}^*(k) \ w(k)| \ e^{j(\phi_k - \theta_k)}\right| \leq \sum_{k=1}^{K} \left|h_{\mathcal{I}}^*(k) w(k)\right|$$

with equality that holds if $\theta_k = \phi_k \ \ \forall k$, where θ_k and ϕ_k are the arguments-phases of h(k),w(k) respectively.
Now,

$$\sum_{k=1}^{K} \left|h_{\mathcal{I}}^*(k) w(k)\right| = \sum_{k=1}^{K} |h_{\mathcal{I}}^*(k)| \, |w(k)| \leq \sum_{k=1}^{K} |h_{\mathcal{I}}^*(k)| = \|\mathbf{h}_{\mathcal{I}}\|_1$$

with equality if $|w(k)| = 1 \ \ \forall_k$.

So we conclude that we need $arg\left(\mathbf{w}\left(\mathcal{I}\right)\right) = arg\left(\mathbf{h}_{\mathcal{I}}\right)$ and the absolute values of all entries of $\mathbf{w}(\mathcal{I})$ to be equal to 1 which results in

:

$$\mathbf{w}\left(\mathcal{I}\right) = e^{jarg(\mathbf{h}_{\mathcal{I}})}.$$

3.1.2 Total Power Constraint

Under the total power constraint scenario where $W_K = \left\{\mathbf{w} \in \mathbb{C}^K : \|\mathbf{w}\| \leq 1\right\}$, we obtain

$$\mathbf{w}(\mathcal{I}) = \underset{\mathbf{w}\in W_K}{\operatorname{argmax}} \left|\mathbf{h}_{\mathcal{I}}^H \mathbf{w}\right|$$

with $\left|\mathbf{h}_{\mathcal{I}}^H \mathbf{w}\right| \leq \|\mathbf{h}_{\mathcal{I}}\| \ \|\mathbf{w}\| \leq \|\mathbf{h}_{\mathcal{I}}\|$.

The first (Cauchy-Schwarz) inequality holds when $\mathbf{w}$ and $\mathbf{h}$ are linearly dependent $\Rightarrow \mathbf{w} = a\mathbf{h},\ a \neq 0$.

The second inequality holds when $\|\mathbf{w}\| = 1$ given the total power constraint ($\|\mathbf{w}\| \leq 1$),
$\Longrightarrow \|a\mathbf{h}_{\mathcal{I}}\| = 1 \Rightarrow |a| \ \|\mathbf{h}_{\mathcal{I}}\| = 1 \Leftrightarrow |a| = \dfrac{1}{\|\mathbf{h}_{\mathcal{I}}\|}$.

It is easy from the above to denote that the optimal beamforming vector is the Maximal-ratio-combining (MRC) vector :

$$\mathbf{w}_{\mathcal{I}} = \frac{\mathbf{h}_{\mathcal{I}}}{\|\mathbf{h}_{\mathcal{I}}\|}$$

and (2.9) becomes

$$\mathcal{I}_{opt} = \underset{\mathcal{I}\in\mathcal{S}}{\operatorname{argmax}} \left|\mathbf{h}_{\mathcal{I}}^H \frac{\mathbf{h}_{\mathcal{I}}}{\|\mathbf{h}_{\mathcal{I}}\|}\right| = \underset{\mathcal{I}\in\mathcal{S}}{\operatorname{argmax}} \|\mathbf{h}_{\mathcal{I}}\| \quad = \underset{\mathcal{I}\in\mathcal{S}}{\operatorname{argmax}} \sqrt{\sum_{i\in\mathcal{I}} |h_i|^2}\ .$$

3.1.3 Antenna selection

As shown from the above sections, in both scenarios (Total Power Constraint, Per-Antenna-Element Power Constraint) the optimal set $\mathcal{I}$ is the one that consists of the indices corresponding to the K largest elements of $|\mathbf{h}|$.

For this purpose, we present the function *select* which as its name indicates, selects the K largest (in magnitude) elements of a vector:

$$\operatorname{select}(\mathbf{u}; k) \triangleq \underset{\substack{\mathcal{I}\subset\{1,2,\ldots,N\} \\ |\mathcal{I}|=k}}{\arg\max} \|\mathbf{u}_{\mathcal{I}}\|_1\ .$$

That is, select computes $|u_1|, |u_2|, ..., |u_N|$ and returns the indices of the largest K values . The computational cost of select is $O(N)$ [26], [27], linear in the number of available transmit antennas N. We conclude that for both antenna power constraint scenarios, the optimal transmit antenna selection when $M = 1$ receive antenna is occupied, is given by

$$\mathcal{I} = \operatorname{select}(\mathbf{h}; K),$$

whose complexity is linear in the number of available transmit antennas N. Given the antenna selection , optimal beamforming vector is then calculated/approximated for total power constraint and per antenna power constraint respectively. In the developments that follow when $M = 2$, *select* is critical in obtaining in polynomial time a polynomial-size collection of antenna subset candidates that include the optimal one.

In the next subsection we illustrate the BER versus the total number of available transmit antennas for both power constraint scenarios, regarding the straightforward case of this section where we occupy one antenna on the receiver.

3.1.4 One receive antenna performance

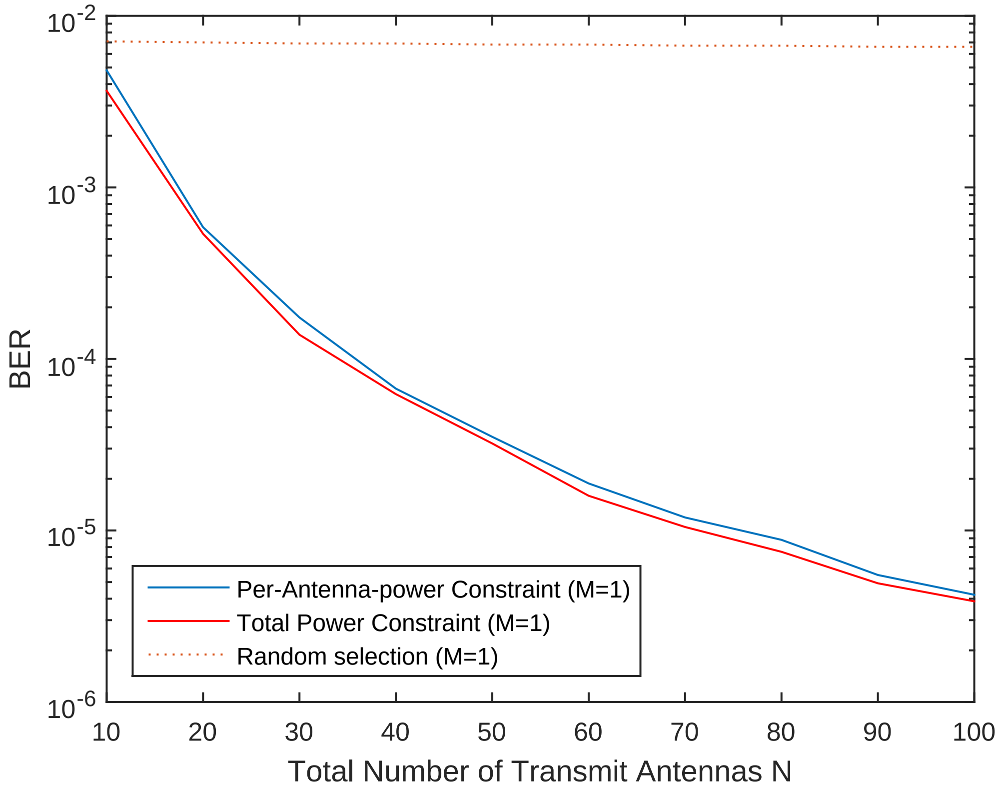

Figure 3.1: Bit error rate versus total number of transmit antennas N for $M = 1$ receive antennas and selection of $K = 6$ transmit antennas, for both per-antenna element and power constraint scenarios.

Concluding with the straight forward scenario where we occupy one antenna on the receiver, we present Fig. 3.1, where we consider both power constraint cases (per-antenna, Total power constraint) for transmissions in the presence of white noise, set the total transmit SNR to -3dB, and plot the bit error rate as a function of the available transmit antennas N and compare them against random antenna selection.

3.2 Two Receive Antennas

When two antennas are occupied on the receiver ($M = 2$), the channel matrix becomes $\mathbf{H} = [\mathbf{h}_1\ \mathbf{h}_2]^H$ and the problem of selecting the optimal subset $\mathcal{I}_{opt}$ becomes more challenging.In this subsection we show that this problem can be solved with $O(N^4)$ complexity for any number of K selected antennas.

First we redefine the problem space by introducing the auxiliary angles $\phi \in [0, \pi/2]$ and $\theta \in (-\pi, \pi]$ and defining the unit-norm vector

$$\mathbf{c}(\phi, \theta) \triangleq \begin{bmatrix} \sin(\phi) \\ e^{j\theta}\cos(\phi) \end{bmatrix}. \tag{3.1}$$

Angles ϕ and θ will help us identify a polynomial number of *locally optimal* antenna selection sets $\mathcal{I}$ consisting the optimal solution of (2.6).

Since $\left\|\mathbf{c}(\phi, \theta)\right\| = 1$, from Cauchy-Schwarz Inequality, we get for any vector $\mathbf{a} \in \mathbb{C}^2$

$$\left|\mathbf{a}^H\mathbf{c}(\phi, \theta)\right| \leq \|\mathbf{a}\| \left\|\mathbf{c}(\phi, \theta)\right\| = \|\mathbf{a}\|. \tag{3.2}$$

The equality above is achieved <u>iff</u> $\mathbf{c}(\phi, \theta)$ is collinear with $\mathbf{a}$ within a phase rotation such that its first element is positive,i.e., if and only if

$$\mathbf{c}(\phi, \theta) = \frac{\mathbf{a}}{\|\mathbf{a}\|} e^{-j\arg(a_1)}. \tag{3.3}$$

Indeed we note that $\left\|\mathbf{c}(\phi, \theta)\right\| = 1$, that $\mathbf{c}$ is collinear with $\mathbf{a}$, and that its first element being $\dfrac{|\mathbf{a}_1|\ e^{j\arg(a_1)}}{\|\mathbf{a}\|}\ e^{-j\arg(a_1)} = \dfrac{|\mathbf{a}_1|}{\|\mathbf{a}\|}$ is positive.

For any $\mathbf{a} \in \mathbb{C}^2$ there always exists a pair of angles $(\phi, \theta) \in \Phi$, where $\Phi \triangleq [0, \pi/2] \times (-\pi, \pi]$, such that (3.3) is satisfied. Therefore from (3.2) we obtain

$$\|\mathbf{a}\| = \max_{(\phi,\theta)\in\Phi} \left|\mathbf{a}^H\mathbf{c}(\phi, \theta)\right|. \tag{3.4}$$

If we substitute $\mathbf{a}$ with $\mathbf{H}_{:,\mathcal{I}}\mathbf{w}$ in (3.4) then our optimization problem in (2.6) becomes

$$\max_{\substack{\mathbf{w}\in W_N \\ \|\mathbf{w}\|_0 = K}} \|\mathbf{H}\mathbf{w}\| = \max_{\mathcal{I}\in\mathcal{S}} \max_{\mathbf{w}\in\mathcal{W}_K} \max_{(\phi,\theta)\in\Phi} \left|\mathbf{w}^H\mathbf{H}_{:,\mathcal{I}}^H\mathbf{c}(\phi, \theta)\right|. \tag{3.5}$$

Defining $\mathbf{u}(\phi, \theta) \triangleq \mathbf{H}^H\mathbf{c}(\phi, \theta)$ and changing the order of the maximizations in (3.5) we get:

$$\max_{(\phi,\theta)\in\Phi} \max_{\mathcal{I}\in\mathcal{S}} \max_{\mathbf{w}\in\mathcal{W}_K} \left|\mathbf{w}^H\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right|. \tag{3.6}$$

Similar to the previous subsection ($M = 1$) we tend to examine the two power constraint scenarios and derive the optimal beamforming vector $\mathbf{w}$ for any pair of angles (ϕ, θ) and any given antenna selection subset $\mathcal{I}$.

3.2.1 Per Antenna Element Power Constraint

Under a per antenna element power constraint scenario the inner maximization in (3.6) becomes as follows

$$\mathbf{w}(\phi,\theta;\mathcal{I}) = \underset{\mathbf{w}\in W_K}{\operatorname{argmax}} \left|\mathbf{w}^H \mathbf{u}_{\mathcal{I}}(\phi,\theta)\right| = \underset{\mathbf{w}\in W_K}{\operatorname{argmax}} \left|\mathbf{u}_{\mathcal{I}}^H(\phi,\theta)\ \ \mathbf{w}\right| = \underset{\mathbf{w}\in W_K}{\operatorname{argmax}} \left|\sum_{k=1}^{K} \mathbf{u}_{\mathcal{I}}^*(\phi,\theta;k)\mathbf{w}(k)\right|$$

However,

$$\left|\sum_{k=1}^{K} \mathbf{u}_{\mathcal{I}}^*(\phi,\theta;k)\mathbf{w}(k)\right| = \left|\sum_{k=1}^{K} \left|\mathbf{u}_{\mathcal{I}}^*(\phi,\theta;k)\right|\ e^{-j\theta_k}\ \ \left|\mathbf{w}(k)\right| e^{j\phi_k}\right| =$$

$$\left|\sum_{k=1}^{K} \left|\mathbf{u}_{\mathcal{I}}^*(\phi,\theta;k)\ \mathbf{w}(k)\right| e^{j(\phi_k-\theta_k)}\right| \le \sum_{k=1}^{K} \left|\mathbf{u}_{\mathcal{I}}^*(\phi,\theta;k)\mathbf{w}(k)\right|,$$

with equality that holds if $\theta_k = \phi_k\ \ \forall k$ where θ_k, ϕ_k are the arguments-phases of $\mathbf{u}(\phi,\theta;k), \mathbf{w}(k)$ respectively.
Now,

$$\sum_{k=1}^{K} \left|\mathbf{u}_{\mathcal{I}}^*(\phi,\theta;k)\mathbf{w}(k)\right| = \sum_{k=1}^{K} \left|\mathbf{u}_{\mathcal{I}}^*(\phi,\theta;k)\right| \left|\mathbf{w}(k)\right| \le \sum_{k=1}^{K} \left|\mathbf{u}_{\mathcal{I}}^*(\phi,\theta;k)\right| = \left\|\mathbf{u}_{\mathcal{I}}(\phi,\theta)\right\|_1$$

with equality if $|\mathbf{w}(k)| = 1\ \ \forall_k$.

So we conclude that we need $arg(\mathbf{w}(\mathcal{I})) = arg(\mathbf{u}_{\mathcal{I}}(\phi,\theta))$ and the absolute values of all entries of $\mathbf{w}(\mathcal{I})$ to be equal to 1.The vector that satisfies the above equalities is

$$\mathbf{w}(\phi,\theta;\mathcal{I}) = e^{j arg(\mathbf{u}_{\mathcal{I}}(\phi,\theta))}. \tag{3.7}$$

Then the optimization problem in(3.6) becomes

$$\max_{(\phi,\theta)\in\Phi}\ \max_{\mathcal{I}\in\mathcal{S}} \left\|\mathbf{u}_{\mathcal{I}}(\phi,\theta)\right\|_1 = \max_{(\phi,\theta)\in\Phi}\ \max_{\mathcal{I}\in\mathcal{S}} \sum_{i\in\mathcal{I}} \left|\mathbf{u}_i(\phi,\theta)\right| \tag{3.8}$$

where, for any given $(\phi,\theta)\in\Phi$, the inner maximization is achieved by the subset that consists of the indices of the K largest elements of $|\mathbf{u}(\phi,\theta)|$.

3.2.2 Total Power Constraint

In the same manner, for any given pair of angles $(\phi, \theta) \in \Phi$, and any given selection subset $\mathcal{I} \in \mathcal{S}$ we obtain

$$\mathbf{w}(\phi, \theta; \mathcal{I}) = \underset{\mathbf{w} \in W_K}{\operatorname{argmax}} |\mathbf{w}^H \mathbf{u}_{\mathcal{I}}(\phi, \theta)| = \underset{\mathbf{w} \in W_K}{\operatorname{argmax}} |\mathbf{u}_{\mathcal{I}}^H(\phi, \theta)\ \mathbf{w}|$$

$$\text{with } |\mathbf{u}_{\mathcal{I}}^H(\phi, \theta)\ \mathbf{w}| \leq \left\|\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\|\ \|\mathbf{w}\| \leq \left\|\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\|.$$

The first (Cauchy-Schwarz) inequality holds when $\mathbf{w}$ and $\mathbf{u}$ are linearly dependent $\Rightarrow \mathbf{w} = a\, \mathbf{u}_{\mathcal{I}}(\phi, \theta),\ a \neq 0$.

The second inequality holds when $\|\mathbf{w}\| = 1$ given the total power constraint ($\|\mathbf{w}\| \leq 1$).

$$\Longrightarrow \left\|a\, \mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\| = 1 \Rightarrow |a|\ \left\|\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\| = 1 \Leftrightarrow |a| = \frac{1}{\left\|\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\|}.$$

Its easy from the above to denote that the optimal beamforming vector is the $K \times 1$ size Maximal-ratio-combining (MRC) vector :

$$\mathbf{w}(\phi, \theta; \mathcal{I}) = \frac{\mathbf{u}_{\mathcal{I}}(\phi, \theta)}{\left\|\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\|} \tag{3.9}$$

resulting in the value

$$\max_{\substack{\mathbf{w} \in \mathbb{C}^K \\ \|\mathbf{w}\|=1}} |\mathbf{w}^H \mathbf{u}_{\mathcal{I}}(\phi, \theta)| = \left\|\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\|. \tag{3.10}$$

Then the optimization problem becomes

$$\max_{(\phi,\theta) \in \Phi}\ \max_{\mathcal{I} \in \mathcal{S}} \left\|\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\| \tag{3.11}$$

Given a fixed pair of angles (ϕ, θ) the inner maximization is equivalent to

$$\begin{aligned} \max_{\mathcal{I} \in \mathcal{S}} \left\|\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\| &= \max_{\mathcal{I} \in \mathcal{S}} \sqrt{\sum_{i \in \mathcal{I}} |\mathbf{u}_i(\phi, \theta)|^2} \equiv \\ \max_{\mathcal{I} \in \mathcal{S}} \left\|\mathbf{u}_{\mathcal{I}}(\phi, \theta)\right\|_1 &= \max_{\mathcal{I} \in \mathcal{S}} \sum_{i \in \mathcal{I}} |\mathbf{u}_i(\phi, \theta)|. \end{aligned} \tag{3.12}$$

From both power constraint scenarios, we observe that the optimal set in (2.8) and (2.10), given a fixed pair (ϕ, θ), is the one that consists of the indices of the K largest values of $|\mathbf{u}(\phi, \theta)|$,i.e.,

$$\mathcal{I}(\phi, \theta) = \text{select}(\mathbf{u}(\phi, \theta); K) = \text{select}\left(\begin{bmatrix} \mathbf{u}_1(\phi, \theta) \\ \mathbf{u}_1(\phi, \theta) \\ \vdots \\ \mathbf{u}_1(\phi, \theta \end{bmatrix}; K \right). \tag{3.13}$$

That is for any given $(\phi, \theta) \in \Phi$, the corresponding antenna selection subset $\mathcal{I}(\phi, \theta)$is obtained with complexity $O(N)$.

However we should explain why(ϕ, θ) simplifies the computation of the solution. We notice that for any $n = 1, 2, ..., N$,

$$|\mathbf{u}_n(\phi, \theta)| = |\mathbf{H}^H_{:,n}\mathbf{c}(\phi, \theta)| = |\mathbf{H}^*_{1,n}\sin(\phi) + \mathbf{H}^*_{2,n}e^{j\theta}\cos(\phi)|, \tag{3.14}$$

i.e., every element of $\left|\mathbf{u}(\phi, \theta)\right|$ is a continuous function, or a surface, of (ϕ, θ). When, due to (2.13), we select the K largest elements of $\left|\mathbf{u}(\phi, \theta)\right|$ at a given point (ϕ, θ) as function *select* requires, we actually compare the surfaces $\left|\mathbf{u}_1(\phi, \theta)\right|, \left|\mathbf{u}_2(\phi, \theta)\right|, \ldots, \left|\mathbf{u}_N(\phi, \theta)\right|$ at a point (ϕ, θ). The optimal selection $\mathcal{I} \in \{1, 2, \ldots, N\}$in (2.5) is met if we scan the entire space Φ and collect the locally optimal selection $\mathcal{I}(\phi, \theta)$ for any point (ϕ, θ).Due to the continuity of the surfaces $\left|\mathbf{u}_n(\phi, \theta)\right|$ we expect that in an area around (ϕ, θ), the selection subset $\mathcal{I}(\phi, \theta)$ will be retained because either the sorting of the surfaces does not change, or the sorting of the surfaces changes but the but the group of the K(number of selected antennas) surfaces with the higher value is retained.

Hence, we expect the formation of regions in Φ within which the optimal selection subset $\mathcal{I}$ is unique.In the sequel we determine all these regions ,show that their number is less than or equal to $6\binom{N}{3}$ and present a polynomial-time algorithm that identifies the selection subsets $\mathcal{I}$ that correspond to these regions.Once we have collected all candidate subsets,for the total-power-constraint scenario, the solution is derived from a polynomial-time exhaustive search among them.

Scanning the space Φ we observe that the selection subset $\mathcal{I}$ does not change unlesss two surfaces intersect (which implies that the sorting of the surfaces $\left|\mathbf{u}(\phi, \theta)\right|$ changes). Therefore to identify all regions that retain their selection subset $\mathcal{I}$, it suffices to examine when two surfaces intersect.We not that this is a necessary, but not sufficient, condition for a change of $\mathcal{I}$, since the intersecting surfaces may correspond to elements of $\left|\mathbf{u}(\phi, \theta)\right|$ that, before they intersect neither of them belonged to $\mathcal{I}$. In this case although the sorting of the surfaces changes, the selection subset $\mathcal{I}$ does not.

Two surfaces, say $\left|\mathbf{u}_n(\phi, \theta)\right|$ and $\left|\mathbf{u}_m(\phi, \theta)\right|$, intersect when $\left|\mathbf{u}_n(\phi, \theta)\right| = \left|\mathbf{u}_m(\phi, \theta)\right|$. We note that, for any $n, m \in \{1, 2, \ldots, N\}$, the intersection between them always exists, since all surfaces meet 0 for some (ϕ, θ). This holds, because for any surface $\left|\mathbf{u}_n(\phi, \theta)\right| = \left|\mathbf{H}^H_{:,n}\mathbf{c}(\phi, \theta)\right|$, we can always obtain a vector $\mathbf{c}(\phi, \theta)$ which is orthogonal to $\mathbf{H}_{:,n}$. As a result, the intersection of two surfaces determines a curve on the (ϕ, θ)-plane whice we define as $\mathcal{L}_{n,m} \triangleq \left\{(\phi, \theta) \in \Phi : \left|\mathbf{u}_n(\phi, \theta)\right| = \left|\mathbf{u}_m(\phi, \theta)\right|\right\}$.

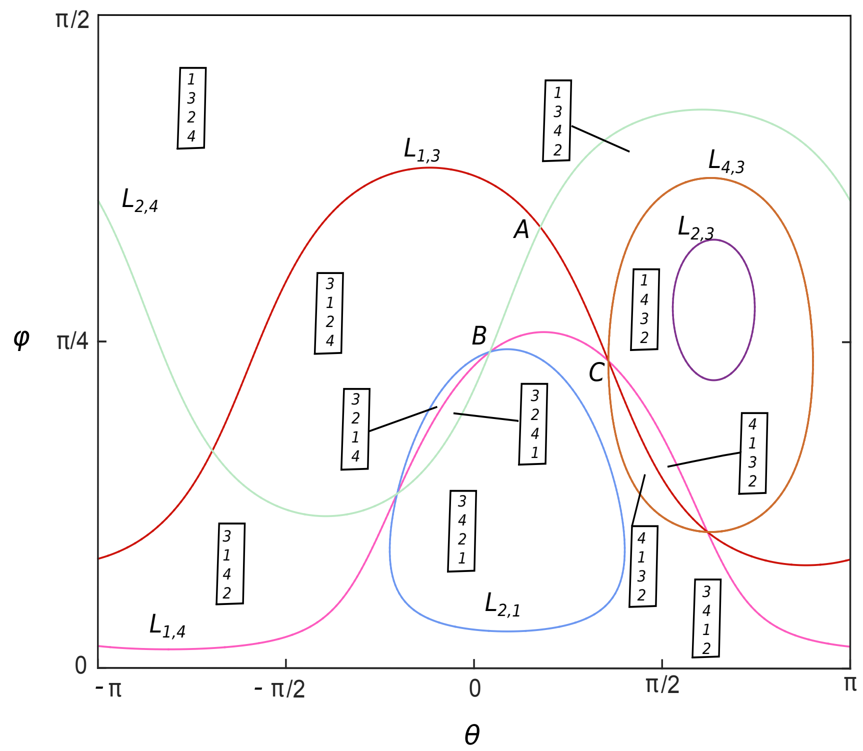

Figure 3.2: An illustration of intersection curves $L_{n,m}$, for $n, m \in \{1, 2, 3, 4\}$ with $n \neq m$, and formed regions, resulting from $N = 4$ surfaces.

We illustrate this in Figure 2.1 where we set $N = 4$ and consider an arbitrary 2×4 channel matrix $\mathbf{H}$, and plot curves $\mathcal{L}_{n,m}$, for any $n, m \in \{1, 2, \ldots, N\}$ with $n \neq m$.

We now observe the regions that are formed. Within each region, the selection subset $\mathcal{I}$ remains the same. We also observe that, in most cases, each region ‘’touches” an intersection of two or three curves. That is, we can identify these regions by examining intersections between curves $L_{n,m}$. In addition, we can concentrate only on three curve intersections and ignore intersections of two curves, since the latter change the sorting of surfaces but do not generate a new subset $\mathcal{I}$ that has not been generated by a neighboring three curve intersection.

3.2.3 Polynomial-Complexity Antenna Selection Algorithm

As mentioned in the previous subsection we start by examining all three-curved intersections. We observe that each such intersection corresponds to a three-surface intersection. Hence, it suffices to examine when three surfaces, say $\left|\mathbf{u}_n(\phi,\theta)\right|, \left|\mathbf{u}_m(\phi,\theta)\right|$ and $\left|\mathbf{u}_l(\phi,\theta)\right|$, where $n, m, l \in \{1, 2, \ldots, N\}$ with $n \neq m$, $m \neq l$, $n \neq l$, intersect. That is, we have to find (ϕ, θ) that satisfies:

$$\left|\mathbf{u}_n(\phi,\theta)\right| = \left|\mathbf{u}_m(\phi,\theta)\right| = \left|\mathbf{u}_l(\phi,\theta)\right| \tag{3.15}$$

By substituting $\mathbf{u}(\phi,\theta)$ with $\mathbf{H}^H\mathbf{c}(\phi,\theta)$ we get:

$$\left|\mathbf{H}^H_{:,n}\mathbf{c}(\phi,\theta)\right| = \left|\mathbf{H}^H_{:,m}\mathbf{c}(\phi,\theta)\right| = \left|\mathbf{H}^H_{:,l}\mathbf{c}(\phi,\theta)\right| \tag{3.16}$$

Given two complex numbers $a, b \in \mathbb{C}$ with $|a| = |b|$ we get $a = e^{j\lambda}\, b$ for some $\lambda \in \mathbb{R}$.Then equation (2.17) becomes:

$$\begin{Bmatrix} e^{j\lambda}\, \mathbf{H}^H_{:,n}\mathbf{c}(\phi,\theta) = \mathbf{H}^H_{:,m}\mathbf{c}(\phi,\theta) \\ \\ e^{j\mu}\, \mathbf{H}^H_{:,n}\mathbf{c}(\phi,\theta) = \mathbf{H}^H_{:,m}\mathbf{c}(\phi,\theta) \end{Bmatrix} \tag{3.17}$$

or, equivalently,

$$\begin{bmatrix} e^{j\lambda}\, \mathbf{H}^H_{:,n} - \mathbf{H}^H_{:,m} \\ \\ e^{j\mu}\, \mathbf{H}^H_{:,n} - \mathbf{H}^H_{:,l} \end{bmatrix} \mathbf{c}(\phi,\theta) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{3.18}$$

for some $\lambda, \mu \in \mathbb{R}$. We define the 2×2 matrix

$$\mathbf{A} = \begin{bmatrix} e^{j\lambda}\, \mathbf{H}^H_{:,n} - \mathbf{H}^H_{:,m} \\ \\ e^{j\mu}\, \mathbf{H}^H_{:,n} - \mathbf{H}^H_{:,l} \end{bmatrix}. \tag{3.19}$$

We note that $\mathbf{c}(\phi,\theta)$ belongs to the nullspace of matrix $\mathbf{A}$. In order to obtain the unit norm vector $\mathbf{c}$ we need $\mathcal{N}(\mathbf{A}) \neq \{\mathbf{0}_{n\times1}\}$, which occurs when $\mathbf{A}$ is singular, .i.e, rank($\mathbf{A}$)=1 or, equivalently, $|\mathbf{A}|$=0.
So if $\mathbf{A}$ is singular which leads us to $|\mathbf{A}| = 0$, we might obtain $\mathbf{c}(\phi,\theta)$, find the coordinates (ϕ,θ) of the triple intersection point, and call function *select* for this particular point. To proceed further we form

$$\mathbf{A}^H = \begin{bmatrix} e^{-j\lambda}\mathbf{H}_{:,n} & e^{-j\mu}\mathbf{H}_{:,n} \end{bmatrix} - \begin{bmatrix} \mathbf{H}_{:,m} & \mathbf{H}_{:,l} \end{bmatrix}. \tag{3.20}$$

Lemma 3. *Suppose* $\mathbf{A}$ *is an invertible square matrix and* $\mathbf{u}, \mathbf{v}$ *are column vectors. Then the matrix determinant lemma states that*

$$\det(\mathbf{A} + \mathbf{u}\mathbf{v}^{\mathrm{T}}) = (1 + \mathbf{v}^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{u}) \det(\mathbf{A}).$$

Proof of Lemma 3 is available at appendix A.3.

In order to find the coordinates of the triple intersection point (ϕ, θ) we set the determinant of the matrix $-\mathbf{A}^H$ to zero. Utilizing Lemma 3 on the same matrix we obtain

$$\begin{aligned} &\left|-\mathbf{A}^H\right| = 0 \\ \Leftrightarrow &\left(1 - \begin{bmatrix} e^{-j\lambda} & e^{-j\mu} \end{bmatrix} \begin{bmatrix} \mathbf{H}_{:,m} & \mathbf{H}_{:,l} \end{bmatrix}^{-1} \mathbf{H}_{:,n}\right) \left|\begin{bmatrix} \mathbf{H}_{:,m} & \mathbf{H}_{:,l} \end{bmatrix}\right| = 0. \end{aligned} \tag{3.21}$$

We note that $\left|\begin{bmatrix} \mathbf{H}_{:,m} & \mathbf{H}_{:,l} \end{bmatrix}\right| \neq 0$. That is because the columns of the matrix $\mathbf{H}_{:,m}$, $\mathbf{H}_{:,l}$ are linearly independent, i.e., the matrix is full rank.

We define the 2×1 vector

$$\mathbf{d} \triangleq \begin{bmatrix} \mathbf{H}_{:,m} & \mathbf{H}_{:,l} \end{bmatrix}^{-1} \mathbf{H}_{:,n} \tag{3.22}$$

and the scalar

$$D \triangleq \frac{1 - \|\mathbf{d}\|^2}{2\,|d_1 d_2|}. \tag{3.23}$$

Therefore , to find $\lambda, \mu \in \mathbb{R}$ and solve for $\mathbf{c}(\phi, \theta)$ in (2.18), it suffices to solve

$$\begin{bmatrix} e^{-j\lambda} & e^{-j\mu} \end{bmatrix} \cdot \mathbf{d} = 1 \tag{3.24}$$

Given that $\left\|\begin{bmatrix} e^{-j\lambda} & e^{-j\mu} \end{bmatrix}\right\| = \sqrt{\left|e^{-j\lambda}\right|^2 + \left|e^{-j\mu}\right|^2} = \sqrt{2}$, when we apply Cauchy-Schwarz Inequality on (2.24) we get

$$\left\|\begin{bmatrix} e^{-j\lambda} & e^{-j\mu} \end{bmatrix} d\right\| \leq \left\|\begin{bmatrix} e^{-j\lambda} & e^{-j\mu} \end{bmatrix}\right\| \|d\|$$

$$\Leftrightarrow 1 \leq \sqrt{2}\|d\|$$

$$\Leftrightarrow \|d\| \geq \frac{1}{\sqrt{2}},$$

which is a necessary condition for (3.24) to have a solution (and, consequently for a triple intersection to exist).If the latter is satisfied, then (3.24) becomes

$$e^{-j\lambda} \begin{bmatrix} 1 & e^{-j(\mu-\lambda)} \end{bmatrix} \cdot \mathbf{d} = 1 \Leftrightarrow \begin{bmatrix} 1 & e^{-j(\mu-\lambda)} \end{bmatrix} \cdot \mathbf{d} = e^{j\lambda} \tag{3.25}$$

implying

$$\left|\begin{bmatrix}1 & e^{-j\psi}\end{bmatrix} \cdot \mathbf{d}\right| = 1 \tag{3.26}$$

where $\psi = \mu - \lambda$. We carry on, by solving (3.26) in order to obtain the values of ψ that make matrix $\mathbf{A}$ singular.

$$\left|\begin{bmatrix}1 & e^{-j\psi}\end{bmatrix} \cdot \mathbf{d}\right| = 1 \Rightarrow \left|\begin{bmatrix}1 & e^{-j\psi}\end{bmatrix} \cdot \mathbf{d}\right|^2 = 1^2$$

$$\Leftrightarrow \left|\begin{bmatrix}1 & e^{-j\psi}\end{bmatrix} \cdot \begin{bmatrix}d1 \\ d2\end{bmatrix}\right|^2 = 1$$

$$\Leftrightarrow \left|d_1 + e^{-j\psi} d_2\right|^2 = 1.$$

$$\Leftrightarrow |d_1|^2 + \left|e^{-j\psi} d_2\right|^2 + 2\mathbb{R}e\left(e^{-j\psi} d_1 d_2^*\right) = 1.$$

$$\Leftrightarrow |d_1|^2 + \underbrace{\left|e^{-j\psi}\right|}_{=1} |d_2|^2 + 2\mathbb{R}e\left(e^{-j\psi}|d_1 d_2| e^{-jangle\left(d_1 d_2^*\right)}\right) = 1.$$

$$\Leftrightarrow \underbrace{|d_1|^2 + |d_2|^2}_{=\|d\|^2} + 2\,|d_1 d_2|\,\mathbb{R}e\left(e^{-j\left(\psi + angle\left(d_1 d_2^*\right)\right)}\right) = 1.$$

$$\Leftrightarrow \cos\left(\psi + angle\left(d_1 d_2^*\right)\right) = \frac{1 - \|d\|^2}{2\,|d_1 d_2|}$$

$$\Leftrightarrow \psi + angle\left(d_1 d_2^*\right) = \pm \cos^{-1} D$$

which results in two values of ψ tha make matrix $\mathbf{A}$ singular, i.e., two intersection points

$$\psi = -angle\left(d_1 d_2^*\right) \pm \cos^{-1} D \tag{3.27}$$

From the solution of (3.27), we can claim that $|D| \leq 1$ is a necessary and sufficient condition to make matrix $\mathbf{A}$ singular. If the condition is not satisfied, then no intersection point exists between the three surfaces (and their corresponding curves). If the condition is satisfied though, after computing ψ in (3.27), we can compute $\lambda \triangleq \arg\left(\begin{bmatrix}1 & e^{-j\psi}\end{bmatrix}\mathbf{d}\right)$ and $\mu = \psi + \lambda$, matrix $\mathbf{A}$ from(3.19), and $\mathbf{c}(\phi, \theta)$ as the unit-norm vector in the null space of $\mathbf{A}$. Finally, from (3.1) we can derive the coordinates of the intersection point (ϕ, θ). Since there are two values of ψ from the solution that make $\mathbf{A}$ singular, we obtain two intersection points.

For any combination of three surfaces, the above procedure, with complexity $\mathcal{O}(1)$, examines if they intersect and, if so, computes the two intesection points (ϕ, θ). Each intersection is a vertex of six regions.Then function $\mathcal{I} = select\left(\mathbf{u}(\phi, \theta); K\right)$ computes, with complexity $\mathcal{O}(N)$, the indices of the K largest surfaces. We recall from (3.15) that, at the intersection point, all three surfaces that intersect have the same value. If $\mathcal{I}$ contains the indices n, m, l of the three surfaces that intersect or it does not contain either of them, then $\mathcal{I}$ is the common optimal selection subset of for all six neighboring regions. Otherwise we have to consider three different cases for the index among n, m, l that belongs or does not belong to $\mathcal{I}$. Overall, for each combination of three surfaces, we obtain at most 6 candidate sets $\mathcal{I}$ with complexity $\mathcal{O}(N)$. Since the total number of surfaces is N, the overall number of candidate sets is upper bounded by $6\binom{N}{3}$.

Finally, regarding the case where a region does not touch any intersection, we can see that such a region touches a curve $\mathcal{L}_{n,m}$ which does not intersect with any other curve. For example, in Figure 2.1, such a curve is $\mathcal{L}2, 3$.Therefore, when we examine the three-surface intersections as described before, we should mark the corresponding curves and, in the end, if a curve, say $\mathcal{L}_{n,m}$, is unmarked, then we should examine it separately as follows. We need to pick a point on $\mathcal{L}_{n,m}$ by solving

$$\left|\mathbf{H}_{:,n}^{H}\mathbf{c}(\phi, \theta)\right| = \left|\mathbf{H}_{:,m}^{H}\mathbf{c}(\phi, \theta)\right| \tag{3.28}$$

or, equivalently,

$$\left(e^{j\lambda}\mathbf{H}_{:,n} - \mathbf{H}_{:,m}\right)\mathbf{c}(\phi, \theta) = 0, \tag{3.29}$$

for some $\lambda \in \mathbb{R}$. Actually, we can set λ to zero, obtain $\mathbf{c}(\phi, \theta)$ as the unit-norm vector in the null space of $\mathbf{H}_{:,n} - \mathbf{H}_{:,m}$, and from 3.1 uniquely determine the intersection point (ϕ, θ).

We now present the algorithm [1] developed and proposed in [22], that computes the optimal antenna selection subset (among N antennas) with complexity $\mathcal{O}\left(N^4\right)$, independently of the selection paremeter K.

Input : $\mathbf{H} \in \mathbb{C}^{2\times N}$ (channel matrix) , $K \in \{1, 2, \ldots, N\}$ (desired selection)

$\mathcal{S}' \leftarrow \{\ \}$ (set of candidates)

$L_{n,m} \leftarrow 0 \ \ \forall_{n,m} \in \{1, 2, \ldots, N\}, n \neq m$

for $\{j_1, j_2, j_3\} \subseteq \{1, 2, \ldots, N\} : j_1 \neq j_2, j_2 \neq j_3, j_1 \neq j_3$

  $\mathbf{d} \leftarrow \left[\mathbf{H}_{:,j_1}\ \mathbf{H}_{:,j_3}\right]^{-1} \mathbf{H}_{:,j_2}$

  $D \leftarrow \dfrac{1 - \|\mathbf{d}\|^2}{2\,|d_1 d_2|}$

  if $|D| \leq 1$ then

    $\mathcal{L}j_1, j_2 \leftarrow 1, \mathcal{L}j_1, j_3 \leftarrow 1, \mathcal{L}j_2 j_3 \leftarrow 1$

    $\psi \leftarrow -\text{angle}\left(d_1 d_2^*\right) \pm \cos^{-1} D$

    $\lambda \leftarrow \text{angle}\left(\begin{bmatrix}1 & e^{-j\psi}\end{bmatrix} \mathbf{d}\right)$

    $\mu \leftarrow \psi + \lambda$

    $\mathbf{c} \leftarrow \text{null}\left(\begin{bmatrix} e^{j\lambda}\ \mathbf{H}^H_{:,j_2} - \mathbf{H}^H_{:,j_1} \\ e^{j\mu}\ \mathbf{H}^H_{:,j_2} - \mathbf{H}^H_{:,j_3} \end{bmatrix}\right)$

    $\mathbf{u} \leftarrow \left|\mathbf{H}^H \mathbf{c}\right|$

    $\mathcal{I} = \text{select}\,(\mathbf{u}; K)$ if multiple entries equal to the Kth order element, include all in $\mathcal{I}$

    if $|\mathcal{I}| = K$

      $\mathcal{S}' \leftarrow \mathcal{S}' \cup \{\mathcal{I}\}$

    elseif $|\mathcal{I}| = K + 1$

      $\mathcal{I}' \leftarrow \mathcal{I} - \{j_1, j_2, j_3\}$

      $\mathcal{S}' \leftarrow \mathcal{S}' \cup \left\{\mathcal{I}' \cup \{j_1, j_2\}, \mathcal{I}' \cup \{j_1, j_3\}, \mathcal{I}' \cup \{j_1, j_3\}\right\}$

    else

      $\mathcal{I}' \leftarrow \mathcal{I} - \{j_1, j_2, j_3\}$

      $\mathcal{S}' \leftarrow \mathcal{S}' \cup \left\{\mathcal{I}' \cup \{j_1\}, \mathcal{I}' \cup \{j_2\}, \mathcal{I}' \cup \{j_3\}\right\}$

for $\{n, m\} : L_{n,m} = 0$, when a curve does not intersect with any other curve

  $\mathbf{c} \leftarrow \text{null}\left(\left[\mathbf{H}^H_{:,n} - \mathbf{H}^H_{:,m}\right]\right)$

  $\mathbf{u} \leftarrow \left|\mathbf{H}^H \mathbf{c}\right|$

  $\mathcal{I} = \text{select}\,(\mathbf{u}; K)$ if multiple entries equal to the Kth order element, include all in $\mathcal{I}$

  if $|\mathcal{I}| = K$

    $\mathcal{S}' \leftarrow \mathcal{S}' \cup \{\mathcal{I}\}$

  else

    $\mathcal{S}' \leftarrow \mathcal{S}' \cup \left\{\mathcal{I}' \cup \{m\}, \mathcal{I}' \cup \{n\}\right\}$

Output : $\mathcal{S}'$

---

[1] While the algorithm is the same with the one proposed at [22], corrections have been made considering the lines where vectors $\mathbf{c}, \mathbf{u}$ are being computed.

3.2.4 Two receive antennas performance

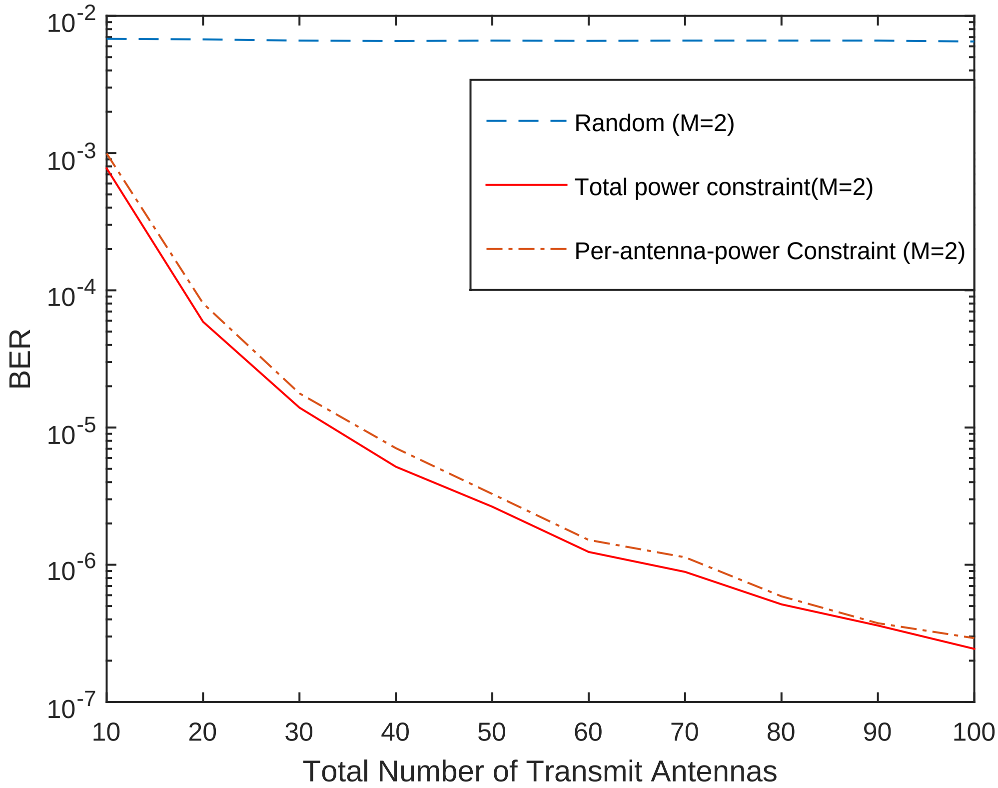

Figure 3.3: Bit error rate versus total number of transmit antennas N for $M = 2$ receive antennas and selection of $K = 6$ transmit antennas, for both per-antenna element and power constraint scenarios.

Regarding the bit error rate performance of this section's case where $M = 2$ receive antennas are occupied, we illustrate the results for total power constraint (optimal solution) and per-antenna-element power constraint (suboptimal solution/approximation) and compare them with random antenna selection. . We can derive from Fig 3.3 that the proposed solution for the per-antenna constraint case offers performance that approaches that of the optimal algorithm considering the total power constraint case.

For the per-antenna-element power constraint case performance to be obtained we used a cyclic-iterative maximazation algorithm in order to achieve an approximate solution for the beamforming vector. In the next chapter, we present and compare such algorithms for the beamformer approximation problem.

Chapter 4

Beamforming Approximation Algorithms

As mentioned in chapter 1, when we enforce a per-antenna-element power constraint on the beamformer, and for any antenna selection $\mathcal{I}$, i.e., when $W_K = \{\mathbf{w} \in \mathbb{C}^k : |w_k| \leq 1, k = 1, 2, \ldots, K\}$, we seek $\mathbf{w}$ that solves

$$\mathbf{w}(\mathcal{I}) \triangleq \arg\max_{\mathbf{w} \in W_K} \left\| \mathbf{H}_{:,\mathcal{I}}\, \mathbf{w} \right\|. \tag{4.1}$$

The problem of identyfying the optimal beamforming vector for the given maximization is a UQP problem that has not been efficiently solved so far. In this work we approach this problem by including semidefinite relaxation, cyclic-maximization as well as gradient projection iterative algorithms.

4.1 Cyclic Maximization

Because no efficient solver exists for (4.1) we consider its relaxed version

$$\max_{\mathbf{w_1}, \mathbf{w_2} \in W_K} \mathbf{w}_1^H \mathbf{H}_{:,\mathcal{I}}^H \mathbf{H}_{:,\mathcal{I}} \mathbf{w}_2. \tag{4.2}$$

As shown in [34], for a fixed beamformer $\mathbf{w}_2$, the vector $\mathbf{w}_1$that maximizes (4.2) is given by $\mathbf{w}_1 = e^{j \arg\left(\mathbf{H}_{:,\mathcal{I}}^H \mathbf{H}_{:,\mathcal{I}} \mathbf{w}_2\right)}$, while, for fixed $\mathbf{w}_1$, the optimal $\mathbf{w}_2$ becomes $\mathbf{w}_2 = e^{j \arg\left(\mathbf{H}_{:,\mathcal{I}}^H \mathbf{H}_{:,\mathcal{I}} \mathbf{w}_1\right)}$. Using this observation, starting fron an initial beamforming vector $\mathbf{w}^(0) =$ principal right singular vector of $\mathbf{H}_{:,\mathcal{I}}$ we perform a cyclic maximization of the form

$$\mathbf{w}^{(t+1)} = e^{j \arg\left(\mathbf{H}_{:,\mathcal{I}}^H \mathbf{H}_{:,\mathcal{I}} \mathbf{w}^{(t)}\right)} \tag{4.3}$$

and continue the iterative algorithm for a predefined number of steps or until $\left\| \mathbf{w}^{(t+1)} - \mathbf{w}^{(t)} \right\| \leq e$. In the simulations that follow we set $e = 0.01$ and the maximum number of iterations to 100.

4.2 Gradient Projection

In this section we ultilize the gradient projection algorithms proposed in [35],which tackle the Unit-modulus Least Squares (ULS) optimization problem.

$$\min_{\mathbf{w} \in \mathcal{C}^N} \|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2 \tag{4.4}$$

We begin by redefining our UQP problem in order to transform it to the ULS scenario.

$$\max_{\mathbf{w} \in \mathcal{C}^N} \mathbf{w}^H \mathbf{H}_{:,\mathcal{I}}^H \mathbf{H}_{:,\mathcal{I}} \mathbf{w} = \max_{\mathbf{w} \in \mathcal{C}^N} \mathbf{w}^H \mathbf{R} \mathbf{w} \tag{4.5}$$

where $\mathbf{R}_{N \times N} = \mathbf{H}_{:,\mathcal{I}}^H \mathbf{H}_{:,\mathcal{I}}$, is easily shown to be a Hermitian and semi-positive definite matrix.

The UQP formulation in (4.4) covers both maximization and minimization of quadratic forms (one can obtain the minimization of the quadratic form in (4.4) by considering $-\mathbf{R}$ in lieu of $\mathbf{R}$). If $\mathbf{R}$ is not positive (semi)definite, we can make it so using the diagonal loading technique (i.e. $\mathbf{R} \leftarrow \mathbf{R} + \lambda \mathbf{I}$ where $\lambda \geq \sigma_n(\mathbf{R})$. Note that such a diagonal loading does not change the solution of UQP as $\mathbf{w}^H(\mathbf{R} + \lambda \mathbf{I})\mathbf{w} = \mathbf{w}^H \mathbf{R}\mathbf{w} + \lambda \|\mathbf{w}\|_2^2$.

Using the above steps we transform

$$\max_{\mathbf{w} \in \mathcal{C}^N} \mathbf{w}^H \mathbf{R}\mathbf{w} \rightarrow \min_{\mathbf{w} \in \mathcal{C}^N} \mathbf{w}^H \mathbf{R}\mathbf{w}$$

WE proceed by showing that following UQP problem

$$\begin{aligned}
&\min_{\tilde{\mathbf{w}} \in \mathcal{C}^{N+1}} \tilde{\mathbf{w}}^H \mathbf{R} \tilde{\mathbf{w}} \\
&\text{subject to } |\tilde{\mathbf{w}}_i|^2 = 1, i = 1, \ldots, N+1;\ \tilde{\mathbf{w}}_{N+1} = 1 \\
&\text{where } \tilde{\mathbf{w}} = \begin{bmatrix} \mathbf{w} \\ 1 \end{bmatrix},
\end{aligned} \tag{4.6}$$

can be expressed as a ULS problem. To see this, we consider the square root decomposition of a semi-positive definite matrix $\mathbf{R} = \mathbf{R}^{H/2}\mathbf{R}^{1/2}$ and apply the following partioning

$$\mathbf{R}^{1/2} = [\mathbf{A}, -\mathbf{y}].$$

Consequently, we have

$$\mathbf{R} = [\mathbf{A}, -\mathbf{y}]^H [\mathbf{A}, -\mathbf{y}] = \begin{bmatrix} \mathbf{A}^H\mathbf{A} & -\mathbf{A}^H\mathbf{y} \\ -\mathbf{y}^H\mathbf{A} & \|\mathbf{y}\|_2^2 \end{bmatrix}$$

and by expanding $\tilde{\mathbf{w}}^H \mathbf{R} \tilde{\mathbf{w}}$ we get

$$\begin{aligned}
\tilde{\mathbf{w}}^H \mathbf{R} \tilde{\mathbf{w}} &= \begin{bmatrix} \mathbf{w}^H e^{-j\theta} & e^{-j\theta} \end{bmatrix} \begin{bmatrix} \mathbf{A}^H\mathbf{A} & -\mathbf{A}^H\mathbf{y} \\ -\mathbf{y}^H\mathbf{A} & \|\mathbf{y}\|_2^2 \end{bmatrix} \begin{bmatrix} \mathbf{w}^H e^{j\theta} \\ e^{j\theta} \end{bmatrix} \\
&= \mathbf{w}^H\mathbf{A}^H\mathbf{A}\mathbf{w} - \mathbf{y}^H\mathbf{A}\mathbf{w} - \mathbf{w}^H\mathbf{A}^H\mathbf{y} + \|\mathbf{y}\|_2^2 \\
&= \|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2
\end{aligned}$$

transforming our initial UQP problem to the equivalent ULS. Given this equivalency, we present the 2 algorithms proposed in [35] that approach the ULS/UQP problem, which we used to achieve a suboptimal approach for our beamforming vector $\mathbf{w}$.

Algorithm 1:

Input : $\mathbf{H} \in \mathbb{C}^{2\times K}$ (channel matrix) , $K \in \{1, 2, \ldots, N\}$ (Desired Antenna Selection)

$\mathbf{R} \leftarrow \mathbf{H}^H\mathbf{H}$

$\mathbf{R} \leftarrow -\mathbf{R} + \lambda_{max}(\mathbf{R})\,\mathcal{I}_K$

$\mathbf{R}^{1/2} = [\mathbf{A}, -\mathbf{y}]$

Set $r = 0, \quad \alpha = \dfrac{\beta}{\lambda_{max}\left(\mathbf{A}^H\mathbf{A}\right)}, \quad \beta \in (0, 1), \quad \mathbf{w}_0 = e^{j\angle\left(\mathbf{A}^\dagger\mathbf{y}\right)}$

Repeat

$\zeta^{(r+1)} = \mathbf{w}^{(r)} + \alpha\mathbf{A}^H\left(\mathbf{y} - \mathbf{A}\mathbf{w}^{(r)}\right)$

$\mathbf{w}^{r+1} = e^{j\angle\left(\zeta^{(r+1)}\right)}$

$r = r + 1$

until convergence

Algorithm 1 is nothing but a GP algorithm what is special isthat the projection step involves a non-convex set the elementwise unit modulus constraint, in particular. In Algorithm 1, is the step size along the opposite direction of the gradient. The motivation of Algorithm 1 is simple: gradient descent has the advantage of scalability, and is able to exploit data (i.e.,**A** and **y**) sparsity. These traits are well-suited for large-scale problems. In addition, projection onto a unit modulus constraint admits a closed-form solution (cf. line 4 in Algorithm 1), and the entire procedure can be carried out very efficiently. On the other hand, the concern is that projection onto a non-convex set may in fact increase the cost value, and thus tends to be problematic in terms of optimization.

Different from other approaches, the proposed algorithm keeps the unit-modulus constraint and uses projected gradient descent instead of unconstrained gradient descent.

Algorithm 2:

Input : $\mathbf{H} \in \mathbb{C}^{2\times K}$ (channel matrix) , $K \in \{1, 2, \ldots, N\}$ (Desired Antenna Selection)

$\mathbf{R} \leftarrow \mathbf{H}^H\mathbf{H}$

$\mathbf{R} \leftarrow -\mathbf{R} + \lambda_{max}(\mathbf{R})\,\mathcal{I}_K$

$\mathbf{R}^{1/2} = [\mathbf{A}, -\mathbf{y}]$

Set $r = 0$, , $\beta \in (0,1)$, $\mathbf{w}_0 = e^{j\angle\left(\mathbf{A}^\dagger\mathbf{y}\right)}$

Repeat

$s^{(r+1)} = \left(\mathbf{w}^{(r)}\right)^H \mathbf{A}^H\mathbf{y}/\left\|\mathbf{A}\mathbf{w}^{(r)}\right\|_2^2$

$\alpha^{(r+1)} = \beta/\lambda_{max}(|s^{r+1}|^2\mathbf{A}^H\mathbf{A})$

$\zeta^{(r+1)} = \mathbf{w}^{(r)} + \alpha^{r+1}\left(s^{r+1}\right)^* \mathbf{A}^H\left(\mathbf{y} - s^{r+1}\mathbf{A}\mathbf{w}^{(r)}\right)$

$\mathbf{w}^{r+1} = e^{j\angle\left(\zeta^{(r+1)}\right)}$

$r = r + 1$

until convergence

Algorithm 2 is a modified instance of Algorithm 1. However, considering the ULS problem $\|\mathbf{y} - \mathbf{A}\mathbf{w}\|_2^2$ and our beamforming trasmission case, we often desire a good beampattern that concentrates power in the directions of interest. To that end, Algorithm 2 introduces an additional scaling variable $s \in \mathbb{C}$ to obtain

$$\begin{aligned} &\min_{\mathbf{w}\in\mathbb{C}^N} \mathbf{w}^H\mathbf{R}\mathbf{w} \\ &\text{subject to } |\mathbf{w}_i|^2 = 1, i = 1, \ldots, N. \end{aligned} \tag{4.7}$$

Variable s can be regarded as an 'automatic normalization' factor, which addresses the aforementioned issue. Optimal s is computed and substituted as a function of $\mathbf{w}$:

$$s_{opt} = \frac{\mathbf{w}^H\mathbf{A}^H\mathbf{y}}{\|\mathbf{A}\mathbf{w}\|^2}.$$

Using this equation its easy to determine the tuning of s in each iteration and insert it in Algorithm 1, leading to Algorithm 2.

Chapter 5

Simulations

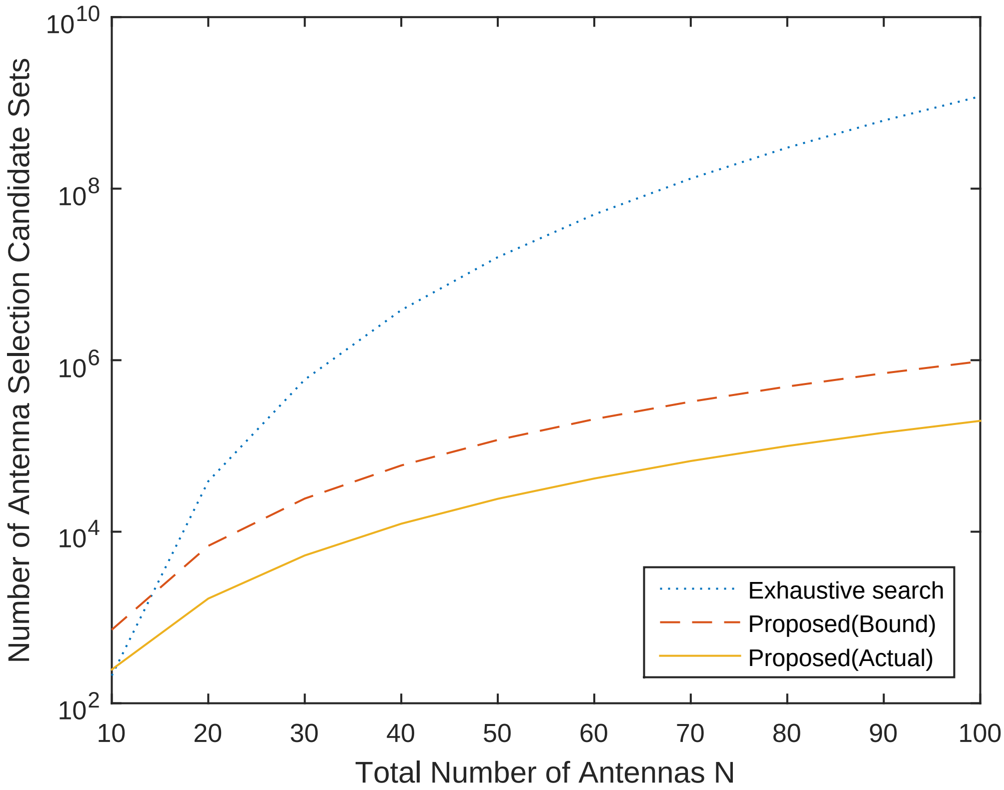

Figure 5.1: Complexity towards the number of candidate antenna selection sets versus the total number of transmit antennas N for $M = 2$ receive antennas and selection of $K = 6$ transmit antennas.

At the beginning of our simulations and specifically in Fig.5.1 we present the complexity of the Antenna Selection Algorithm (in terms of number of candidate AS sets) and plot it versus the corresponding upper bound $6\binom{N}{3}$, as well as the exhaustive search complexity $\binom{N}{K}$.

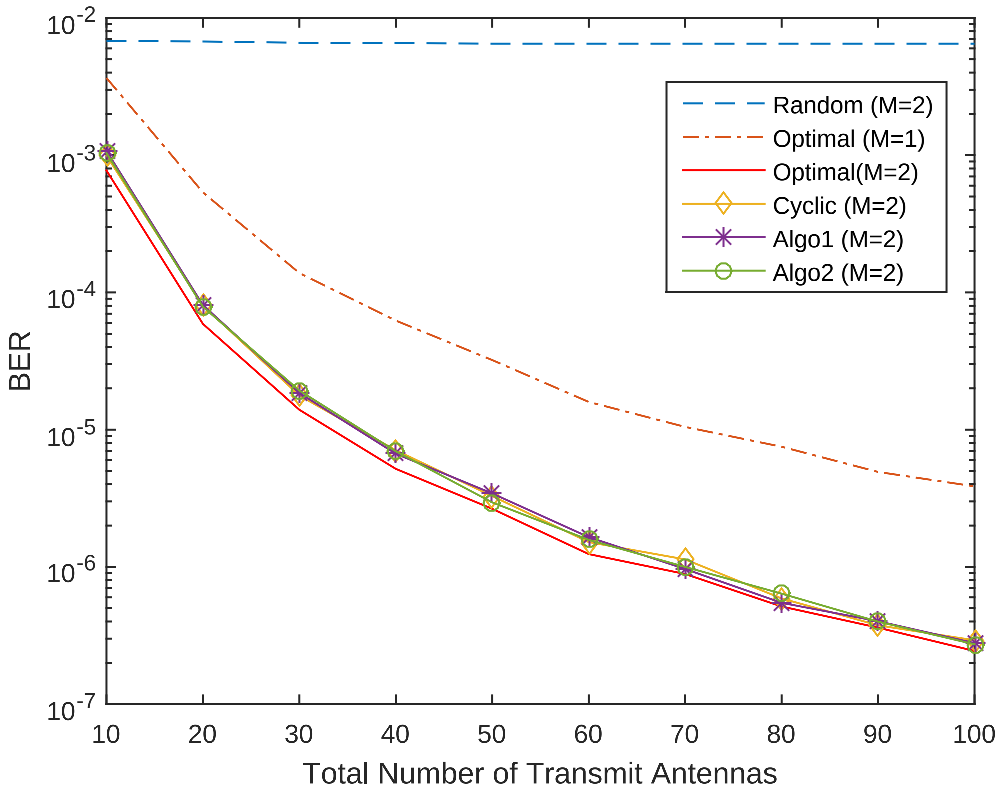

Figure 5.2: Bit error rate versus total number of transmit antennas N for $M \leq 2$ receive antennas and selection of $K = 6$ transmit antennas.

For each one of those candidate sets, we implemented the suboptimal approaches examined in chapter 3 regarding the beamforming vector approximation for the per-antenna-element power constraint scenario.Each algorithm provided an approximation of the beamforming vector $\mathbf{w}$, which we use to calculate our BER metric. The same metric was computed and plotted for the straight-forward case where $M = 1$ receive antenna, and the optimal total power constraint scenario where $M = 2$.

The presented results are averages over 1000 i.i.d. Rayleigh channel realizations. We can observe from Fig.5.2 that the results of the iterative suboptimal Algorithms (Cyclic,GP) discussed in chapter 3 considering the per antenna power constraint scenario match those of the optimal algorithm corresponding to the total power constraint case. Achieving equivalent results, regarding the bit error rate we now tend to examine the complexity of the three iterative aglorithms discussed in chapter 3 to tackle the per-antenna-element power constraint suboptimal problem.

In figures 5.3,5.4 that follow we illustrate the per-iteration complexity of each algorithm considering the number of multiplications and additions required for each instance. Finally, considering figures 5.5, 5.6 we present the average number of multiplications/additions required by each algorithm to converge, thus the complexity required for each algorithm to provide the approximation result of the beamforming vector $\mathbf{w}$.

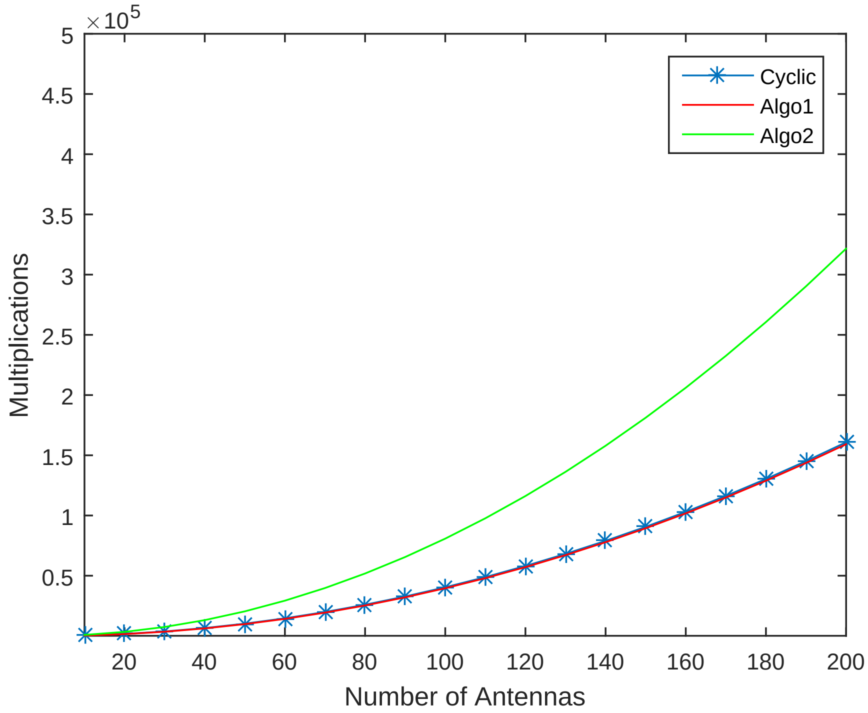

Figure 5.3: Per-iteration complexity of each algorithm towards the number of real-value multiplications versus total number of transmit antennas N for $M \leq 2$ receive antennas.

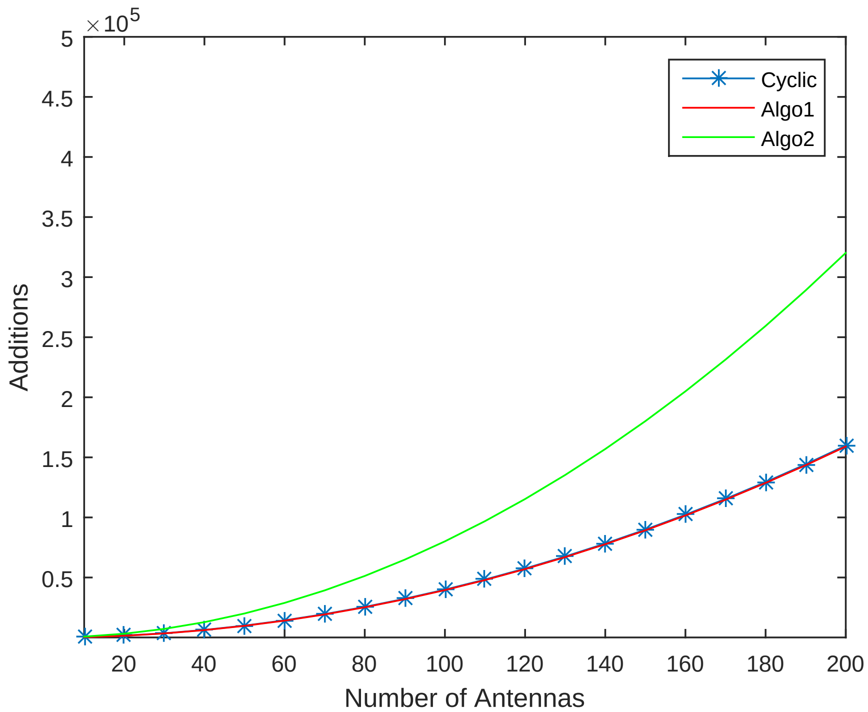

Figure 5.4: Per-iteration complexity of each algorithm towards the number of real-value additions versus total number of transmit antennas N for $M \leq 2$ receive antennas.

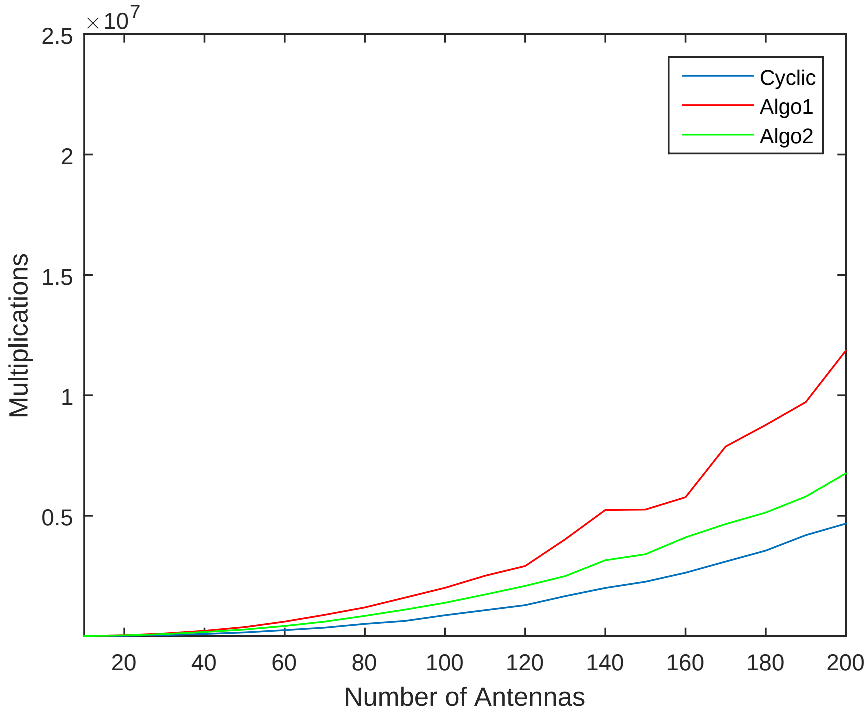

Figure 5.5: Average complexity complexity of each algorithm towards the number of real-value multiplications versus total number of transmit antennas N for $M \leq 2$ receive antennas.

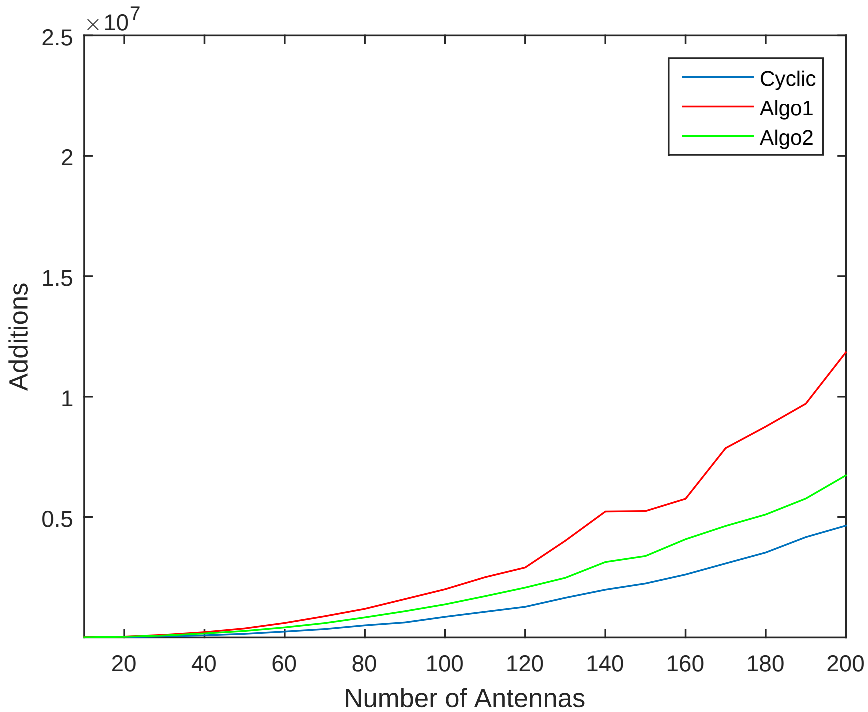

Figure 5.6: Average complexity complexity of each algorithm towards the number of real-value additions versus total number of transmit antennas N for $M \leq 2$ receive antennas.

Chapter 6

Conclusion

In this thesis, we have considered the problem of maximum-SNR transmit antenna selection and beamforming for a MIMO system that consists of a large number N of transmit antennas. We examined the algorithm that for $M = 2$ occupied antennas on the receiver provides a solution to the antenna selection with quadratic complexity. It was shown that the solution is provided in polynomial time in N, independently of the number K of selected antennas (i.e., even if K grows with N). Interestingly, the set of transmit antenna selection subsets that is computed by this algorithm contains the optimal antenna subset for both total power constraint and per-antenna-element power constraint scenarios. Moreover, under the per-antenna-element constraint scenario,and given a fixed antenna selection, we considered the beamforming approximation problem. We noted and examined the suboptimal nature of this ULS/UQP problem, and tested various iterative suboptimal algorithms (Cyclic maximazation, Gradient projection) to obtain a nearly-optimal solution for the beamforming vector. Finally, we provided insight on the performance of each iterative algorithm examined , as well as the efficiency of each one when applied to our case where we had to deal with a large number of antenna candidate subsets and compute the beamforming vector for each one of them.

Bibliography

[1] A. F.Molisch andM. Z.Win, *"MIMO systems with antenna selection,"* IEEE Microwave Mag., vol. 5, no. 1, pp. 4656, Mar. 2004.

[2] S. Sanayei and A. Nosratinia, *"Antenna selection in MIMO systems,"* IEEE Commun. Mag., vol. 42, no. 10, pp. 6873, Oct. 2004.

[3] Y. Jiang and M. K. Varanasi, *"The RF-chain limited MIMO systempart I: Optimum diversity-multiplexing tradeoff,"* IEEE Trans. Wireless Commun., vol. 8, no. 10, pp. 52385247, Oct. 2009.

[4] T. L. Marzetta, *"Noncooperative cellular wireless with unlimited number of base station antennas,"* IEEE Trans. Wireless Commun., vol. 9, no. 11, pp. 35903600, Nov. 2010.

[5] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, *"Scaling up MIMO: Opportunities and challenges with very large arrays,"*IEEE Signal Process. Mag., vol. 30, no. 1, pp. 4060, Jan. 2013.

[6] J.Hoydis, S. ten Brink, andM.Debbah, *"Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?,"* IEEE J. Sel. Areas Commun., vol. 31, no. 2., pp. 160171, Feb. 2013.

[7] A. Gorokhov, D. A. Gore, and A. J. Paulraj, *"Receive antenna selection for MIMO spatial multiplexing: Theory and algorithms,"* IEEE Trans. Signal. Process., vol. 51, no. 11, pp. 27962807, Nov. 2003.

[8] A. Gorokhov, D. A. Gore, and A. J. Paulraj, *"Receive antenna selection for MIMO flat-fading channels: Theory and algorithms,"* IEEE Trans. Inf. Theory, vol. 49, no. 10, pp. 26872696, Oct. 2003.

[9] A. F. Molisch, M. Z. Win, Y.-S. Choi, and J. H. Winters, *"Capacity of MIMO systems with antenna selection,"*IEEE Trans. Wireless Commun., vol. 4, no. 4, pp. 17591772, Jul. 2005.

[10] M. Ghavari-Alkansari and A. B. Gershman, *"Fast antenna subset selection in MIMO systems,"* IEEE Trans. Signal. Process., vol. 52, no. 2, pp. 339347, Feb. 2004.

[11] A. Dua, K. Medepalli, and A. J. Paulraj, *"Receive antenna selection in MIMO systems using convex optimization,"*IEEE Trans. Wireless Commun., vol. 5, no. 9, pp. 23532357, Sep. 2006.

[12] L. Zhou and M. Shimizu, *"Fast recursive algorithm for efficient transmit antenna selection in spatial multiplexing systems,"* in Proc. IEEE VTC Fall, Anchorage, AK, USA, Sep. 2009, pp. 15.

[13] R. S. Blum, Z. Xu, and S. Sfar, *"A near optimal joint transmit and receive antenna selection algorithm forMIMO systems,"* in Proc. IEEE Radio Wireless Symp., San Diego, CA, USA, Jan. 2009, pp. 554557.

[14] C. Chen, *"A computationally efficient near-optimal algorithm for capacity- maximization based joint transmit and receive antenna selection,"* IEEE Commun. Lett., vol. 14, no. 5, pp. 402405, May 2010.

[15] B. H. Wang, H. T. Hui, and M. S. Leong, *"Global and fast receiver antenna selection for MIMO systems,"*IEEE Trans. Commun., vol. 9, no. 9, pp. 25052510, Sep. 2010.

[16] R. Vaze and H. Ganapathy, *"Sub-modularity and antenna selection in MIMO systems,"* IEEE Commun. Lett., vol. 16, no. 9, pp. 14461449, Sep. 2012.

[17] W. Junling, A. I. Perez-Neira, and G. Meiguo, *"A concise joint transmit/receive antenna selection algorithm,"*China Commun., vol. 10, pp. 9199, Mar. 2013.

[18] H. A. Seleh andW. Hamouda, *"Transmit antenna selection for decision feedback detection in MIMO fading channels,"* IEEE Trans. Wireless Commun., vol. 8, pp. 44404444, Sep. 2009.

[19] R. W. Heath, S. Sandhu, and A. J. Paulraj, *"Antenna selection for spatial multiplexing systems with linear receivers,"*IEEE Commun. Lett., vol. 5, no. 4, pp. 142144, Apr. 2001.

[20] D. G. Gore, R. W. Heath, and A. J. Paulraj, *"Transmit selection in spatial multiplexing systems,"*IEEE Commun. Lett., vol. 6, no. 11, pp. 491493, Nov. 2002.

[21] K. T. Phan and C. Tellambura, *"Receive antenna selection based on union-bound minimization using convex optimization,"* IEEE Signal Process. Lett., vol. 14, no. 9, pp. 609612, Sep. 2007.

[22] Maria Gkizeli,George N. Karystinos, *"Maximum-SNR Antenna Selection Among A Large Number of Transmit Antennas"*, IEEE journal of selected topics in signal processing, vol. 8, No.5,October 2014.

[23] M. Soltanalian, B. Tang, J. Li, and P. Stoica,*"Joint design of the receive filter and transmit sequence for active sensing,"* IEEE Signal Process. Lett., vol. 20, no. 5, pp. 423426, May 2013.

[24] M. Soltanalian and P. Stoica, *"Designing unimodular codes via quadratic optimization,"* IEEE Trans. Signal Process., vol. 62, no. 3, pp. 12211234, Mar. 2014.

[25] John Tranter, Nicholas D. Sidiropoulos, Xiao Fu, Ananthram Swami, *"Fast Unit-Modulus Least Squares With Applications in Beamforming"*, IEEE transactions on signal processing, vol.65, No. 11, June 2017.

[26] M. Blum, R.W. Floyd, V. Pratt, R. L. Rivest, and R. E. Tarjan, *"Time bounds for selection"*, J. Comput. Syst. Sci., vol. 7, pp. 448-461, Aug.1973.

[27] T.H.Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *" Introduction to Algorithms"*, 2nd ed. Cambridge, MA, USA: MIT Press, 2001.

Appendix A

A.1 Proof of lemma 1

The first step is to derive the maximum-SNR filter.

$E\{\mathbf{n}\} = \mathbf{0}, \quad \mathbf{R} = E\{\mathbf{n}\mathbf{n}^H\} = \sigma^2\mathbf{I}.$

Considering filter $\mathbf{f}$ we have the filter-output SNR

$$SNR = \frac{E\left\{\left|\mathbf{f}^H\mathbf{G}\mathbf{w}x\right|^2\right\}}{E\left\{\left|\mathbf{f}^H\mathbf{n}\right|^2\right\}} = \frac{\left(\mathbf{f}^H\mathbf{G}\mathbf{w}\right)\left(\mathbf{f}^H\mathbf{G}\mathbf{w}\right)^H E\left\{|x|^2\right\}}{\mathbf{f}^H\mathbf{R}\,\mathbf{f}} \quad .$$

$\mathbf{R}$ is Hermitian: $\mathbf{R}^H = (E\left\{\mathbf{n}\mathbf{n}^H\right\})^H = E\left\{\left(\mathbf{n}\mathbf{n}^H\right)^H\right\} = E\left\{\mathbf{n}\mathbf{n}^H\right\} = \mathbf{R}.$

$\mathbf{R}$ is Positive-definite: $\forall \mathbf{a} \in \mathbb{C}^m - \{\mathbf{0}\}, \mathbf{a^H}\mathbf{R}\,\mathbf{a} = \mathbf{a}^H\sigma^2\mathbf{I}\;\;\mathbf{a} = \sigma^2\left\|\mathbf{a}\right\|^2 > 0.$

Therefore, $\mathbf{R} = \mathbf{R}^{\frac{1}{2}}\left(\mathbf{R}^{\frac{1}{2}}\right)^H, \left(\mathbf{R}^{\frac{1}{2}}\right)^H = \mathbf{R}^{\frac{1}{2}} \qquad \mathbf{R}, \mathbf{R}^{\frac{1}{2}}$ non singular

Let $\mathbf{s} = \mathbf{R}^{\frac{1}{2}}\mathbf{f} \Leftrightarrow \mathbf{f} = \mathbf{R}^{-\frac{1}{2}}\mathbf{s}.$

Then,

$$SNR = \frac{(\mathbf{s}^H\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w})(\mathbf{w}^H\mathbf{G}^H\mathbf{R}^{-\frac{1}{2}}\mathbf{s})}{\left\|\mathbf{s}\right\|^2}E\{|x|^2\} \leq \frac{\left\|\mathbf{s}\right\|^2\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^2}{\left\|\mathbf{s}\right\|^2}E\{|x|^2\} = E\{|x|^2\}\left\|\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}\right\|^2$$

with equality <u>iff</u> $\mathbf{s} = a\,\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w} \;\Leftrightarrow\; \mathbf{R}^{\frac{1}{2}}\mathbf{f} = a\,\mathbf{R}^{-\frac{1}{2}}\mathbf{G}\mathbf{w}$

$\Leftrightarrow\; \mathbf{f} = a\,\mathbf{R}^{-1}\mathbf{G}\mathbf{w},\; a \neq 0$

The next step now is to show that the mimimum mean square error filter (MMSE) is a maximum-SNR filter.

Filter output error $e = \mathbf{f}^H\mathbf{y} - x.$

We denote $\mathbf{R}_y = E\left\{\mathbf{y}\mathbf{y}^H\right\}.$

The $M \times M$ matrix $\mathbf{R}_y$ is Hermitian and Positive-definite (similar proof to the one above for noise autocorrelation matrix $\mathbf{R}$).

Then, $\mathbf{f} = \underset{\mathbf{f}}{\operatorname{argmin}}\, E\left\{|e|^2\right\}.$

$$E\left\{|e|^2\right\} = E\left\{\left|\mathbf{f}^H\mathbf{y} - x\right|^2\right\} = E\left\{\left(\mathbf{f}^H\mathbf{y} - x\right)\left(\mathbf{f}^H\mathbf{y} - x\right)^H\right\}$$

$$= E\left\{\mathbf{f}^H\mathbf{y}\mathbf{y}^H\mathbf{f}\right\} - \mathbf{f}^H E\left\{\mathbf{y}x^*\right\} - E\left\{\mathbf{y}^H x\right\} f + E\left\{|x|^2\right\}$$

$$= \mathbf{f}^H\mathbf{R}_y\mathbf{f} - \mathbf{f}^H E\left\{\mathbf{y}x^*\right\} - E\left\{\mathbf{y}^H x\right\}\mathbf{f} + E\left\{|x|^2\right\}.$$

$$\bigtriangledown_{\mathbf{f}}\ \left[E\left\{|e|^2\right\}\right] = 2\mathbf{R}_y\mathbf{f} - 2E\left\{\mathbf{y}x^*\right\} = 2\left(\mathbf{R}_y\mathbf{f} - E\left\{(\mathbf{G}\mathbf{w}x + \mathbf{n})x^*\right\}\right)$$

$$= 2\left(\mathbf{R}_y\mathbf{f} - \mathbf{G}\mathbf{w}E\{|x|^2\} - E\{x^*\mathbf{n}\}\right)$$

$$= 2\left(\mathbf{R}_y\mathbf{f} - \mathbf{G}\mathbf{w}E\{|x|^2\}\right).$$

In order to achieve the minimum we get:

$$\bigtriangledown_{\mathbf{f}}\ [E\left\{|e|^2\right\}] = 0 \Leftrightarrow \mathbf{R}_y\mathbf{f} = \mathbf{G}\mathbf{w}E\left\{|x|^2\right\} \Leftrightarrow \mathbf{f} = E\left\{|x|^2\right\}\ \mathbf{R}_y^{-1}\mathbf{G}\mathbf{w}.$$

$$\mathbf{R}_y = E\left\{\mathbf{y}\mathbf{y}^H\right\} = E\left\{(\mathbf{G}\mathbf{w}x + \mathbf{n})(\mathbf{G}\mathbf{w}x + \mathbf{n})^H\right\}$$

$$= E\left\{\mathbf{G}\mathbf{w}x\, x^*\mathbf{w}^H\mathbf{G}^H\right\} + \underbrace{E\left\{\mathbf{G}\mathbf{w}x\right\}E\left\{\mathbf{n}^H\right\}}_{=0} + \underbrace{E\left\{\mathbf{n}\right\}E\left\{x^*\mathbf{w}^H\mathbf{G}^H\right\}}_{=0} + \underbrace{E\left\{\mathbf{n}\mathbf{n}^H\right\}}_{=\mathbf{R}}$$

$$= E\left\{|x|^2\right\}\ \mathbf{G}\mathbf{w}(\mathbf{G}\mathbf{w})^H + \mathbf{R} \Longrightarrow \mathbf{R}_y^{-1} = \mathbf{R}^{-1} - \frac{\mathbf{R}^{-1}E\left\{|x|^2\right\}\ \mathbf{G}\mathbf{w}\ (\mathbf{G}\mathbf{w})^H\mathbf{R}^{-1}}{1 + (\mathbf{G}\mathbf{w})^H\mathbf{R}^{-1}E\left\{|x|^2\right\}\mathbf{G}\mathbf{w}}.$$

To obtain the MMSE filter we determine

$$\mathbf{R}_y^{-1}\,\mathbf{Gw} = \mathbf{R}^{-1}\,\mathbf{Gw} - \frac{E\left\{|x|^2\right\}\ \mathbf{Gw}\ \mathbf{R}^{-1}\left(\mathbf{Gw}\right)^H}{1+E\left\{|x|^2\right\}\left(\mathbf{Gw}\right)^H\mathbf{R}^{-1}\mathbf{Gw}}\ \mathbf{R}^{-1}\mathbf{Gw}$$

$$= \mathbf{R}^{-1}\,\mathbf{Gw}\left(1 - \frac{E\left\{|x|^2\right\}\ \mathbf{Gw}\ \mathbf{R}^{-1}\left(\mathbf{Gw}\right)^H}{1+E\left\{|x|^2\right\}\left(\mathbf{Gw}\right)^H\mathbf{R}^{-1}\mathbf{Gw}}\right)$$

$$= \frac{1}{1+E\left\{|x|^2\right\}\left(\mathbf{Gw}\right)^H\mathbf{R}^{-1}\mathbf{Gw}}\ \mathbf{R}^{-1}\mathbf{Gw},$$

providing the result

$$\mathbf{f} = \frac{E\left\{|x|^2\right\}}{1+E\left\{|x|^2\right\}(\mathbf{Gw})^H\mathbf{R}^{-1}\mathbf{Gw}}\ \mathbf{R}^{-1}\mathbf{Gw}.$$

Hence, the MMSE filter $\mathbf{f}$ is a maximum-SNR filter.

A.2 Proof of lemma 2

We begin with the statement that the singular values of a $M \times N$ matrix $\mathbf{X}$ are the square roots of the eigenvalues of the $N \times N$ matrix $\mathbf{X}^*\mathbf{X}$ (where $*$ stands for the transpose-conjugate matrix if it has complex coefficients, or the transpose if it has real coefficients).

Hence, $\sigma_i\left(\mathbf{H}\right) = \sqrt{\lambda_i(\mathbf{H}^H\mathbf{H})}$
We denote the $N \times N$ matrix $\mathbf{A} = \mathbf{H}^H\mathbf{H}$

$\mathbf{H}^H\mathbf{H} = \left(\mathbf{H}^H\mathbf{H}\right)^H$ which means $\mathbf{A}$ is a hermitian matrix (real positive eigen values).

For an arbitrary complex valued vector $\mathbf{w} : \|\mathbf{w}\| = 1$ it is applied that

$$\lambda_{\max} = \max_{\|\mathbf{w}\|=1} \mathbf{w}^H\mathbf{A}\,\mathbf{w}$$

There exists unitary matrix $\mathbf{U}$: $\mathbf{U}^H\mathbf{A}\mathbf{U} = \mathbf{D} \Rightarrow \mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{U}^H$,

where $\mathbf{D} = diag(\lambda_1, \lambda_2, .., \lambda_N)$.

Let $\mathbf{y} = \mathbf{U}^H\mathbf{w}$. Since $\|\mathbf{w}\| = 1 \Rightarrow \|\mathbf{y}\| = \mathbf{w}^H\underbrace{\mathbf{U}^H\mathbf{U}}_{\mathrm{I}}\mathbf{w} = \|\mathbf{w}\| = 1$.

Then, $\max\limits_{\|\mathbf{w}\|=1} \mathbf{w}^H \mathbf{A}\, \mathbf{w} = \max\limits_{\|\mathbf{y}\|=1} \mathbf{y}^H \mathbf{D}\, \mathbf{y} = \sum_{i=1}^{N} \lambda_i |\mathbf{y}_i|^2$

$\sum_{i=1}^{N} \lambda_i |\mathbf{y}_i|^2 \leq \lambda_{\max} \sum_{i=1}^{N} |\mathbf{y}_i|^2 = \lambda_{\max} \|\mathbf{y}\| = \lambda_{\max}$

Hence, $\quad \lambda_{\max}(\mathbf{H}^H\mathbf{H}) = \max\limits_{\|\mathbf{w}\|=1} \mathbf{w}^H \mathbf{H}^H \mathbf{H}\, \mathbf{w} = \max\limits_{\|\mathbf{w}\|=1} \|\mathbf{H}\mathbf{w}\|^2$

$\Rightarrow \max\limits_{\|\mathbf{w}\|=1} \|\mathbf{H}\mathbf{w}\| = \sqrt{\lambda_{\max}(\mathbf{H}^H\mathbf{H})} = \sigma_{\max}(\mathbf{H}),$

meaning $\left\|\mathbf{H}_{:,\mathcal{I}}\mathbf{w}\right\| \leq \sigma_{\max}(\mathbf{H})$ with equality iff $\mathbf{w}$ corresponds to the "principal right singular vector" (the one that corresponts to $\sigma_{\max}(\mathbf{H}_{:,\mathcal{I}})$).

A.3 Proof of lemma 3

First the proof of the special case where A = I, follows from the equality:

$$\begin{pmatrix} \mathbf{I} & 0 \\ \mathbf{v}^{\mathrm{T}} & 1 \end{pmatrix} \begin{pmatrix} \mathbf{I} + \mathbf{u}\mathbf{v}^{\mathrm{T}} & \mathbf{u} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \mathbf{I} & 0 \\ -\mathbf{v}^{\mathrm{T}} & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{u} \\ 0 & 1 + \mathbf{v}^{\mathrm{T}}\mathbf{u} \end{pmatrix}.$$

The determinant of the left hand side is the product of the determinants of the three matrices. Since the first and third matrix are triangle matrices with unit diagonal, their determinants equal to 1. The determinant of the middle matrix is our desired value. The determinant of the right hand side is simply $(1 + \mathbf{v}^T\mathbf{u})$. So we have the result:

$$\det(\mathbf{I} + \mathbf{u}\mathbf{v}^{\mathrm{T}}) = (1 + \mathbf{v}^{\mathrm{T}}\mathbf{u}).$$

Then the general case can be found as:

$$\det(\mathbf{A} + \mathbf{u}\mathbf{v}^{\mathrm{T}}) = \det(\mathbf{A}) \det(\mathbf{I} + (\mathbf{A}^{-1}\mathbf{u})\mathbf{v}^{\mathrm{T}})$$

$$= (1 + \mathbf{v}^{\mathrm{T}}(\mathbf{A}^{-1}\mathbf{u})) \det(\mathbf{A}).$$