Low-Complexity Near-Optimal Detection Algorithms for MIMO Systems

by

Kazem Izadinasab

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2020

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner: Aria Nosratinia, Professor
Department of Electrical and Computer Engineering
University of Texas at Dallas

Supervisor: Oussama Damen, Professor
Department of Electrical and Computer Engineering
University of Waterloo

Internal Member: Guang Gong, Professor
Department of Electrical and Computer Engineering
University of Waterloo

Patrick Mitran, Professor
Department of Electrical and Computer Engineering
University of Waterloo

Internal-External Member: Stephen Vavasis, Professor
Department of Combinatorics and Optimization
University of Waterloo

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

As the number of subscribers in wireless networks and their demanding data rate are exponentially increasing, multiple-input multiple-output (MIMO) systems have been scaled up in the 5G where tens to hundreds of antennas are deployed at base stations (BSs). However, by scaling up the MIMO systems, designing detectors with low computational complexity and close to the optimal error performance becomes challenging. In this dissertation, we study the problem of efficient detector designs for MIMO systems.

In Chapter 2, we propose efficient detection algorithms for small and moderate MIMO systems by using lattice reduction and subspace (or conditional) detection techniques. The proposed algorithms exhibit full receive diversity and approach the bit error rate (BER) of the optimal maximum likelihood (ML) solution. For quasi-static channels, the complexity of the proposed schemes is cubic in the system dimension and is only linear in the size of the QAM modulation used. However, the computational complexity of lattice reduction algorithms imposes a large burden on the proposed detectors for large MIMO systems or fast fading channels.

In Chapter 3, we propose detectors for large MIMO systems based on the combination of minimum mean square error decision feedback equalization (MMSE-DFE) and subspace detection tailored to an appropriate channel ordering. Although the achieved diversity order of the proposed detectors does not necessarily equal the full receive diversity for some MIMO systems, the coding gain allows for close to ML error performance at practical values of signal-to-noise ratio (SNR) at the cost of a small computational complexity increase over the classical MMSE-DFE detection. The receive diversity deficiency is addressed by proposing another algorithm in which a partial lattice reduction (PLR) technique is deployed to improve the diversity order.

Massive multiuser MIMO (MU-MIMO) is another technology where the BS is equipped with hundreds of antennas and serves tens of single-antenna user terminals (UTs). For the uplink of massive MIMO systems, linear detectors, such as zero-forcing (ZF) and minimum mean square error (MMSE), approach the error performances of sophisticated nonlinear detectors. However, the exact solutions of ZF and MMSE involve matrix-matrix multiplication and matrix inversion operations which are expensive for massive MIMO systems. In Chapter 4, we propose efficient truncated polynomial expansion (TPE)-based detectors that achieve the error performance of the exact solutions with a computational complexity proportional to the system dimensions.

The millimeter wave (mmWave) massive MIMO is another key technology for 5G cellular networks. By using hybrid beamforming techniques in which a few numbers of radio frequency (RF) chains are deployed at the BSs and the UTs, the fully-digital precoder (combiner) is approximated as a product of analog and digital precoders (combiners). In Chapter 5, we consider a signal detection scheme using the equivalent channel consisting of the precoder, mmWave channel, and combiner. The available structure in the equivalent channel enables us to achieve the BER of the optimal ML solution with a significant reduction in the computational complexity.

Acknowledgements

I would like to express my sincere gratitude to my PhD advisor, Professor Oussama Damen, for his guidance, patience, and continued support throughout my PhD studies. This thesis would have not been possible without his encouragement and guidance.

My appreciation also extends to my thesis committee members: Professor Aria Nosratinia, Professor Stephen Vavasis, Professor Guang Gong, and Professor Patrick Mitran, for their insightful comments and recommendations on this thesis.

I would also like to thank my amazing friends for all the joy and memories that we shared. Having such incredible people in my life made my studies at the University of Waterloo a wonderful experience.

Finally, the most deserved acknowledgment to my lovely parents for their unconditional love and support throughout my life. Thank you both for your encouragement, which has always given me the strength and motivation to reach my goals in all aspects of my life.

Dedication

Dedicated to my lovely family.

Table of Contents

List of Figures

List of Tables

List of Algorithms

List of Abbreviations

3G	3rd Generation
4G	4th Generation
5G	5th Generation
AM	Alternating Minimization
APP	A-Posteriori Probability
BER	Bit Error Rate
BS	Base Station
CLPS	Closest Lattice Point Search
CSI	Channel State Information
DFE	Decision Feedback Equalizer
DL	Deep Learning
FE	Full Expansion
FSD	Fixed-complexity Sphere Decoder
FLOP	Floating Point Operation
LLL	Lenstra-Lenstra-Lovász
LLR	Log-Likelihood Ratio
LORD	Layered Orthogonal Lattice Detector
LRA	Lattice Reduction Aided
LSD	List Sphere Decoder
MAP	Maximum A-Posteriori Probability

MIMO	Multiple-Input Multiple-Output
ML	Maximum Likelihood
MMSE	Minimum Mean Square Error
mmWave	Millimeter Wave
MU-MIMO	Multiuser MIMO
NN	Neural Network
PLR	Partial Lattice Reduction
PM	Partial Marginalization
QAM	Quadrature Amplitude Modulation
QRD	QR Decomposition
RFSD	Regularized Fixed-complexity Sphere Decoder
SD	Sphere Decoding
SE	Single Expansion
SINR	Signal-to-Interference-plus-Noise Ratio
SISO	Soft-Input Soft-Output
SNR	Signal-to-Noise Ratio
SVD	Singular Value Decomposition
TPE	Truncated Polynomial Expansion
UPA	Uniform Planar Array
UT	User Terminal
V-BLAST	Vertical-Bell Laboratories Layered Space-time
VR	Visibility Region
XL-MIMO	Extra-Large Scale MIMO
ZF	Zero Forcing

List of Notations

$\mathbf{x}$	Boldface small letters denote vectors
$\mathbf{X}$	Boldface capital letters denote matrices
$\mathbf{x}^T$, $\mathbf{X}^T$	Transpose of $\mathbf{x}$, $\mathbf{X}$
$\mathbf{x}^H$, $\mathbf{X}^H$	Conjugate (or Hermitian) transpose of $\mathbf{x}$, $\mathbf{X}$
$\mathbf{X}^{-1}$	Inverse of matrix $\mathbf{X}$
$\|\|\mathbf{x}\|\|$	L2 norm of $\mathbf{x}$
$\|\|\mathbf{X}\|\|_F$	Frobenius norm of $\mathbf{X}$
$\det(\mathbf{X})$	Determinant of matrix $\mathbf{X}$
$\text{Trace}(\mathbf{X})$	Trace of matrix $\mathbf{X}$
$x_{i,j}$	(i,j)-th entry of matrix $\mathbf{X}$
$\|\mathcal{X}\|$	Cardinality of a set $\mathcal{X}$
$\langle \mathbf{a}, \mathbf{b} \rangle$	The inner product of vector $\mathbf{a}$ and vector $\mathbf{b}$
$\mathbf{I}_M$	Identity matrix of dimension M
$\mathcal{C}^N$	N-dimensional complex space
$\log_2(\cdot)$	Logarithm in base 2
$\lceil \cdot \rfloor$	Nearest integer function
$\Re\{\cdot\}$	The real part of a vector or a matrix
$\Im\{\cdot\}$	The imaginary part of a vector or a matrix
$\mathcal{O}(\cdot)$	Big O notation

Chapter 1

Introduction

In the past decades, multiple-input multiple-output (MIMO) systems with multiple antennas at the transmitter and receiver sides have gained significant interest in wireless communications. MIMO systems can provide considerable improvements in capacity, link reliability, and power efficiency when compared to single-input single-output systems [28, 67]. As a result, they have been widely implemented in modern wireless communication standards, such as long-term evolution (LTE) and WiFi (IEEE 802.11n/ac) [15, 67]. However, the wireless products and standards in the 3rd generation (3G) and 4th generation (4G) of wireless systems have implemented MIMO technology by deploying small and moderate numbers of antennas mainly in the range of $(2-8)$ antennas. MIMO systems with such numbers of antennas have small or moderate spectral efficiency and are not able to meet the high data rate demands of ever-increasing subscribers with their data-hungry applications. Therefore, in the 5th generation (5G) of wireless systems and beyond, the numbers of antennas are scaled up to tens or hundreds of antennas in order to achieve spectral efficiency of tens to hundreds of bps/Hz [15]. In the literature, such systems are referred to as large-scale or massive MIMO systems.

Depending on the application, different configurations can be considered for MIMO systems,

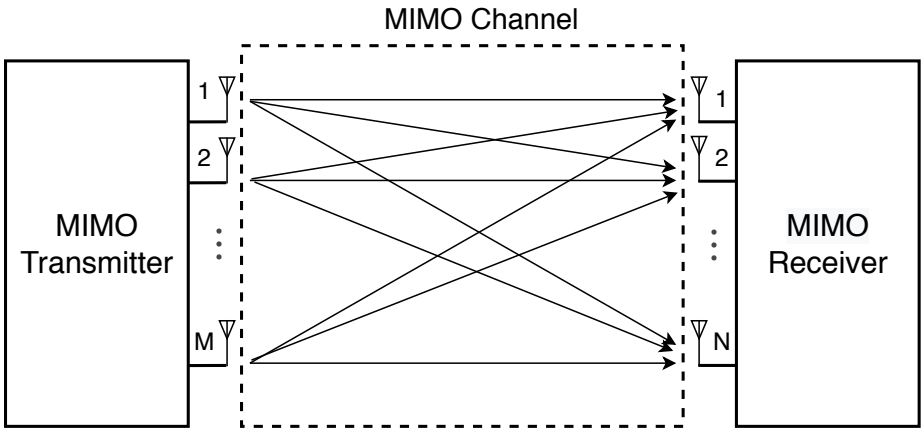


Figure 1.1: Point-to-point MIMO system.

such as point-to-point MIMO and multiuser MIMO (MU-MIMO) configurations [15]. Fig.1.1 indicates a point-to-point MIMO system where several antennas are deployed at both terminals. When the terminals are equipped with a large number of antennas, we will have a large point-to-point MIMO system which can be utilized for high-speed wireless backhaul connectivity between base stations (BSs). In MU-MIMO systems, the BS with multiple antennas simultaneously serves several user terminals (UTs). The UTs can be small devices, such as single-antenna mobile/smartphones or medium to large terminals, like laptops or TVs [15]. MU-MIMO systems provide multiplexing gain to all users and robustness to the propagation environment. However, the number of antennas and consequently, the spectral efficiency for such systems is moderate. To meet the high data rate requirements of the 5G wireless networks, massive MU-MIMO systems have been proposed where the BS is equipped with hundreds of antennas simultaneously communicating with tens of single-antenna UTs [5, 43, 67]. Fig.1.2 shows a multi-cell massive MU-MIMO.

The optimal recovery of the transmitted signals at the receiver, which is corrupted by channel fading, noise, and interference, is a challenging issue in MIMO systems. It is known that the MIMO detection problem can be translated to the closest lattice point search (CLPS) problem, which is NP-hard [1, 22]. The optimal maximum likelihood (ML) detection [22] of the

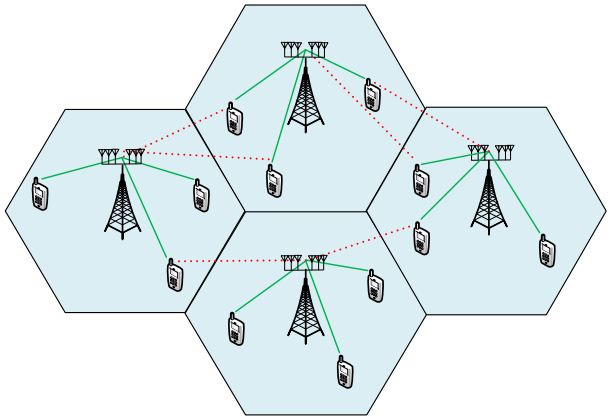


Figure 1.2: Massive MU-MIMO system.

transmitted signals imposes a huge computational complexity on the system, which increases exponentially with the number of transmit antennas. Moreover, other well-known detection algorithms and techniques that work well for small MIMO systems may not be practical for large MIMO systems as they may not scale well with the dimensions of the system [15]. Therefore, designing detectors with low computational complexity and close to optimal performance is of great importance for moderate and large-scale MIMO systems.

In an effort to reduce the complexity compared to the ML solution, suboptimal receivers, such as zero-forcing (ZF), decision feedback equalizer (DFE), and minimum mean square error (MMSE) [26, 37, 72], have been extensively studied in the literature. However, the bit error rate (BER) performances of these detectors are still far worse than that of the ML, especially at high signal-to-noise ratios (SNRs) and for systems with comparable MIMO channel dimensions. A typical example of such systems is the point-to-point MIMO or the MU-MIMO system when the numbers of antennas at both terminals are comparable. For the uplink of massive MU-MIMO systems, since the number of antennas at the BS is much larger than the number of UTs, linear detectors, such as ZF and MMSE, perform sufficiently close to the ML solution [67, 71]. We will discuss detection techniques for massive MU-MIMO systems later in this section.

Lattice reduction aided (LRA) techniques [53, 99, 106] have been proposed to improve the

BER performances of suboptimal detectors. In these algorithms, efficient lattice reduction subroutines, such as Lenstra-Lenstra-Lovász (LLL) algorithm, are used. Although the maximum receive diversity is guaranteed by LRA techniques [94], there is still a performance gap, i.e., a coding gain or horizontal shift in the error curves compared to the ML solution. Moreover, the computational complexity of reduction algorithms imposes a large burden on the detectors, especially when the dimension of the system increases.

The idea of splitting the transmitted symbol vector into two subsets where an exhaustive search over a given constellation is performed for one subset and the detection process is completed by a suboptimal detector for the other subset, a.k.a. subspace detection or conditional optimization detection [90], was initially considered for under-determined systems in [24]. In the literature, conditional detection has also been referred to as subspace detection [14] or partial marginalization (PM) [60]. Fixed-complexity sphere decoder (FSD) [7], parallel detection [63], Chase detector [97], layered orthogonal lattice detector (LORD) [91], and channel puncturing scheme [86] are all based on the same idea. The idea of conditional optimization detection in conjunction with LRA techniques has been considered in [74]. The FSD was proposed to deal with the variable complexity of the tree-based sphere decoding (SD) algorithms. In the FSD, the branches are fully expanded for only a certain number of levels in the corresponding detection tree (typically obtained via the QR decomposition (QRD) of the channel matrix).

The FSD scheme can achieve the optimal performance by performing some full search levels over the constellation points according to the FSD ordering. However, the required number of full search levels makes the scheme computationally prohibitive when the dimension of the system increases or the size of the constellation is large. In [63] and [97], the parallel and the Chase detectors are respectively introduced where a full search over the constellation along with a specific channel ordering for detecting one symbol is conducted. These orderings rearrange the channel matrix to find the best submatrix for conditional optimization detection and restrict the performance improvement of the conditional MIMO detection, especially for soft-output MIMO

systems.

The PM scheme in [60], or its extension for higher-order constellations [78], is a detection scheme for soft-output MIMO detection which approximates the exact *a-posteriori* probability (APP). The PM scheme marginalizes the posterior density for the received signal exactly over a subset of the transmitted bits with the lower SNR. The summation in the max-log approximation for the remaining bits is approximated by a ZF-DFE solution. Their proposed ordering (or permutation) of the channel matrix is exactly the FSD ordering proposed in [7, 52]. In order to reduce the computational complexity of the ZF-DFE process, the achievable information rate (AIR) based PM (AIR-PM) detector is proposed in [44].

In the subspace detection scheme proposed in [14], low-reliable data streams are considered in several sub-MIMO systems by exchanging the columns of the channel matrix in order to improve the error performance. Then, exhaustive APP detections are performed for the sub-MIMO systems. The log-likelihood ratios (LLRs) of the overlapped part of the signal space are summed while for other streams only a single APP is performed. In [86], a subspace detection scheme is proposed in which a punctured QRD is employed instead of regular QRD in order to reduce computational complexity and allow for joint detection of a subset of decoupled streams. The QRD version of their proposed scheme is similar to the complex version of LORD [91].

Although the works in [90] - [86] are essentially based on the same idea, they differ in some aspects, including the column ordering, column back substitution strategy, and hard/soft-output detection. Each choice of these options affects the error performance and computational complexity of the system such that in some cases they are far away from the optimal error performance or achieve a near-optimal performance at the cost of huge computational complexity. There exist other MIMO detection schemes like the work in [12], which is an ML tree search detection algorithm based on differential metrics and their recursive calculations and does not require the QRD and matrix inversion. Although the complexity of the proposed detector in [12]

approaches to a constant at high SNRs, it suffers from a variant complexity at low SNRs. Moreover, the efficiency of the proposed scheme in [12] reduces for high order modulations even for small MIMO systems. As a result, a near-optimal MIMO detection scheme with low computational complexity is still a challenging issue for moderate to large-scale MIMO systems.

In massive MU-MIMO systems, the design of practical signal detectors for the uplink is a vital challenge. The optimal ML detector is unfeasible for such systems with large dimensions. However, when the ratio of the antennas at the BS to the number of UTs is large, the signal processing is simplified such that the error performances of linear detectors, such as ZF and MMSE, are close enough to that of the optimal ML solution [67, 71]. However, the exact solutions of ZF and MMSE involve $\mathcal{O}(NK^2 + K^3)$ [1] operations because of a matrix-matrix multiplication - for the computation of the Gram channel matrix - and matrix inversion, respectively, with the computational complexity of $\mathcal{O}(NK^2)$ and $\mathcal{O}(K^3)$ where N is the number of receive antennas at the BS, and K is the number of UTs. Such operations become very expensive for massive MIMO as N and/or K become very large. Several techniques have been proposed in the literature in order to reduce the computational complexity of these two operations. In an effort to avoid direct matrix inversion, approximate algorithms using the truncated Neumann series expansion have been proposed in [31, 32, 101, 102, 108, 113]. The approximate inversion methods are useful when the ratio of the BS antennas to the UTs is large, while an exact inversion is preferred for small ratios. Further improvements have been achieved by replacing the Neumann series approximation with other iterative algorithms, such as the Gauss-Seidel [21], successive over-relaxation [32], Jacobi [92], Richardson [30], and conjugate gradient [45] methods. All these methods reduce the complexity of the matrix inversion from $\mathcal{O}(K^3)$ to $\mathcal{O}(K^2)$, but they still suffer from the complexity of $\mathcal{O}(NK^2)$ due to the Gram matrix calculation [13].

Another class of approximate methods is the truncated polynomial expansion (TPE) in which

[1] $\mathcal{O}(\cdot)$ is the big O notation that we use to evaluate how the computational complexity of detection algorithms grows as the dimensions of the system grow.

the matrix inversion is approximated with J finite terms of the Taylor series expansion. Such methods reduce the overall complexity to a factor which is proportional to the number of BS antennas and users, i.e., $\mathcal{O}(JKN)$. This is achieved by an iterative implementation of a series of matrix-vector multiplications instead of the matrix-matrix multiplication. However, the convergence speed of the TPE detector highly depends on a normalization factor in the polynomial coefficients. The optimal normalization factor, in the sense of convergence speed, requires the calculations of the largest and smallest eigenvalues of the Gram matrix [89]. The calculations of the Gram matrix and its eigenvalues increase the computational complexity to $\mathcal{O}(NK^2 + K^3)$. An approximate method has been proposed in [89] for the calculations of the eigenvalues of the Gram matrix. This method reduces the complexity of the eigenvalues calculations to $\mathcal{O}(K^2)$. However, it still needs the calculations of the entries of the Gram matrix. Therefore, its overall computational complexity is $\mathcal{O}(NK^2 + K^2)$. Consequently, efficient detectors with low computational complexity are required for the uplink of massive MIMO systems.

1.1 Summary of the Thesis

In this thesis, we deal with the problem of signal detection for MIMO systems. The objective is to find low-complexity near-optimal detectors for such systems. We start with small and moderate MIMO systems in Chapter 2. The proposed detectors exhibit essentially ML performance as demonstrated in simulations. Moreover, it is shown that the proposed algorithms guarantee full receive diversity. In the first algorithm, similar to what has been proposed in [74], a total of M numbers of LLL reductions are performed on M submatrices, but the exhaustive searches for detecting one symbol are conducted M times. Then, the solution that has the smallest Euclidean distance to the received signals is selected. The main cost resides in performing the lattice reduction M times. As a result, we propose a second algorithm to reduce the computational complexity due to M times lattice reduction. In the first algorithm, to find the best

submatrix for LRA conditional detection, a separate lattice reduction is done on each submatrix. However, the submatrices differ only in one column. Hence, one expects that their lattice reduction unimodular matrices are very close to each other. The second algorithm uses the unitary transformation matrix obtained from the reduction of one of the submatrices for transforming all the other submatrices. For our proposed algorithms, the overall computational complexity of the preprocessing and detection stages is $\mathcal{O}(M^4/N_B + M^3|\mathcal{X}|))$ where N_B is the number of signal vectors for which the channel is considered constant. For quasi-static channels, where the cost of the preprocessing stage can be negligible over the whole packet over which the channel is constant, the computational complexity of the detector is linear in the size of the QAM constellation, i.e., $\mathcal{O}(M^3|\mathcal{X}|))$. However, the cost is prohibitive for fast fading channels. Moreover, for large MIMO systems, the computational complexity of the lattice reduction algorithms imposes a large burden on the detectors.

In the first part of Chapter 3, we remove the lattice reduction algorithm and propose a detection algorithm that is based on the combination of MMSE-DFE and subspace detection tailored to the V-BLAST ordering on the whole MIMO channel. The proposed detector achieves the same diversity order as the FSD scheme with the FSD ordering and one full search level. However, the search over all possible submatrices makes the proposed detector outperform the FSD scheme. Although the diversity order of the proposed detector does not necessarily achieve the full ML diversity for some MIMO systems, its coding gain compared to the FSD scheme with one full search level allows for close to ML performance at practical values of SNR. The BER performance is close to that of the ML while the FSD scheme needs a larger number of full searches to achieve the optimal error performance. The proposed algorithm has a linear complexity in the constellation size and a complexity of $\mathcal{O}(M^4)$ in the system dimension M because of the cost of M successive QRD of all M possible subdivisions of the backward filter obtained by the MMSE-DFE filtering, even by a fast implementation procedure as described in [8, 37, 114] for example. We show how the cost can be reduced to a cubic complexity in M using the Givens

rotations method. Because of the structure of the submatrices and the fact that successive submatrices have $M-2$ columns in common, there is no need to perform the QRD from scratch. For the proposed detection algorithm, there is a performance gap to the ML BER at high SNRs when we increase the dimension of the system and also the constellation order. Hence, we propose a second algorithm that offers a modification on the first algorithm at the cost of a small increase in the computational complexity yet linear in $\mathcal{X}$ and cubic in M. In the first algorithm, the whole channel matrix is ordered whereas the remaining submatrix obtained by removing one column at a time is not necessarily in the optimal order which results in performance degradation for larger systems with high order constellations. To improve the performance, one can perform a new greedy ordering on each submatrix; however, this increases the complexity of all QRD to $\mathcal{O}(M^4)$ due to destroying the upper triangular structure of the submatrices. Alternatively, we suggest conducting the one-step V-BLAST ordering such that only the first index of the optimal ordering of the corresponding submatrix is found. Then, the remaining of the detection is the same as the proposed detector in the first algorithm. The overall computational complexity of the second algorithm is also $\mathcal{O}(|\mathcal{X}|.M^3)$. The proposed detectors can naturally be exploited for soft-output MIMO detectors when concatenated with a soft-input soft-output (SISO) decoder for MIMO systems with error-correcting codes. This detector gives a better list of candidate codewords compared to the state of the art in [7, 63, 97]. It is shown that when the proposed detector is accompanied by a SISO decoder, the error performance is close to that of the list sphere decoder (LSD) which uses a much larger list size. In addition, the proposed detector naturally lends itself to parallel implementation and offers a unified framework with a fixed-complexity that is applicable for all configurations of MIMO systems including under-determined MIMO systems.

In the second part of Chapter 3, we propose a detection scheme in order to improve the diversity order of conditional detection with one full search level. We deploy the partial lattice reduction (PLR) [64] which utilizes the lattice reduction with a tunable block size over the channel matrix. With this technique, the diversity order improvement of the SIC detection equals to the

block size of the PLR algorithm. The proposed detector outperforms the-state-of-the-art MIMO detectors in terms of diversity order and error performance with comparable computational complexity. By applying several efficient techniques, and also selecting an appropriate block size for the PLR algorithm, a computational complexity cost of $\mathcal{O}(|\mathcal{X}|.M^3)$ can be achieved for the proposed detector.

In Chapter 4, we propose an efficient detector for the uplink of massive MU-MIMO systems. For such systems, suboptimal linear detectors, such as ZF and MMSE, exhibit a BER performance close to that of the optimal ML solution. Two main operations in these linear detectors are matrix inversion and matrix-matrix multiplication for the calculation of the Gram channel matrix. These operations are expensive for massive MIMO systems as they require $\mathcal{O}(KN^2 + K^3)$ operations where N is the number of antennas at the BS, and K is the number of UTs. Therefore, we resort to the class of TPE methods as they avoid matrix inversion and alternatively approximate it with J finite terms of the Taylor series expansion. Moreover, by using an iterative implementation, the expensive matrix-matrix multiplication is replaced with a series of matrix-vector multiplications. This reduces the overall computational complexity of the TPE-based detector to a factor which is proportional to the number of antennas at the BS and also the number of UTs, i.e., $\mathcal{O}(JKN)$. However, the optimal TPE-based detector, in the sense of convergence speed, requires the calculations of the largest and smallest eigenvalues of the Gram matrix [89] which increase the computational complexity to $\mathcal{O}(NK^2 + K^3)$. In this chapter, by using the approximate extreme eigenvalues of the Gram matrix of massive MIMO systems, we propose a constant normalization factor which depends on the dimensions of the system. This normalization factor is calculated with $\mathcal{O}(1)$ operations and consequently, the overall complexity of the proposed TPE-based detector is $\mathcal{O}(JNK)$. The proposed TPE-based detector exhibits an error performance similar to that of the TPE with the optimal polynomial coefficients for small loading factors $\beta = \frac{K}{N}$. A modification is considered for systems with larger β's in order to remove the error performance degradation at high SNRs which happens due to the inaccuracy of

the extreme eigenvalues approximations for such systems. An efficient algorithm based on the power method is proposed for calculating the extreme eigenvalues of such systems. By selecting an appropriate value for J, our proposed detectors achieve the BER performances of the ZF or MMSE detectors while the computational complexity is significantly reduced.

In Chapter 5, we propose a detection algorithm for millimeter wave (mmWave) massive MIMO systems with hybrid beamforming. Hybrid beamforming is considered to be the viable realization of massive MIMO systems in the mmWave bands as it offers low implementation costs by utilizing a few RF chains at the BSs and/or the UTs. Hybrid beamforming structures are designed by approximating the optimal fully-digital beamformer as a product of analog and digital beamformers under Euclidean distance measure. Although such hybrid beamformers exhibit relatively good performance in terms of spectral efficiency, their BER performance is poor for some mmWave massive MIMO systems due to the approximation error. In this chapter, we propose a signal detection scheme using the equivalent channel consisting of the precoder, mmWave channel, and combiner in order to improve the error performance of such systems. This equivalent channel is square with the dimension of up to 8 equal to the number of data streams, i.e., N_s, in mmWave massive MIMO systems. This enables us to utilize the subspace MMSE-DFE based detector proposed in Chapter 3. However, because of the available structures in the equivalent channel matrix, the computational complexity can be reduced compared to the detectors in Chapter 3. It is shown that the complexity of the proposed detector for the worst-case scenario is $\mathcal{O}(N_s^3|\mathcal{X}|)$. On the other hand, the overall computational complexity of hybrid beamforming techniques, such as orthogonal matching pursuit (OMP) [6], or its variants [66], and alternating minimization (AM) [80, 109], are respectively $\mathcal{O}(M^2 N_{\text{RF}} N_s)$ and $\mathcal{O}(M(N + N_{\text{RF}}^2)N_s)$ where N_{RF} is the number of RF chains. Consequently, for massive MIMO systems, the computational complexity of our proposed detector is in the same order as that of the hybrid beamforming technique. Our proposed detector achieves the optimal error performance while the computational complexity is reduced compared to the ML or other detection schemes for

MIMO systems.

1.2 MIMO Detection Background

In this section, we introduce the system model as well as the main detection schemes that we use in this thesis. We consider a MIMO system with M transmit and N receive antennas. The received signal is

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1.1}$$

where $\mathbf{x}$ is the $M \times 1$ transmitted symbols vector chosen from a given constellation, $\mathcal{X}$, with the average energy of E_s. Moreover, $\mathbf{H}$ is the $N \times M$ channel matrix whose entries are independent and identically distributed (i.i.d.) following a complex Gaussian distribution with zero mean and unit variance, and $\mathbf{n}$ is the complex Gaussian noise vector with zero-mean and i.i.d. entries of variance N_0. We further assume that the channel state information is perfectly known only at the receiver. The ML solution to (1.1) is given by

$$\hat{\mathbf{x}} = \underset{\mathbf{x} \in \mathcal{X}^M}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2. \tag{1.2}$$

This solution imposes a huge computational complexity on the system as it increases exponentially with the number of transmit antennas. Hence, it is infeasible for large constellations or large MIMO systems. SD is a more efficient implementation of the ML solution. The idea in SD is to search over only lattice points located in a sphere with a certain radius around the received vector [38] (See Fig. 1.3). Although SD is feasible for small lattice dimensions, its exponential complexity makes this scheme impractical for systems with larger dimensions.

Therefore, approximate detection methods have been proposed in the literature in order to reduce the computational complexity. ZF technique is a simple linear approximate method which

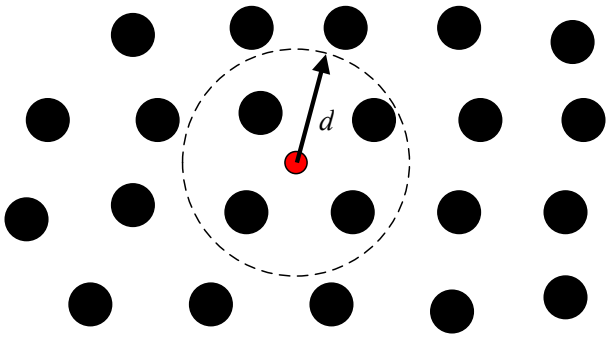


Figure 1.3: Sphere decoding with radius d.

detects the signal vector as the closest lattice point to $\hat{\mathbf{x}}_{\text{ZF}}$, where we have

$$\hat{\mathbf{x}}_{\text{ZF}} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{y} \tag{1.3}$$

$$= \mathbf{x} + (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{n}, \tag{1.4}$$

where $\mathbf{H}^H$ is the Hermitian matrix of $\mathbf{H}$. Then, the vector of closest constellation points to $\hat{\mathbf{x}}_{\text{ZF}}$ is declared as the detected signal vector. ZF completely suppress the mutual interference between the layers of the detection. This method reduces the computational complexity of the detection, but it has a poor error performance due to noise amplification. When $\mathbf{H}$ has small singular values, the effective noise is magnified significantly resulting in a poor error performance. In order to improve over this issue, MMSE is utilized which considers the noise term in the design of linear detector such that

$$\hat{\mathbf{x}}_{\text{MMSE}} = (\mathbf{H}^H\mathbf{H} + \alpha\mathbf{I}_M)^{-1}\mathbf{H}^H\mathbf{y}, \tag{1.5}$$

where $\alpha = \frac{N_0}{E_s}$. In contrast to ZF and MMSE that detect all entries of the transmitted signals vector in parallel, there are other techniques that perform successive interference cancellation (SIC) in which the signals are detected one after another [26, 33, 37]. One way to implement SIC is to utilize the DFE technique by performing the QRD on the channel matrix. The MMSE

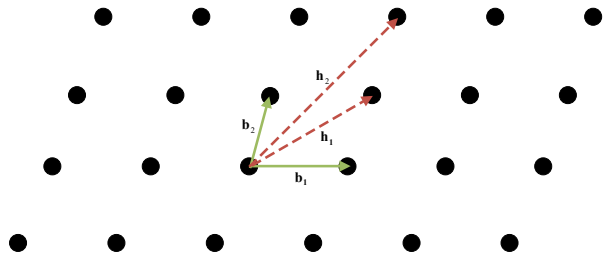


Figure 1.4: Lattice reduction for a two-dimensional lattice.

counterpart of this SIC scheme, which is referred to as MMSE-DFE, is similarly implemented when the channel matrix is replaced with the augmented (or regularized) channel matrix [22]

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{H} \\ \sqrt{\alpha}\mathbf{I}_M \end{bmatrix} = \tilde{\mathbf{Q}}\mathbf{R}, \tag{1.6}$$

where $\tilde{\mathbf{Q}} \in \mathcal{C}^{(N+M)\times M}$ is a matrix with orthonormal columns, and $\mathbf{R} \in \mathcal{C}^{M\times M}$ is an upper triangular matrix with real and positive diagonal elements. Multiplying $\mathbf{F} = \mathbf{Q}^H$ as the forward matrix on the received signal vector, where $\mathbf{Q}$ is the first N rows of $\tilde{\mathbf{Q}}$, gives

$$\begin{aligned} \mathbf{y}' &= \mathbf{Q}^H\mathbf{y} \\ &= \mathbf{R}\mathbf{x} + \mathbf{w}, \end{aligned} \tag{1.7}$$

where $\mathbf{w} = \mathbf{Q}^H\mathbf{n} - [\mathbf{R} - \mathbf{Q}^H\mathbf{H}]\mathbf{x}$. Next, using the upper triangular structure of $\mathbf{R}$, SIC can be easily implemented by working from the last row upward. The drawback of this method is error propagation because the detection of each symbol depends on the detection of previously detected symbols.

LRA techniques [53, 99, 106] have been proposed to improve the BER performances of sub-optimal detectors. Lattice reduction algorithms try to find a basis in which the vectors are short and approximately orthogonal. See Fig. 1.4 for more illustration of a two-dimensional lattice reduction. LLL [61] is a well-known lattice reduction algorithm which has a polynomial com-

plexity in the dimension of the lattice. See Appendix A for more details. It is worth mentioning that the original LLL algorithm was proposed for real-valued lattices while its complex-valued version has been offered in [29]. A complex-valued basis $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, ..., \mathbf{b}_M]$ is considered as LLL-reduced if

$$|\Re\{(p_{i,j}\}| \leq 1/2, \quad |\Im\{p_{i,j}\}| \leq 1/2, \quad \text{for} \quad 1 \leq j < i \leq M, \tag{1.8}$$

$$\| \hat{\mathbf{h}}_i + p_{i,i-1}\hat{\mathbf{h}}_{i-1} \|^2 \geq \kappa \| \hat{\mathbf{h}}_{i-1} \|^2 \quad 1 < i \leq M, \tag{1.9}$$

where $1/2 < \kappa \leq 1$ for complex LLL reduction and also $\mathbf{H} = \hat{\mathbf{H}}\mathbf{P}^T$ is the Gram-Schmidt orthogonalization where $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, ..., \hat{\mathbf{h}}_M]$ and $\mathbf{P} = [p_{i,j}]$ is a lower triangular matrix. In the LLL algorithm, for a given channel matrix, a basis with smaller orthogonality defect factor is determined. The orthogonality defect factor for matrix $\mathbf{B}$ is defined as

$$\delta(\mathbf{B}) \triangleq \frac{\|\mathbf{b}_1\|^2 \|\mathbf{b}_2\|^2 ... \|\mathbf{b}_M\|^2}{\det(\mathbf{B}^H\mathbf{B})}, \tag{1.10}$$

in which $\mathbf{b_i}$'s are the columns of $\mathbf{B}$. Note that $\delta(\mathbf{B}) \geq \mathbf{1}$ with equality if and only if the basis is orthogonal. In practice, for small to medium dimensions, the cost of the LLL reduction for a square matrix is cubic in the dimension even though in theory it is higher.

The LRA detection algorithms perform a reduction on the channel matrix such that

$$\mathbf{B} = \mathbf{H}\mathbf{U}, \tag{1.11}$$

where $\mathbf{B}$ is the reduced channel matrix and $\mathbf{U}$ is an $M \times M$ unimodular transformation matrix, i.e. it contains only complex entries with integer real and imaginary components and its determinant is ± 1, $\pm i$. Hence, one may write

$$\mathbf{y} = \mathbf{B}\mathbf{x}' + \mathbf{n}, \tag{1.12}$$

where $\mathbf{x}' = \mathbf{U}^{-1}\mathbf{x}$. $\hat{\mathbf{x}}'$ is detected using a suboptimal detectors like ZF or DFE. Then, the vector of closest constellation points to $\hat{\mathbf{x}} = \mathbf{U}\hat{\mathbf{x}}'$ is declared as the detected signal vector. LRA algorithms provide better performances than suboptimal detectors alone. They also guarantee achieving the full receive diversity [94] which is equal to the number of receive antennas. However, the computational complexity of the lattice reduction algorithms impose a large burden on large MIMO systems.

Conditional (subspace) detection is the idea of splitting the optimization sets into two subsets where an exhaustive search over a given constellation is performed for one subset. Then, the detection process is completed by a suboptimal detector in order to detect the entries of the other subset. In this thesis, the term conditional detection is used interchangeably with subspace detection. In this technique, the optimization set is split in such a way that the transmitted signal vector and the channel matrix are divided into two subvectors and submatrices, respectively. Hence, one can write (1.1) as

$$\begin{aligned} \mathbf{y} &= \mathbf{H}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{x}_2 + \mathbf{n}, \\ \mathbf{x}_1 &\in \mathcal{X}^{M-p}, \;\; \mathbf{x}_2 \in \mathcal{X}^{p}, \end{aligned} \tag{1.13}$$

where $\mathbf{H}_1$ and $\mathbf{H}_2$ are chosen in such a way that they have $M - p$ and p columns, respectively. In this method, an exhaustive search for finding subset $\mathbf{x}_2$ is conducted. Then, for a given $\mathbf{x}_2$, subvector $\mathbf{x}_1$ is computed by either linear detectors, such as ZF or MMSE, or SIC schemes like DFE, or both. From the complexity perspective, the most interesting case is when $p = 1$ as the number of exhaustively searched points is linear in the size of the constellation.

The FSD scheme can be considered as a special case of the family of conditional detectors. The FSD algorithm consists of two stages; full expansion (FE) and single expansion (SE). In the FE stage, all the constellation points are considered for the detection while in the SE stage, only one branch is expanded. The FSD algorithm is tailored to the FSD channel ordering in which a certain number of columns, p, corresponding to the signals with the maximum post-processing

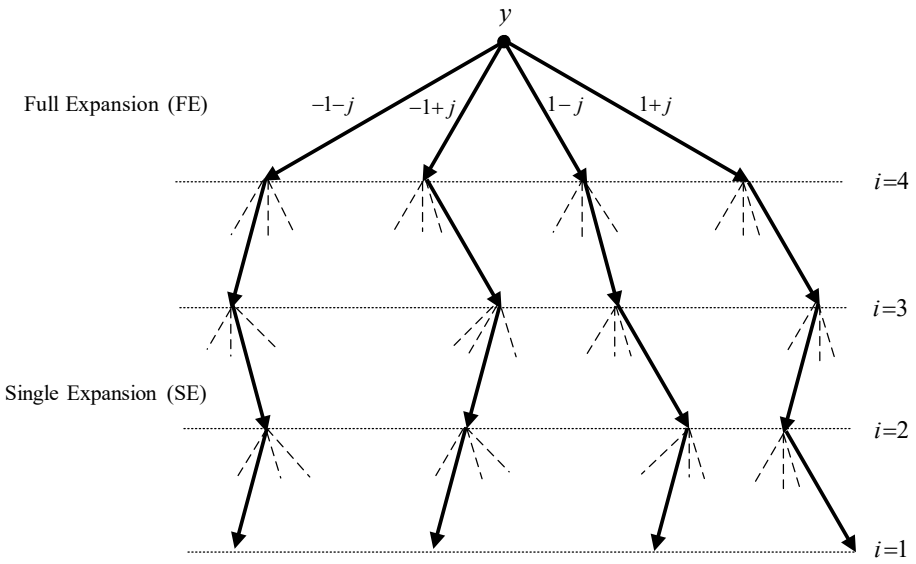


Figure 1.5: The FSD tree search for a 4×4 MIMO system with 4-QAM constellation.

noise is used for the FE stage and the remaining columns corresponding to the signals with minimum post-processing noise are sorted in the SE stage. Fig. 1.5 shows the tree representation of the FSD algorithm for a 4×4 MIMO system with 4-QAM constellation when $p = 1$.

The FSD can approach the optimal performance if at least a certain number of search levels are fully expanded along with the FSD ordering [7, 52]. For a square channel matrix, $p_{min} \geq \sqrt{M} - 1$ full search levels are required to achieve the optimal receive diversity [52]. Therefore, the detection computational complexity is in the order of $\mathcal{O}(|\mathcal{X}|^{p_{min}})$ which could still be prohibitive when a high order constellation with the size of $|\mathcal{X}|$ is used or the number of transmit antennas is large.

Chapter 2

Detection for Small and Moderate-Scale MIMO Systems

2.1 Introduction

The optimal recovery of the transmitted signal at the receiver corrupted by channel fading, noise and interference is a challenging issue for MIMO systems. To avoid the high computational complexity of the optimal ML detection, suboptimum detectors, such as ZF, DFE, MMSE, and V-BLAST detector [26, 33, 37], have been studied in the literature. Although these simple detection algorithms reduce the computational complexity, their BER performances are far worse than those of the ML, especially at high SNRs. LRA algorithms [53, 99, 106] have been proposed where efficient lattice reduction sub-routines, such as LLL algorithm [61], in conjunction with detectors like ZF and DFE provide better performances than ZF or DFE alone. On the other hand, LRA conditional detection was proposed in [74] where the orthogonality defect factor has been utilized as a metric for choosing the best LLL reduced submatrix for the conditional detection. This combination of lattice reduction and conditional optimization further reduces the

gap to the ML performance. In order to save in the computations and have the desirable case of linear complexity in the size of QAM constellation, only one symbol is detected using exhaustive search.

In this chapter [1], first, we propose a scheme which brings the BER performance closer to the ML solution with a computational complexity which is linear in the size of a given constellation. In the first algorithm, we search over all possible LLL reduced submatrices instead of considering the orthogonality defect factor to choose the best LLL reduced submatrix for conditional detection. Then, to reduce the complexity of the proposed detection algorithm, we introduce another scheme which performs lattice reduction only for one submatrix and the corresponding transformation matrix is used for transforming other submatrices. While the complexity is reduced, the performance remains essentially ML. We also discuss the computational complexity and error performance of the proposed detectors. Simulation results show the superiority of the proposed schemes compared to those presented in the literature for small and moderate-scale MIMO systems.

2.2 System Model

We consider a MIMO system with M transmit and N receive antennas. The received signal $\mathbf{y}$ is

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{2.1}$$

where $\mathbf{x}$ is the vector of transmitted symbols whose entries are chosen from a QAM constellation, $\mathcal{X}$, with average energy of E_s. Moreover, $\mathbf{H}$ is the $N \times M$ channel matrix whose elements are i.i.d. following a complex Gaussian distribution with zero mean and unit variance. Also, $\mathbf{n}$ represents the noise vector which is a complex Gaussian random variable with zero mean and

[1] The results of this chapter have been published in [51].

i.i.d. entries of variance N_0. We consider the conditional detection

$$\begin{aligned} \mathbf{y} &= \mathbf{H}_1\mathbf{x}_1 + \mathbf{H}_2\mathbf{x}_2 + \mathbf{n}, \\ \mathbf{x}_1 &\in \mathcal{X}^{M-p}, \;\; \mathbf{x}_2 \in \mathcal{X}^{p}. \end{aligned} \tag{2.2}$$

If $P = M - p$ denotes the number of columns of the first submatrix, we are interested in $P = M-1$ (or $p = 1$) because the exhaustive search is conducted only for detecting one symbol, and thus, it has a linear complexity in the size of the corresponding QAM constellation [74]. For $p = 1$, the submatrix selection set, $S = \{\mathbf{H}_{[1]}, \mathbf{H}_{[2]}, ..., \mathbf{H}_{[M]}\}$, is the set of all submatrices of $\mathbf{H}$ with $M - 1$ columns in which $\mathbf{H}_{[i]}$ is the submatrix of $\mathbf{H}$ obtained by removing its i-th column. Hence, (2.2) can be written as

$$\begin{aligned} \mathbf{y} &= \mathbf{H}_{[i]}\mathbf{x}_{[i]} + \mathbf{h}_i x_i + \mathbf{n}, \\ \mathbf{x}_{[i]} &\in \mathcal{X}^{M-1}, \;\; x_i \in \mathcal{X}; \\ & i = 1, 2, \ldots, M. \end{aligned} \tag{2.3}$$

The performance of the conditional detection method mainly depends on the orthogonality of the first submatrix. The orthogonality defect factor for a submatrix, e.g., $\mathbf{H}_{[M]}$ is defined as

$$\delta(\mathbf{H}_{[M]}) \triangleq \frac{\|\mathbf{h}_1^2\|\|\mathbf{h}_2^2\|...\|\mathbf{h}_{M-1}^2\|}{\det(\mathbf{H}_M^H\mathbf{H}_M)}, \tag{2.4}$$

in which $\mathbf{h}_1, \ldots, \mathbf{h}_{M-1}$ are columns of $\mathbf{H}_{[M]}$. In the LLL algorithm, for a given channel matrix, a basis with smaller orthogonality defect factor is determined. By performing the LLL algorithm on submatrix $\mathbf{H}_M$, we have

$$\mathbf{B}_M = \mathbf{H}_{[M]}\mathbf{U}_M, \tag{2.5}$$

where $\mathbf{U}_M$ is a $(M-1) \times (M-1)$ unimodular transformation matrix. In [74], the submatrix which gives the smallest orthogonality defect factor, i.e.,

$$\delta_{min} = \min_{i=1,\ldots,M} \delta(\mathbf{B}_i), \tag{2.6}$$

is used for conditional detection where $\mathbf{B}_i = \mathbf{H}_{[i]}\mathbf{U}_i$ is the LLL reduced matrix of submatrix $\mathbf{H}_{[i]}$. Without loss of generality, we assume that $\mathbf{B}_M$ is the submatrix corresponding to δ_{min}. Hence, conditional detection is performed using

$$\mathbf{y} = \mathbf{B}_M \mathbf{x}'_{[M]} + h_M x_M + \mathbf{n}, \tag{2.7}$$

However, according to the BER performance results reported in [74] there exists a performance gap between such a conditional setting for $P = M - 1$ and the ML solution. For example, for an 8×8 MIMO system, when $P = 7$, there is about 1.5-dB gap to the ML performance and the optimal performance is achieved for $P = 6$ ($p = 2$). In the next section, we propose two LRA conditional detection schemes with $p = 1$ and a different submatrix selection strategy to close this gap.

2.3 LRA Conditional Optimization Detector

In [74], the submatrix giving the smallest orthogonality defect factor is considered as the best submatrix for conditional optimization and the exhaustive search is performed for detecting one symbol for $P = M - 1$. For such a setting there exists a BER performance gap to the ML performance. To reduce this gap, we first propose to perform M times (considering all submatrice of size $N \times (M-1)$) LRA conditional detection with $P = M - 1$ and then choose the solution that corresponds to the best metric. This is explained in more details next.

Algorithm 2.1 The Proposed LRA Conditional Optimization Detector

1: Consider all possible submatrices $\mathbf{H}_{[i]} \in \mathbf{S}$, $i = 1, 2, ..., M$.
2: $\mathbf{h}_i$ is the remaining column of matrix $\mathbf{H}$ other than $\mathbf{H}_{[i]}$.
3: Perform the LLL reduction: $\mathbf{B}_i = \mathbf{H}_{[i]}\mathbf{U}_i$.
4: Perform the conditional detection: $\mathbf{y} = \mathbf{B}_i\mathbf{x}'_{[i]} + \mathbf{h}_i x_i + \mathbf{n}$; $\mathbf{x}'_{[i]} = \mathbf{U}_i^{-1}\mathbf{x}_{[i]}$

1. Fix $\hat{x}_i^{(j)} = a_j; a_j \in \mathcal{X};\ \ j = 1, 2, ..., |\mathcal{X}|$.
2. Detect $\hat{\mathbf{x}}_{[i]}^{(j)}$: By having $\mathbf{B}_i$, ZF or DFE is performed on $\mathbf{y} - \mathbf{h}_i\hat{x}_i^{(j)}$ and is denoted by $\tilde{\mathbf{x}}_{[i]}^{(j)}$. Then, $\hat{\mathbf{x}}_{[i]}^{(j)} = \mathbf{U}_i\hat{\mathbf{x}}_{[i]}^{\prime(j)} = \mathbf{U}_i \left\lceil \tilde{\mathbf{x}}_{[i]}^{(j)} \right\rfloor$ where $\lceil \cdot \rfloor$ is the nearest integer function.

5: Calculate $d_k = \|\mathbf{y} - \mathbf{B}_i\hat{\mathbf{x}}_{[i]}^{\prime(j)} - \mathbf{h}_i\hat{x}_i^{(j)}\|^2$; $k = (i-1)|\mathcal{X}| + j$
6: Find $d_{min} = \min_{k=1,..,M|\mathcal{X}|} d_k$ and declare the corresponding $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_{[i]}^{(j)}, \hat{x}_i^{(j)}]$ after reordering as detected symbol vector.

2.3.1 Algorithm 2.1

We consider all possible submatrices of the set S and for each one, conditional detection is performed. Then, we choose the solution that has the smallest Euclidean distances to the received signals. Note that in this scheme, similar to what has been proposed in [74], a total M numbers of LLL reductions are performed on the M submatrices, but the exhaustive searches for detecting one symbol are conducted M times. Consequently, the computational complexity is still linear in the size of the QAM constellation. The main cost resides in performing the lattice reduction M times (with a cost, in practice, of $\mathcal{O}(M^4)$). This cost can be negligible for quasi-static channels, however, prohibitive for fast fading ones. We note that to obtain the LLL reduction of matrix $\mathbf{H}_{[i]}$, we start by $\mathbf{H}_{[i]}\mathbf{U}_{[i-1]}$. Since the difference of these two matrices is only one column, the LLL computations will take a lot less than starting from scratch. In the next subsection, we propose another algorithm to reduce the computational complexity due to lattice reductions but still giving essentially ML performance as demonstrated in simulations.

Algorithm 2.2 The Proposed LRA Conditional Optimization Detector

1: Consider $M-1$ first columns of $\mathbf{H}$, denoted by $\mathbf{H}_{[M]}$.
2: Perform the LLL reduction: $\mathbf{B}_M = \mathbf{H}_{[M]}\mathbf{U}_M$.
3: For all other submatrices: $\mathbf{B}_i \triangleq \mathbf{H}_{[i]}\mathbf{U}_M, \;\; i = 1, 2, , ..., M-1$.
4: $\mathbf{h}_i$ is the remaining column of matrix $\mathbf{H}$ other than $\mathbf{H}_{[i]}$.
5: Perform the conditional detection: $\mathbf{y} = \mathbf{B}_i\mathbf{x}'_{[i]} + \mathbf{h}_i x_i + \mathbf{n}$; $\mathbf{x}'_{[i]} = \mathbf{U}_i^{-1}\mathbf{x}_{[i]}$

   1. Fix $\hat{x}_i^{(j)} = a_j; a_j \in \mathcal{X}; \;\; j = 1, 2, ..., |\mathcal{X}|$.
   2. Detect $\hat{\mathbf{x}}_{[i]}^{(j)}$: By having $\mathbf{B}_i$, ZF or DFE is performed on $\mathbf{y} - \mathbf{h}_i\hat{x}_i^{(j)}$ and is denoted by $\tilde{\mathbf{x}}_{[i]}^{(j)}$. Then, $\hat{\mathbf{x}}_{[i]}^{(j)} = \mathbf{U}_i\hat{\mathbf{x}}'^{(j)}_{[i]} = \mathbf{U}_i \left\lceil \tilde{\mathbf{x}}_{[i]}^{(j)} \right\rfloor$.

6: Calculate $d_k = \|\mathbf{y} - \mathbf{B}_i\hat{\mathbf{x}}'^{(j)}_{[i]} - \mathbf{h}_i\hat{x}_i^{(j)}\|^2$; $k = (i-1)|\mathcal{X}| + j$
7: Find $d_{min} = \min\limits_{k=1,..,M|\mathcal{X}|} d_k$ and declare the corresponding $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_{[i]}^{(j)}, \hat{x}_i^{(j)}]$ after reordering as detected symbol vector.

2.3.2 Algorithm 2.2

In order to find the best submatrix for LRA conditional detection, a separate lattice reduction is done on each submatrix. However, some of these submatrices are related to each other because for $P = M-1$, there are submatrices that differ only in one column. Hence, one expects that their lattice reduction unimodular matrices are very close to each other. Algorithm 2.2 uses the unitary transformation matrix obtained from the reduction of one of the submatrices for transforming all the other submatrices. In fact, we obtain $\mathbf{U}_M$ using the LLL algorithm as

$$\mathbf{B}_M = \mathbf{H}_{[M]}\mathbf{U}_M, \tag{2.8}$$

and we use $\mathbf{U}_M$ for the other submatrices:

$$\mathbf{B}_i \triangleq \mathbf{H}_{[i]}\mathbf{U}_M, \quad i = 1, 2, ..., M-1. \tag{2.9}$$

2.4 Computational Complexity

In this subsection, the computational complexity of the proposed algorithms is discussed. Similar to the LRA conditional detection in [74], when $P = M - 1$, Algorithm 2.1 requires M lattice reductions and M consecutive exhaustive search over the QAM constellation. The complexity is still linear in the size of the QAM constellation. In practice, for small to medium dimensions, LLL reduction cost is cubic in the dimension (even though in theory it is higher). For example, using the average complexity results for the LLL reduction in [36, 76], for an 8×8 MIMO system, the LLL reduction of each submatrix requires around 2×10^4 floating point operations (FLOPs) on average. In Algorithm 2.1, if we consider all submatrices and perform a separate LLL reduction for each, a total 1.6×10^5 FLOPs are needed. However, we can save on FLOPs by considering an already reduced submatrix as a starting point for LLL reduction of the next submatrix. Algorithm 2.2 needs only one LLL reduction, i.e., 2×10^4 FLOPs on average, and the reductions for other submatrices are replaced with matrix multiplications requiring $21,952$ FLOPs [17]. Hence, the overall numbers of FLOPs for the reduction and matrix multiplications are about 2.4×10^4. To obtain further saving in computational complexity, one can implement the LLL reduction in an efficient way such that the lattice reductions of different submatrices are conducted in parallel because the submatrices differ only in one column and their reductions are closely related to each other.

2.5 Diversity Order Analysis

To improve the reliability of MIMO systems, it must be ensured that the symbols that are sent through multiple paths fade independently which is known as diversity technique. In [40, 72], it is shown that the diversity order of ZF and the MMSE for separately encoded systems like the V-BLAST transmission scheme is $N - M + 1$. However, this diversity order is far away from the

maximum receive diversity. In this subsection, the diversity order of the proposed detectors is discussed. In the following theorem, we show that the proposed detectors achieve the maximum receive diversity order under the V-BLAST transmission scheme.

Theorem 1 *The LRA-conditional detectors using Algorithm 2.1 or Algorithm 2.2, under V-BLAST transmission scheme achieve the maximum receive diversity order, i.e.*

$$d \triangleq \lim_{snr \to \infty} -\frac{\log P_e}{\log snr} = N, \tag{2.10}$$

where P_e is the probability of error and snr is the average SNR at each receive antenna.

Proof: See the proof in Appendix B. □

2.6 Numerical Results

In this section, we consider several MIMO system in order to investigate the error performance of the proposed algorithms. Moreover, we use MMSE-DFE filtering [23] in the preprocessing stage of the simulations. We evaluate the performance of the system by comparing the BER of SD, MMSE-LLL-DFE, MMSE-DFECD and MMSE-LLL-DFECD[2] detection methods [74].

In our simulation figures, we present BER performances for different values of E_b/N_0. For a given constellation, we defined E_s as the average energy per symbol [3]. Since the fading coefficients are assumed to be independent with unit variance, the average signal energy per receive antenna is $E_s.M$. Therefore, the total energy of $E_s.M.N$ is collected by N receive antennas. $R.M.k$ information bits are carried where R is the coding rate and $k = \log_2 |\mathcal{X}|$. Note that for

[2] CD refers to conditional detection.

[3] In some works, the average energy per symbol is normalized by M such that the total transmitted energy is the same for different MIMO configurations.

uncoded systems we have $R = 1$. As a result, the signal energy per transmitted information bit at the receiver is defined as [41]

$$E_b = \frac{E_s MN}{RMk}, \tag{2.11}$$

or we can express it as

$$E_b/N_0(\text{dB}) = snr(\text{dB}) + 10\log_{10}\frac{N}{RMk}, \tag{2.12}$$

where $snr = \frac{E_s.M}{N_0}$ is the average SNR at each receive antenna. We use this term for our BER simulation results throughout the thesis.

Fig. 2.1 and Fig. 2.2 show the performance results for an 8×8 MIMO system for 4-QAM and 16-QAM, respectively. As was expected, a better performance than [74], with a gain of 1.5 dB at the BER of 10^{-5} for both constellations, is obtained which is very close to the ML solution. Fig. 2.3 shows the performance results for a 16×16 MIMO system for the 4-QAM constellation. For this system also a better performance than [74], with a gain of 5 dB at the BER of 10^{-5} is achieved. Moreover, the full receive diversity of the proposed detectors, the slope of BER figure at high SNRs, can be seen in all figures. It can be concluded that the use of exhaustive search for all possible submatrices with LRA detection algorithm makes the performance essentially ML, especially at high SNRs.

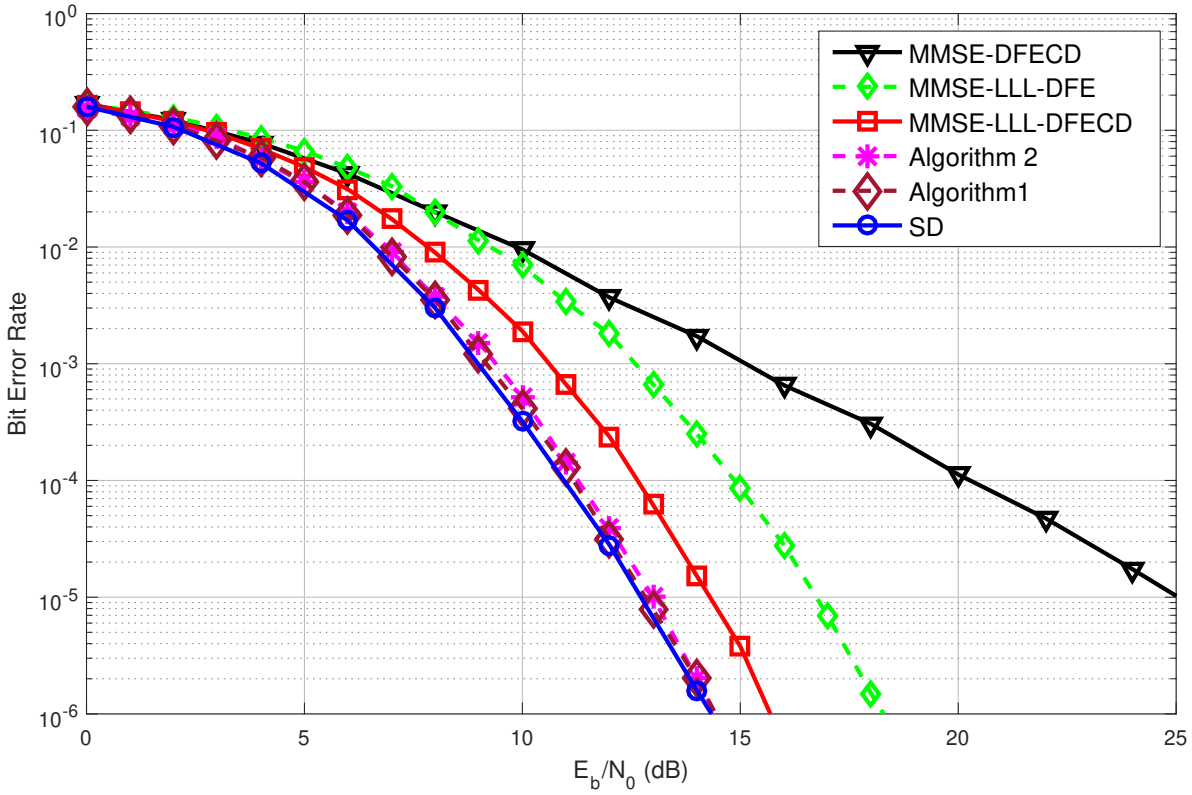


Figure 2.1: BER performance of the LRA conditional detection by using Algorithm 2.1 and Algorithm 2.2 for an 8×8, 4-QAM MIMO system.

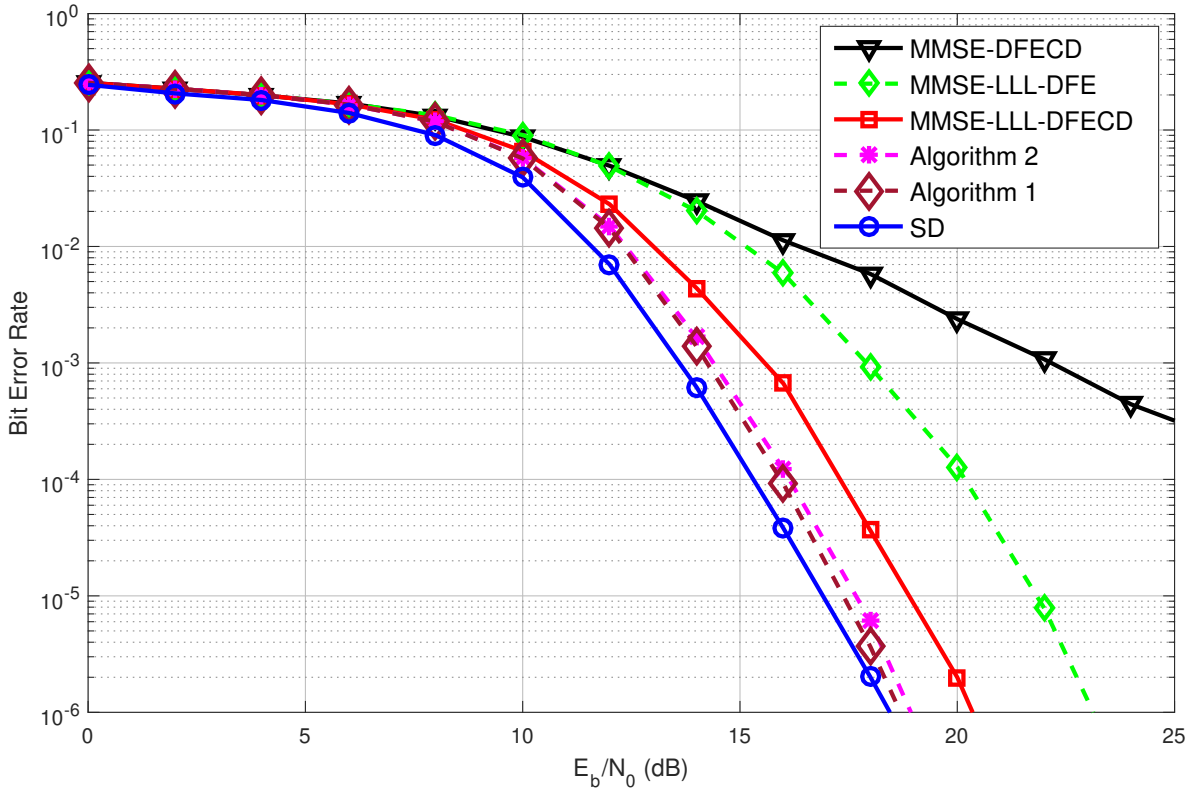


Figure 2.2: BER performance of the LRA conditional detection by using Algorithm 2.1 and Algorithm 2.2 for an 8×8, 16-QAM MIMO system.

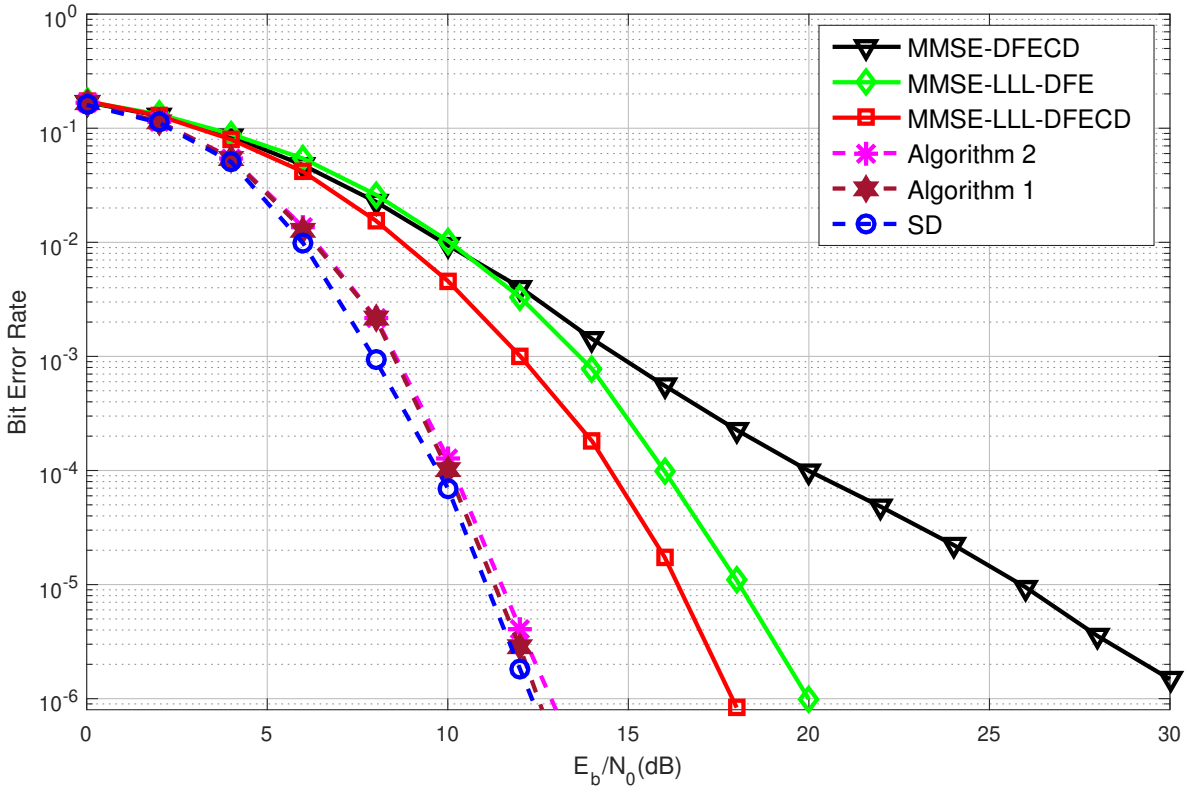


Figure 2.3: BER performance of the LRA conditional detection by using Algorithm 2.1 and Algorithm 2.2 for a 16×16, 4-QAM MIMO system.

2.7 Conclusion and Discussion

In this chapter, we proposed two detection algorithms for MIMO systems based on the combination of lattice reduction and conditional detection techniques. The first algorithm brings the BER performance close to the ML solution (essentially ML as verified by simulations) while it has the same computational complexity order as the previously reported LRA conditional detectors in the literature. The computational complexity of the preprocessing stage of the first algorithm is $\mathcal{O}(M^4)$ due to M times lattice reduction of M submatrices. Moreover, its real-time detection stage has a computational complexity of $\mathcal{O}(M^3|\mathcal{X}|)$ because of $M|\mathcal{X}|$ times conditional detections. In the second algorithm, we reduced the computational complexity of M times lattice reduction. The second algorithm performs the lattice reduction only for one submatrix, and the corresponding transformation matrix is used for transforming other submatrices. This reduces the computational complexity of the preprocessing stage compared to the first algorithm. While the computational complexity is largely reduced compared to the first algorithm, the BER performance degradation is negligible. For quasi-static channels, the computational complexity of the preprocessing stage in both algorithms becomes negligible, as it can be shared among many channel realizations, and consequently, the overall complexity will be $\mathcal{O}(M^3|\mathcal{X}|)$. However, it becomes prohibitive for fast fading channels as the computational complexity of the preprocessing stage can be large. Moreover, for large MIMO systems, the deployment of lattice reduction algorithms is not of interest as they impose a computational complexity burden on the system. In the next chapter, we propose low-complexity near-optimal conditional based detectors, which can be deployed for large MIMO systems.

Chapter 3

Detection for Large-Scale MIMO Systems

3.1 Introduction

In Chapter 2, we proposed near-optimal LRA conditional detectors for MIMO systems. However, the computational complexity of reduction algorithms imposes a large burden on the detectors, especially when the dimension of the system increases.

In this chapter [1], we propose detection algorithms which are mainly based on the idea of conditional (subspace) detection. In the first part of this chapter, we propose an MMSE-DFE based detection algorithm that performs exhaustive search on one symbol in a round-robin way. This provides two advantages: First, the BER performance of the proposed scheme is almost optimal at the cost of a small computational complexity increase over the classical MMSE-DFE detection (further reduction in complexity is discussed). Second, the proposed detector can naturally be exploited for soft-output MIMO detectors when concatenated with a SISO decoder for MIMO systems with error correcting codes. This detector gives a better list of candidate codewords compared to the state of the art in [7, 63, 97]. It is shown that when the proposed detector is

[1] The results of this chapter have been published in [46–50].

accompanied by a SISO decoder, the error performance is close to that of the LSD which uses a much larger list size. In addition, the proposed detector naturally lends itself to parallel implementation and offers a unified framework with a fixed-complexity which is applicable for all configurations of MIMO systems including under-determined MIMO systems.

In the second part of this chapter, we propose a partially reduced conditional detector for large-scale MIMO systems. Conditional detection with only one exhaustive search level may degrade the diversity order and error performance of the system. The PLR was proposed in [64] for SIC detection where the reduction with a tunable block size is performed over the channel matrix. The diversity improvement of SIC with the PLR algorithm equals to the block size of the PLR scheme. As a result, in order to improve the diversity order, the PLR is performed on a subchannel of non-exhaustive levels. Additionally, by considering other remaining non-reduced columns for conditional detection, the error performance is improved.

3.2 System Model

We consider the system model in (1.1) for MIMO systems

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{3.1}$$

where M and N are the transmit and receive antennas, respectively. In this chapter, we deploy several main techniques which are briefly explain in the following. We consider the conditional detection system model for $p = 1$ in (2.3)

$$\begin{aligned} \mathbf{y} &= \mathbf{H}_{[i]}\mathbf{x}_{[i]} + \mathbf{h}_i x_i + \mathbf{n}, \\ \mathbf{x}_{[i]} &\in \mathcal{X}^{M-1}, \;\; x_i \in \mathcal{X}; \\ & i = 1, 2, \ldots, M. \end{aligned} \tag{3.2}$$

We also implement the SIC scheme by utilizing MMSE-DFE technique in which the QRD is performed on the augmented (or regularized) channel matrix

$$\tilde{\mathbf{H}} = \begin{bmatrix} \mathbf{H} \\ \sqrt{\alpha}\mathbf{I}_M \end{bmatrix} = \tilde{\mathbf{Q}}\mathbf{R}, \tag{3.3}$$

where $\alpha = \frac{N_0}{E_s}$. Moreover, $\tilde{\mathbf{Q}} \in \mathcal{C}^{(N+M)\times M}$ is a matrix with orthonormal columns and $\mathbf{R} \in \mathcal{C}^{M\times M}$ is an upper triangular matrix with real and positive diagonal elements. Multiplying $\mathbf{F} = \mathbf{Q}^H$ as the forward matrix on the received signal vector, where $\mathbf{Q}$ is the first N rows of $\tilde{\mathbf{Q}}$, gives

$$\begin{aligned} \mathbf{y}' &= \mathbf{Q}^H\mathbf{y} \\ &= \mathbf{R}\mathbf{x} + \mathbf{w}, \end{aligned} \tag{3.4}$$

where $\mathbf{w} = \mathbf{Q}^H\mathbf{n} - [\mathbf{R} - \mathbf{Q}^H\mathbf{H}]\mathbf{x}$.
In the first part of this chapter, we propose detection schemes based on the integration of MMSE-DFE filtering and best submatrix selection for conditional optimization detection.

In the second part of this chapter, in addition to two above techniques, we deploy the PLR technique in order to improve the receive diversity order of conditional optimization detectors. In the PLR scheme, the last K columns of the basis $\mathbf{H}$ are projected on the orthogonal complement of the first $M - K$ columns [64].
If we denote this projection by $[\pi_{M-K}(\mathbf{h}_{M-K+1}), \pi_{M-K}(\mathbf{h}_{M-K+2}), ..., \pi_{M-K}(\mathbf{h}_M)]$, we have

$$\pi_{M-K}(\mathbf{h}_n) = \mathbf{h}_n - \sum_{i=1}^{M-K} p_{n,i}\hat{\mathbf{h}}_i, \tag{3.5}$$

for $n = M - K + 1, ..., M$ where $\mathbf{H} = \hat{\mathbf{H}}\mathbf{P}^T$ is the Gram-Schmidt orthogonalization in which $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, ..., \hat{\mathbf{h}}_M]$ and $\mathbf{P} = [p_{i,j}]$ is a lower triangular matrix. Then, the projected vectors are input to the LLL reduction algorithm. If $\mathbf{U}_K$ denotes the corresponding $K \times K$ unimodular

matrix, one writes the partially reduced basis as [64]

$$\mathbf{B} = \begin{bmatrix} \mathbf{H}_{N\times(M-K)}, & \mathbf{H}_{N\times K} \end{bmatrix} \begin{bmatrix} \mathbf{I}_{M-K} & \mathbf{0}_{(M-K)\times K} \\ \mathbf{0}_{K\times(M-K)} & \mathbf{U}_K \end{bmatrix}.$$

Note that directly reducing the last K columns of the channel matrix may not work because they might be reduced perfectly, but their interference on the other channel columns can be important which could affect the performance of the system. Hence, the projected vectors on the orthogonal complements are reduced to make it suitable for SIC detection.

3.3 Part I: MMSE-DFE Conditional Optimization Detector

3.3.1 Algorithm 3.1

The columns of the channel matrix are ordered by V-BLAST greedy ordering. We denote the set of all $N \times (M-1)$ submatrices of $\mathbf{B} = \mathbf{R}$, the backward filter matrix, as $S_{\mathbf{B}} = \{\mathbf{B}_{[1]}, \mathbf{B}_{[2]}, ..., \mathbf{B}_{[M]}\}$ where $\mathbf{B}_{[i]}$ denotes the $N \times (M-1)$ submatrix obtained by excluding column i. The proposed detector transforms the received signal using the forward matrix obtained from the QRD of the whole ordered channel matrix. Then, multiple conditional detection is conducted such that for each submatrix of the selection set, an exhaustive search over the constellation points is performed in order to detect one symbol and the remaining symbols are detected by DFE ($p = 1$, giving a list of $|\mathcal{X}|$ symbols vector candidates along with their metrics for each submatrix). Among all possible submatrices, choose the one that has the minimum Euclidean distance. Note that before performing the minimum Euclidean test, the candidates are mapped back to the original constellation. The detailed steps are summarized in the table containing Algorithm 3.1. Fig. 3.1 also shows the structure of the proposed detector (Algorithm 3.1).

Algorithm 3.1 The Proposed MMSE-DFE Conditional Optimization Detector

1: Order the columns of matrix $\mathbf{H}$ according to the V-BLAST ordering; denote this ordered matrix as $\mathbf{H}_O$.
2: Perform MMSE-DFE filtering by QRD on the augmented matrix of $\mathbf{H}_O$, i.e., $\tilde{\mathbf{H}}_O = \tilde{\mathbf{Q}}\mathbf{R}$. Obtain the forward and backward matrices, i.e., $\mathbf{Q}^H$ and $\mathbf{B} = \mathbf{R}$, respectively (first QRD).
3: $\mathbf{y}' = \mathbf{Q}^H\mathbf{y} = \mathbf{B}\mathbf{x} + \mathbf{w}$.
4: Perform the conditional optimization detection on $\mathbf{y}'$;

1. Consider all $N \times (M-1)$ possible submatrices of $\mathbf{B}$, i.e., $\mathbf{B}_{[i]} \in S_{\mathbf{B}}$, $i = 1, 2, ..., M$ and $S_{\mathbf{B}} = \{\mathbf{B}_{[1]}, \mathbf{B}_{[2]}, ..., \mathbf{B}_{[M]}\}$.
2. $\mathbf{b}_i$ is the i-th column of matrix $\mathbf{B}$.
3. $\mathbf{y}' = \mathbf{B}_{[i]}\mathbf{x}_{[i]} + \mathbf{b}_i x_i + \mathbf{w}$.
4. Perform QRD on $\mathbf{B}_{[i]}$, i.e., $\mathbf{B}_{[i]} = \mathbf{Q}^{[i]}\mathbf{R}^{[i]}$ (second QRD).
5. For all points of the constellation:
   (a) Fix $\hat{x}_i^{(j)} = a_j, \quad a_j \in \mathcal{X}; \;\; j = 1, 2, ..., Q = |\mathcal{X}|$.
   (b) $\mathbf{z} = {\mathbf{Q}^{[i]}}^H(\mathbf{y}' - \mathbf{b}_i\hat{x}_i^{(j)})$.
   (c) $\hat{\mathbf{x}}_{[i]}^{(j)}$ is detected using DFE on $\mathbf{z}$ and matrix $\mathbf{R}^{[i]}$ and is sliced to the original constellation points.
   (d) Calculate
   $d_k = \|\mathbf{y}' - \mathbf{B}_{[i]}\hat{\mathbf{x}}_{[i]}^{(j)} - \mathbf{b}_i\hat{x}_i^{(j)}\|^2$;
   $k = (i-1)|\mathcal{X}| + j$.

5: Find $d_{min} = \min_{k=1,..,M.Q} d_k$ and declare the corresponding $\hat{\mathbf{x}} = [\hat{\mathbf{x}}_{[i]}^{(j)}, \hat{x}_i^{(j)}]$ after reordering as detected symbol vector.

The above algorithm has a linear complexity in the constellation size and a complexity of $\mathcal{O}(M^4)$ in the system dimension M (the cost of M successive QRD of all M possible subdivisions of the backward filter even by a fast implementation procedure as described in [8, 37, 114] for example). We will discuss how this cost can be reduced to a cubic complexity in M.

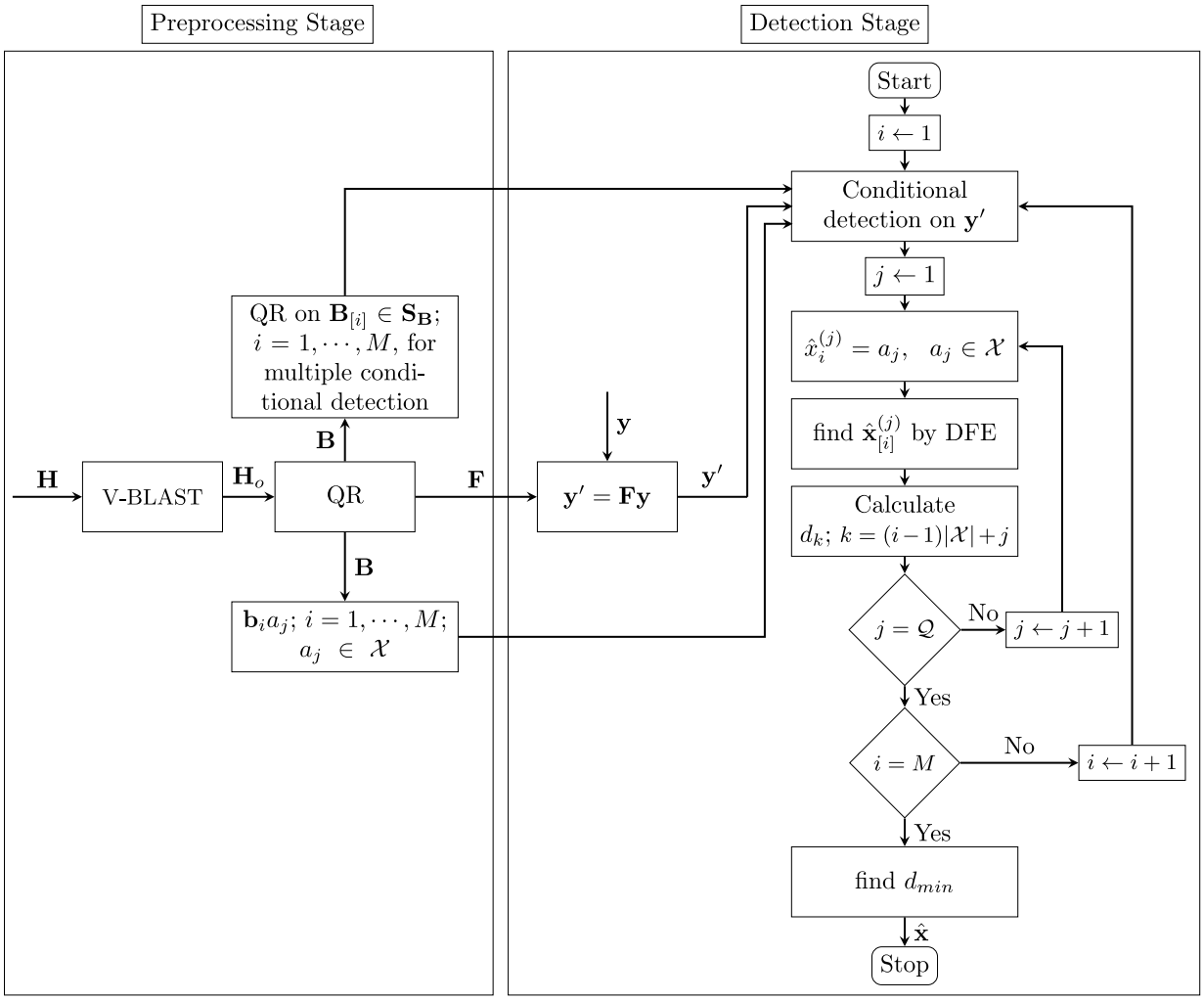


Figure 3.1: Structure of the Proposed Detector (Algorithm 3.1).

3.3.2 Algorithm 3.2

One-step V-BLAST Ordering on Submatrices

As some simulations results show in Section 3.3.6, when we increase the dimension of the system and also the constellation order, the error performance of the proposed detector degrades such that there is a performance gap to the ML BER at high SNRs. For example, for an 8×8 MIMO system and 64-QAM constellation the proposed detector exhibits about 1-dB BER gap to the ML error performance at BER of 10^{-5}. We offer a modification of the proposed MMSE-

DFE conditional optimization detector that exhibits an essentially ML BER performance with a computational complexity which is linear in Q and cubic in M. In Algorithm 3.1, the whole channel matrix is ordered whereas the remaining submatrix obtained by removing one column at a time is not necessarily in the optimal order which results in a performance degradation for larger systems with high order constellations. To improve the performance, one can perform a new greedy ordering on each submatrix; however, this increase the complexity to $\mathcal{O}(M^4)$ due to destroying the upper triangular structure of the submatrices. Alternatively, we suggest conducting the one-step V-BLAST ordering such that only the first index of the optimal ordering of the corresponding submatrix is found. Then, the remaining of the detection is the same as the proposed detector in Algorithm 3.1.

3.3.3 Computational Complexity

Computational Saving in V-BLAST Ordering

In V-BLAST greedy ordering, the signal to-interference-plus-noise ratio (SINR) for each layer of the detection is maximized. We consider the scheme introduced in [104] which is the MMSE extension of the V-BLAST ordering. With respect to the MMSE filtering in (3.3), the estimation error covariance matrix is

$$\alpha \left(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}}\right)^{-1} = \alpha \left(\mathbf{R}^H \mathbf{R}\right)^{-1}. \tag{3.6}$$

Assuming perfect interference cancellation, the estimation error is given by $\frac{\alpha}{|r_{kk}|^2}$. Hence, in each step, the optimal ordering chooses the column that maximizes $|r_{kk}|$.

For the proposed detector, we obtain the V-BLAST ordering for the whole matrix of $\mathbf{R}$ using the scheme in [104]: An efficient sub-optimal MMSE sorted QR algorithm with a computational complexity of $\mathcal{O}(M^3)$ is followed by a post-sorting algorithm with a complexity of $\mathcal{O}(M^3)$.

Hence, by applying the post-sorting algorithm, the scheme in [104] gives the QRD of the

optimally ordered augmented channel matrix with a complexity of $\mathcal{O}(M^3)$. Moreover, since $\mathbf{R}^{-1}$ is a byproduct of the scheme in [104], the matrix in (3.6) can easily be computed.

Computational Saving in the QRD

In the proposed detector, by using the scheme in [104], we first calculate the QRD of the augmented channel matrix, $\tilde{\mathbf{H}}$

$$\tilde{\mathbf{H}} = \tilde{\mathbf{Q}}\mathbf{R},$$

where $\mathbf{R}$ is an $M \times M$ upper triangular matrix:

$$\mathbf{B} = \mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & \dots & r_{1M} \\ 0 & r_{22} & r_{23} & \dots & r_{2M} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & r_{MM} \end{bmatrix}.$$

Then, we calculate the QRDs of all M submatrices of $\mathbf{R}$ with $M-1$ columns, where one column is removed at a time (corresponding to the transmitted symbol over which an exhaustive search is performed). Since $\mathbf{R}$ is upper triangular, the QRD of the $M-1$ first columns of $\mathbf{R}$, i.e., $\mathbf{B}_{[M]}$, is already upper triangular. However, other submatrices in the set need to be transformed into upper triangular forms. Because of the structure of the submatrices and the fact that successive submatrices have $M-2$ columns in common, there is no need to perform the QRDs from scratch. Givens rotations method is an appropriate candidate for conducting the QRDs because each rotation modifies only two rows of the operating matrix and also the order of rotations could be interchanged for parallelized decompositions [34]. For instance, consider

the following submatrix:

$$\mathbf{B}_{[1]} = \begin{bmatrix} r_{12} & r_{13} & \dots & r_{1M} \\ r_{22} & r_{23} & \dots & r_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & r_{MM} \end{bmatrix}.$$

In order to make r_{22} equal to zero we apply rotation $\mathbf{G}_1$ on $\mathbf{B}_{[1]}$, i.e., $\mathbf{G}_1\mathbf{B}_{[1]}$, where $\mathbf{G}_1$ defined as

$$\mathbf{G}_1 = \begin{bmatrix} \alpha & \beta & \dots & 0 \\ -\beta & \alpha & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}.$$

where $|\alpha|^2 + |\beta|^2 = 1$ and $-\beta r_{12} + \alpha r_{22} = 0$. Other entries are also made equal to zero by obtaining the corresponding rotation matrices. Finally, we will have an upper triangular matrix $\mathbf{R}^{[1]} = \mathbf{G}_{M-1} \dots \mathbf{G}_2\mathbf{G}_1\mathbf{B}_{[1]}$ and $\mathbf{Q}^{[1]^H} = \mathbf{G}_{M-1} \dots \mathbf{G}_2\mathbf{G}_1$ (Note that if we utilize Givens rotations technique instead of the traditional QRD schemes, $\mathbf{Q}^{[1]^H}$ is used in step 4.4 of Algorithm 3.1 in order to have an upper triangular matrix). $M-1$ rotations are required to make $\mathbf{B}_{[1]}$ upper triangular. Each submatrix requires up to $M-1$ rotations, equal to the number of non-zero entries. Each rotation changes only two rows of the submatrix and it requires up to $4(M-1)$ multiplications and $2(M-2)$ additions. Consequently, $\mathcal{O}(M^3)$ operations are required for all submatrices to be transformed into upper triangular forms. Moreover, since the first QRD needs $\mathcal{O}(M^3)$ operations, the total required operations associated with QRDs for the proposed detector is $\mathcal{O}(M^3)$.

Therefore, the overall computational complexity for QRDs and V-BLAST ordering, which are done in the pre-processing stage, is in the order of $\mathcal{O}(M^3)$. This cost can be negligible for quasi-static fading channels where the channel matrix is fixed over the whole block length.

One can further reduce the search complexity as follows. For each column removed from

the backward filter matrix $\mathbf{R}$, instead of performing a full search over all the constellation points for the corresponding symbol, one does a pruned search over the constellation $\mathcal{X}$ such that only $1 \leq Q \leq |\mathcal{X}|$ nearest constellation points are considered (via zig-zagging around the ZF-DFE solution or equivalently the Schnorr-Euchner enumeration procedure [22]). While such a pruning strategy largely degrades the performance of FSD type algorithms [7], its effect on our algorithm can be negligible since performing conditional optimization detection over all the columns compensates for this degradation as will be demonstrated in Section 3.3.6 (e.g., it is shown that $Q = 4$ is sufficient to achieve ML performance in a 16-QAM - 8×8 MIMO system, see Fig. 3.8).

Remark 1 *Another possible saving in implementation complexity can be obtained by mimicking the SD algorithm [22] as one searches over the different candidates as follows. Assume one does the exhaustive search on column k, then for each value of $x_k \in \mathcal{X}$, the algorithm calculates the DFE solution for the remaining symbols $x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_M$ using the corresponding upper triangular submatrix. One can define a threshold (or a radius) that is initialized by the first Euclidean distance calculated, and then updated each time the algorithm finds a closer point. To save complexity, one then stops the DFE algorithm whenever the associated metric goes above the threshold. While this trick can save some complexity, it has the disadvantages of making the complexity a random variable and not tailoring itself naturally to parallel and soft-output implementations. Note that our original algorithm as described above has a fixed complexity (linear in Q and cubic in M), and naturally lends itself to parallel and soft-output implementations as described in the sequel.*

In the sequel, we discuss how the one-step V-BLAST ordering for the different submatrices, i.e., Algorithm 3.2, can be conducted in an efficient manner with $\mathcal{O}(M^2)$ operations in total in order to avoid unnecessary computations. The error covariance matrix in (3.6) is a byproduct of the scheme in [104] and needs $\mathcal{O}(M^3)$ operations to be calculated. For the one-step V-BLAST orderings of the submatrices, we need to look at the diagonal elements of the error covariance

matrix corresponding to these submatrices. In each case, the diagonal element with the smallest value is selected. In [84], the relationship between the inverse of a matrix and a submatrix is discussed. Given the inverse matrix in (3.6), it can be shown that for each submatrix we need $\mathcal{O}(M)$ operations to calculate the $M-1$ diagonal elements of the corresponding inverse matrix, and accordingly the one-step V-BLAST ordering for that submatrix. Therefore, the one-step V-BLAST orderings of all submatrices need $\mathcal{O}(M^2)$ operations. By taking the V-BLAST ordering on the whole matrix into account, the total computational complexity is in the order of $\mathcal{O}(M^3)$. Moreover, one can show that the discussion in Subsection 3.3.3, regarding the computational complexity of the QRDs, is still valid for this one-step V-BLAST ordering MMSE-DFE conditional optimization detector. Because in the worst case we have $\mathcal{O}(M)$ elements to get zeroed, hence, the computational complexity of the QRD for all submatrices still requires $\mathcal{O}(M^3)$ operations. Therefore, the overall computational complexity for QRDs and the orderings is still in the order of $\mathcal{O}(M^3)$. In Section 3.3.6, we refer to this proposed scheme as Algorithm 3.2. The computational complexity of the proposed algorithms is compared with that of the FSD scheme in Table 3.1 in terms of number of FLOPs. We assume six FLOPs per complex multiplication and two FLOPs per complex addition and the matrix operations from [34]. The following calculations are considered in the table:

- Multiplication of $N \times M$ and $M \times K$ complex matrices requires $(8M-2)NK$ FLOPs.
- The V-BLAST ordering (or similarly the FSD ordering) along with the first QRD needs $12M^3+16M^2N+\frac{83}{3}M^2+10MN+38M-14N-60$ FLOPs [104].
- According to Subsection 3.3.3, all second QRDs for Algorithm 3.1 require $\frac{56}{3}M^3-45M^2+\frac{91}{3}M-4$ FLOPs.
- Given the inverse matrix in (3.6), which is a byproduct of the algorithm in [104], all one-step V-BLAST orderings in Algorithm 3.2 require $14(M-1)^2$ FLOPs [84].

Table 3.1: Computational Complexity In Terms of FLOPs.

Detector	Operation		Number of FLOPs	Example: 8 × 8 MIMO	
				64-QAM	256-QAM
FSD	Preprocessing:	FSD ordering of $\mathbf{H}$+QRD of $\tilde{\mathbf{H}}_o$ [104]	$12M^3 + 16M^2N + \frac{83}{3}M^2 + 10MN + 38M - 14N - 60$	475,630 ($p=2$)	7,356,900 ($p=2$)
		$\mathbf{B}_2\mathbf{x}_2; \mathbf{x}_2 \in \mathcal{X}^p$	$(8Mp - 2M) \times \lvert\mathcal{X}\rvert^p$		
	Detection:	$\mathbf{y}' = \mathbf{Q}^H\mathbf{y}$	$8MN - 2M$	1,647,088 ($p=2$)	26,345,968 ($p=2$)
		SIC	$(4(M-p)(M-p+1)+2p)\lvert\mathcal{X}\rvert^p$		
		Euclidean Distance Test	$(4(M-p)(M-p+1)+8M-2) \times \lvert\mathcal{X}\rvert^p$		
MMSE-DFE Conditional Optimization Detector (Algorithm 3.1)	Preprocessing:	V-BLAST ordering of $\mathbf{H}$+QRD of $\tilde{\mathbf{H}}_o$ [104]	$12M^3 + 16M^2N + \frac{83}{3}M^2 + 10MN + 38M - 14N - 60$	48,371 ($Q=64$)	122,100 ($Q=256$)
		All second QRDs: $\mathbf{R}^{[i]}, \mathbf{Q}^{[i]}$	$\frac{56}{3}M^3 - 45M^2 + \frac{91}{3}M - 4$		
		$\mathbf{b}_i a_j$	$6M \times M.\mathcal{Q}$		
	Detection:	$\mathbf{y}' = \mathbf{Q}^H\mathbf{y}$	$8MN - 2M$	493,040 ($Q=64$)	1,970,672 ($Q=256$)
		$\mathbf{Q}^{[i]^T}(\mathbf{y}' - \mathbf{b}_i a_j)$	$(8M^2 - 8M + 2)M.\mathcal{Q}$		
		SIC	$(4(M-1)M+2)M.\mathcal{Q}$		
		Euclidean Distance Test	$(4(M-1)M+8M-2) \times M.\mathcal{Q}$		
MMSE-DFE Conditional Optimization Detector (Algorithm 3.2)	Preprocessing:	V-BLAST ordering of $\mathbf{H}$+QRD of $\tilde{\mathbf{H}}_o$ [104]	$12M^3 + 16M^2N + \frac{83}{3}M^2 + 10MN + 38M - 14N - 60$	60,453 ($Q=16$)	66,597 ($Q=32$)
		All one-step V-BLAST orderings	$14(M-1)^2$		
		All second QRDs: $\mathbf{R}^{[i]}, \mathbf{Q}^{[i]}$	$112M^3 - 368M^2 + 376M - 120$		
		$\mathbf{b}_i a_j$	$6M \times M.\mathcal{Q}$		
	Detection:	$\mathbf{y}' = \mathbf{Q}^H\mathbf{y}$	$8MN - 2M$	123,632 ($Q=16$)	246,768 ($Q=32$)
		$\mathbf{Q}^{[i]^T}(\mathbf{y}' - \mathbf{b}_i a_j)$	$(8M^2 - 8M + 2)M.\mathcal{Q}$		
		SIC	$(4(M-1)M+2)M.\mathcal{Q}$		
		Euclidean Distance Test	$(4(M-1)M+8M-2) \times M.\mathcal{Q}$		

- All second QRDs in Algorithm 3.2 need up to $112M^3 - 368M^2 + 376M - 120$ FLOPs.

It is worth mentioning that at the first glance, Algorithm 3.2 might seem more complex than Algorithm 3.1. However, the error performance improvement of Algorithm 3.2 enables us to reduce the search size over the constellation points. This computational complexity saving is more than the complexity increase due to one-step V-BLAST orderings on the submatrices. Therefore, the overall computational complexity of Algorithm 3.2 is less than that of Algorithm 3.1 (See Table 3.1).

3.3.4 Diversity Order Analysis

In this section, first, we extend the FSD to under-determined MIMO systems and discuss the optimal FSD ordering for under-determined MIMO systems that we later use in our simulations is Section 3.3.6. Then, we discuss the diversity order of the proposed MMSE-DFE conditional optimization detectors.

Fixed-complexity Sphere Decoder Ordering for Under-determined MIMO Systems:

As previously reported in the literature, e.g., [19, 79, 96], the diversity optimality of the conventional FSD ordering cannot be insured for under-determined MIMO systems. When $M > N$, the matrix $\mathbf{Q} = \mathbf{H}^H\mathbf{H}$ is a pseudo-Wishart matrix[2] and is not a full-rank matrix as $M-N$ eigenvalues are zero [68]. For under-determined MIMO systems, if the Gram channel matrix is used for the FSD ordering, it does not satisfy the condition of [52, Lemma 2] because it is a positive semi-definite matrix[3]. As a result, the corresponding FSD ordering is not optimal.

However, if we use the regularized (or augmented) channel matrix in (3.3) and accordingly the Gram matrix $\tilde{\mathbf{Q}} = \tilde{\mathbf{H}}^H\tilde{\mathbf{H}} = \mathbf{H}^H\mathbf{H} + \alpha\,\mathbf{I}_M$ and the conventional ordering in [7], a diversity optimal ordering is achieved which is referred to hereafter as the RFSD ordering. We will deal with this fact in the following theorem.

Theorem 2 *For a UD-MIMO system (when $M > N$) if the RFSD ordering is utilized, the diversity order of the SE stage is given by*

$$d_{SE}^{UD} = (M-N)(p_{eq}+1) + (p_{eq}+1)^2 \tag{3.7}$$

where $p_{eq} = p - (M-N)$ and $p \geq (M-N)$.

Proof: See Appendix C. □

Now, we analyze the diversity order of the proposed MMSE-DFE conditional optimization detectors. To this end, we will use results from antenna selection [54] and the diversity order of the FSD [52] or RFSD scheme. One proves the following:

[2]The pseudo-Wishart distribution occurs when the Hermitian matrix generated from a complex Gaussian ensemble is not full-rank [69].

[3]Although it has been mentioned as positive semi-definite in [52], by following the footsteps of the proof of the lemma therein one can easily understand that it must be positive definite instead because otherwise the inverse of the matrix does not exist.

Proposition 1 *In a MIMO system with N receive and M transmit antennas, under V-BLAST transmission scheme, the diversity order of the proposed MMSE-DFE conditional optimization detector is*

$$d = min(N, 2(N - M) + 4), \tag{3.8}$$

where,

$$d \triangleq -\lim_{\rho \to \infty} \frac{\log P_e}{\log snr}, \tag{3.9}$$

in which P_e is the error probability of the detector and $snr = \frac{ME_s}{N_0}$ is the average SNR at a given receive antenna.

Proof: See Appendix D. □

Proposition 1 asserts that the proposed detector achieves the same diversity order as the FSD scheme with the FSD ordering and $p = 1$ full search level. However, the search over all possible submatrices makes the proposed detector outperform the FSD scheme as shown in the simulation results section. Although the diversity order of the proposed detector does not necessarily achieve the full ML diversity for some MIMO systems, it is improved over that of ZF and the MMSE which exhibit a diversity order of $N - M + 1$ under V-BLAST transmission scheme [40, 72]. Moreover, the coding gain of the proposed detector compared to the FSD scheme with $p = 1$ allows for close to ML performance at practical values of SNR. It is shown in Section 3.3.6 that the BER performance is close to that of the ML while the FSD ordering needs a larger number of full searches to achieve the optimal error performance.

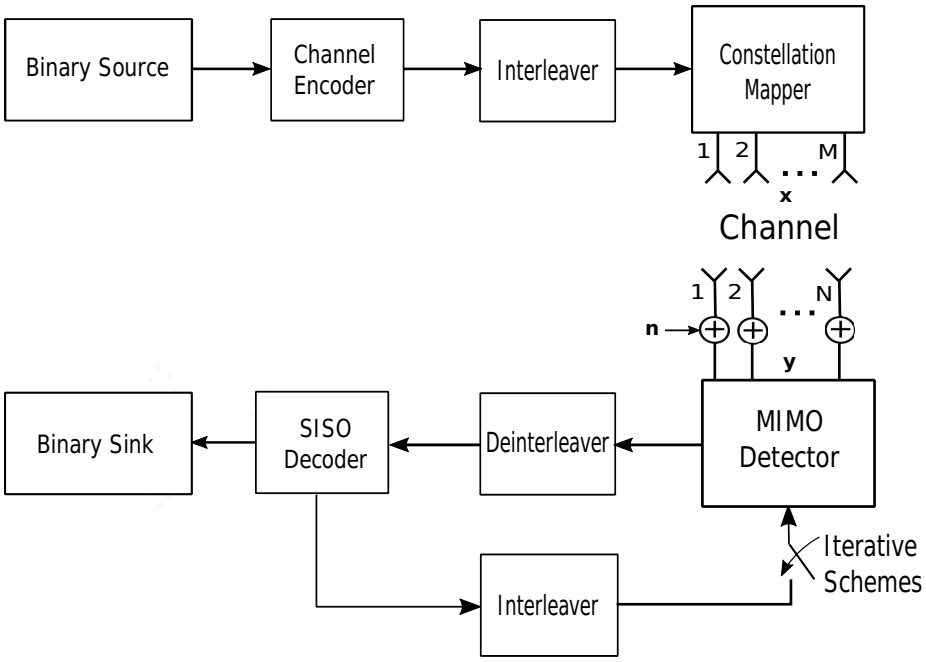


Figure 3.2: Soft-Output MIMO Detector and Decoder.

3.3.5 Soft-Output MMSE-DFE Conditional Optimization Detector

In this section, we discuss how the proposed detector can be utilized for soft-output detection without imposing more computational complexity. The error performance of our detector will be shown by simulations to outperform that of the LSD soft-output detector. First, we review the system model for soft-output MIMO detection. Then, we provide some details on the soft-output version of the proposed MMSE-DFE conditional optimization detector.

Soft-Output MIMO Detection

The concatenation of a SISO MIMO detector and a SISO channel decoder can significantly improve the performance of MIMO systems. In such systems, in order to implement the iterative joint detection and decoding, it is required that the soft information be exchanged between the MIMO detector and SISO decoder (See Fig.3.2 for more details). Using *a-posteriori* probabilities (APPs), soft-output MIMO detection can be implemented by an exhaustive or limited list

size search. However, the optimal exhaustive search is not practical even for moderate MIMO dimensions due to the high computational complexity. LSD schemes have been proposed in the literature [41] to construct a suboptimal list. The list of candidates are constructed by a modified SD scheme by selecting the lattice points with the smallest distance to the received point. However, such an LSD scheme is unstable in the list size when the received signal is outside of the finite lattice constellation. This issue is highly probable, especially at low SNRs or for large dimensional MIMO systems. In this work, we compare our results to the shifted LSD proposed [11] in which the constructed list is centered around the ML point (or any other detected point [75]) instead of the received signal to combat the instability issue.

At the transmitter $M.\log_2|\mathcal{X}|$ bits are sent over the channel per each channel use after channel encoding and interleaving. The APP is usually expressed in terms of log-likelihood ratio (LLR). The sign of the LLR is used for making a decision if the bit is one or zero. Moreover, the magnitude of LLR indicates the reliability of the decision such that the LLR values close to zero indicate unreliable decisions. The modulator associates each group of $\log_2|\mathcal{X}|$ data bits to a constellation point using Gray or natural mapping techniques [11, 41]. The LLR value of the j-th bit is expressed as

$$L(b_j|\mathbf{y}) = \ln\frac{P[b_j=+1|\mathbf{y}]}{P[b_j=-1|\mathbf{y}]};\quad j=1,2,...,M.\log_2|\mathcal{X}|, \tag{3.10}$$

where $b_j \in \{-1,+1\}$ is the representation of the logical zero and one mapped to -1 and $+1$, respectively. Then, if we assume an additive zero mean white circularly symmetric complex Gaussian noise, by using Bayes' rule, the soft-output values $L(b_j|\mathbf{y})$ can be written as [41]

$$L(b_j|\mathbf{y}) = \ln\frac{\sum_{\mathbf{b}\in\mathbb{B}_j^{+1}}\exp(-\frac{1}{N_0}\parallel\mathbf{y}-\mathbf{Hx}\parallel^2+\frac{1}{2}\mathbf{b}_{[j]}^T.\mathbf{L}_{A,[j]})}{\sum_{\mathbf{b}\in\mathbb{B}_j^{-1}}\exp(-\frac{1}{N_0}\parallel\mathbf{y}-\mathbf{Hx}\parallel^2+\frac{1}{2}\mathbf{b}_{[j]}^T.\mathbf{L}_{A,[j]})}, \tag{3.11}$$

where $\mathbb{B}_j^{+1}$ and $\mathbb{B}_j^{-1}$ are the set of $2^{M.\log_2|\mathcal{X}|-1}$ bits such that $b_j=+1$ and $b_j=-1$, respectively.

Moreover, $\mathbf{b}_{[j]}$ denotes the subvector of $\mathbf{b}$ by omitting its j-th element, and $\mathbf{L}_{A,[j]}$ is also the vector of *a-priori* LLR values of $\mathbf{b}$ after omitting $\mathbf{b}_j$. For the first iteration in iterative schemes (or for non-iterative scenario) $\mathbf{L}_A$ is set to zero because all bits are assumed to be equally probable. For the exhaustive APP, the set size is $2^{M.\log_2|\mathcal{X}|-1}$ while for the LSD scheme the set size is equal to the size of the list. Further simplifications can be achieved by using max-log approximation [82] as follows

$$\begin{aligned} L(b_j|\mathbf{y}) \approx & \max_{\mathbf{b}\in\mathbb{B}_j^{+1}}\{-\frac{1}{N_0}\parallel \mathbf{y}-\mathbf{Hx}\parallel^2+\frac{1}{2}\mathbf{b}_{[j]}^T.\mathbf{L}_{A,[j]}\} \\ & -\max_{\mathbf{b}\in\mathbb{B}_j^{-1}}\{-\frac{1}{N_0}\parallel \mathbf{y}-\mathbf{Hx}\parallel^2+\frac{1}{2}\mathbf{b}_{[j]}^T.\mathbf{L}_{A,[j]}\}. \end{aligned} \tag{3.12}$$

Then, after de-interleaving, the APPs are fed as *a-priori* to the SISO channel decoder.

In order to construct the list, we set the radius based on the statistical formula obtained in [11] and list the N_p closest point to the center point. If the number of points is less than N_p, we increase the radius until the desired list size is achieved.

Soft-output MMSE-DFE Conditional Optimization Detector

Our proposed algorithm lends itself naturally to soft-output implementation because its hard-output version selects the symbol vector with the best metric among a list of $M.|\mathcal{X}|$ candidates (for each selected column of $\mathbf{R}$ one performs an exhaustive search on the associated symbol followed by a DFE on the remaining ones). Therefore, in the soft-output version, one can consider these $M.|\mathcal{X}|$ candidates as the list for calculating the LLR values from (3.12). The resulted list is more "balanced" than that obtained from LSD in the sense that one makes sure that for each bit b_j, there is at least two codewords in the list, one that has $b_j=-1$ and the other has $b_j=+1$ whereas for the LSD, one has to resort to ad-hoc approaches when there is no codeword in the list that contains the particular bit value at $+1$ or -1 [41]. The performance of soft-output

MMSE-DFE conditional optimization detector is evaluated in Section 3.3.6 where it is shown that the soft-output MMSE-DFE conditional optimization detector with a smaller list size has a better error performance than that of the soft-output FSD with a larger list size.

3.3.6 Simulations Results: Part I

In this section, we compare our algorithm to the state of the art detectors for several $N \times M$ MIMO systems with different input QAM constellations, with and without outer error correcting codes. For the simulation results, we evaluate the error performance of our proposed detectors for different configurations of small to large-scale MIMO systems. We consider MIMO systems such as square 4×4, 8×8, 16×16, and 32×32, a non-square over-determined 8×6, and an under-determined 5×6 along with V-BLAST transmission scheme in order to investigate the BER performance results of the proposed MMSE-DFE conditional optimization detector. The SD, MMSE-DFE, and FSD schemes are reported for comparison. Moreover, the MMSE-DFE filtering is performed for all simulations.

In Fig. 3.3, the classical MMSE-DFE has a diversity order of one. For this system, the proposed detector achieves the full diversity and near ML performance for both 4-QAM and 16-QAM constellations. Note that p_{min} is the minimum number of full search levels that the FSD scheme needs to achieve the optimal diversity order.

Fig. 3.4 shows that for an 8×8 MIMO system, the proposed detector achieves near-ML error performance and a diversity of $min(8, 4) = 4$ for both 4-QAM and 16-QAM constellations. An improvement of about 2-dB over the FSD scheme with $p = 1$ is obtained for both constellations at the BER of 10^{-5}. In Fig. 3.5, for a 16×16 system, the diversity of the proposed detector is $min(16, 4) = 4$ and the improvement over the FSD with $p = 1$ is about 4.5-dB. For this system, although the FSD with $p = 3$ achieves the optimal diversity, it has a performance gap to the ML performance for practical values of SNRs (we note that even a larger number of full search levels,

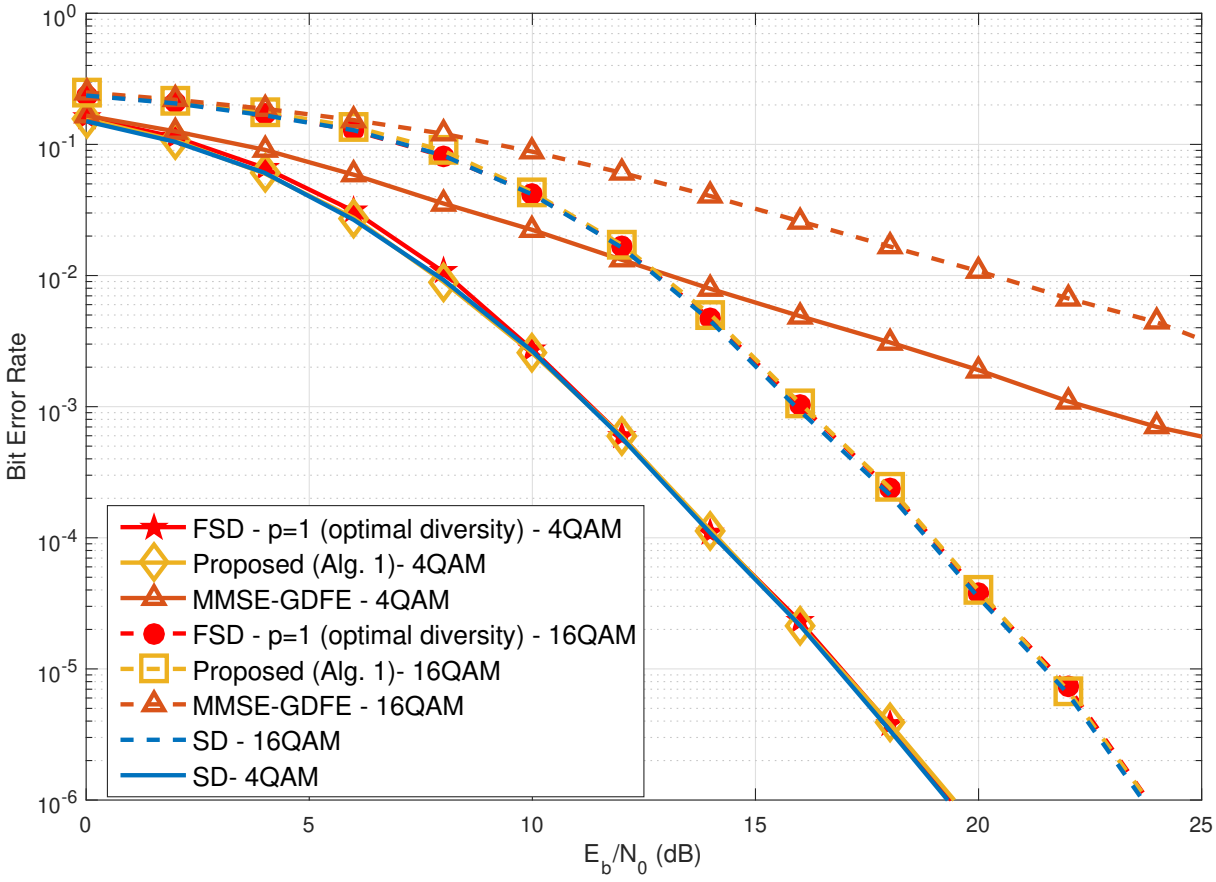


Figure 3.3: Uncoded BER performance of the proposed MMSE-DFE conditional optimization detector in Algorithm 3.1 for a 4×4 MIMO system for both 4-QAM and 16-QAM; $p_{min} = 1$.

$p = 4$ cannot achieve the optimal performance at the reported SNRs). The proposed detector with a detection complexity proportional to $16 \times 4 = 64$ outperforms the FSD with $p = 4$ and the detection complexity proportional to $4^4 = 256$. We also investigate the error performance for non-square MIMO systems. According to Fig. 3.6, the diversity order of the proposed scheme for the non-square 8×6 MIMO system is $min(8, 8) = 8$.

For the under-determined 5×6 MIMO system in Fig. 3.7, the diversity result of conventional FSD ordering is not valid, hence, we use the result of the RFSD ordering. As this figure shows, the diversity order of the proposed detector is $min(5, 2) = 2$ for $p = 1$ and the improvement over the RFSD with $p = 1$ is around 2-dB at the BER of 10^{-5} and similar to the RFSD with $p = 2$, it

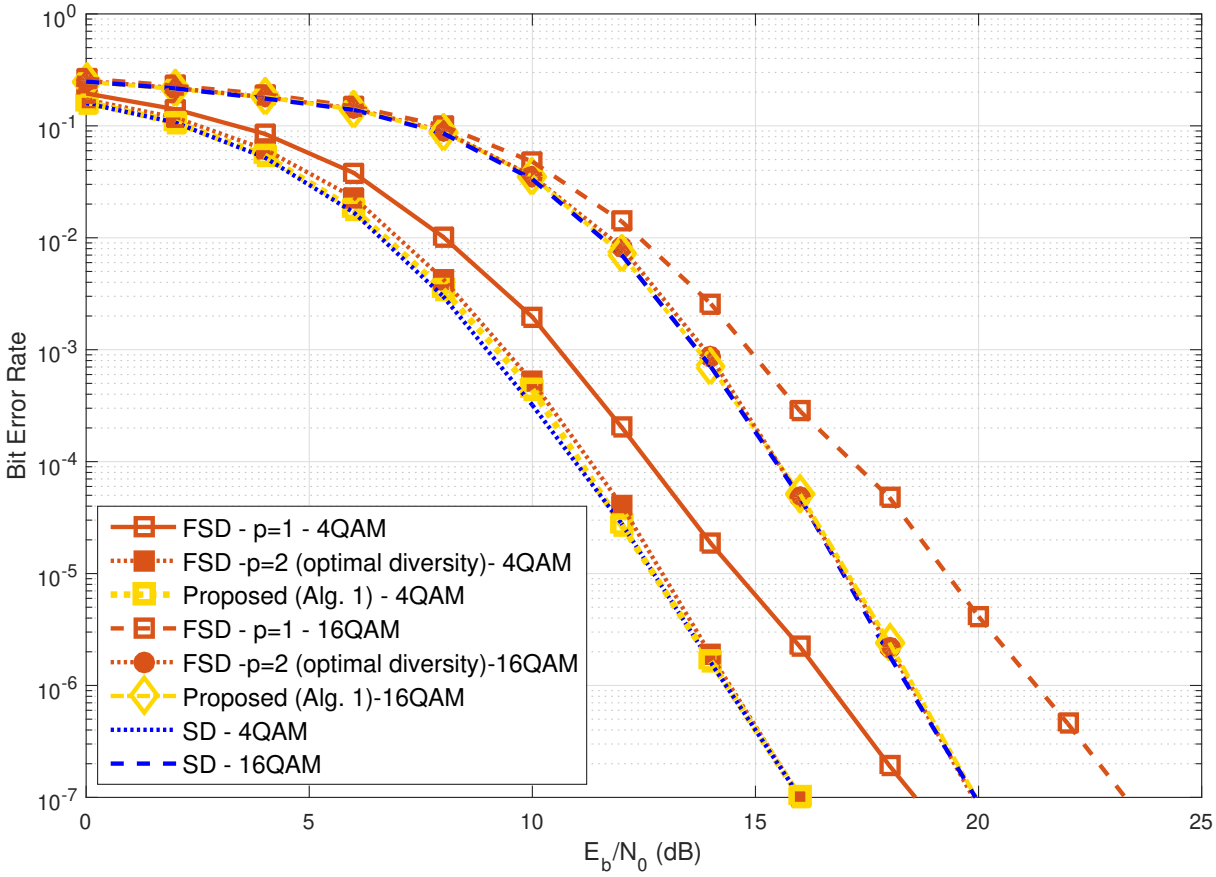


Figure 3.4: Uncoded BER performance of the proposed MMSE-DFE conditional optimization detector in Algorithm 3.1 for an 8×8 MIMO system for both 4-QAM and 16-QAM; $p_{min} = 2$.

achieves the ML BER.

Fig. 3.8 indicates the BER performance of the proposed detector for an 8×8 MIMO system with 16-QAM constellation when different search sizes, Q, over the constellation are considered as discussed in Subsection 3.3.3. According to this figure, $Q = 4$ suffices to achieve a near-ML error performance. The detection complexity for the proposed detector with $Q = 4$ is proportional to $8 \times 4 = 32$ while that of the FSD with $p = 2$ is proportional to $16^2 = 256$.

Fig. 3.9 and 3.10 indicate the BER performances for an 8×8 MIMO system, respectively with 64-QAM and 256-QAM constellations. As these figures show, for the proposed detector in Algorithm 3.1, there is a gap to the ML error performance while when the modification,

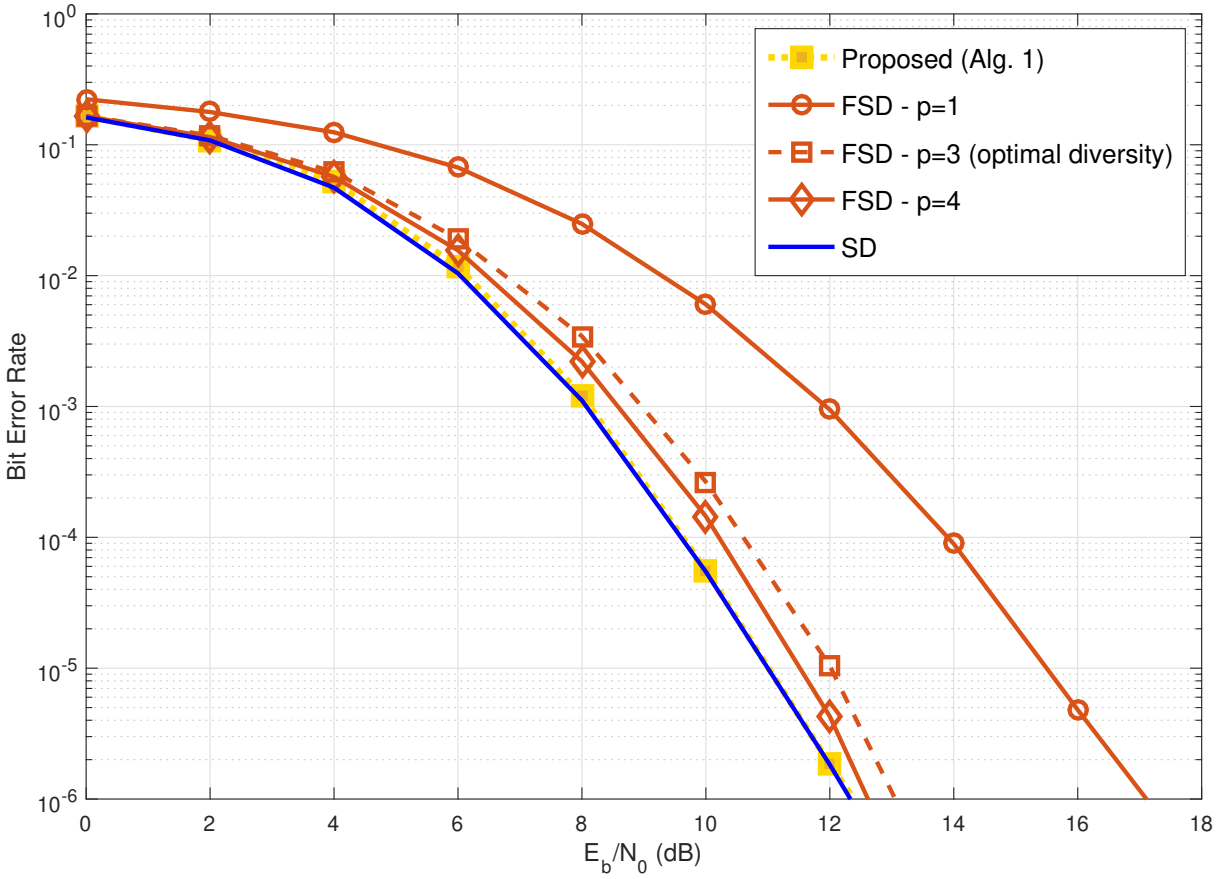


Figure 3.5: Uncoded BER performance of the proposed MMSE-DFE conditional optimization detector in Algorithm 3.1 for a 16×16 MIMO system for 4-QAM constellation; $p_{min} = 3$

discussed in Subsection 3.3.2 and referred as Algorithm 3.2, is applied it achieves a near-ML BER even with smaller search sizes, Q, over the constellation. According to Fig. 3.9, for 64-QAM constellation, Algorithm 3.2 with $Q = 16$ suffices to achieve a near-ML error performance. The detection complexity for this detector with $Q = 16$ is proportional to $8 \times 16 = 128$ while that of the FSD with $p = 2$ is proportional to $64^2 = 4,096$. For 256-QAM constellation in Fig. 3.10, Algorithm 3.2 with $Q = 32$ suffices to achieve a near-ML BER. The detection complexity of this detector with $Q = 32$ is proportional to $8 \times 32 = 256$ while the FSD with $p = 2$ has a complexity about $256^2 = 65,536$.

In order to evaluate the proposed soft-output MMSE-DFE conditional optimization detector,

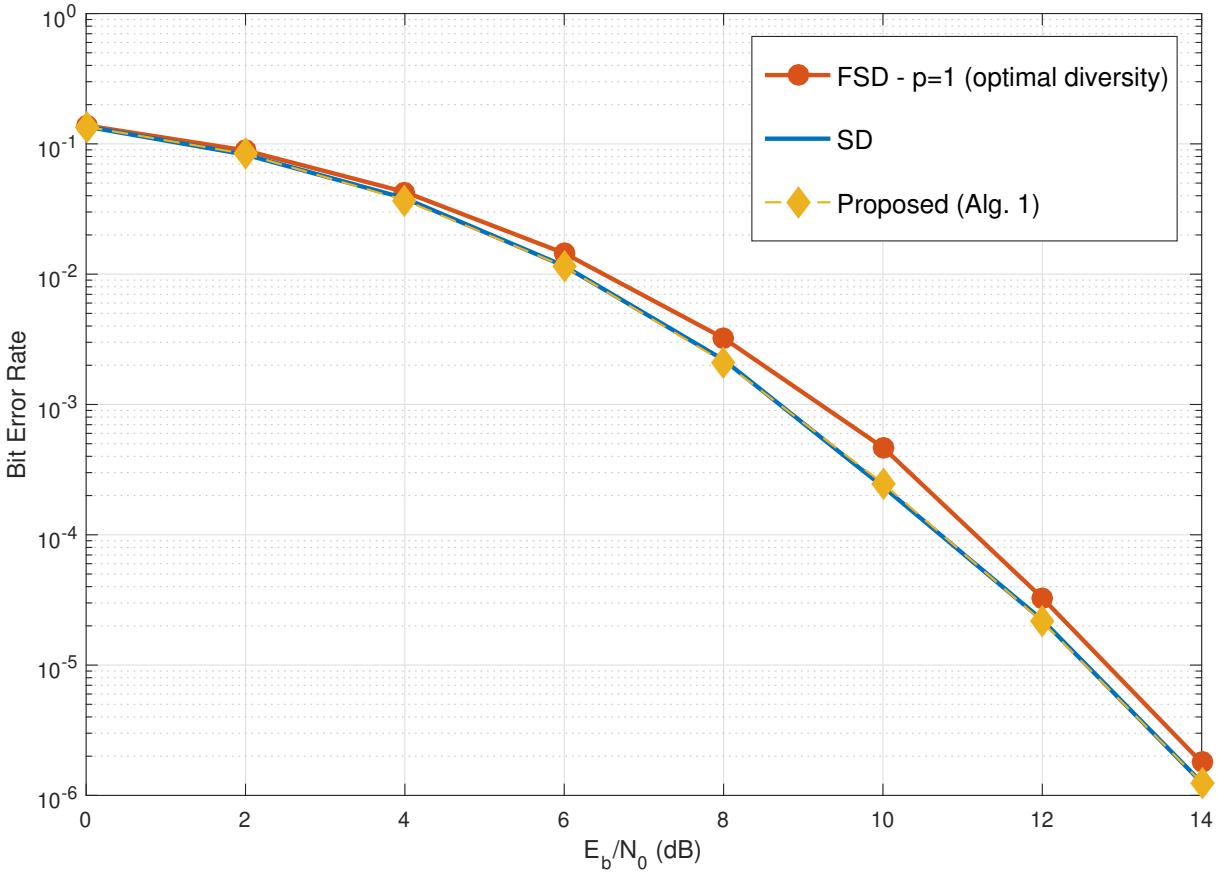


Figure 3.6: Uncoded BER performance of the proposed MMSE-DFE conditional optimization detector in Algorithm 3.1 for an 8×6 MIMO system for 4-QAM constellation; $p_{min} = 1$.

we consider a 4×4 MIMO system, where one uses the standard rate-$1/2$ convolutional code, with generator polynomial $(5, 7)$ with the 16-QAM constellation and Gray mapping. Fig. 3.11 indicates the error performances for different schemes. For this system, the list size of the proposed soft-output detector is equal to $4 \times 16 = 64$ whereas the LSD scheme in [11] has a list size of 512 and the soft-output FSD with $p = 1$ and $p = 2$ respectively have list sizes of 16 and 256. As this figure shows the proposed scheme achieves the error performance of the conventional LSD scheme while a considerable computational complexity is saved by the natural extension of the proposed scheme to the soft-output detector. Moreover, the proposed detector with a list size of 64 outperforms the soft-output FSD with list size of 256.

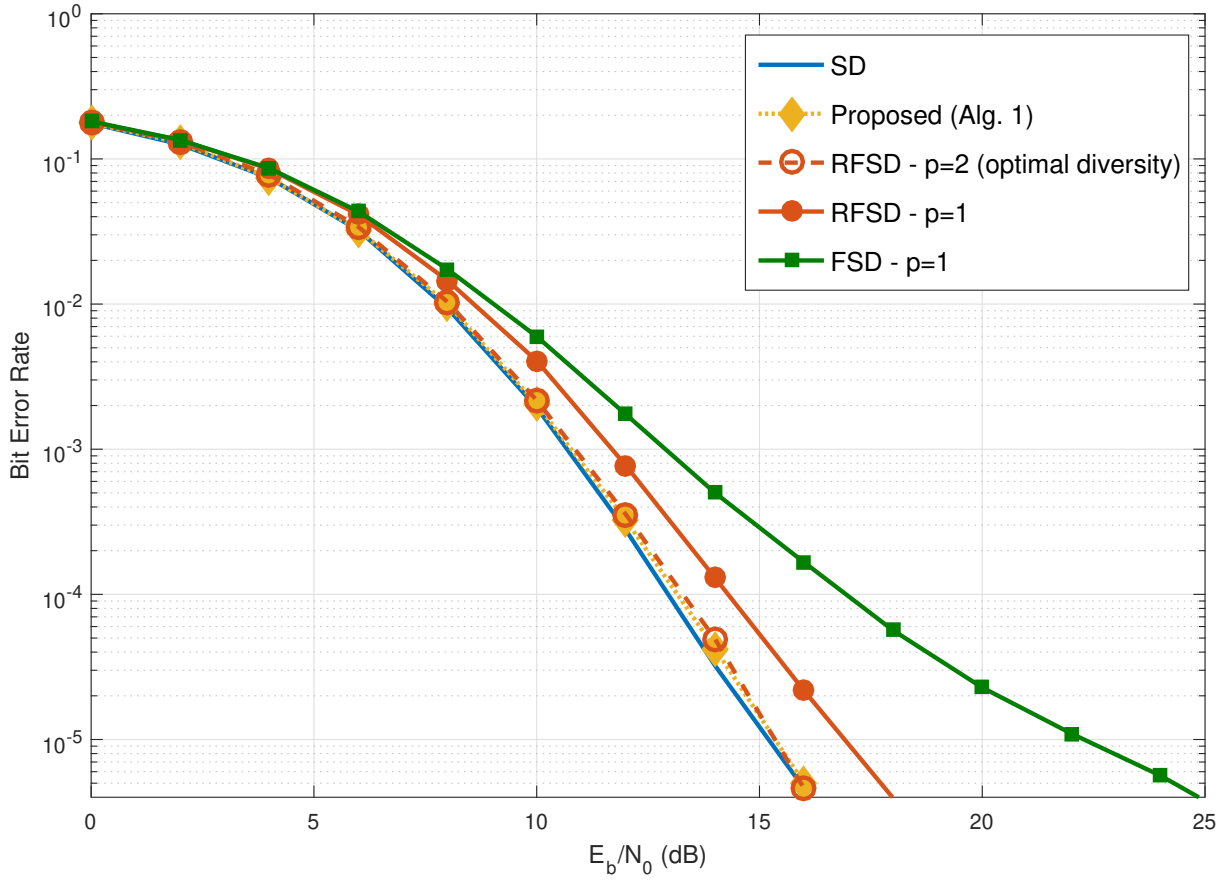


Figure 3.7: Uncoded BER performance of the proposed MMSE-DFE conditional optimization detector in Algorithm 3.1 for a 5×6 (under-determined) MIMO system for 4-QAM; $p_{min} = 2$.

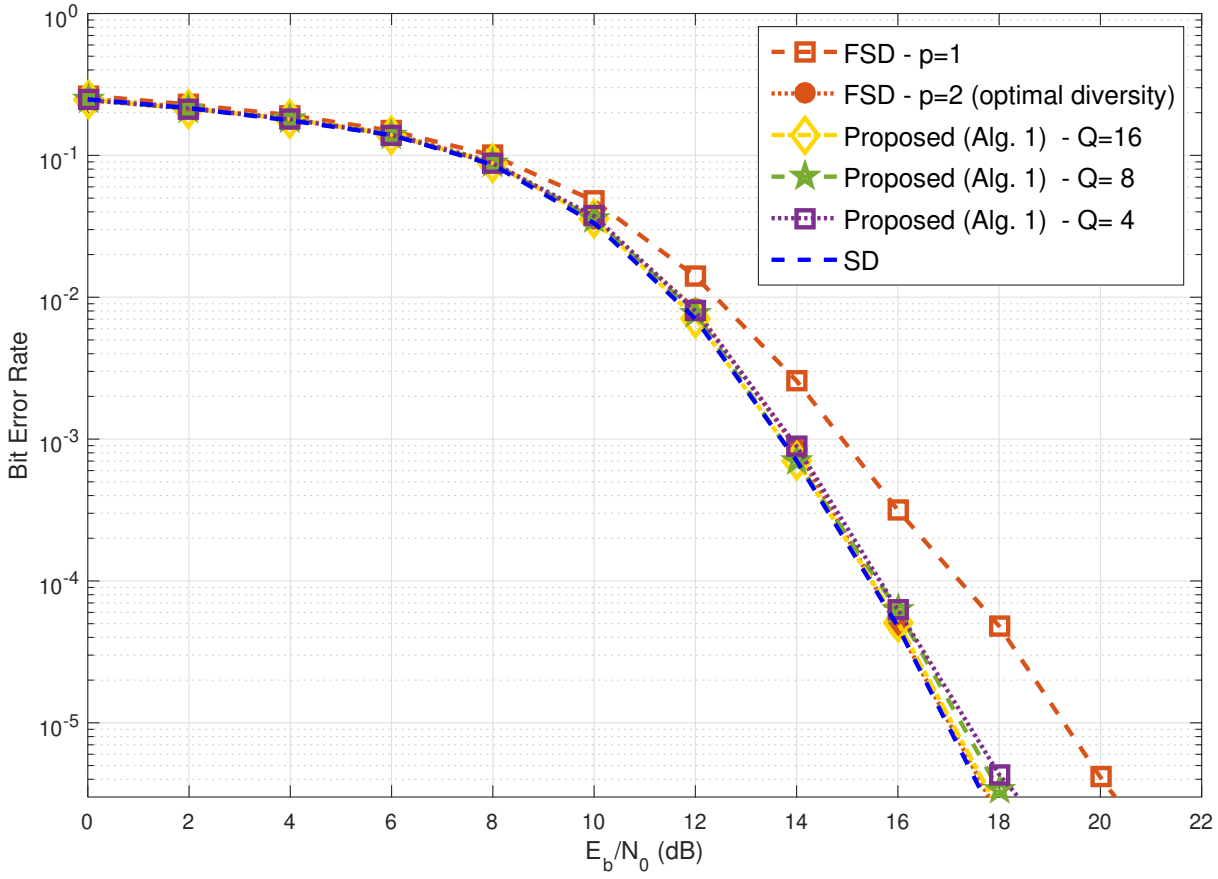


Figure 3.8: Uncoded BER performance of the proposed MMSE-DFE conditional optimization detector in Algorithm 3.1 with different search sizes Q over the constellation for an 8×8 MIMO system for 16-QAM; $p_{min} = 2$.

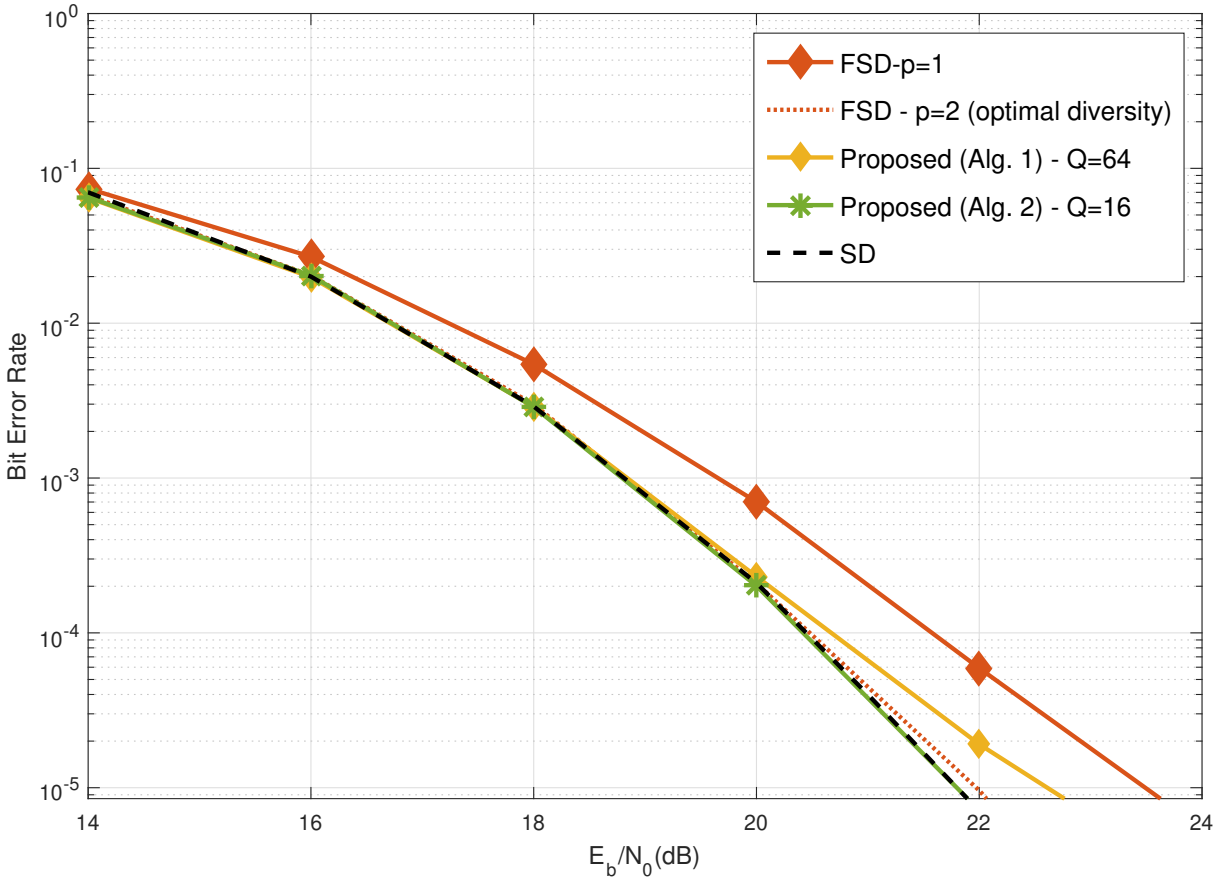


Figure 3.9: Uncoded BER performance of the proposed MMSE-DFE conditional optimization detector in Algorithm 3.1 and its modification in Algorithm 3.2 for an 8×8 MIMO system for 64-QAM; $p_{min} = 2$.

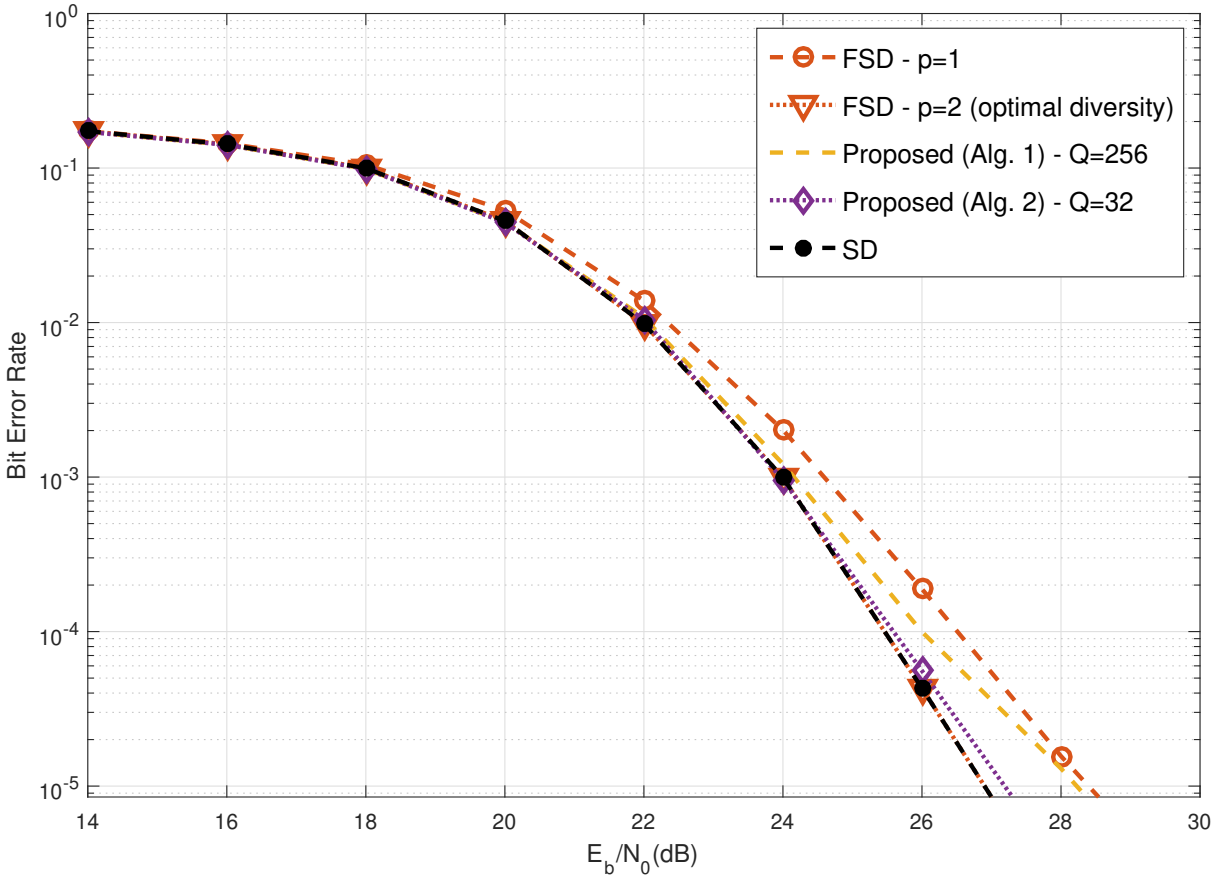


Figure 3.10: Uncoded BER performance of the proposed MMSE-DFE conditional optimization detector in Algorithm 3.1 and its modification in Algorithm 3.2 for an 8×8 MIMO system for 256-QAM; $p_{min} = 2$.

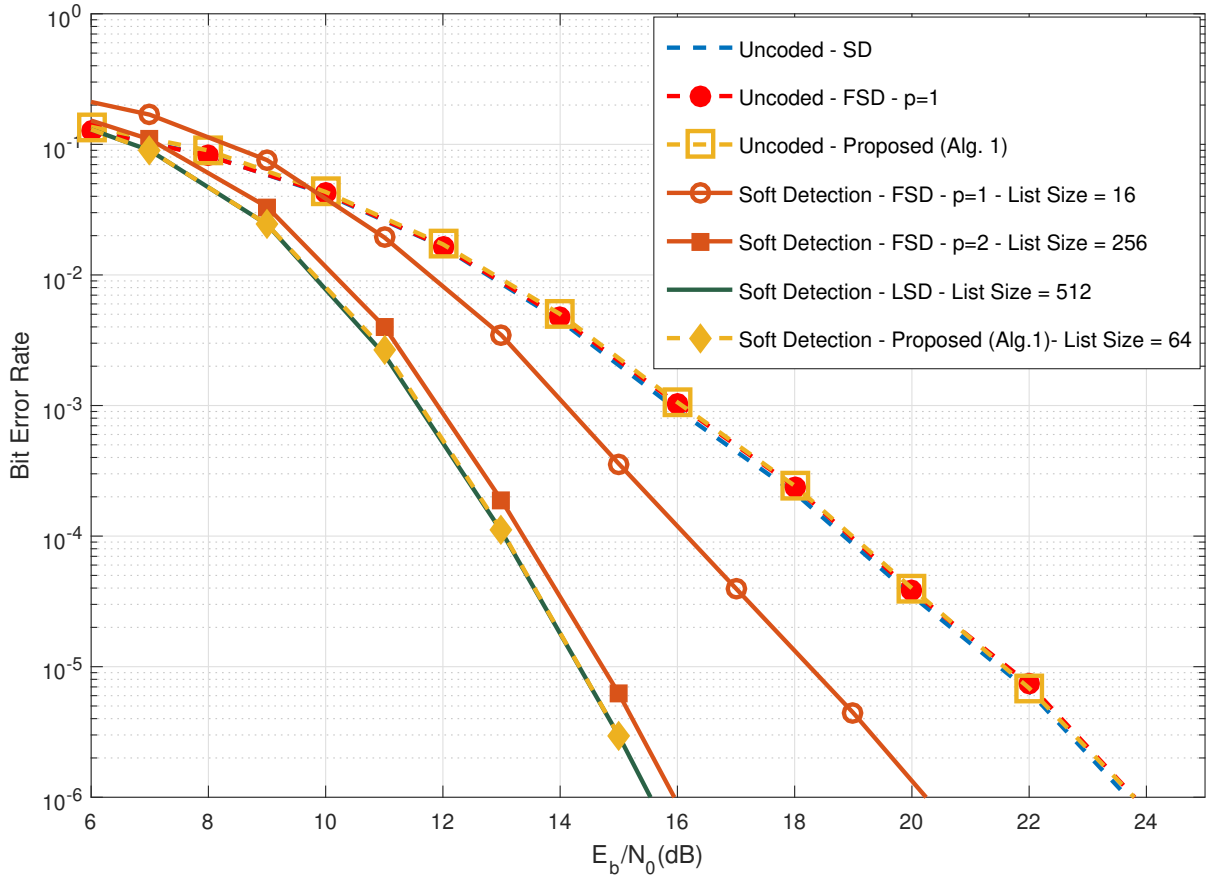


Figure 3.11: BER performance of the proposed soft/hard-output MMSE-DFE conditional optimization detector in Algorithm 3.1 for a 4×4 MIMO system for 16-QAM.

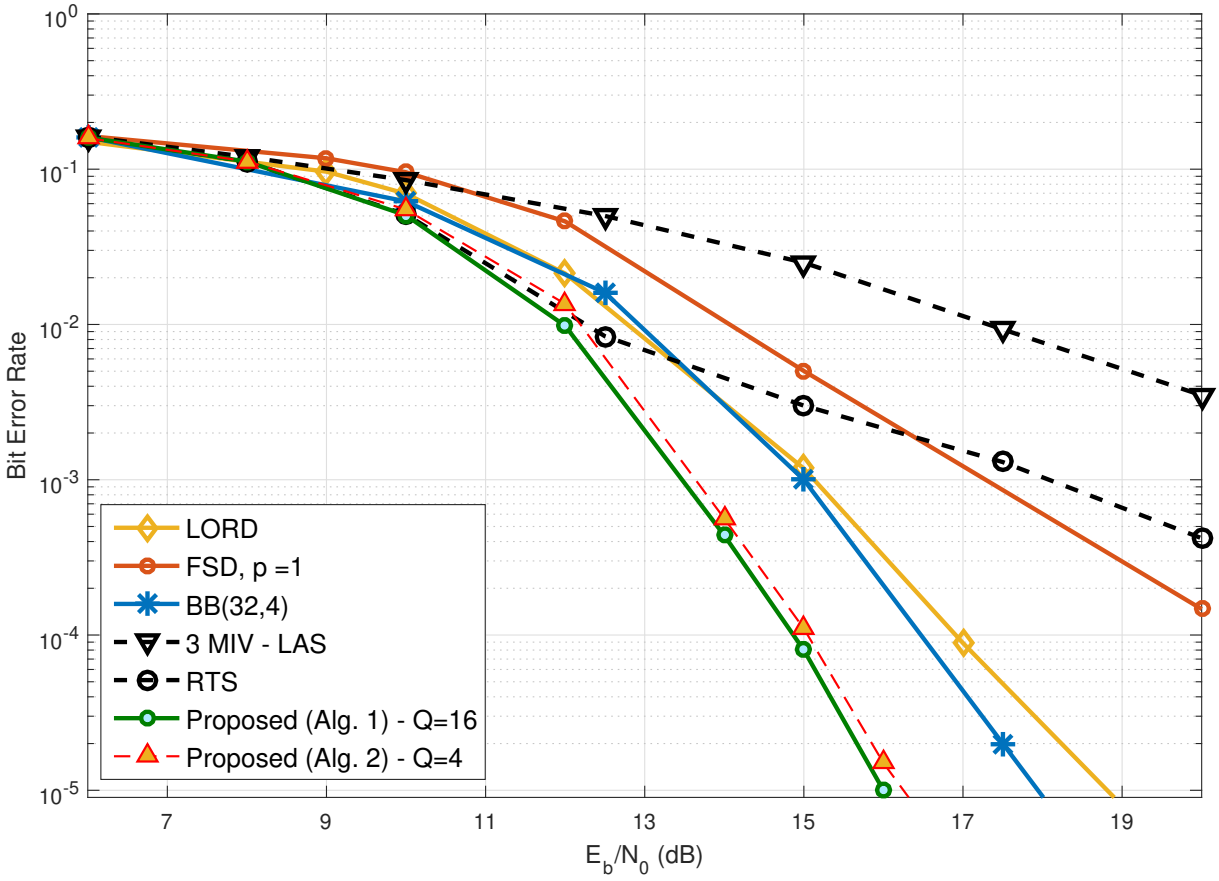


Figure 3.12: Uncoded BER performance of the proposed MMSE-DFE conditional optimization detector in Algorithm 3.1 and its modification in Algorithm 3.2 for a 32×32 MIMO system for 16-QAM.

Fig. 3.12 compares the BER performances of our proposed detectors, for a 32×32 large-scale MIMO system with 16-QAM constellation, with those of several state-of-the-art detectors for large-scale MIMO systems, such as the branch and bound (BB) quadratic programming (QP)-based detector, i.e. BB(L_d,M_w) in [27], reactive tabu search (RTS) [93], multiple initial vectors local ascent search (MIV-LAS) [62] which is a LAS algorithm with three initial inputs [27], and LORD scheme [91]. In [27], some algorithms are developed from the conventional QP detector in which the ML problem is rewritten as a QP optimization problem. They use some techniques, such as BB tree search along with width and depth reductions in the tree search, to further improve their proposed detectors. The computational complexity of the QP-based

detector is $\mathcal{O}(nN_\nu M^3)$ when $M = N$, where n is the number of iterations for the interior-point algorithm. Moreover, N_ν is a function of the controlled BB algorithm in which depth and width reductions are considered. According to [27], $N_\nu \approx L_d M_w$ at low SNRs and $N_\nu \leq L_d M_w$ at high SNRs where L_d and M_w are depth and width sizes of the search tree, respectively. They refer to this scheme as BB(L_d,M_w). On the other hand, the computational complexity of our proposed detectors is in the order of $\mathcal{O}(Q.M^3)$. In the QP-based scheme, the value of N_ν depends on SNR; hence; the scheme exhibits a variable complexity over a range of SNR while our proposed detectors are fixed-complexity schemes for all range of SNR. Moreover, as Fig. 3.12 shows, Algorithm 3.2 with $Q = 4$ has almost the same error performance as Algorithm 3.1 with $Q = 16$. They both outperform the BB(32, 4) scheme (with $n = 4$ iterations) reported in [27] by about 2-dB. The computational complexity of the QP-based scheme is proportional to $N_\nu n M^3 = N_\nu \times 4 \times 32^3$ and that of our proposed detector with $Q = 4$ is proportional to $QM^3 = 4 \times 32^3$. Therefore, with proportionally less computational complexity , our proposed detector outperforms the QP-based detector in [27].

The subspace detection scheme in [86] is a reduced-complexity version of LORD; hence, we compare our proposed detector with LORD which has a better error performance than that of the scheme in [86]. Like LORD, our proposed detector is a QRD-based detector that considers all columns for exhaustive search over the constellation points to detect the corresponding symbol. However, they differ in the some aspects: In LORD scheme, a cyclically shifting strategy is used for column back substitution where the symbols corresponding to the problematic columns may be detected at the early steps of the detection which will result in a performance degradation which is evidenced in Fig. 3.12 . However, in our proposed scheme, the remaining columns after selecting a column for conditional detection are in a relative ordering which is expected to improve the error performance. In our proposed detection scheme, the computational complexity of all operations including the QRD and the V-BLAST orderings are in the order $\mathcal{O}(M^3)$ while that of LORD is in the order of $\mathcal{O}(M^4)$ [86]. According to the figure, our proposed detectors

outperforms LORD scheme by about 2-dB with a smaller computational complexity than that of LORD. Moreover, the proposed idea in Algorithm 3.2 allow us to reduce the search size over the constellation which in turn reduces the computational complexity of the detection stage with a small increase in the computational complexity of the preprocessing stage. Therefore, our proposed detector, especially Algorithm 3.2, with a significantly reduced computational complexity outperforms all the aforementioned schemes for this large-scale MIMO system.

3.4 Part II: Partially Reduced Conditional Optimization Detector

In this section, we propose a partially reduced conditional optimization detector motivated by the work in [64] in order to improve the diversity order of conditional optimization detector with $p = 1$. In the proposed conditional based detector we use $p = 1$ as well. However, the diversity order degradation due to restricting the number of full search levels to one is compensated by performing the PLR on the non-exhaustive search levels. The proposed detector is detailed in the following section.

3.4.1 Algorithm 3.3

We continue with the model in (3.1). The detailed steps of the proposed detector are summarized in the table containing Algorithm 3.3. The proposed detector consists of the following steps:

- *Stage 1- Preprocessing:*
  - *Channel Ordering:* Motivated by the work in [7, 52], the channel matrix is ordered such that the columns corresponding to the signal with maximum post-processing noise is put at the end of the ordered channel matrix. The remaining $M - 1$ columns

Algorithm 3.3 The Proposed PLR Conditional Optimization Detector

1: Order the columns of matrix $\mathbf{H}$ according to the FSD ordering for $p = 1$; denote this ordered matrix as $\mathbf{H}_O$.
2: Perform MMSE-DFE filtering by QR decomposition on the augmented matrix of $\mathbf{H}_O$, i.e., $\tilde{\mathbf{H}}_O = \tilde{\mathbf{Q}}\mathbf{R}$. Obtain the forward and backward matrices, i.e., $\mathbf{Q}^H$ and $\mathbf{R}$, respectively (first QRD).
3: For a given block size K for the PLR, consider $\mathbf{R} = [\mathbf{R}_{1_{N\times(M-K-1)}}, \mathbf{R}_{2_{N\times K}}, \mathbf{r}_M]$ where $\mathbf{R}_1 = [\mathbf{r}_1, \ldots, \mathbf{r}_{M-K-1}]$ and $\mathbf{R}_2 = [\mathbf{r}_{M-K}, \ldots, \mathbf{r}_{M-1}]$.
4: Perform the PLR algorithm on $\mathbf{R}_2$ using the orthogonal complement $[\mathbf{R}_1, \mathbf{r}_M]$. The reduced submatrix is $\mathbf{B_2} = \mathbf{R}_2\mathbf{U}_2$. We have $\mathbf{R}_{\text{PLR}} = [\mathbf{R}_{1_{N\times(M-K-1)}}, \mathbf{B}_{2_{N\times K}}, \mathbf{r}_M]$.
5: $\mathbf{y}' = \mathbf{Q}^H\mathbf{y} = \mathbf{R}\mathbf{x} + \mathbf{w} = \mathbf{R}_{\text{PLR}}\mathbf{x}_{\text{PLR}} + \mathbf{w}$; $\mathbf{x}_{\text{PLR}} = [\mathbf{x}_1, \mathbf{x}'_2, x_M]$.
6: Perform the conditional optimization detection on $\mathbf{y}'$;

1. Consider all $M - K$ possible orderings of $\mathbf{S}_O$.
2. For each i-th ordering in $\mathbf{S}_O$, perform QRD on the corresponding ordered matrix of $\mathbf{R}_{\text{PLR}}$. Denote the equivalent factorized matrices as $\{\mathbf{Q}^{(i)}, \mathbf{R}^{(i)}\}$ where $i = 1, 2, \ldots, |\mathbf{S}_O| = M - K$ (second QRD).
3. $\tilde{\mathbf{y}}^{(i)} = {\mathbf{Q}^{(i)}}^H\mathbf{y}'$.
4. For all points of the constellation:
   (a) Fix $\hat{x}_M^{(i,j)} = a_j, \quad a_j \in \mathcal{X}; \ j = 1, 2, ..., Q = |\mathcal{X}|$.
   (b) By Using DFE, $[\hat{\mathbf{x}}_{[M],1}^{(i,j)}, \hat{\mathbf{x}}_{[M],2}^{(i,j)}]$ is detected and is sliced to the original constellation points where $\hat{\mathbf{x}}_{[M],2}^{(i,j)} = \mathbf{U}_2\lceil\hat{\mathbf{x}}_{[M],2}^{(i,j)\prime}\rfloor$.
   (c) Calculate $d_k = \|\mathbf{y}' - \mathbf{R}^{(i)}\hat{\mathbf{x}}^{(i,j)}\|^2$ where $k = (i-1)|\mathcal{X}| + j$ and $\hat{\mathbf{x}}^{(i,j)} = [\hat{\mathbf{x}}_{[M],1}^{(i,j)}, \hat{\mathbf{x}}_{[M],2}^{(i,j)}, \hat{x}_M^{(i,j)}]$.

7: Find $d_{min} = \min_{k=1,..,(M-K).Q} d_k$ and declare the corresponding $\hat{\mathbf{x}}^{(i,j)}$ after reordering as detected symbol vector.

are ordered according to the V-BLAST ordering. This ordering is known as the FSD ordering. Without loss of generality, we assume that the corresponding ordering is $[1, 2, ..., M]$. Hence, one may rewrite the ordered channel matrix as

$$\mathbf{H}_O = [\mathbf{h}_1, \ldots, \mathbf{h}_M]. \tag{3.13}$$

– *MMSE-DFE Filtering*: The QRD is applied to the augmented matrix of the ordered channel in (3.13). Hence, we have

$$\mathbf{y}' = \mathbf{R}\mathbf{x} + \mathbf{w}, \tag{3.14}$$

where

$$\mathbf{R} = [\mathbf{R}_{1_{N\times(M-K-1)}}, \mathbf{R}_{2_{N\times K}}, \mathbf{r}_M], \tag{3.15}$$

where $\mathbf{R}_1 = [\mathbf{r}_1, \ldots, \mathbf{r}_{M-K-1}]$ and $\mathbf{R}_2 = [\mathbf{r}_{M-K}, \ldots, \mathbf{r}_{M-1}]$.

– *PLR*: Submatrix $\mathbf{R}_2$ with K columns is partially reduced by projecting its columns to the orthogonal complement, i.e. $[\mathbf{R}_1, \mathbf{r}_M]$ with $M-K$ columns and performing PLR algorithm explained in Subsection 3.2. Hence, we have $\mathbf{B}_2 = \mathbf{R}_2\mathbf{U}_2$ and

$$\mathbf{R}_{\text{PLR}} = [\mathbf{R}_{1_{N\times(M-K-1)}}, \mathbf{B}_{2_{N\times K}}, \mathbf{r}_M]. \tag{3.16}$$

– *Column Selection Strategy*: We search over all non-reduced columns, i.e. $[\mathbf{R}_1, \mathbf{r}_M]$, for conditional detection. There exist $M-K$ possibilities. When the conditional detection is performed on a column other than $\mathbf{r}_M$, we keep $\mathbf{r}_M$ at the beginning of the matrix because this column corresponds to the signal with the maximum post-processing noise. We denote the set of all possible orderings as $\mathbf{S}_O$. Note that the ordering of columns of $\mathbf{B}_2$ are fixed for all possible conditional detections. For example, for a 5×5 channel matrix with $K=2$ the set of orderings is

$$\mathbf{S}_O = \Big\{\{1,2,3,4,5\}, \{5,1,3,4,2\}, \{5,2,3,4,1\}\Big\}.$$

– *QRD for Submatrices*: After the PLR, the reduced matrix $\mathbf{B}_2$ may not be upper triangular. By using some matrix operations we can obtain the QRD. Moreover, for conditional detection on a column other than $\mathbf{r}_M$, we need to make some elements zero in order to obtain an upper triangular matrix for SIC scheme. We denote the corresponding equivalent QRD matrices for each ordering in $\mathbf{S}_O$ as $\{\mathbf{Q}^{(i)}, \mathbf{R}^{(i)}\}$ where $i = 1, 2, \ldots, |\mathbf{S}_O| = M - K$. One may write

$$\mathbf{R}^{(i)} = [\mathbf{R}^{(i)}_{1_{N\times(M-K-1)}}, \mathbf{R}^{(i)}_{2_{N\times K}}, \mathbf{r}^{(i)}_M]. \tag{3.17}$$

- *Stage 2 - Detection:*

  – *Conditional Detection:* For all possible orderings of $\mathbf{S}_O$, we perform the conditional detection. Hence, we have

$$\begin{aligned}\tilde{\mathbf{y}}^{(i)} =& {\mathbf{Q}^{(i)}}^H \mathbf{y}' \\ =& \mathbf{R}^{(i)}_1 \mathbf{x}^{(i)}_{[M],1} + \mathbf{R}^{(i)}_2 \mathbf{x}^{(i)\prime}_{[M],2} + \mathbf{r}^{(i)}_M x^{(i)}_M + \mathbf{z}^{(i)},\end{aligned} \tag{3.18}$$

  where, $\mathbf{z}^{(i)} = {\mathbf{Q}^{(i)}}^H \mathbf{w}$ and $i = 1, 2, \ldots, M - K$.

  An exhaustive search over the constellation for finding subset $x^{(i)}_M$ is conducted. Then, the DFE detection is performed to find the subset $[\hat{\mathbf{x}}^{(i)}_{[M],1}, \hat{\mathbf{x}}^{(i)}_{[M],2}]$ where $\hat{\mathbf{x}}^{(i)}_{[M],2}\mathbf{U}_2\lceil\hat{\mathbf{x}}^{(i)\prime}_{[M],2}\rfloor$ in which $\lceil\cdot\rfloor$ is the function that maps the real and imaginary parts of its argument to the nearest integers. In each step, the detected signal is sliced back to the original constellation. If we denote the candidate list for the detected symbol vector after reordering by $\mathcal{L}$, we have:

$$\begin{aligned}\mathcal{L} &= \{\bar{\mathbf{x}}^{(i,j)}\}; \\ i &= 1, \ldots, (M-K); \\ j &= 1, \ldots, |\mathcal{X}|.\end{aligned} \tag{3.19}$$

Hence, the list size is $(M-K)|\mathcal{X}|$.

- *Stage 3- Post-processing:*
  - *Minimum Euclidean Distance Evaluation:*

    Among all possible candidates for $\hat{\mathbf{x}}$ in (3.19), we choose the one that gives the best performance, i.e., the minimum Euclidean distance to the received signal. Hence, we have

$$\hat{\mathbf{x}} = \underset{\mathbf{x}\in\mathcal{L}}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{H}\mathbf{x}\|^2. \tag{3.20}$$

It is worth mentioning that by considering the candidate list in (3.19), the proposed detector can be naturally extended for soft-output detection of MIMO systems.

3.4.2 Computational Complexity

In the proposed detector, the computational complexity of the preprocessing and detection stages depends on the block size in the PLR scheme, i.e. K. The number of tested columns for the proposed detector is $M-K$, hence, the number of visited constellation points is $(M-K)|\mathcal{X}|$. Therefore, a larger value for K decreases the computational complexity of the detection stage. This reduction is transferred to the preprocessing stage where the reduction algorithm and the QRD are performed. In practice, the LLL reduction cost is cubic in the dimension for small-to-moderate dimensions. However, a larger K implies a more complex reduction algorithm which can be computationally prohibitive for large block sizes in the PLR.

Moreover, if K increases the number of column permutations for conditional detection decreases which consequently reduces the computational complexity of the QRD updates for submatrices. After the PLR, the QRD on the reduced submatrix is required. In a similar manner to what was discussed in Subsection 3.3.3, Givens rotations method is an appropriate technique

to make the matrix upper triangular. This technique can also be used for updating the QRD for submatrices with different channel orderings. It can be shown that the overall computational complexity of all QRD operations is in the order of $\mathcal{O}(M^3)$ when $M = N$. Moreover, the algorithm proposed in [104] can be modified to give the FSD ordering along with the QRD in the MMSE-DFE filtering with $\mathcal{O}(M^3)$ operations. As a result, for small values of K, the overall computational complexity of the preprocessing stage is $\mathcal{O}(M^3)$.

3.4.3 Simulation Results: Part II

In this section, we provide some computer simulations in order to investigate the performance of the proposed partially reduced conditional detector. In the provided simulation results,

- MMSE-DFE filtering is applied for all detection schemes.
- The number of exhaustive levels is one, i.e. $p = 1$, for all conditional detectors.

Fig. 3.13 shows the BER performance for a 32×32 MIMO system with 16-QAM constellation. The error performance results for different block sizes in the PLR have been provided. The proposed detector with $K = 5$ outperforms the FSD and LORD by about 6-dB and 1.5-dB, respectively. By comparing the BER result of the proposed detector with that of the FSD scheme, it is clear that the diversity order is improved. It is worth mentioning that the FSD scheme with p exhaustive search levels gives a diversity order of $(p + 1)^2$ for square MIMO systems which is equal to 4 in our simulations. Note that the channel puncturing based scheme in [86] is a reduced-complexity version of LORD scheme and its error performance has a gap to that of LORD. Hence, we compare our proposed detector with LORD.

Fig. 3.14 shows the BER performance for a 64×64 MIMO system with 16-QAM constellation. The same behaviour is observed for the proposed detector. The proposed detector with $K = 5$ outperforms the FSD and LORD by about 2.5-dB and 1.5-dB, respectively.

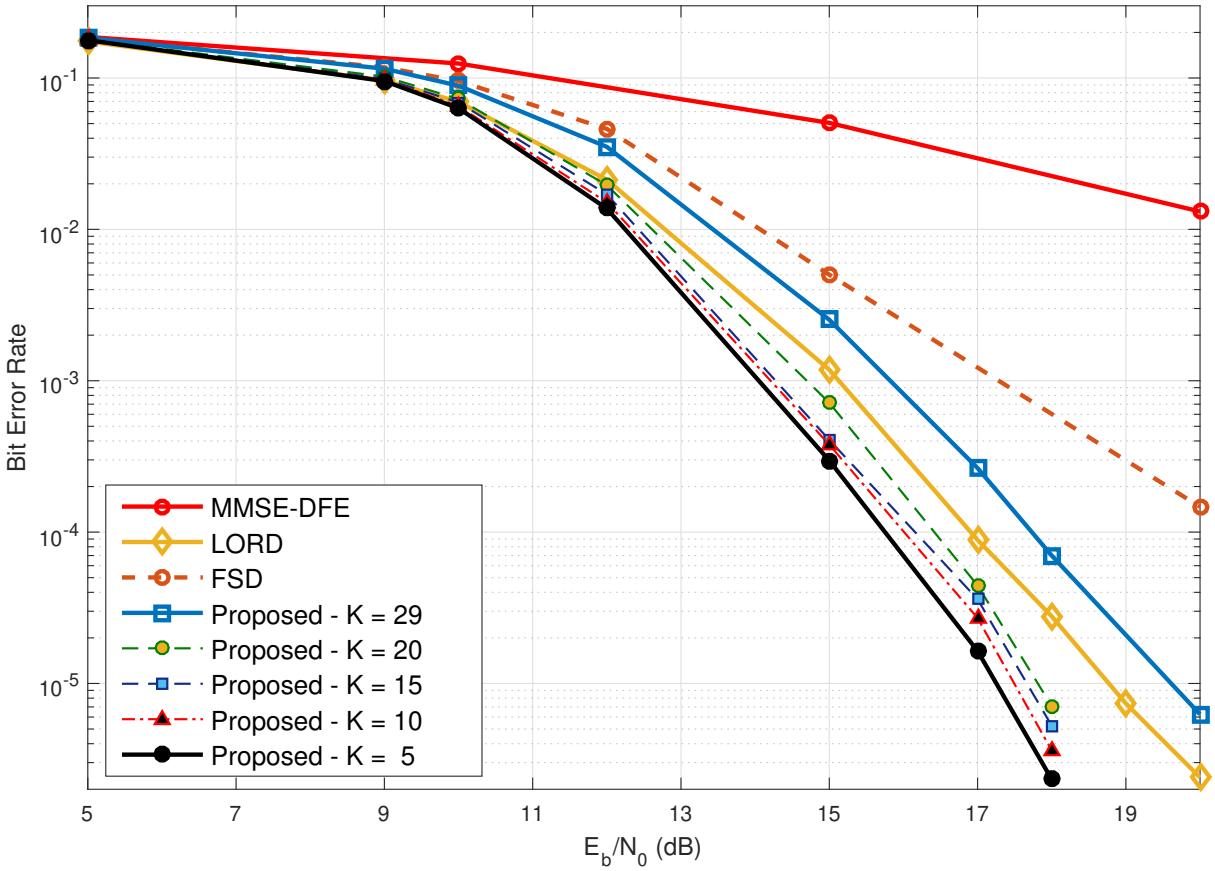


Figure 3.13: BER performance of the proposed PLR conditional optimization detector in Algorithm 3.3 for a 32×32 MIMO system for 16-QAM.

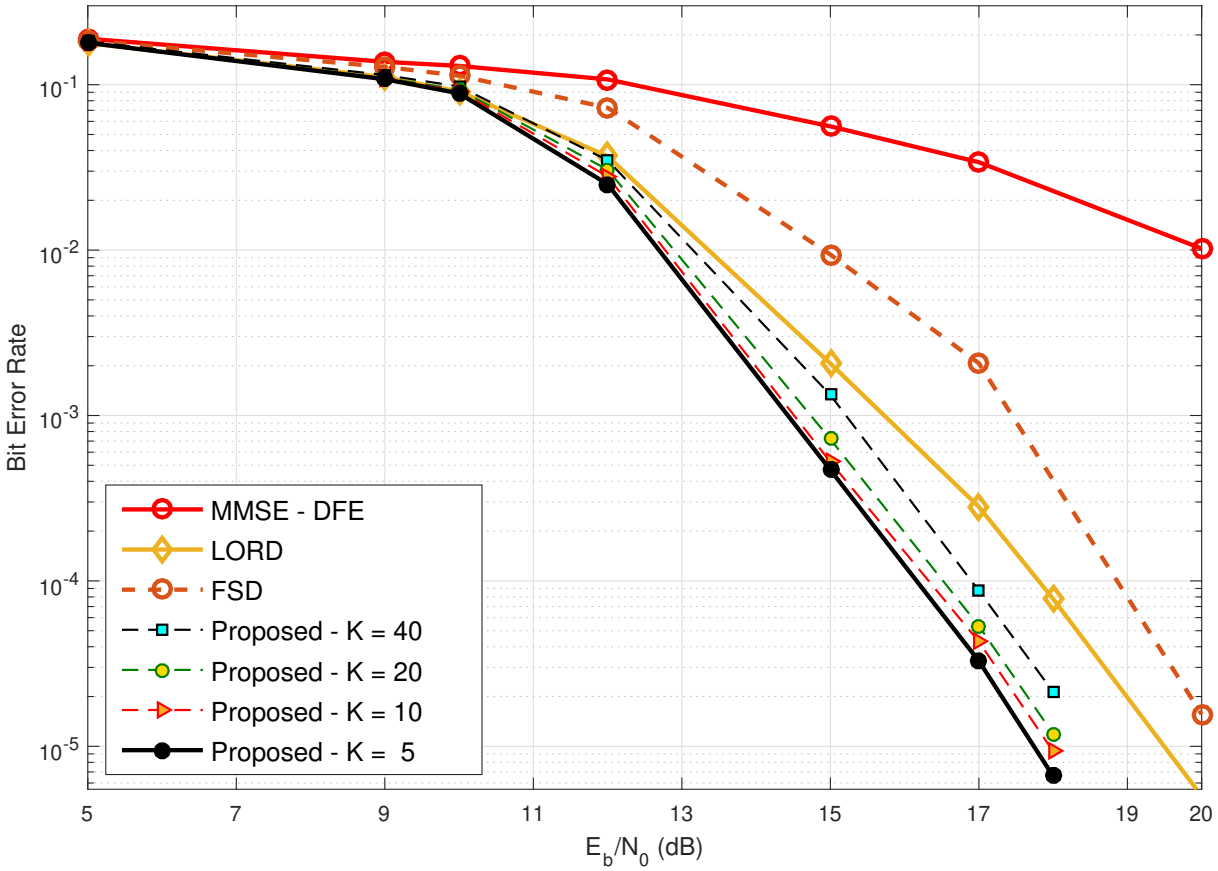


Figure 3.14: BER performance of the proposed PLR conditional optimization detector in Algorithm 3.3 for a 64×64 MIMO system for 16-QAM.

3.5 Conclusion and Discussion

In Chapter 2, we proposed low-complexity near-optimal LRA conditional optimization detectors for MIMO systems. Although their computational complexity is of $\mathcal{O}(M^3|\mathcal{X}|)$ for small and moderate-scale MIMO systems, they become impractical for large-scale MIMO systems due to the computational complexity of the lattice reduction algorithms. As a result, in this chapter, we proposed efficient detection schemes that are applicable to large-scale MIMO systems as well.

In the first part of this chapter, we proposed low-complexity near-optimal detectors that do not use lattice reduction algorithms. In our proposed detectors, the integration of MMSE-DFE filtering and best submatrix selection for conditional optimization detection tailored to an appropriate channel ordering achieves a near-ML error performance with a fixed complexity which is linear in the size of the constellation and cubic in the system dimension, i.e., $\mathcal{O}(M^3|\mathcal{X}|)$. Our proposed detector outperforms the state-of-the-art detection algorithms for MIMO systems with different configurations, such as square, under-determined, and over-determined MIMO systems. Moreover, the proposed detector lends itself naturally to provide soft-outputs where simulation results have shown improvement in performance over the conventional list detector albeit using a much smaller list size. Our algorithm has also a straightforward parallel implementation that further reduces its running time.

In the second part, we proposed an efficient detector for large-scale MIMO systems with the computational complexity of $\mathcal{O}(M^3|\mathcal{X}|)$. Our proposed detector is mainly based on the conditional optimization detection and PLR algorithm tailored to an appropriate channel ordering. The conditional detection for finding one symbol through an exhaustive search over the constellation points is considered. Then, part of the channel is reduced by the PLR technique in order to compensate for the degradation of diversity due to limiting the exhaustive search levels to one. To improve the error performance, all other remaining non-reduced columns are also considered for conditional detection. In the proposed detector, part of the detection complexity is

transferred to the preprocessing stage where the PLR algorithm is implemented. Consequently, for quasi-static channels where the channel is constant over a long block, the cost of PLR can be negligible. Moreover, for fast fading channels, since the lattice reduction is performed on a block of the channel matrix columns, by appropriately selecting the block size in the PLR technique, the computational complexity of $\mathcal{O}(M^3)$ for the preprocessing stage can be achieved. Our proposed detector exhibits error performance improvement over those of the state-of-the-art detection schemes for large-scale MIMO systems.

Chapter 4

Detection for Uplink Multiuser Massive MIMO Systems

4.1 Introduction

We consider a practical multiuser MIMO system where a BS equipped with an antenna array simultaneously serves several cheap single-antenna UTs [71]. To meet the high data rate demand of the different users, massive MIMO technology has been proposed where the BS is equipped with a large number of antennas and serves tens of single-antenna UTs [59, 67, 71, 85]. This also simplifies the signal processing such that linear detectors, such as ZF and MMSE, in the uplink perform sufficiently close to sophisticated non-linear detectors [67, 71].

However, the exact solutions of ZF and MMSE involve $\mathcal{O}(NK^2 + K^3)$ operations with a matrix-matrix multiplication for the computation of the Gram matrix and a matrix inversion, respectively, with the computational complexity of $\mathcal{O}(NK^2)$ and $\mathcal{O}(K^3)$ where N is the number of receive antennas at the BS and K is the number of UTs. Such operations become extremely expensive for massive MIMO systems as N and/or K are very large. Several techniques have

been proposed in the literature in order to reduce the computational complexity of these two operations. In an effort to avoid the direct matrix inversion, approximate algorithms with the help of a truncated Neumann series expansion have been proposed [3, 101, 102, 108]. Further improvements have been achieved by replacing the Neumann series approximation with other iterative algorithms, such as the Gauss-Seidel (GS) [21], successive over-relaxation (SOR) [32], Jacobi [92], and Richardson [30] methods. All these methods reduce the complexity of the matrix inversion from $\mathcal{O}(K^3)$ to $\mathcal{O}(K^2)$, but they still suffer from complexity $\mathcal{O}(NK^2)$ due to the Gram matrix calculation [13].

Another class of approximate methods uses truncated polynomial expansion (TPE) in which the matrix inversion is approximated with J finite terms of the Taylor series expansion [9, 42, 57, 73, 110]. Such methods reduce the overall complexity to a factor which is proportional to the number of BS antennas and the number of users, i.e., $\mathcal{O}(JKN)$. This is achieved by an iterative implementation of a series of matrix-vector multiplications instead of the matrix-matrix multiplication. However, the convergence speed of the TPE detector highly depends on a normalization factor in the polynomial coefficients. The optimal normalization factor, in the sense of convergence speed, requires the calculations of the largest and smallest eigenvalues of the Gram matrix [89]. Thus, calculating the Gram matrix and its eigenvalues increases the computational complexity to $\mathcal{O}(NK^2 + K^3)$.

In this work, we propose a TPE-based detector where a constant normalization factor calculated using the dimensions of the system is tuned for massive MIMO systems with small loading factors $\beta = \frac{K}{N}$. We use the approximate eigenvalues of the Gram matrix of the massive MIMO channel for tuning the normalization factor. The proposed TPE-based detector exhibits an error performance similar to that of the TPE with optimal polynomial coefficients. However, for large loading factors of massive MIMO systems, there exists an error performance degradation at high SNRs due to the inaccuracy of the approximate eigenvalues for such systems. An efficient algorithm based on the power method is proposed for calculating the extreme eigenvalues of

such systems. This method is also utilized for massive MIMO channels with spatial correlations at the BS. The resultant approximate extreme eigenvalues are used for tuning the normalization factor such that the TPE-based detector can achieve the error performance of ZF or MMSE which requires an expensive exact matrix inversion. The overall computational complexity of our proposed TPE-based detector is $\mathcal{O}(JNK)$ for both small and large loading factors.

In the literature, there exist detection schemes with a computational complexity of $\mathcal{O}(mNK)$ where m is the number of iterations of the schemes. Approximate message passing (AMP)-based detectors [55, 98, 103, 112] offer a computational complexity of $\mathcal{O}(mNK)$ by deploying matrix-vector multiplications instead of the matrix-matrix multiplication. However, AMP-based detectors require the knowledge of noise variance and consequently, inappropriate noise variance adjustments can lead to performance degradation. Moreover, they suffer from severe performance degradation for massive MIMO systems with large loading factors, or a realistic scenario when the MIMO channels are spatially correlated such that for some scenarios they may not converge even with a large number of iterations. In contrast to AMP-based detectors, our proposed TPE-based detector ensures convergence for both small and large loading factors of massive MIMO systems for both scenarios of uncorrelated and correlated massive MIMO channels.

The optimized coordinate descent (OCD)-based detector [100] offers a fast converging performance with a computational complexity of $\mathcal{O}(mNK)$. In this scheme, a series of coordinate-wise updates are performed to solve the main optimization problem. Although OCD-based detectors exhibit fast converging performance, the main challenge of this scheme, or coordinate descent (CD)-based techniques in general, resides in the data dependency of successive updates such that the update of each coordinate, which corresponds to a user, depends on the previous coordinates' updates. Such a dependency prevents fully-parallel implementation [100] and also increases the processing delay of each user to a number which is proportional to the number of users and iterations, i.e., mK. Pipeline interleaving is deployed to address this challenge by simultaneously processing multiple coordinates. However, this approach results in a significant

hardware overhead [100]. On the other hand, our proposed TPE-based detector may need a larger number of iterations to achieve the error performance of ZF or MMSE. However, the detections for all users are simultaneously calculated in each iteration resulting in a processing delay which is independent of the number of users and scales only with the number of iterations. This property enables our proposed TPE-based detector to be implemented in a fully-parallel manner with architecture pipelining.

4.2 System Model

We consider an uplink multiuser massive MIMO system with N receive antenna at the BS and K single-antenna UTs. The transmitted signal vector is $\mathbf{x} \in \mathcal{X}^{K\times 1}$ whose entries are chosen from a given constellation, $\mathcal{X}$, such that $\mathbb{E}[\mathbf{x}\mathbf{x}^H] = E_s\mathbf{I}_K$. The entries of the channel matrix $\mathbf{H} = \frac{1}{\sqrt{N}}[\mathbf{h}_1, \mathbf{h}_2, \ldots, \mathbf{h}_K] \in \mathbb{C}^{N\times K}$ are assumed to be i.i.d. $\sim \mathcal{CN}(0, 1/N)$. Hence, the received signal is

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{4.1}$$

where $\mathbf{n} \in \mathbb{C}^{N\times 1}$ is the noise vector at the BS with i.i.d. entries following $\mathcal{CN}(0, N_0)$. We denote $\beta = \frac{K}{N}$ as the loading factor of the massive MIMO system. Simple solutions for multiuser uplink massive MIMO are the ZF detector

$$\hat{\mathbf{x}}_{\text{ZF}} = \mathbf{W}_{\text{ZF}}\mathbf{y} = (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H\mathbf{y}, \tag{4.2}$$

and the MMSE detector

$$\hat{\mathbf{x}}_{\text{MMSE}} = \mathbf{W}_{\text{MMSE}}\mathbf{y} = (\mathbf{H}^H\mathbf{H} + \frac{1}{snr}\mathbf{I}_K)^{-1}\mathbf{H}^H\mathbf{y}, \tag{4.3}$$

where $snr = \frac{E_s}{N_0}$ is the transmit SNR. Both detectors require a matrix inversion with $\mathcal{O}(K^3)$ operations and a matrix-matrix multiplication with $\mathcal{O}(NK^2)$ operations.

An alternative approach to the inverse calculation is the TPE since the inverse of the Gram matrix can be expressed as a matrix polynomial [73] where the low-order terms are the most dominant ones.

Lemma 3 ([73]) *For any positive definite Hermitian matrix* $\mathbf{X}$*,*

$$\mathbf{X}^{-1} = \alpha(\mathbf{I} - (\mathbf{I} - \alpha\mathbf{X}))^{-1} = \alpha \sum_{l=0}^{\infty} (\mathbf{I} - \alpha\mathbf{X})^l \tag{4.4}$$

where the second equality holds when $0 < \alpha < \frac{2}{\lambda_{max}(\mathbf{X})}$ *such that* $\lambda_{max}(\mathbf{X})$ *is the largest eigenvalue of* $\mathbf{X}$*.* α *is referred to as the normalization factor.*

By using this lemma and $\mathbf{X} = \mathbf{H}^H\mathbf{H}$, the ZF can be approximated as

$$\begin{aligned}\mathbf{W}_{\text{ZF}} &= (\mathbf{H}^H\mathbf{H})^{-1}\mathbf{H}^H \\ &\approx \sum_{l=0}^{J-1} \left(\alpha \sum_{n=l}^{J-1} \binom{n}{l} (-\alpha)^l \right) (\mathbf{H}^H\mathbf{H})^l \mathbf{H}^H. \end{aligned} \tag{4.5}$$

Similarly, using $\mathbf{X} = \mathbf{H}^H\mathbf{H} + \frac{1}{snr}\mathbf{I}_K$, the MMSE detector can be expanded as

$$\mathbf{W}_{\text{MMSE}} \tag{4.6}$$

$$= (\mathbf{H}^H\mathbf{H} + \frac{1}{snr}\mathbf{I}_K)^{-1}\mathbf{H}^H \tag{4.7}$$

$$\approx \sum_{l=0}^{J-1} \left(\alpha \sum_{n=l}^{J-1} \binom{n}{l} (1 - \frac{1}{snr})^{n-l} (-\alpha)^l \right) (\mathbf{H}^H\mathbf{H})^l \mathbf{H}^H. \tag{4.8}$$

We denote $\mathbf{W}_{\text{TPE}}$ as the TPE detector such that

$$\mathbf{W}_{\text{TPE}} = \sum_{l=0}^{J-1} w_l \Big(\mathbf{H}^H\mathbf{H}\Big)^l \mathbf{H}^H, \tag{4.9}$$

where

$$w_l = \alpha \sum_{n=l}^{J-1} \binom{n}{l} (1 - \frac{1}{snr})^{n-l} (-\alpha)^l, \tag{4.10}$$

and J is the TPE order. Note that when $N \gg K$, the matrix-matrix multiplication is even more expensive than the matrix inversion. Hence, the matrix-matrix multiplication is avoided by implementing an iterative computations of the J terms where a series of matrix-vector multiplications are performed to detect the transmitted symbols vector. Hence, one can write

$$\hat{\mathbf{x}}_l = \begin{cases} w_0 \mathbf{H}^H \mathbf{y}, & l = 0 \\ w_l \mathbf{H}^H (\mathbf{H} \hat{\mathbf{x}}_{l-1}), & 1 \leq l \leq J-1. \end{cases} \tag{4.11}$$

From (4.9) and (4.11), it can be shown that the calculation of

$$\hat{\mathbf{x}}_{\text{TPE}} = \mathbf{W}_{\text{TPE}} \mathbf{y}, \tag{4.12}$$

requires $(16J - 8)KN - (2J - 2)N + (J - 2)K$ FLOPs. Moreover, the calculations of w_l's needs some additional operations. For a given α, all coefficients require $J^2 + 25J - 12$ and $J^2 + 37J - 16$ FLOPs respectively for ZF and MMSE versions of the TPE detector.

4.3 Proposed Low-Complexity TPE-based Detectors

The normalization factor α has a key impact on the convergence of the TPE detector. The convergence condition in Lemma 3 results in the following condition

$$0 < \alpha < \frac{2}{\lambda_{max}(\mathbf{G})}, \tag{4.13}$$

for ZF and

$$0 < \alpha < \frac{2}{\lambda_{max}(\mathbf{G}) + \frac{1}{snr}}, \tag{4.14}$$

for MMSE where $\lambda_{max}(\mathbf{G})$ denotes the maximum eigenvalue of $\mathbf{G} = \mathbf{H}^H\mathbf{H}$. This normalization factor is used to shift the eigenvalues of $\mathbf{G}$ to the convergence area as they may lie outside of the area. A coarse choice for the normalization factor can be $\alpha = \frac{2}{\text{Trace}(\mathbf{G})}$ as one can write

$$\sum_n \lambda_n(\mathbf{G}) = \sum_n g_{n,n} = \text{Trace}(\mathbf{G}) > \lambda_{max}(\mathbf{G}), \tag{4.15}$$

where $\lambda_n(\mathbf{G})$ is the n-th eigenvalue of $\mathbf{G}$. However, the convergence of the TPE with this value of α is slow. This approximation needs $8NK-2$ operations for the calculations of the summation of the diagonal entries of $\mathbf{G}$. It is shown in [88] that the fastest convergence happens when the two extreme cases $\alpha\lambda_{min}(\mathbf{X})$ and $\alpha\lambda_{max}(\mathbf{X})$ are equally distant to unity. Hence, one can write

$$\alpha_{\text{opt}} = \frac{2}{\lambda_{min}(\mathbf{G}) + \lambda_{max}(\mathbf{G})}, \tag{4.16}$$

for ZF and

$$\alpha_{\text{opt}} = \frac{2}{\lambda_{min}(\mathbf{G}) + \lambda_{max}(\mathbf{G}) + \frac{1}{snr}}. \tag{4.17}$$

for MMSE. However, the calculation of $\mathbf{G}$ itself requires $\mathcal{O}(NK^2)$ operations. In addition, the calculations of $\lambda_{min}(\mathbf{G})$ and $\lambda_{max}(\mathbf{G})$ require $\mathcal{O}(K^3)$ operations. An approximate method is proposed in [89] by offering intervals for the eigenvalues of $\mathbf{G}$. This method reduces the complexity of the extreme eigenvalues calculations to $\mathcal{O}(K^2)$. However, it still needs the calculations of the entries of $\mathbf{G}$.

4.3.1 Proposed Normalization Factor for Small Loading Factors

In an effort to reduce the computational complexity pertaining to the calculations of the optimal normalization factor, we here propose to set the normalization factor based on approximations for the extreme eigenvalues of $\mathbf{G}$. Since $\mathbf{G}$ is a complex central Wishart matrix, when N and K grows we have [65]

$$\begin{aligned} \lambda_{max}(\mathbf{G}) &\approx (1+\sqrt{\beta})^2, \\ \lambda_{min}(\mathbf{G}) &\approx (1-\sqrt{\beta})^2. \end{aligned} \tag{4.18}$$

The approximations in (4.18) are accurate when K and N are large and the loading factor is small. Suppose we are using the ZF version of the TPE detector. If we select

$$\alpha = \frac{2}{(1-\sqrt{\beta})^2 + (1+\sqrt{\beta})^2} \tag{4.19}$$

$$= \frac{1}{(1+\beta)}, \tag{4.20}$$

then this normalization factor satisfies the convergence condition in Lemma 3 as one can write

$$\alpha = \frac{2}{(1-\sqrt{\beta})^2 + (1+\sqrt{\beta})^2} \tag{4.21}$$

$$\approx \frac{2}{\lambda_{min}(\mathbf{G}) + \lambda_{max}(\mathbf{G})} < \frac{2}{\lambda_{max}(\mathbf{G})}. \tag{4.22}$$

The computational complexity of calculating the proposed normalization factor in (4.20) is of $\mathcal{O}(1)$ as it only needs one addition and one division. Moreover, as will be shown in Section 4.6, by using this normalization factor, the TPE converges fast such that its behaviour is similar to that of the TPE detector with the optimal normalization factor in (4.16).

Algorithm 4.1 The Power Method for Extreme Eigenvalues

Input: $\mathbf{H}$, $\mathbf{y}$, J

Output: $\lambda_{max}(\mathbf{G})$, $\lambda_{min}(\mathbf{G})$

$\mathbf{x}_0 = \mathbf{H}^H\mathbf{y}$

$\mathbf{x}_{max} = \mathbf{x}_0$, $\mathbf{x}_{min} = \mathbf{x}_0$

for $l = 1, \cdots, J-2$ **do**

  $\mathbf{v}_{max} = \mathbf{H}^H\mathbf{H}\mathbf{x}_{max}$

  $\mathbf{x}_{max} = \mathbf{v}_{max}$

  if $l = 1$ **then**

    $\hat{\lambda}_{max} = \frac{(\mathbf{H}^H\mathbf{H}\mathbf{x}_{max})^H\mathbf{x}_{max}}{\mathbf{x}_{max}{}^H\mathbf{x}_{max}}$

  end if

  $\mathbf{v}_{min} = (\mathbf{H}^H\mathbf{H} - \hat{\lambda}_{max}\mathbf{I})\mathbf{x}_{min}$

  $\mathbf{x}_{min} = \mathbf{v}_{min}$

end for

$\lambda_{max}(\mathbf{G}) = \frac{(\mathbf{H}^H\mathbf{H}\mathbf{x}_{max})^H\mathbf{x}_{max}}{\mathbf{x}_{max}{}^H\mathbf{x}_{max}}$

$\lambda_{min}(\mathbf{G}) = \frac{(\mathbf{H}^H\mathbf{H}\mathbf{x}_{min})^H\mathbf{x}_{min}}{\mathbf{x}_{min}{}^H\mathbf{x}_{min}}$

4.3.2 Massive MIMO Systems with Large Loading Factors

The approximations for the extreme eigenvalues of $\mathbf{G}$ in (4.18) are accurate when K and N are large, and β is small. For large β's, the actual extreme eigenvalues may differ from the approximations which will affect the error performance of the system. As will be shown in Section 4.6, for a large β, e.g, $\beta = \frac{16}{64} = 0.25$, the BER performance of the proposed TPE detector diverges from that of the TPE detector with α_{opt}. As a result, we propose another approximate method with low-computational complexity for calculating the extreme eigenvalues of $\mathbf{G}$ for such systems. These approximate extreme eigenvalues are used for tuning the normalization factor for massive MIMO systems with large loading factors.

The Power Method for Extreme Eigenvalues Calculations

The power method [34] is an iterative algorithm for approximating the largest eigenvalue of a matrix with linearly independent eigenvectors and also with a dominant eigenvalue. The algorithm

starts with a non-zero initialization vector iteratively multiplied by the matrix, i.e.,

$$\begin{aligned} \mathbf{x}_1 &= \mathbf{G}\mathbf{x}_0, \\ \mathbf{x}_2 &= \mathbf{G}\mathbf{x}_1 = \mathbf{G}^2\mathbf{x}_0, \\ &\vdots \\ \mathbf{x}_m &= \mathbf{G}\mathbf{x}_{m-1} = \mathbf{G}^m\mathbf{x}_0. \end{aligned} \tag{4.23}$$

For large powers of m, a good approximation of the dominant eigenvector of $\mathbf{G}$ is obtained. The corresponding eigenvalue is obtained by the Rayleigh quotient

$$\lambda_{max}(\mathbf{G}) \approx \frac{(\mathbf{G}\mathbf{x}_m)^H \mathbf{x}_m}{\mathbf{x}_m{}^H \mathbf{x}_m}. \tag{4.24}$$

Here, we propose initiating the algorithm with $\mathbf{x}_0 = \mathbf{H}^H\mathbf{y}$ and also limiting the number of iterations of the power method in (4.23) to $m = J-2$ in order to use the already calculated TPE terms in (4.11) for approximating the largest eigenvalue. By comparing (4.11) and (4.24), one can write

$$\lambda_{max}(\mathbf{G}) \approx \frac{\hat{\mathbf{x}}_{J-1}^H \hat{\mathbf{x}}_{J-2}}{\hat{\mathbf{x}}_{J-2}^H \hat{\mathbf{x}}_{J-2}}. \tag{4.25}$$

With these choices of initialization and number of iterations, the power method only needs $16K+2$ additional FLOPs for computing $\lambda_{max}(\mathbf{G})$. We note that since we utilize the TPE terms for the calculation of $\lambda_{max}(\mathbf{G})$, it can be implemented in parallel to the TPE terms calculations without imposing further processing delays on the system.

In the following, we discuss how efficiently we can obtain an approximation for $\lambda_{min}(\mathbf{G})$. As $\mathbf{G}$ is a positive definite matrix, matrix $\mathbf{G}' = \mathbf{G} - \lambda_{max}(\mathbf{G})\mathbf{I}_K$ is a negative definite matrix with the dominant eigenvalue $\lambda_{min}(\mathbf{G}) - \lambda_{max}(\mathbf{G})$. As a result, by giving matrix $\mathbf{G}'$ to the power

method, an approximation of $\lambda_{min}(\mathbf{G})$ can be achieved after shifting the output by $\lambda_{max}(\mathbf{G})$, i.e.,

$$\lambda_{min}(\mathbf{G}) \approx \frac{((\mathbf{G} - \lambda_{max}\mathbf{I})\mathbf{x}_m)^H \mathbf{x}_m}{{\mathbf{x}_m}^H \mathbf{x}_m} + \lambda_{max}(\mathbf{G}). \tag{4.26}$$

For this approach, the approximation of $\lambda_{min}(\mathbf{G})$ requires a prior knowledge of $\lambda_{max}(\mathbf{G})$. The straightforward solution is to calculate $\lambda_{min}(\mathbf{G})$ after obtaining $\lambda_{max}(\mathbf{G})$ in (4.24) using two separate instances of the power method. However, this approach increases the processing delay by the number of iterations of the power method for calculating $\lambda_{min}(\mathbf{G})$. As a result, we propose to calculate both $\lambda_{min}(\mathbf{G})$ and $\lambda_{max}(\mathbf{G})$ simultaneously where a coarse approximation of $\lambda_{max}(\mathbf{G})$, which is obtained at the early steps of the power method, is used for $\lambda_{min}(\mathbf{G})$ approximation. Algorithm 4.1 contains the detailed steps of the proposed approach.

Remark 2 *For massive MIMO systems with large loading factors, the number of TPE terms, i.e., J, is such that an accurate approximation of $\lambda_{max}(\mathbf{G})$ is attained. However, for some channel realizations, the achieved $\lambda_{min}(\mathbf{G})$ might be different from the actual smallest eigenvalue. This mismatch can be reduced by increasing the number of iterations or selecting an appropriate initialization at the cost of an increase in computational complexity. However, for our application, as will be shown in Section 4.6, such a mismatch does not affect the error performance of the system.*

4.3.3 Spatially Correlated Massive MIMO Channels

In realistic wireless communication environments, the error performance of the uplink of massive MIMO systems is affected by the spatial correlation between the antennas at the BS. We consider

the specially correlated channel model in [16, 20] such that

$$\mathbf{H}_{\text{sc}} = \mathbf{R}^{1/2}\mathbf{H}, \tag{4.27}$$

where $\mathbf{R} \in \mathbb{R}^{N\times N}$ is the correlation matrix defined as

$$\mathbf{R} = \begin{bmatrix} 1 & \rho & \dots & \rho^{N-1} \\ \rho & 1 & \dots & \rho^{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{N-1} & \rho^{N-2} & \dots & 1 \end{bmatrix}.$$

where ρ is the correlation coefficient. In Section 4.6, we evaluate the error performance of our proposed TPE-based detector for the spatially correlated MIMO channels. In our evaluation, we replace the channel matrix $\mathbf{H}$ in (4.1) with $\mathbf{H}_{sc}$ in (4.27). We also normalize $\mathbf{H}_{sc}$ in (4.27) by the norm of $\mathbf{R}^{1/2}$ in order to have a consistent SNR adjustment with the uncorrelated MIMO channel scenario. We also note that for such channels, we use the TPE-based detectors with the proposed normalization factor using Algorithm 4.1 as the extreme eigenvalues approximations in (4.18) are not valid for such channels.

4.4 Computational Complexity Analysis

In this section, we analyze the computational complexity of the proposed TPE-based detectors. We assume six and two FLOPs, respectively, per complex multiplication/division and complex addition/subtraction. We note that the multiplication of matrices of size $K \times N$ and $N \times M$ requires $(8N-2)KM$ FLOPs [34]. For a given massive MIMO system, one can show that:

- The calculation of α requires 8 FLOPs.

Table 4.1:
Computational Complexity In Terms of Complex Multiplications

Detector	Complex Multiplications	Complexity Order	Complex Multiplications (N=128, $K=16$, 16-QAM)
ZF/MMSE	$\frac{1}{2}NK^2+\frac{1}{2}K^3+\frac{3}{2}NK+\frac{5}{2}K^2$	$\mathcal{O}(NK^2+K^3)$	22,144
AMP [13]	$m(2NK+N+K)$	$\mathcal{O}(2mNK)$	16,960 $(m=4)$
OCD [100]	$m(2NK+K)$	$\mathcal{O}(2mNK)$	12,336 $(m=3)$
CG	$(m+1)NK+6mK+m$	$\mathcal{O}((m+1)NK)$	10,628 $(m=4)$
Proposed TPE-based ($m=J$, α_{constant})	$2mNK-NK+mK$	$\mathcal{O}((2m-1)NK)$	18,512 $(m=J=5)$

- For a given α, the calculations of all w_l's require $J^2+25J-12$ and $J^2+37J-16$ FLOPs, respectively, for ZF and MMSE versions of the TPE.
- The iterative computations of J terms in (4.11) require $(16J-8)KN-(2J-2)N+(J-2)K$ FLOPs.

Consequently, the computational complexity of our proposed TPE-based detector when the constant normalization factor in (4.20) is used is $\mathcal{O}(JKN)$.

For massive MIMO systems with large loading factors, one also needs to add the computational complexity of Algorithm 4.1. One can show that:

- By using the already calculated TPE terms in (4.11), the calculation of $\lambda_{max}(\mathbf{G})$ only needs $16K+2$ FLOPs.
- The calculation of $\hat{\lambda}_{max}$ needs $16K+2$ FLOPs.
- We need to calculate

$$(\mathbf{G}-\hat{\lambda}_{max}\mathbf{I})^l\mathbf{H}^H\mathbf{y}=\sum_{k=0}^{l}\binom{n}{l}(-\hat{\lambda}_{max})^k\mathbf{G}^{l-k}\mathbf{H}^H\mathbf{y}, \tag{4.28}$$

  for $l=1,\cdots,J-1$. For all iterations, the calculations of the required powers of $\hat{\lambda}_{max}$ and the coefficients need $12(J-2)$ FLOPs.

- Given the coefficients, the calculation of the expression in (4.28) needs $\frac{J(J-1)}{2}$ additions and $\frac{J(J+1)}{2} - 1$ multiplications for all iterations. Hence, it needs $K(4J^2 + 2J - 6)$ FLOPs.
- The calculation of $\hat{\lambda}_{min}$ needs $2K + 4$ FLOPs.

As a result, compared to the constant normalization factor, Algorithm 4.1 increases the computational complexity by $\mathcal{O}(J^2K)$. Hence, for both cases of small and large loading factors of massive MIMO systems, the overall computational complexity of our proposed TPE-based detectors is $\mathcal{O}(JKN)$.

In Table 5.1, we compare the complexity of different detection schemes. Our proposed TPE-based detector has the same computational complexity of $\mathcal{O}(mNK)$ as AMP-based and OCD-based detectors if we consider $m = J$ for our TPE-based detector. The last column of the table indicates the required number of complex multiplications for a 128×16 massive MIMO with 16-QAM modulation when different schemes with sufficient number of iterations used to achieve an error performance close to the ZF or MMSE solution. In Section 4.6, for different massive MIMO systems, we compare the error performance of our proposed TPE-based detector with those of AMP-based and OCD-based detectors.

The AMP-based detector fails to converge for large loading factor of massive MIMO systems or spatially correlated massive MIMO channels while with our proposed TPE-based detector, convergence to the ZF or MMSE detector is ensured. Moreover, an inappropriate adjustment of the noise variance can degrade the error performance of the AMP-based detector. However, the ZF version of the proposed TPE-based detector does not require the knowledge of noise variance.

The OCD-based detector is able to converge fast for different scenarios of massive MIMO systems. However, the main drawback of the OCD-based scheme lies in the delay of processing associated with this scheme as the update of each coordinate (corresponding to each user) depends on the updates of previous coordinates. This dependency increases the processing delay to a number which is proportional to the number of users and iterations, i.e., mK, and also prevents

fully-parallel implementation of the scheme. However, although our proposed TPE-based detector may need a larger number of iterations to approach the error performance of ZF or MMSE, it can be implemented in a fully-parallel manner as in each iteration, the estimates of all users are calculated simultaneously resulting in a processing delay which is proportional only to the number of iterations (or TPE terms), i.e, $m = J$.

4.5 Discussion on the Conjugate Gradient-Based Detector

Conjugate gradient (CG) is an efficient iterative algorithm to solve the problem of signal detection for the uplink of massive MIMO systems. One key advantage of the CG-based detector is that it converges at K iterations and it can even be terminated with a fewer number of iterations than K while the solution is sufficiently close to the exact solution [107]. However, one disadvantage of the CG-based detector is that it does not provide the post-processing SINR information for the calculations of LLR values for the soft-output version of the detector. In [107], a CG-based soft-output detection scheme is proposed for massive MIMO systems. It is worth mentioning that in some works in the literature, for example [45], it is assumed that the Gram matrix (or its regularized version) is given as an input to the CG-based detector which requires a matrix-matrix multiplication. However, one can easily verify that the corresponding matrix-matrix multiplication can be replaced with matrix-vector multiplications. Therefore, as Table 5.1 shows, the overall computational complexity of the CG-based detector is $\mathcal{O}(mKN)$. In the following, we describe the steps of the CG-based detector.

We can rewrite the ZF solution in (4.2) as

$$\begin{aligned} \hat{\mathbf{x}}_{\text{ZF}} &= \mathbf{G}^{-1}\mathbf{H}^{H}\mathbf{y} \\ &= \mathbf{G}^{-1}\mathbf{b}, \end{aligned} \tag{4.29}$$

or the MMSE solution in (4.3) as

$$\begin{aligned}\hat{\mathbf{x}}_{\text{MMSE}} &= (\mathbf{G} + \frac{1}{snr}\mathbf{I}_K)^{-1}\mathbf{H}^H\mathbf{y} \\ &= (\mathbf{G} + \frac{1}{snr}\mathbf{I}_K)^{-1}\mathbf{b},\end{aligned} \tag{4.30}$$

where $\mathbf{b} = \mathbf{H}^H\mathbf{y}$. These solutions can be seen as the solutions of linear equations. For example, for the ZF solution, we have

$$\mathbf{G}\mathbf{x} = \mathbf{b}. \tag{4.31}$$

This linear equation can be transformed into the following quadratic optimization function

$$\min_{\mathbf{x}\in\mathcal{C}^{K\times 1}} f(\mathbf{x}) \triangleq \left(\frac{1}{2}\mathbf{x}^H\mathbf{G}\mathbf{x} - \mathbf{b}^H\mathbf{x}\right). \tag{4.32}$$

Since $\mathbf{G}$ is a symmetric and positive definite matrix, the gradient of $f(\mathbf{x})$ at the optimal point $\hat{\mathbf{x}}$ is zero. If $\mathcal{D} \triangleq \{\mathbf{d}^{(0)}, \mathbf{d}^{(1)}, \cdots, \mathbf{d}^{(K-1)}\}$ denotes the CG with respect to $\mathbf{G}$, we have ${\mathbf{d}^{(i)}}^H\mathbf{G}\mathbf{d}^{(j)} = 0; \forall i \neq j$. The detailed steps of the CG-based detector for the ZF solution have been described in Algorithm 4.2. One can show that this algorithm requires $(m+1)NK + 6mK + m$ complex multiplications and $(m+1)NK - mN + (6m-1)K - 3m$ complex additions. Therefore, the overall computational complexity is $\mathcal{O}(8(m+1)NK)$ in terms of FLOPs. In Section 4.6, we compare the error performance of the CG-based detector with that of our proposed TPE-based detector.

4.6 Simulation Results

In this section, we consider several massive MIMO systems in order to investigate the error performance of the proposed TPE-based detectors. For all simulations of the TPE-based detectors in this section, we refer to the proposed constant normalization factor in (4.20) and using Algorithm

Algorithm 4.2 The Conjugate Gradient-Based Detector

Input: $\mathbf{H}$, $\mathbf{y}$

Output: $\hat{\mathbf{x}}$

1: **Initialization:** $\mathbf{b} = \mathbf{H}^H\mathbf{y}$, $\hat{\mathbf{x}} = \mathbf{0}$, $\mathbf{r}_0 = \mathbf{b}$, $\mathbf{p}_0 = \mathbf{r}_0$
2: **for** $k = 1, \ldots, K$ **do**
3: $\mathbf{e}_{k-1} = \mathbf{H}^H\mathbf{H}\mathbf{p}_{k-1}$
4: $\alpha_k = ||\mathbf{r}_{k-1}||^2/(\mathbf{p}_{k-1}^H\mathbf{e}_{k-1})$
5: $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k-1} + \alpha_k\mathbf{p}_{k-1}$
6: $\mathbf{r}_k = \mathbf{r}_{k-1} - \alpha_k\mathbf{e}_{k-1}$
7: $\beta_k = ||\mathbf{r}_k||^2/||\mathbf{r}_{k-1}||^2$
8: $\mathbf{p}_k = \mathbf{r}_k + \beta_k\mathbf{p}_{k-1}$
9: **end for**

4.1, respectively, as α_{constant} and α_{power}.

Fig. 4.1 shows the average mean square error (MSE) between the exact inverse of $\mathbf{G}$ and its approximation using the TPE for $10,000$ channel realizations. Different normalization factors and TPE orders are considered in order to investigate the convergence of the TPE. Our proposed normalization factor α_{constant} exhibits an MSE similar to that of the TPE with α_{opt}. Moreover, $\alpha = \frac{2}{\text{Trace}(\mathbf{G})}$ has a very slow convergence. Also, $\alpha = \frac{1}{\lambda_{max}(\mathbf{G})}$ shows a better convergence speed, but it still is far from the optimal convergence speed. The simple choice of $\alpha = 1$ diverges at large TPE orders when the loading factor is large. It is worth mentioning that the provided results in Fig. 4.1 is the average MSE for $10,000$ channel realizations while the worst-case MSE is also important, especially for the BER of wireless communication systems, where one coarse approximation can result in a poor error performance; especially at high SNRs.

As Fig. 4.2-4.5 show, the BER performance of the TPE-based detector with α_{constant} is similar to the case when α_{opt} is used for 128×16, 256×16, and 512×16 massive MIMO systems. Moreover, $J = 5$, $J = 4$, and $J = 3$ are sufficient for these systems, respectively, to approach the BER performances of ZF and MMSE detectors with the exact inversion. However, for a 64×16 massive MIMO system, although the TPE-based detector with α_{constant} performs similar to the TPE-based detector with α_{opt} at low and moderate SNRs, there is a performance gap at

high SNRs. It happens due to the inaccuracy of the extreme eigenvalues approximations in (4.18) for systems with large loading factors.

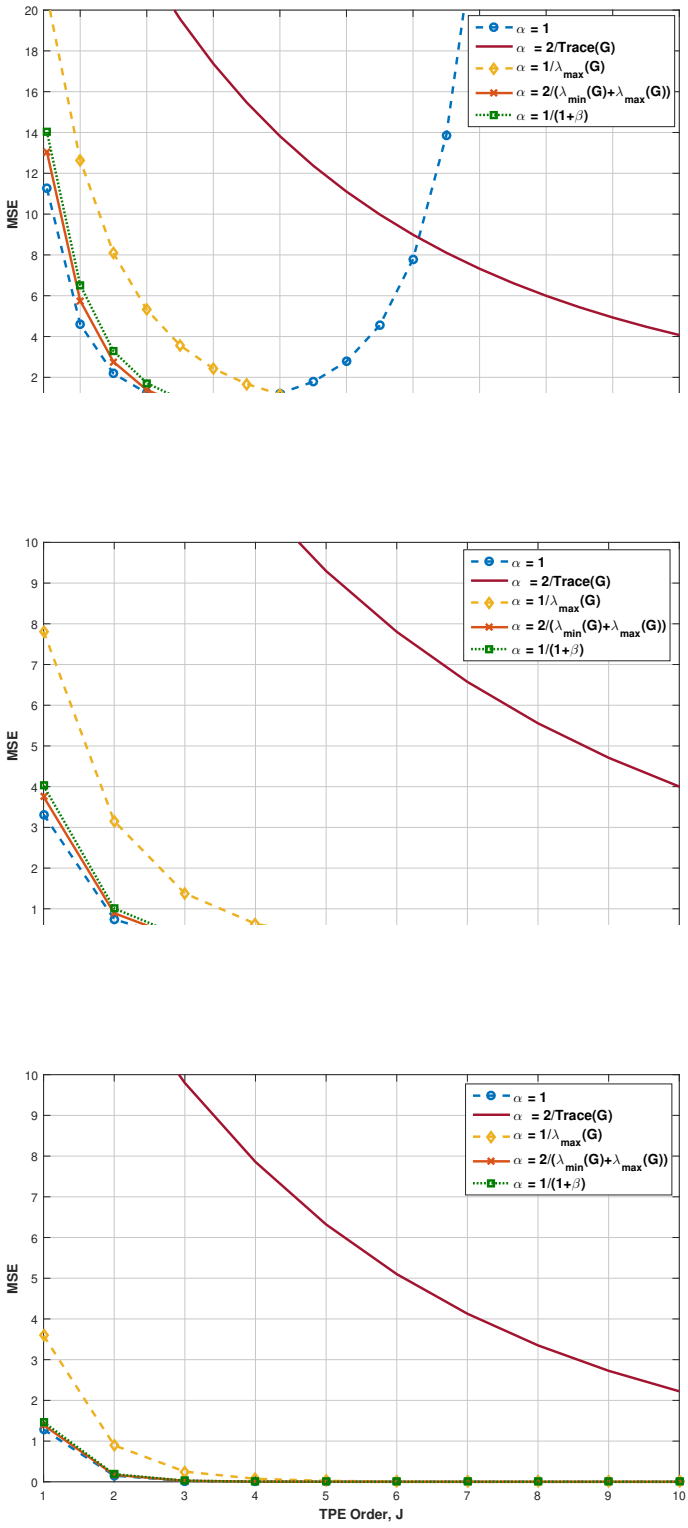
In Fig. 4.6, we use the TPE-based detector with α_{power} in order to resolve the convergence issue when α_{constant} is used for the 64×16 massive MIMO systems which has a large loading factor. As Fig. 4.6a shows, the TPE-based detector converges with α_{power} and it requires $J = 10$ to approach the BER of ZF or MMSE for 16-QAM modulation. In Fig. 4.6b, we evaluate our proposed TPE-based detector for the higher order modulation 64-QAM. For this system, a similar convergence behaviour is observed, and α_{power} with $J = 12$ is required to approach the BER of the ZF or MMSE detector.

In Fig. 4.7, we compare our proposed TPE-based detector with AMP-based and OCD-based detectors. We consider a 64×16 massive MIMO system with a large loading factor of $\beta = 0.25$, and a 256×16 massive MIMO system with a small loading factor of $\beta = 0.0625$. The proposed TPE-based detector converges for these two systems. However, although the AMP-based detector converges with $m = 3$ for the massive MIMO system with the small β in Fig. 4.7b, it fails to converge for the massive MIMO system with the large β in Fig. 4.7a such that it cannot approach the BER of MMSE, and after a certain point increasing the number of iterations will not improve the BER performance.

According to Fig. 4.7, the OCD-based detector converges to the BER of MMSE with $m = 6$ and $m = 3$ iterations, respectively, for 64×16 and 256×16 massive MIMO systems while the proposed TPE-based detector requires $m = J = 10$ and $m = 4$ iterations, respectively. However, the processing delay associated with the OCD-based detector for these two systems with the mentioned required number of iterations is proportional to $mK = 96$ and $mK = 48$, respectively, while those of our proposed TPE-based detector is proportional to $m = 10$ and $m = 4$, respectively. Such smaller processing delays are achieved because of the fully-parallel implementation capability of our proposed TPE-based detector.

In Fig. 4.8, we consider a 128×16 massive MIMO system with two different correlation coefficients of $\rho = 0.2$ and $\rho = 0.3$. For these two systems, the AMP-based detector suffers from a severe performance degradation such that it is not able to approach the BER of MMSE detector even with a large number of iterations. The proposed TPE-based detector approaches the BER of the MMSE detector, respectively with $J = 12$ and $J = 16$. Although the OCD-based detector can achieve such BER performances with approximately half of the iterations required for the TPE-based detector, the processing delays of the OCD-based detector for these two systems are 96 and 128, respectively, while those of our proposed TPE-based detectors are 12 and 16, respectively.

In Fig. 4.9, we compare our proposed TPE-based detector with the CG-based detector for a 128×16 massive MIMO system with 16-QAM modulation. As Fig. 4.9a shows, for this system with the uncorrelated channel matrix, the proposed TPE-based detector with α_{constant} requires $m = J = 5$ while the CG-based detector requires $m = 4$ in order to approach the ZF or MMSE solution. For the system with the spatial correlation of $\rho = 0.2$ between BS antennas in Fig. 4.9b, the CG-based detector converges with $m = 6$ while our proposed TPE-based detector needs $m = J = 12$ in order to approach the ZF or MMSE solution. Similar to our proposed TPE-based detectors, in each iteration, the CG-based detector updates the detected signal vector for all users simultaneously. As a result, it can be considered as a promising alternative to our proposed TPE-based detectors.



(c) $N = 256$, $K = 16$, $\beta = \frac{K}{N} = \frac{16}{256}$

Figure 4.1: Convergence evaluation of TPE-based detector with different normalization factors.

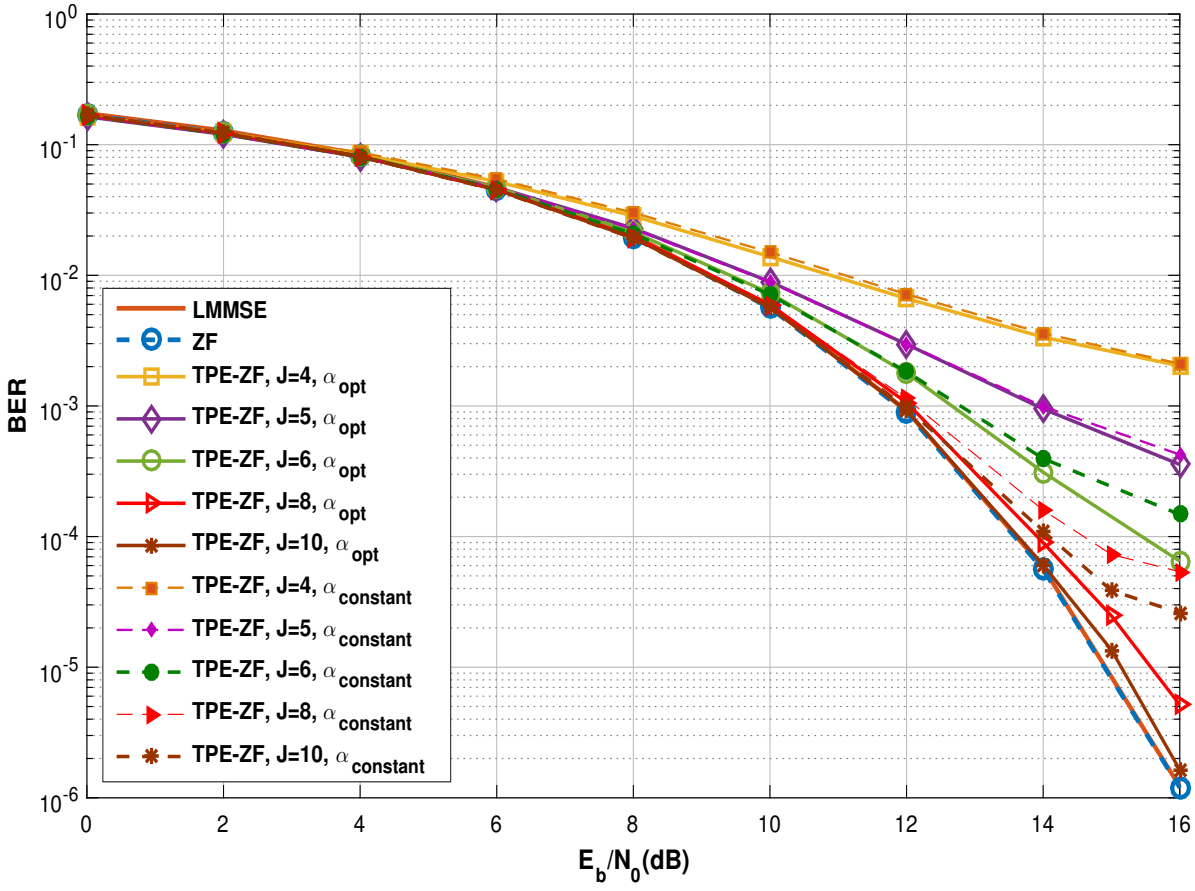


Figure 4.2: BER performances of the TPE-based detector using α_{constant} for a 64×16 massive MIMO systems with 16-QAM modulation; $\beta = \frac{K}{N} = \frac{16}{64}$.

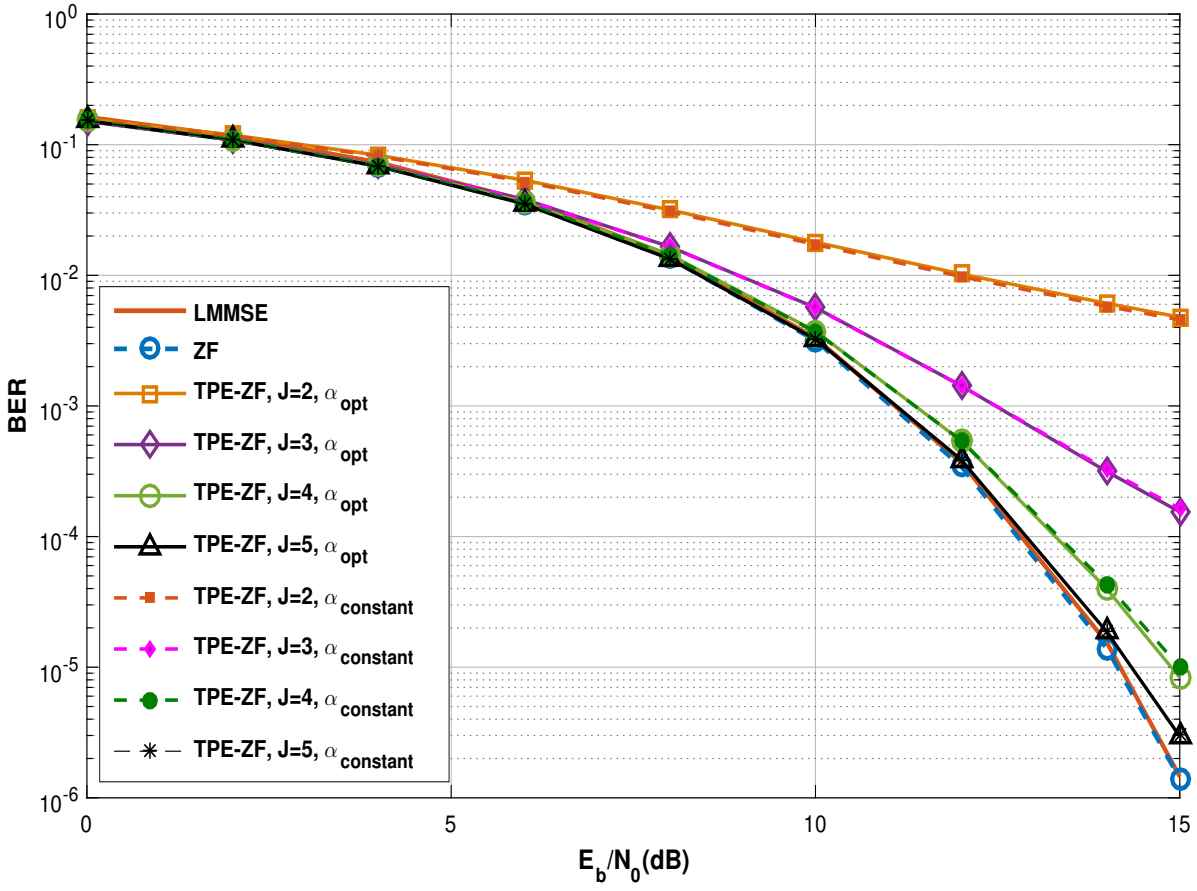


Figure 4.3: BER performances of the TPE-based detector using α_{constant} for a 128×16 massive MIMO systems with 16-QAM modulation; $\beta = \frac{K}{N} = \frac{16}{128}$.

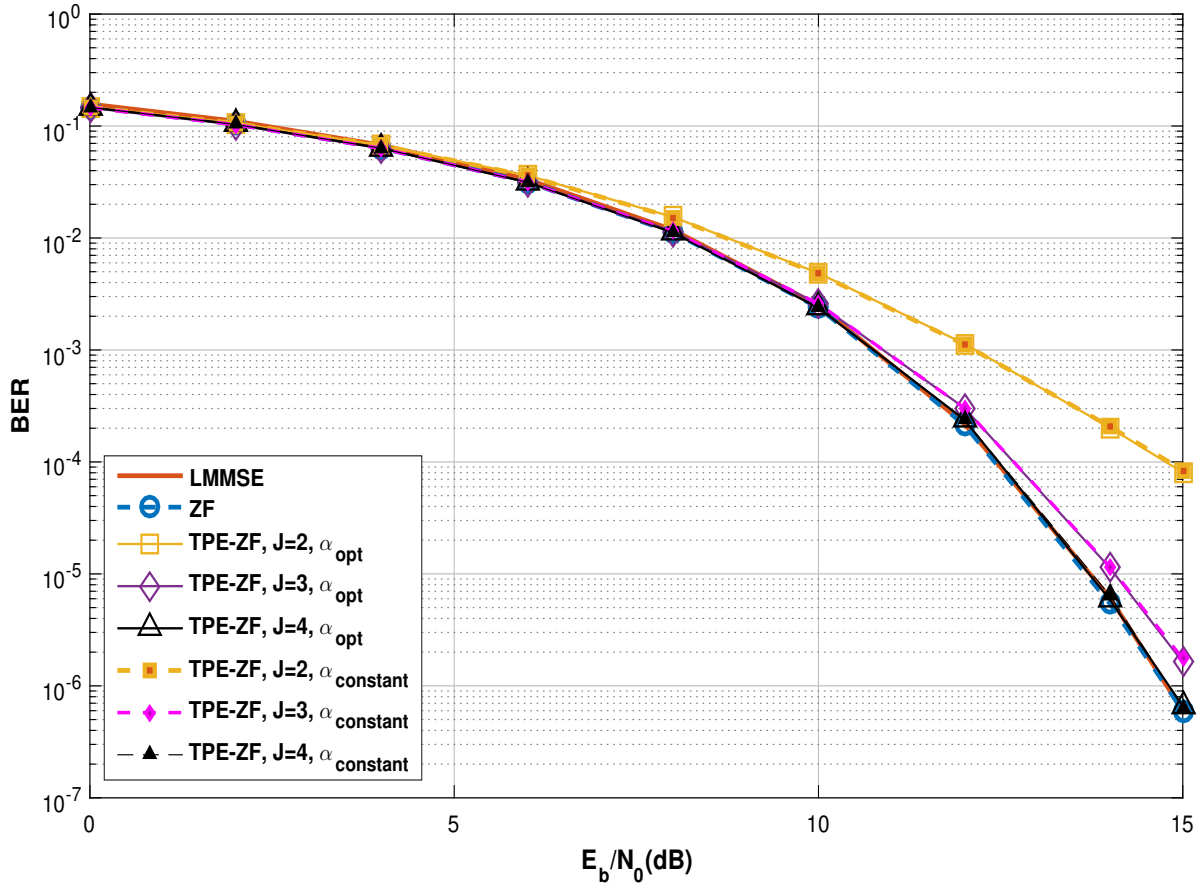


Figure 4.4: BER performances of the TPE-based detector using α_{constant} for a 256×16 massive MIMO systems with 16-QAM modulation; $\beta = \frac{K}{N} = \frac{16}{256}$.

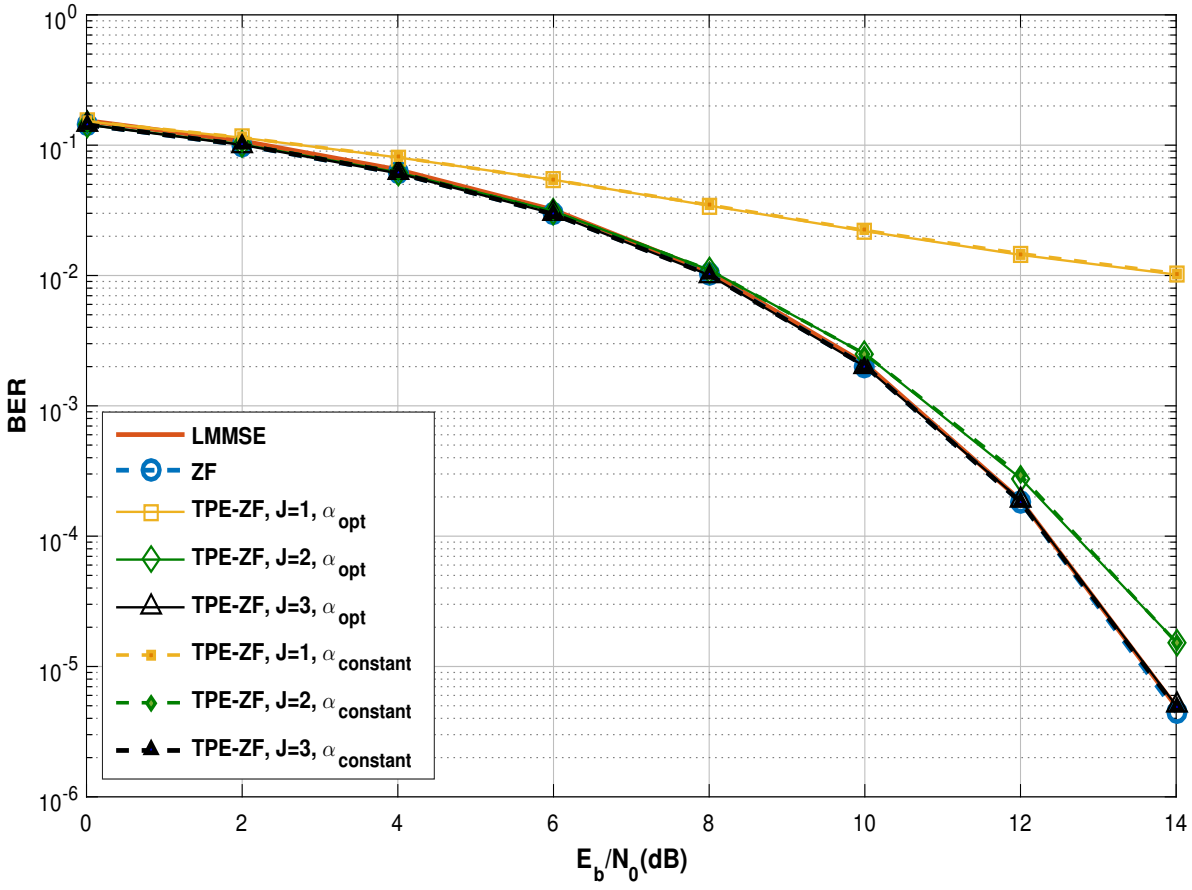
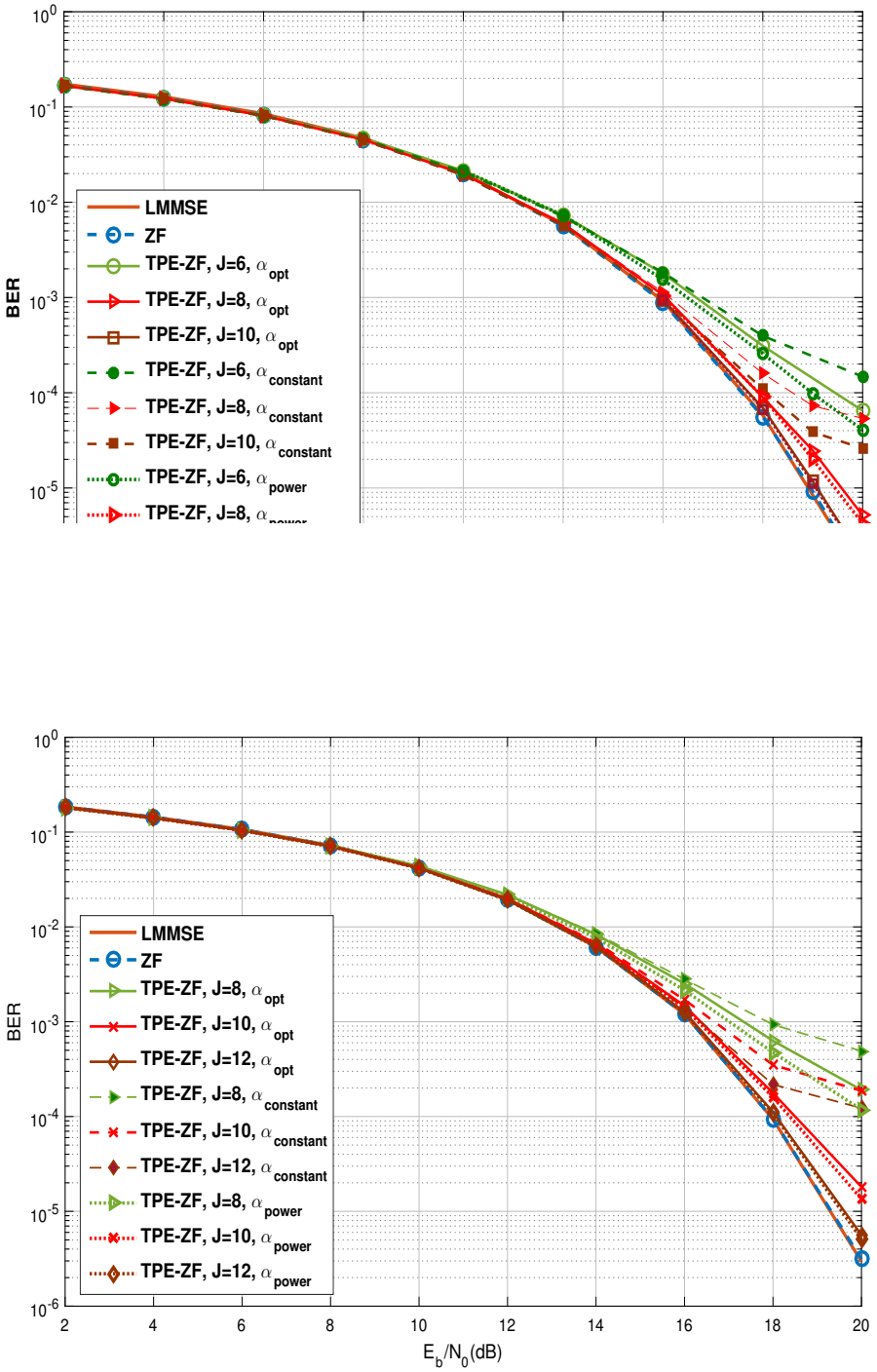
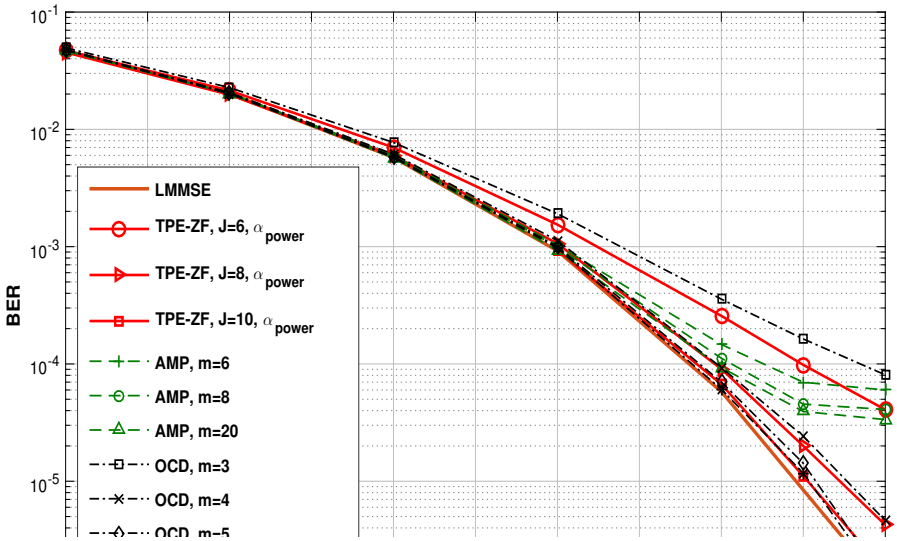


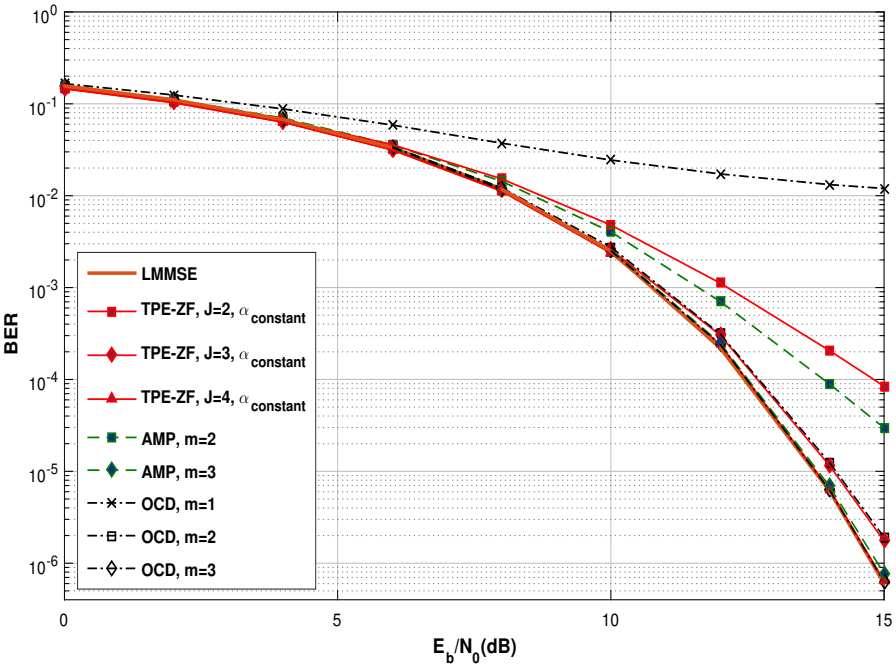
Figure 4.5: BER performances of the TPE-based detector using α_{constant} for a 512×16 massive MIMO systems with 16-QAM modulation; $\beta = \frac{K}{N} = \frac{16}{512}$.



(b) 64×16 - 64QAM

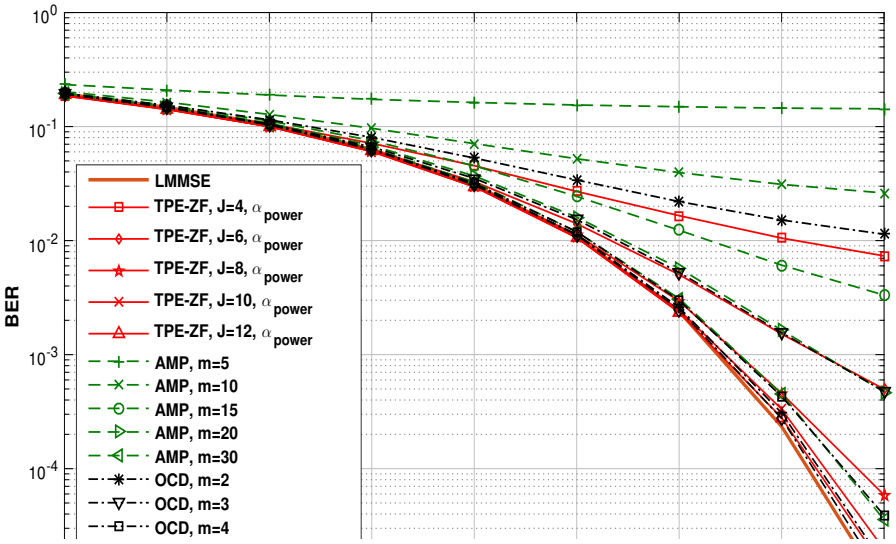
Figure 4.6: BER performances of the TPE-based detector for a 64×16 massive MIMO system with 16-QAM and 64-QAM modulations. For this system, the loading factor is large, $\beta = \frac{16}{64} = 0.25$.

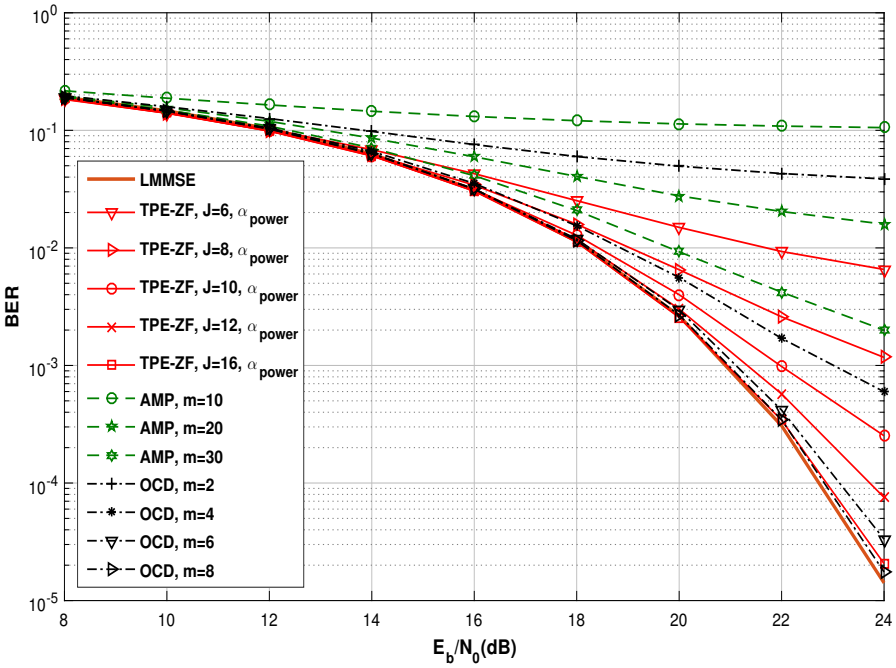




(b) 256×16 - 16QAM

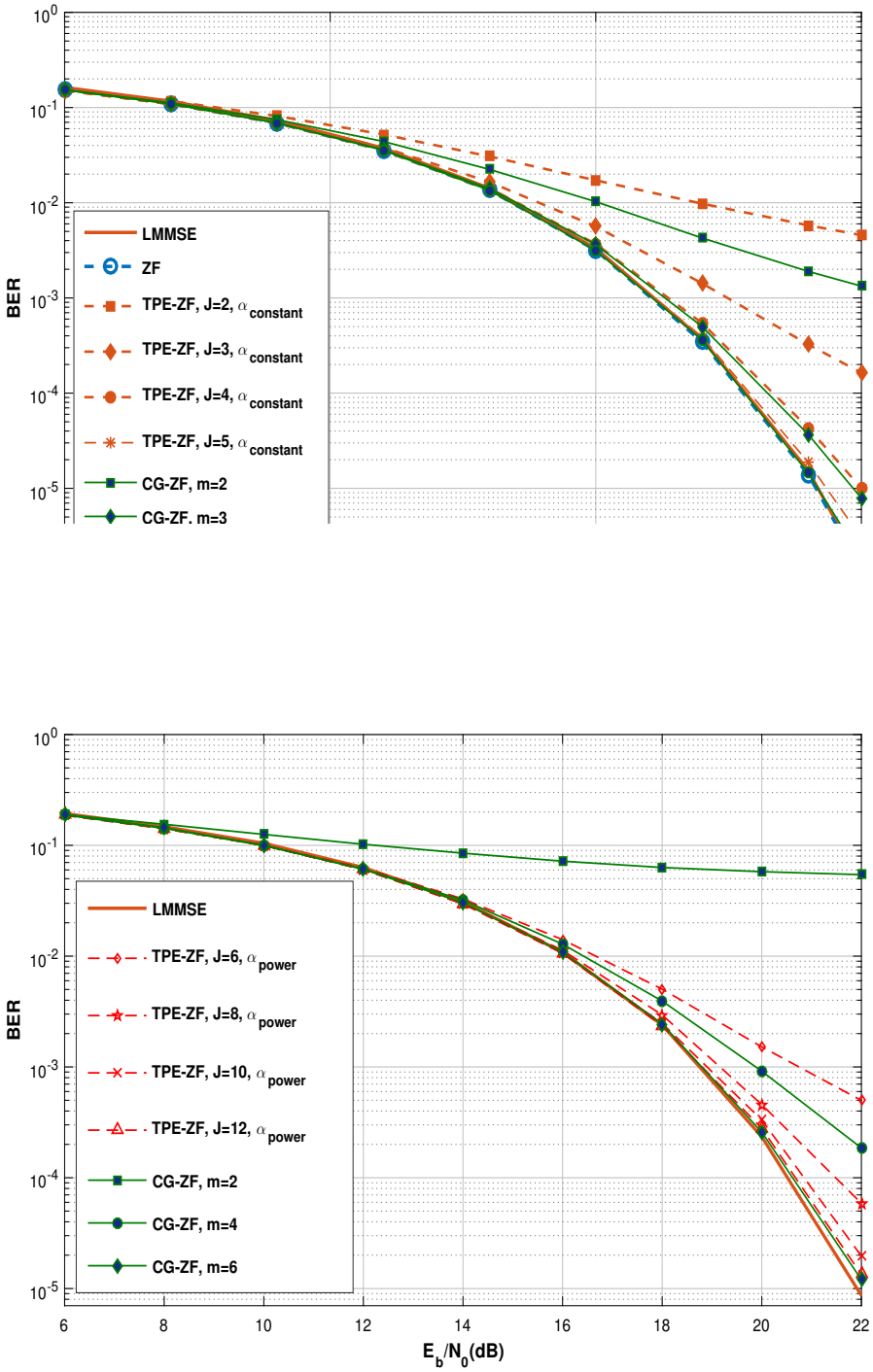
Figure 4.7: BER performances comparison of the proposed TPE-based detector with AMP-based and OCD-based detectors for (a) 64×16 and (b) 256×16 massive MIMO systems with 16-QAM modulation.





(b) 128×16 - 16QAM, $\rho = 0.3$

Figure 4.8: BER performances comparison of the proposed TPE-based detector with AMP-based and OCD-based detectors for a spatially correlated 128×16 massive MIMO channel with (a) $\rho = 0.2$ and (b) $\rho = 0.3$ with 16-QAM modulation. The normalization factor of the TPE-based detector is obtained using Algorithm 4.1.



(b) 128×16 - 16QAM, $\rho = 0.2$

Figure 4.9: BER performances comparison of the proposed TPE-based detector with the CG-based detector for a 128×16 massive MIMO channel with 16-QAM modulation; (a) uncorrelated and (b) spatially correlated with $\rho = 0.2$.

4.7 Conclusion and Discussion

In massive MIMO systems, the number of antennas at the BS is much larger than the number of UTs. As a result, linear detectors, such as ZF and MMSE, are used for the uplink as they exhibit BER performances close to those of non-linear detectors. However, the exact solutions for ZF and MMSE involve matrix-matrix multiplication and matrix inversion operations, respectively, with $\mathcal{O}(N^2K)$ and $\mathcal{O}(K^3)$ operations which become computationally expensive for massive MIMO systems. Therefore, efficient detectors are vital for the uplink of massive MIMO systems. In this chapter, with the help of TPE-based detectors, we proposed an efficient detection scheme for the uplink multiuser massive MIMO systems. The advantage of the TPE-based detectors is that the inverse of the Gram channel matrix can be written as a weighted summation of J finite terms. This enables us to implement the detection scheme in an iterative manner such that the matrix-matrix multiplication and matrix inversion are replaced with a finite series of matrix-vector multiplications. Consequently, for a given set of optimal weights, the computational complexity of the TPE-based detector is $\mathcal{O}(JKN)$. However, the convergence speed of the TPE-based detectors depends on a normalization factor which is optimally calculated from the maximum and minimum eigenvalues of the Gram channel matrix. The exact calculations of these eigenvalues, or the proposed approximate methods in the literature, require $\mathcal{O}(K^3)$ operations. In our proposed TPE-based detector, we offer a constant normalization factor for small loading factors of massive MIMO systems calculated from the dimensions of the system with a complexity of $\mathcal{O}(1)$. For massive MIMO systems with large loading factors, an efficient approximate algorithm for calculating the extreme eigenvalues was proposed. The proposed TPE-based detector achieves the error performance of ZF and MMSE for both small and large loading factors as well as spatially correlated massive MIMO channels with a computational complexity that is proportional to the number of BS antennas and users.

Chapter 5

Detection for Millimeter Wave Massive MIMO Systems

5.1 Introduction

The mmWave has gained a great interest in the 5G cellular networks because of the higher bandwidth communication channels it offers by using the spectrum from 30 GHz to 300 GHz compared to the wireless systems operating at carrier frequencies below 6 GHz [39]. The increase in the carrier frequency results in a higher free-space pathloss. Massive MIMO technology enables stacking a large number of antenna elements providing beamforming gain to combat the increased pathloss [6]. A broad range of technologies of the 5G, such as machine to machine, Internet of vehicles, device to device, backhaul, access, and small cell, will utilize mmWave massive MIMO [2].

To reduce the implementation costs of massive MIMO systems in the mmWave bands, hybrid beamforming techniques have been proposed in the literature. Hybrid beamforming has demonstrated applicability in a wide range of applications, including autonomous driving cars

and machine to machine technology [2]. In hybrid beamforming techniques, the optimal fully-digital beamformers are approximated using a fewer number of RF chains at BSs and/or UTs at the price of some performance degradation. Such approximations are evaluated using the Euclidean distance between the hybrid beamformers and the optimal fully-digital ones. Minimizing the Euclidean distance, which under some strict assumptions is translated to minimizing the mutual information loss due to hybrid beamforming, does not directly imply its efficacy in terms of more practical metrics like error performance. One reason is that the spectral efficiency performance depends on the largest eigenvalues of the channel while the error performance is practically dictated by the stream with the lowest SNR. Therefore, such approximations for hybrid beamforming can result in poor error performances.

Consequently, in this chapter, we improve the error performance of hybrid beamforming mmWave massive MIMO systems. We consider a signal detection scheme using an equivalent channel consisting of the precoder, mmWave channel, and combiner. Then, the proposed detection schemes proposed in Chapter 3 are applied for the equivalent MIMO system. Our proposed detectors achieve an error performance close to that of the optimal ML detector with significantly reduced computational complexity by exploiting the available structure in the equivalent channel. In [4], SD is proposed for mmWave massive MIMO systems by considering the diagonal matrix of the singular values as an effective channel. However, our equivalent channel matrix considers the effect of off-diagonal elements on the detection process. Moreover, SD suffers from varying complexity with an exponential average complexity whereas our proposed detector has fixed polynomial complexity with fully-parallel implementation with reduced complexity.

5.2 System Model

We consider a single-user MIMO system with large antenna arrays and a limited number of RF chains in an mmWave environment where the transmitter and receiver are equipped with M

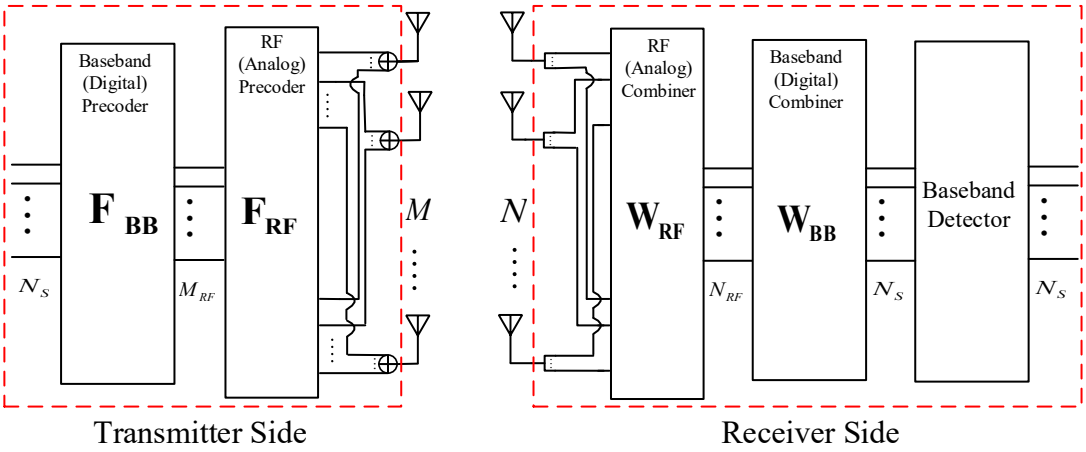


Figure 5.1: System model of hybrid beamforming and detection for mmWave massive MIMO systems.

and N antenna array elements, respectively, as shown in Fig. 5.1. The transmitter sends N_s independent data streams to the receiver with the help of N_{RF} RF chains. The data vector $\mathbf{s}$ is of size $N_s \times 1$ where $\mathbb{E}[\mathbf{s}\mathbf{s}^H] = \frac{1}{N_s}\mathbf{I}_{N_s}$. The data vector is processed by a hybrid beamforming technique where a digital (baseband) precoder $\mathbf{F}_{\text{BB}}$ of size $N_{\text{RF}} \times N_s$ followed by an analog (RF) precoder $\mathbf{F}_{\text{RF}}$ of size $M \times N_{\text{RF}}$ are applied. Therefore, the vector $\mathbf{x} = \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\mathbf{s}$ is sent over the channel where the total transmitted power is normalized such that $\|\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\|_F^2 = N_s$. The received vector of size $N \times 1$ is given by

$$\mathbf{y} = \sqrt{\rho}\mathbf{H}\mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}}\mathbf{s} + \mathbf{n}, \tag{5.1}$$

where ρ is the average received power and $\mathbf{n}$ is the white Gaussian noise vector with $\mathcal{CN}(0, N_0)$ i.i.d. entries. $\mathbf{H}$ is the fading channel matrix which is widely modeled in the mmWave literature by the clustered model [6, 66, 80, 109]. In the clustered model, the channel is represented as the sum of all propagation paths that are scattered in N_{cl} clusters with each cluster contributing N_{ray}

rays, and it is given by

$$\mathbf{H} = \sqrt{\frac{MN}{N_{\text{cl}} N_{\text{ray}}}} \sum_{i,l} \alpha_{il} \mathbf{a}_{\text{r}}(\phi_{il}^{\text{r}}, \theta_{il}^{\text{r}}) \mathbf{a}_{\text{t}}(\phi_{il}^{\text{t}}, \theta_{il}^{\text{t}})^{H}, \tag{5.2}$$

where α_{il} is the complex gain of the j-th ray in the i-th cluster, and $\mathbf{a}_{\text{t}}(\phi_{il}^{\text{t}}, \theta_{il}^{\text{t}})$ is the transmit antenna array response vector of length M for given azimuth and elevation angles of departure, respectively, denoted by ϕ_{il}^{t} and θ_{il}^{t}. Similarly, $\mathbf{a}_{\text{r}}(\phi_{il}^{\text{r}}, \theta_{il}^{\text{r}})$ is the receive antenna array response vector of length N for given azimuth and elevation angles of arrival denoted by ϕ_{il}^{r} and θ_{il}^{r}, respectively. We consider uniform planar array (UPA) at the transmitter and the receiver.

We assume that the received signal vector is processed by two linear combining stages of the hybrid combining followed by a (non-linear) detection stage. Particularly, the received vector is processed by $\mathbf{W}_{\text{RF}}$, i.e., the analog combiner, then by $\mathbf{W}_{\text{BB}}$, the digital precoder and thereby the output of combiners is

$$\tilde{\mathbf{y}} = \sqrt{\rho} \mathbf{W}_{\text{BB}}^{\mathbf{H}} \mathbf{W}_{\text{RF}}^{\mathbf{H}} \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \mathbf{s} + \mathbf{W}_{\text{BB}}^{\mathbf{H}} \mathbf{W}_{\text{RF}}^{\mathbf{H}} \mathbf{n}. \tag{5.3}$$

We note here that both the analog precoder and combiner have RF hardware constraints where their entries have unit magnitudes since they are implemented by constant magnitude phase shifters. After that, the signal combining stage is followed by a detection stage to estimate the transmitted data vector $\hat{\mathbf{s}}$.

The optimal fully-digital precoder and combiner under mutual information performance metric is given by [87]

$$\mathbf{F}_{\text{opt}} = \bar{\mathbf{V}}_{N_s}, \mathbf{W}_{\text{opt}} = \bar{\mathbf{U}}_{N_s}, \tag{5.4}$$

where $\bar{\mathbf{V}}_{N_s}$ and $\bar{\mathbf{U}}_{N_s}$ are the N_s largest right and left singular vectors of the channel matrix,

respectively. Most prior works on hybrid beamforming designs adopted the approach of decomposing the optimal beamforming (precoding/combining) matrix into two matrices, i.e., analog and digital beamforming ones [6, 56, 66, 80, 109] using the following approximation

$$\begin{aligned} \underset{\mathbf{F}_{\text{RF}},\mathbf{F}_{\text{BB}}}{\text{argmin}} \quad & \| \mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}} \|_F^2 \\ \text{subject to} \quad & \mathbf{F}_{\text{RF}} \in \mathbb{U}^{M \times N_{\text{RF}}}, \\ & \| \mathbf{F}_{\text{RF}}\mathbf{F}_{\text{BB}} \|_F^2 = N_s, \end{aligned} \tag{5.5}$$

where $\mathbb{U}^{M \times N_{\text{RF}}}$ is the space of $M \times N_{\text{RF}}$ matrices that have unit magnitude entries. The same problem formulation is applied to the receiver side for approximating $\mathbf{W}_{\text{opt}}$ as $\mathbf{W}_{\text{RF}}\mathbf{W}_{\text{BB}}$. We note that there is no closed form solution for this non-convex optimization problem and all the prior works have provided sub-optimal solutions.

Orthogonal matching pursuit (OMP) and its variants [6, 66], and alternating minimization (AM) are well-known techniques for hybrid beamforming [80, 109]. The Euclidean distance has been widely used for hybrid beamforming designs for maximizing the spectral efficiency. However, the Euclidean distance is not necessarily the best metric to be utilized in measuring the closeness of the hybrid beamformers to the optimal one as to minimize the BER. This directly results in high errors in approximating the optimal beamformer by the hybrid ones which in turn causes inter-stream interference affecting the stream with the lowest SNR and aggravating the total BER performance as will be shown using computer simulations in Section 5.5. Therefore, we propose adding a low-complexity detection stage after the hybrid beamforming (precoding/combining) stage to enhance the BER performance of the system.

5.3 Signal Detection

5.3.1 Equivalent Channel Matrix

Consider the output of the combiner in (5.3). The equivalent channel of size $N_s \times N_s$ consisting of the precoder, mmWave channel, and combiner, i.e.,

$$\mathbf{H}_{\text{eq}} = \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{H} \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}} \tag{5.6}$$

is considered for the detection process. Hence, one can write

$$\tilde{\mathbf{y}} = \sqrt{\rho} \mathbf{H}_{\text{eq}} \mathbf{s} + \mathbf{n}', \tag{5.7}$$

where $\mathbf{n}' = \mathbf{W}_{\text{BB}}^H \mathbf{W}_{\text{RF}}^H \mathbf{n}$. When the optimal fully-digital singular value decomposition (SVD)-based precoder and combiner are used, this equivalent channel will be $\mathbf{\Sigma}_1$, the $N_s \times N_s$ partition of $\mathbf{\Sigma}$, obtained by the SVD of $\mathbf{H}$, i.e., $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^H$. For this diagonal matrix, the optimal solution can be obtained using a symbol by symbol demodulation. However, for the hybrid beamforming structure, there is a difference between the optimal precoder/combiner and the hybrid precoder/combiner due to the approximations. We denote $\mathbf{e}_{\mathbf{F}}$ and $\mathbf{e}_{\mathbf{W}}$ as the error matrices due to such approximations, respectively. Hence, one can write

$$\mathbf{e}_{\mathbf{F}} = \mathbf{F}_{\text{opt}} - \mathbf{F}_{\text{RF}} \mathbf{F}_{\text{BB}}, \tag{5.8}$$

$$\mathbf{e}_{\mathbf{W}} = \mathbf{W}_{\text{opt}} - \mathbf{W}_{\text{RF}} \mathbf{W}_{\text{BB}}. \tag{5.9}$$

Because of these error matrices, the off-diagonal entries of the equivalent channel in (5.6) will not be zero. These non-zero off-diagonal entries are problematic for signal detection, especially for the symbols corresponding to the last columns of the equivalent channel as the diagonal

entries are in descending order as a result of the SVD operation. The error matrices depend on the ability of hybrid beamforming in approximating the optimal precoder/combiner. L and N_{RF} are two parameters which affect the approximation. As we will discuss in Section 5.5, for OMP, the approximation becomes worse whenever N_{RF} decreases or L increases. The error performance of the AM hybrid beamforming technique degrades with decreasing L. The quality of the approximation for each technique affects the error gap between a simple detector and the optimal ML solution. In the next subsection, we address the design of low-complexity signal detectors that achieve the optimal ML error performance.

5.3.2 Detection for the Equivalent MIMO System

In practical mmWave massive MIMO systems, the number of data streams, i.e., N_s is up to 8 [81]. Consequently, the conventional techniques for small and moderate MIMO systems can be utilized for the equivalent channel matrix. However, because of the near-diagonal structure of the equivalent channel in (5.6), it is expected that the conventional detection techniques, such as the FSD [7] and multiple subspace detection proposed in Chapter 3, can be extensively simplified.

We consider the regularized (or augmented) channel matrix, i.e.,

$$\tilde{\mathbf{H}}_{\text{eq}} = \begin{bmatrix} \mathbf{H}_{\text{eq}} \\ \sqrt{\alpha}\mathbf{I}_{N_s} \end{bmatrix}, \tag{5.10}$$

where $\alpha = \frac{N_0}{\rho}$, and we perform the QRD such that

$$\tilde{\mathbf{H}}_{\text{eq}} = \tilde{\mathbf{Q}}\mathbf{R}, \tag{5.11}$$

where $\tilde{\mathbf{Q}}$ is a $2N_s \times N_s$ matrix with orthonormal columns, and $\mathbf{R}$ is an $N_s \times N_s$ upper triangular matrix with positive diagonal entries. By considering $\mathbf{Q}$ as the $N_s \times N_s$ part of $\tilde{\mathbf{Q}}$, consisting of

the first N_s rows, one writes [22]

$$\mathbf{y}' = \mathbf{Q}^H \mathbf{y} = \mathbf{R}\mathbf{s} + \mathbf{w}, \tag{5.12}$$

where $\mathbf{w} = \mathbf{Q}^H \mathbf{n}' - [\mathbf{R} - \mathbf{Q}^H \mathbf{H}_{\text{eq}}]\mathbf{s}$. SIC is implemented using the upper triangular structure of $\mathbf{R}$. In Algorithm 5.1, we propose a multiple subspace detector for the subspace detection of one symbol tailored to an appropriate channel ordering for SIC with complexity saving considerations. After ordering the columns, first, a search over the constellation symbols is performed. Then, SIC is conducted for the detection of the remaining symbols. In the case of multiple detections, i.e., $c > 1$, we examine c worst columns in a round-robin fashion.

By changing the parameters of the algorithm, it will be converted to some detection schemes with reduced complexity. For example, for $c = 1$, it is equivalent to the FSD scheme with $p = 1$, and for $c = N_s$, other than the ordering, it is similar to the multiple subspace detection proposed in Chapter 3. Moreover, by skipping the subspace detection, the algorithm is equivalent to the MMSE-SIC scheme along with the V-BLAST ordering. Consequently, as will be shown later in Section 5.5, by properly selecting the parameters in this algorithm, an appropriate detection scheme can be implemented to achieve a near-optimal error performance while the computational complexity is maintained as low as possible.

For the case of multiple detections, i.e., $c > 1$, if we want to simply swap the next column with the previously examined one, another QRD operation - with a complexity of $\mathcal{O}(N_s^3)$ - will be required due to intermediate fill-in [83] in the Givens rotations process as the current column is usually located at the first columns of the optimal ordering. However, if we order the columns such that the c worst columns place at the c first columns of the ordered matrix, and we perform QRD on this matrix, for each next QRD operation we require up to $N_s - 1$ rotations, equal to the number of non-zero entries below the main diagonal, in order for the matrix to get the upper triangular form. Each rotation changes two rows of the matrix and needs up to $4(N_s - 1)$

Algorithm 5.1 The Proposed Multiple MMSE-DFE Subspace Detector

Inputs: $\mathbf{H}_{\text{eq}} \in \mathcal{C}^{N_s \times N_s}$, $\mathbf{y} \in \mathcal{C}^{N_s \times 1}$, $\mathcal{X}$.
Parameters: $1 \leq c \leq N_s$.

Channel Ordering:

1. Find c worst columns of matrix $\mathbf{H}_{\text{eq}}$ for multiple subspace detection of one symbol.
2. If $c = 1$ keep this worst column at the end of the ordered matrix denoted as $\mathbf{H}_{\text{O}}$. Otherwise, i.e., if $c > 1$ put these c worst columns at the c first column of the ordered matrix for the sake of complexity saving for multiple QRD operations.
3. Then, place the $N_s - c$ best remaining columns in order at the remaining columns of $\mathbf{H}_{\text{O}}$.

Perform QRD on $\tilde{\mathbf{H}}_{\text{O}}$; $\tilde{\mathbf{H}}_{\text{O}} = \tilde{\mathbf{Q}}\mathbf{R}$.
$\mathbf{y}' = \mathbf{Q}^H\mathbf{y} = \mathbf{R}\mathbf{s} + \mathbf{w}$.
For all $i = 1, 2, \ldots, c$ perform the subspace detection on $\mathbf{y}'$(multiple detection)

1. $\mathbf{y}' = \mathbf{R}_{[i]}\mathbf{s}_{[i]} + \mathbf{r}_i s_{[i]} + \mathbf{w}$.
2. If $c > 1$ perform QRD on $[\mathbf{R}_{[i]}, \mathbf{r}_{[i]}]$; i.e., $[\mathbf{R}_{[i]}, \mathbf{r}_{[i]}] = \mathbf{Q}^{[i]}\mathbf{B}^{[i]}$.
3. For all points of the constellation:
   (a) If $c > 1$ obtain $\mathbf{y}'' = {\mathbf{Q}^{[i]}}^H\mathbf{y}'$.
   (b) Fix $\hat{s}_i^{(j)} = a_j$, $a_j \in \mathcal{X}$; $j = 1, ..., Q = |\mathcal{X}|$.
   (c) Perform SIC to detect $\hat{\mathbf{s}}_{[i]}^{(j)}$.
   (d) Calculate $d_k = \|\mathbf{y}' - \mathbf{R}_{[i]}\hat{\mathbf{s}}_{[i]}^{(j)} - \mathbf{r}_i\hat{s}_i^{(j)}\|^2$;$k = (i-1)|\mathcal{X}| + j$.

Find $d_{min} = \min_{k=1,..,c.Q} d_k$ and declare the reordered corresponding $\hat{\mathbf{s}} = [\hat{\mathbf{s}}_{[i]}^{(j)}, \hat{s}_i^{(j)}]$ as detected symbol vector.

Output: $\hat{\mathbf{s}}$

multiplications and $2(N_s - 1)$ additions. As a result, $\mathcal{O}(N_s^2)$ operations are required for all c matrices to be transformed into the upper triangular form. Moreover, the QRD along with the orderings can be obtained by a slight modification of the sorted QRD scheme in [104] with $\mathcal{O}(N_s^3)$ operations (See Table 5.1 for more details). Additionally, by rounding small off-diagonal entries of the equivalent channel to zero, sparse QRD techniques [83] can be used for further simplifications of the proposed detector.

Table 5.1: Computational Complexity In Terms Of FLOPs

Detector	Operation		Number of FLOPs	$N_s = 8$	
				64-QAM	
FSD	Preprocessing:	FSD ordering+ QRD [104]	$28N_s^3 + \frac{113}{3}N_s^2 + 24N_s - 60$	$475,630$	
		$\mathbf{R}_2\mathbf{x}_2; \mathbf{x}_2 \in \mathcal{X}^p$ for p worst column of $\mathbf{R}$	$(8N_sp - 2N_s).\lvert\mathcal{X}\rvert^p$	$(p=2)$	
	Detection:	$\mathbf{y}' = \mathbf{Q}^H\mathbf{y}$	$(8N_s - 2)N_s$	$1,647,088$	
		SIC	$(4(N_s-p)(N_s-p+1)+2p).\lvert\mathcal{X}\rvert^p$	$(p=2)$	
		Euclidean Distance Test	$(4N_s^2 + 4N_s - 2).\lvert\mathcal{X}\rvert^p$		
Multiple MMSE-SIC Subspace Detector (Algorithm 5.1)	Preprocessing:	Ordering + QRD [104]	$28N_s^3 + \frac{113}{3}N_s^2 + 24N_s - 60$		
		QRD of $\mathbf{R}^{[i]}; i = 1, \ldots, c$ (if $c > 1$)	$c.(6N_s^2 - 14N_s + 8)$	$19,951$	$26,935$
		$\mathbf{r}_i a_j; i = 1, \ldots, c$	$6N_s.c.\lvert\mathcal{X}\rvert$	$(c=1)$	$(c=3)$
	Detection:	$\mathbf{y}' = \mathbf{Q}^H\mathbf{y}$	$8N_s^2 - 2N_s$		
		$\mathbf{y}'' = \mathbf{Q}^{[i]\,H}\mathbf{y}'; i = 1, \ldots, c$ (if $c > 1$)	$(8N_s^2 - 2N_s).c$	$33,264$	$100,288$
		SIC	$(4N_s^2 - 4N_s + 2).c.\lvert\mathcal{X}\rvert$	$(c=1)$	$(c=3)$
		Euclidean Distance Test	$(4N_s^2 + 4N_s - 2).c.\lvert\mathcal{X}\rvert$		

5.4 Computational Complexity

In this section, we discuss the computational complexity of the proposed algorithm. The main operations in Algorithm 5.1 are QRD and ordering both requiring $\mathcal{O}(N_s^3)$ operations. For multiple subspace detection, Algorithm 5.1 requires an additional $\mathcal{O}(N_s^2)$ operation for multiple QRD. By bringing each column to the last column of the matrix, the resultant matrix has up to $N_s - 1$ elements below the main diagonal to get annihilated to have the upper triangular form. By using Givens rotations, each element requires $\mathcal{O}(N_s)$ operations. Hence, the matrix needs $\mathcal{O}(N_s^2)$ operation, and all c matrices require $\mathcal{O}(N_s^2)$ operations. Table 5.1 shows the detailed complexity of the proposed algorithm compared to the FSD scheme in terms of FLOPs. We assume six and two FLOPs per complex multiplication and addition, respectively. It is worth mentioning that the complexity of the proposed detector is $\mathcal{O}(N_s^3|\mathcal{X}|)$ for the worst-case scenario.

The computational complexity of the OMP hybrid beamforming technique or its variants is $\mathcal{O}(M^2 N_{\text{RF}} N_s)$ as it is mainly dominated by the successive orthogonal projections (pseudo-inverse) [18]. AM hybrid beamforming techniques require $\mathcal{O}(M N_{\text{RF}}^2 N_s)$ operations mainly due to partial SVD computations using rank revealing QRD [35]. One has to add the computation

complexity of calculating the optimal beamformers; this requires $\mathcal{O}(MNN_s)$ operations when using rank revealing QRD to calculate $\mathbf{U}$ and $\mathbf{V}$ [35]. Hence, the overall computational complexity of the hybrid beamforming processes is $\mathcal{O}(M(N + N_{\text{RF}}^2)N_s)$.

5.5 Simulation Results

In this section, we provide some simulation results to investigate the error performance of the proposed scheme for both OMP and AM hybrid beamforming techniques. For AM, we set the number of iterations to N_{RF} to have the same complexity order as OMP. We consider the clustered channel model given in (5.2). The azimuth angles of departure and arrival are i.i.d. with Laplacian distribution with means assumed to be uniformly distributed between 0 and 2π, and the angular spreads are 7.5° [6]. For our simulations, we use $L \in \{12, 30\}$, $N_s \in \{4, 6, 8\}$, with 64-QAM or 256-QAM modulation. In all simulations, we perform hybrid beamforming for both precoding and combining. We compare our proposed scheme with the optimal linear MMSE receiver in the sense of minimizing error [6, 87]. This optimal MMSE receiver is equivalent to our system model when $\mathbf{W}_{\text{opt}} = \bar{\mathbf{U}}_{N_s}$ is used for the combiner followed by MMSE applied on the output of the combiner.

Fig. 5.2 shows the error performance for $N_s = 4$, $N_{\text{RF}} = 6$ and $L = 12$. In this figure, the MMSE-SIC detector with the optimal V-BLAST ordering achieves a performance close to that of the SD for both AM and OMP hybrid algorithms. Note that the MMSE-SIC with V-BLAST ordering can be implemented by skipping the subspace detection in Algorithm 5.1. By using the proposed detection scheme, gains of about 2-dB and 4-dB are achieved, respectively, for OMP and AM at the BER of 10^{-4}.

Fig. 5.3 shows the error performance for the system with $Ns = N_{\text{RF}} = 4$ and $L = 30$. In this figure, Algorithm 5.1 with $c = 1$ achieves the ML error performance for OMP. However, for the AM technique, the system exhibits a good performance even with the MMSE detector, and the

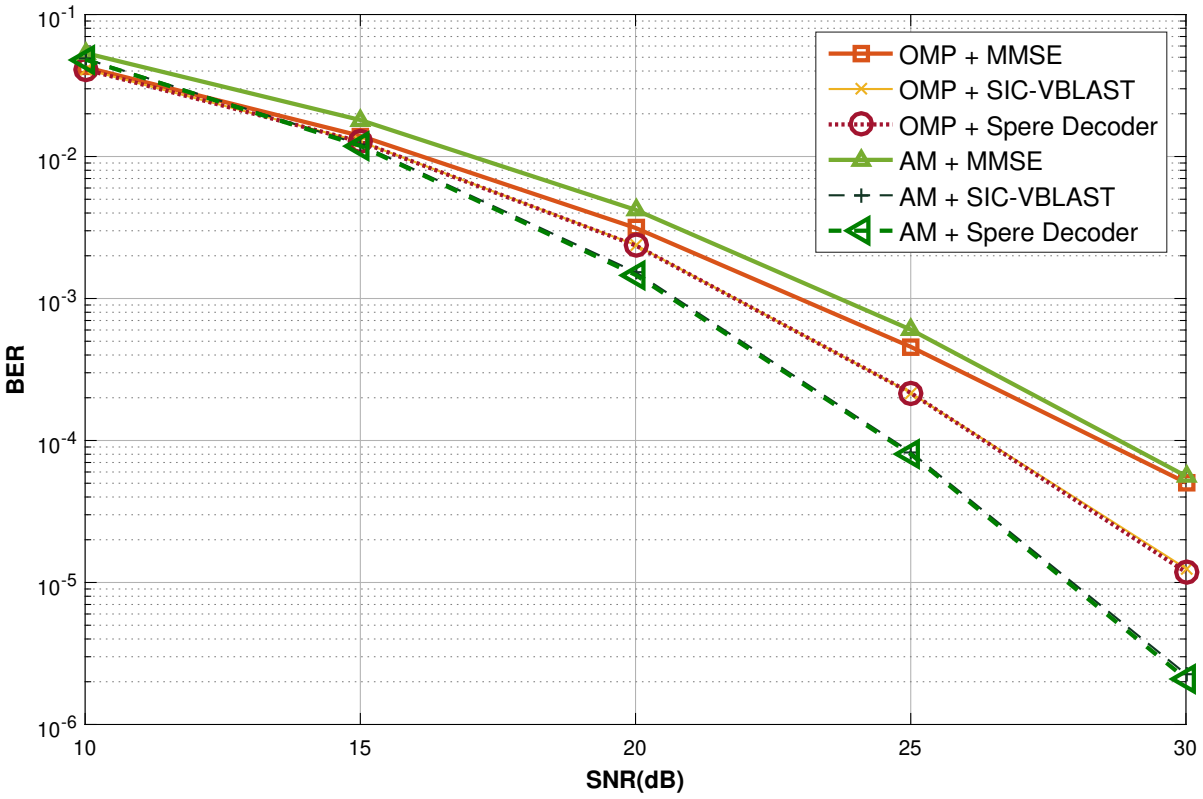


Figure 5.2: BER performance of hybrid precoding and beamforming with detection in Algorithm 5.1 for a 64×16 UPA mmWave MIMO system for $N_s = 4$, $N_{RF} = 6$, $L = 12$, and 256-QAM.

SD cannot add improvement.

In Fig. 5.4, $N_s = N_{\text{RF}} = 6$ and $L = 12$. The proposed algorithm significantly improves the error performance for both OMP and AM techniques. For AM, the SIC-VBLAST suffices to achieve the SD error performance while OMP needs Algorithm 5.1 with $c = 1$. Hence, by using the proposed scheme, the computational complexity is reduced compared to the FSD scheme for a conventional 6×6 MIMO system where $p = 2$ full-search levels are required to achieve the optimal error performance.

In Fig. 5.5, we increase the number of streams to $N_s = N_{\text{RF}} = 8$. The proposed algorithm with $c = 1$ improves the error performance for both OMP and AM techniques. $c = 1$ is sufficient for the AM technique to achieve the SD error performance. For OMP, however, by selecting

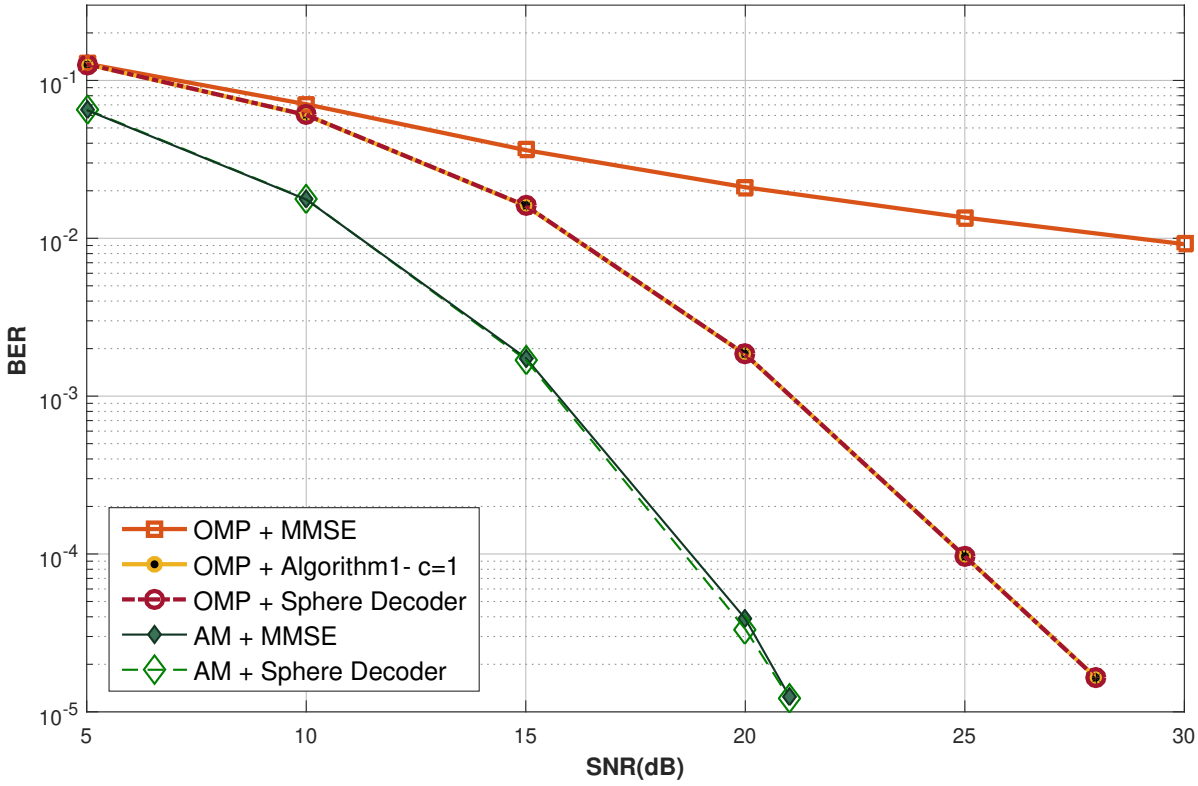


Figure 5.3: BER performance of hybrid precoding and beamforming with detection in Algorithm 5.1 for a 64×16 UPA mmWave MIMO system for $N_s = 4$, $N_{RF} = 4$, $L = 30$, and 256-QAM.

$c = 1$ in the proposed algorithm, there is about 1-dB gap at high SNRs. This gap is removed by performing multiple subspace detections be selecting $c = 3$. As Table 5.1 shows, for this system, the computational complexity of the proposed algorithm is significantly reduced, even with $c = 3$, compared to the FSD scheme.

From the simulation results, it is observed that the OMP technique has a poor error performance when L is large, or N_{RF} is minimal, i.e., $N_{\text{RF}} = N_s$. And, the error performance of the AM technique is highly susceptible to the parameter L such that for a large L, the gap between MMSE and the SD becomes smaller even for the edge-case $N_s = N_{\text{RF}}$ while for a small L, there will be a performance gap. Note that in Fig. 5.5, although L is large, a performance gap is observed for AM. For this system, since N_s is large, small eigenvalues can degrade the error

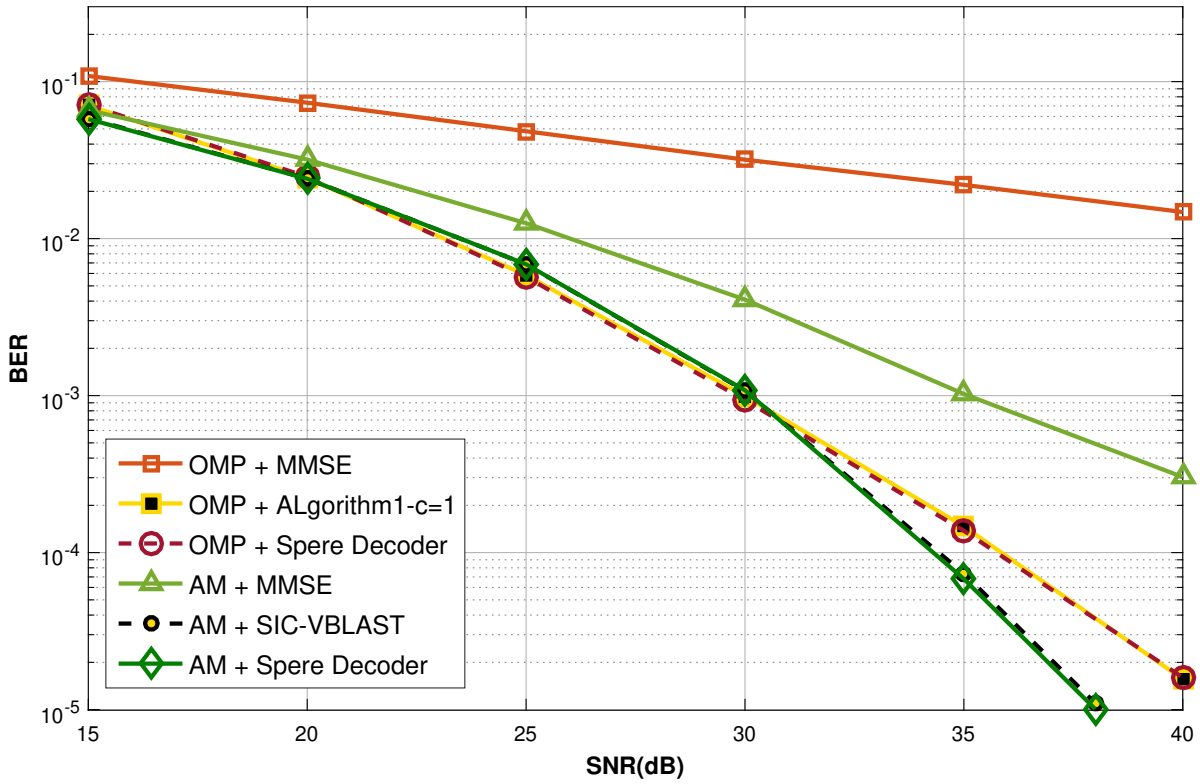


Figure 5.4: BER performance of hybrid precoding and beamforming with detection in Algorithm 5.1 for a 64×16 UPA mmWave MIMO system for $N_s = 6$, $N_{\mathrm{RF}} = 6$, $L = 12$, and 256-QAM.

performance. Hence, Algorithm 5.1 improves the BER.

Note that for some mmWave systems, we may achieve the optimal error performance with further complexity reduction if we consider smaller search size over the constellation, i.e., Q for the subspace detection in Algorithm 5.1.

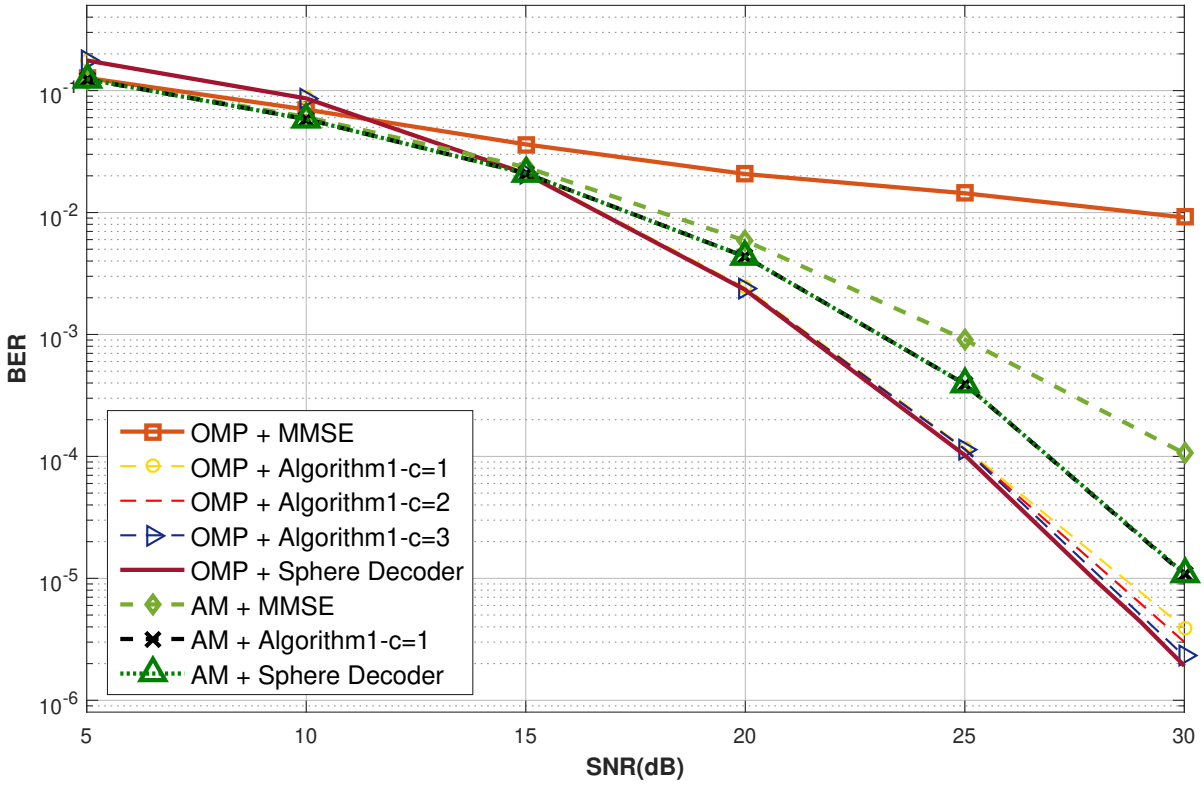


Figure 5.5: BER performance of hybrid precoding and beamforming with detection in Algorithm 5.1 for a 64×16 UPA mmWave MIMO system for $N_s = 8$, $N_{\mathrm{RF}} = 8$, $L = 30$, and 64-QAM.

5.6 Conclusion and Discussion

In this chapter, we proposed a detection scheme for hybrid beamforming mmWave massive MIMO systems in order to improve the error performance of the systems. The approximation error introduced due to the hybrid beamforming structures can largely degrade the error performance of the systems. As a result, first, we discussed the effect of the mmWave channel parameters and hybrid beamforming settings on the equivalent channel consisting of the precoder, mmWave channel, and combiner. Then, we proposed a signal detection scheme for the equivalent channel matrix. By considering the dimension of the equivalent channel matrix, the detection algorithms that we proposed for MIMO systems, especially the ones in Chapter 3, were utilized for the equivalent MIMO system. By using computer simulations, it was shown that the error performance of the system can be significantly improved with reduced computational complexity compared to the state-of-the-art MIMO detection schemes. Such complexity reductions were achieved by exploiting the available structure in the equivalent channel. Our proposed detector approaches the optimal ML error performance and has a fixed and polynomial complexity with parallel implementation.

The complexity of our proposed detector for the worst-case scenario is $\mathcal{O}(N_s^3|\mathcal{X}|)$. In the recently released specifications for mmWave massive MIMO systems, the number of data streams, i.e., N_s is up to 8, and the 16-QAM or 64-QAM modulation is used [81]. On the other hand, the overall computational complexity of OMP and AM hybrid beamforming techniques is $\mathcal{O}(MNN_s)$. Consequently, our proposed detector will not impose a burden on the system as its computational complexity order is the same as that of the hybrid beamforming techniques.

Chapter 6

Conclusions and Future Works

6.1 Conclusions

Signal detection is of great importance for MIMO systems as they have become an essential technology for wireless communication systems. However, designing signal detectors with low computational complexity and close to optimal performance becomes challenging as the system dimension increases.

We introduced tow efficient LRA conditional detectors. The proposed LRA conditional detectors exhibit close to optimal error performance and achieve maximum receive diversity. For quasi-static channels, the overall computational complexity of the proposed LRA conditional detectors is cubic in the system dimension and linear in the size of the constellation for small and moderate-scale MIMO systems. However, for large-scale MIMO systems, the computational complexity of the lattice reduction algorithms imposes a large burden on the system. To address this challenge, we proposed two MMSE-DFE subspace detectors for large-scale MIMO systems. In the proposed detectors, the search over all columns of the channel matrix along with an appropriate channel ordering has the advantages of improving the performance of the

hard-output version of the detector and refining the list of candidates for efficient implementation of the soft-output detector for MIMO systems with error-correcting codes. In particular, it was shown that the hard-output detectors exhibit near-optimal error performance with a computational complexity that is cubic in system dimension and linear in the constellation size. Moreover, the soft-output system is comparable to that of the LSD but with a much smaller list size, and hence, smaller complexity than the latter. However, the proposed MMSE-DFE detectors may not achieve the full receive diversity for some MIMO systems. A subspace detector with PLR was proposed in order to improve the diversity order of subspace detectors restricted to one level of full search over the constellation points. The proposed subspace detector with PLR is of interest, especially for quasi-static fading channels as the PLR algorithm is performed in the preprocessing stage which is shared among many channel realizations. All proposed detectors discussed above can be used for point-to-point and the uplink of multipoint-to-point MIMO systems that the system dimensions are comparable.

For the uplink of massive MIMO systems in which the number of antennas at the BS is much larger than the number of served UTs, we proposed a low-complexity TPE-based detector in order to avoid the expensive matrix-matrix multiplication and matrix inversion operations associated with the ZF and MMSE detectors. The proposed TPE-based detector achieves the error performance of the ZF and MMSE detectors while its computational complexity is proportional to the system dimensions.

Massive MIMO is also a key technology for communication systems in the mmWave band. A viable realization for such systems is hybrid beamforming in which the implementation cost is reduced by limiting the number of expensive RF chains at the cost of performance losses, which can be huge in some cases. We proposed a detection scheme for such systems to compensate for the error performance degradation due to the approximation errors associated with the hybrid beamforming techniques.

In summary, in this thesis, we proposed low-complexity near-optimal detectors for different applications of MIMO systems. By exploiting appropriate techniques, we designed novel algorithms to save in computational complexity and provide practical yet efficient detectors.

6.2 Future Works

Extra-Large MIMO Systems

One possible future direction for the works studied in this thesis is to extend the proposed detectors for extra large-scale MIMO (XL-MIMO) systems. XL-MIMO is a promising technology for the 6th generation of wireless communications [10] where extra large-scale antenna arrays are integrated into large infrastructures like the walls of stadiums or large shopping malls [70]. In such systems, as the dimensions of the system become very large, non-stationarities start to appear [25]. This results in the appearance of a visibility region (VR) where users can see only a portion of the antenna array at the BS. Each user has its specific VR, and depending on the position of the users and their surrounding environment, their VRs can be partially or completely overlapped [105]. Users' signals are detected using subarrays or distributed units with local processors that are connected to a central unit. When there are overlapped VRs, multiple users are served by consecutive subarrays which will create inter-user interference. Considering the non-stationary properties, signal detection techniques are needed to be adapted for such systems. Since the dimensions of the system become very large in XL-MIMO systems, designing signal detectors at the receiver with low computational complexity is a challenging issue especially when the system is serving a large number of users. Our proposed TPE-based detector in Chapter 4 is a promising candidate for such systems as its computational complexity scales linearly with the number of BS antennas and users.

Application of Machine/Deep Learning

One promising future research direction for the works introduced in this thesis is the application of the state-of-the-art deep learning (DL) techniques to assist with a further reduction of the computational complexity or the improvement of the techniques associated with our proposed MIMO detectors. Model-driven DL approaches are commonly used for designing iterative algorithms using NNs where an iterative algorithm is unfolded such that each layer of the NN represents one iteration of the algorithm. Then, the parameters are optimized and additional parameters are learned by DL in order to improve the performance of the system. In our proposed detectors in this thesis, we have used several iterative algorithms, including lattice reduction, channel ordering, SIC, and TPE. Hence, a model-driven approach can be deployed for the implementation of these iterative techniques.

Joint Detection and Hybrid Beamforming

We proposed a detection scheme in order to improve the error performance of mmWave massive MIMO systems. In the literature, hybrid beamforming structures are designed to be as close as possible to the optimal fully-digital solutions. For such designs, we proposed a detection scheme by considering the equivalent channel consisting of the precoder, mmWave channel, and combiner. Hence, in our proposed detection scheme, the hybrid beamforming structure and detection are separately designed. One future research direction for this line of work is to consider the joint design of hybrid beamforming and detection for minimizing the BER of the system.

References

[1] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger. Closest point search in lattices. *IEEE Transactions on Information Theory*, 48(8):2201–2214, 2002.

[2] I. Ahmed, H. Khammari, A. Shahid, A. Musa, K. S. Kim, E. De Poorter, and I. Moerman. A survey on hybrid beamforming techniques in 5G: Architecture and system model perspectives. *IEEE Communications Surveys Tutorials*, 20(4):3060–3097, 2018.

[3] M. A. M. Albreem, A. A. El-Saleh, and M. Juntti. On approximate matrix inversion methods for massive MIMO detectors. In *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, pages 1–6, 2019.

[4] M. Alouzi, F. Chan, and C. D'Amours. Sphere decoding for millimeter wave massive MIMO. In *IEEE Vehicular Technology Conference*, pages 1–6, 2019.

[5] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang. What will 5G be? *IEEE Journal on Selected Areas in Communications*, 32(6):1065–1082, 2014.

[6] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath. Spatially sparse precoding in millimeter wave MIMO systems. *IEEE Transactions on Wireless Communications*, 13(3):1499–1513, Mar. 2014.

[7] L. G. Barbero and J. S. Thompson. Fixing the complexity of the sphere decoder for MIMO detection. *IEEE Transactions on Wireless Communications*, 7(6):2131–2142, 2008.

[8] J. Benesty, Y. Huang, and J. Chen. A fast recursive algorithm for optimum sequential signal detection in a BLAST system. *IEEE Transactions on Signal Processing*, 51(7):1722–1730, 2003.

[9] A. Benzin, G. Caire, Y. Shadmi, and A. M. Tulino. Low-complexity truncated polynomial expansion DL precoders and UL receivers for massive MIMO in correlated channels. *IEEE Trans. Wireless Commun.*, 18(2):1069–1084, 2019.

[10] E. Björnson, L. Sanguinetti, H. Wymeersch, J. Hoydis, and T. L. Marzetta. Massive MIMO is a reality – what is next? five promising research directions for antenna arrays, 2019.

[11] J. Boutros, N. Gresset, L. Brunel, and M. Fossorier. Soft-input soft-output lattice sphere decoder for linear channels. In *IEEE Global Communications Conference*, volume 3, pages 1583–1587, Dec. 2003.

[12] M. Chang and W. Chang. Maximum-Likelihood detection for MIMO systems based on differential metrics. *IEEE Transactions on Signal Processing*, 65(14):3718–3732, July 2017.

[13] J. Chen. A low complexity data detection algorithm for uplink multiuser massive MIMO systems. *IEEE Journal on Selected Areas in Communications*, 35(8):1701–1714, 2017.

[14] Y. Chen and S. ten Brink. Near-capacity MIMO subspace detection. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications*, pages 1733–1737, Sep. 2011.

[15] A. Chockalingam and B. Sundar Rajan. *Large MIMO systems*, page 16–24. Cambridge University Press, 2014.

[16] J. Choi, S. R. Kim, and I. Choi. Statistical eigen-beamforming with selection diversity for spatially correlated OFDM downlink. *IEEE Trans. Veh. Technol.*, 56(5):2931–2940, 2007.

[17] H. Cohen. *A Course in Computational Algebraic Number Theory*. Springer-Verlag New York, Inc., New York, NY, USA, 1993.

[18] P. Courrieu. Fast computation of Moore-Penrose inverse matrices. *Neural Information Processing - Letters and Reviews, KAIST Press*, 8(2):25–29,, Aug. 2005.

[19] T. Cui and C. Tellambura. An efficient generalized sphere decoder for rank-deficient MIMO systems. *IEEE Communications Letters*, 9(5):423–425, 2005.

[20] Da-Shan Shiu, G. J. Foschini, M. J. Gans, and J. M. Kahn. Fading correlation and its effect on the capacity of multielement antenna systems. *IEEE Trans. Commun.*, 48(3):502–513, 2000.

[21] L. Dai, X. Gao, X. Su, S. Han, C. I, and Z. Wang. Low-complexity soft-output signal detection based on Gauss–Seidel method for uplink multiuser large-scale MIMO systems. *IEEE Transactions on Vehicular Technology*, 64(10):4839–4845, 2015.

[22] M. O. Damen, H. El Gamal, and G. Caire. On maximum-likelihood detection and the search for the closest lattice point. *IEEE Transactions on Information Theory*, 49(10):2389–2402, 2003.

[23] M. O. Damen, H. El Gamal, and G. Caire. MMSE-GDFE lattice decoding for under-determined linear channels. In *Conference on Information Sciences and Systems*, pages 1–5, Mar. 2004.

[24] M.O. Damen, K.A. Meraim, and S. Burykh. Iterative QR detection for BLAST. In *Allerton Conference of Communications*, Oct 2000.

[25] E. De Carvalho, A. Ali, A. Amiri, M. Angjelichinoski, and R. W. Heath. Non-stationarities in extra-large scale massive MIMO, 2019.

[26] M. Debbah, B. Muquet, M. de Courville, M. Muck, S. Simoens, and P. Loubaton. A MMSE successive interference cancellation scheme for a new adjustable hybrid spread OFDM system. In *IEEE Vehicular Technology Conference*, volume 2, pages 745–749, 2000.

[27] A. Elghariani and M. Zoltowski. Low complexity detection algorithms in large-scale MIMO systems. *IEEE Transactions on Wireless Communications*, 15(3):1689–1702, March 2016.

[28] G. J. Foschini. Layered space-time architecture for wireless communication in a fading environment when using multi-element antennas. *Bell Labs Technical Journal*, 1(2):41–59, 1996.

[29] Y. H. Gan, C. Ling, and W. H. Mow. Complex lattice reduction algorithm for low-complexity full-diversity MIMO detection. *IEEE Transactions on Signal Processing*, 57(7):2701–2710, Jul. 2009.

[30] X. Gao, L. Dai, Y. Ma, and Z. Wang. Low-complexity near-optimal signal detection for uplink large-scale MIMO systems. *Electronics Letters*, 50(18):1326–1328, 2014.

[31] X. Gao, L. Dai, C. Yuen, and Y. Zhang. Low-complexity MMSE signal detection based on richardson method for large-scale MIMO systems. In *IEEE Vehicular Technology Conference*, pages 1–5, 2014.

[32] X. Gao, L. Dai, J. Zhang, S. Han, and C. L. I. Capacity-approaching linear precoding with low-complexity for large-scale MIMO systems. In *IEEE International Conference on Communications*, pages 1577–1582, 2015.

[33] G. D. Golden, C. J. Foschini, R. A. Valenzuela, and P. W. Wolniansky. Detection algorithm and initial laboratory results using V-BLAST space-time communication architecture. *Electronics Letters*, 35(1):14–16, 1999.

[34] G. H. Golub and C. F. Van Loan. *Matrix Computations*. Johns Hopkins Univ. Press, Baltimore, MD, USA, 3rd edition, 1996.

[35] M. Gu and S. C. Eisenstat. Efficient algorithms for computing a strong rank-revealing QR factorization. *SIAM Journal on Scientific Computing*, 17(4):848–869, 1996.

[36] H. Najafi. *Applications of Lattices over Wireless Channels*. PhD thesis, 2012.

[37] B. Hassibi. An efficient square-root algorithm for BLAST. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, volume 2, pages 737–740, 2000.

[38] B. Hassibi and H. Vikalo. On the sphere-decoding algorithm. I. expected complexity. *IEEE Transactions on Signal Processing*, 53(8):2806–2818, 2005.

[39] R. W. Heath, N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed. An overview of signal processing techniques for millimeter wave MIMO systems. *IEEE Journal of Selected Topics in Signal Processing*, 10(3):436–453, 2016.

[40] A. Hedayat and A. Nosratinia. Outage and diversity of linear receivers in flat-fading MIMO channels. *IEEE Transaction on Signal Processing*, 55(12):5868–5873, 2007.

[41] B. M. Hochwald and S. ten Brink. Achieving near-capacity on a multiple-antenna channe. *IEEE Transactions on Communications*, (3):389–399.

[42] J. Hoydis, M. Debbah, and M. Kobayashi. Asymptotic moments for interference mitigation in correlated fading channels. In *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pages 2796–2800, 2011.

[43] J. Hoydis, S. ten Brink, and M. Debbah. Massive MIMO in the UL/DL of cellular networks: How many antennas do we need? *IEEE Journal on Selected Areas in Communications*, 31(2):160–171, 2013.

[44] S. Hu and F. Rusek. A soft-output MIMO detector with achievable information rate based partial marginalization. *IEEE Transactions on Signal Processing*, 65(6):1622–1637, March 2017.

[45] Y. Hu, Z. Wang, X. Gaol, and J. Ning. Low-complexity signal detection using CG method for uplink large-scale MIMO systems. In *IEEE International Conference on Communication Systems*, pages 477–481, 2014.

[46] M. K. Izadinasab and O. Damen. Transmit antenna selection for rank-deficient spatial multiplexing systems. In *2018 29th Biennial Symposium on Communications (BSC)*, pages 1–5, 2018.

[47] M. K. Izadinasab and O. Damen. Near-optimal MIMO detectors based on MMSE-GDFE and conditional detection. In *IEEE International Conference on Communications (ICC)*, pages 1–6, 2019.

[48] M. K. Izadinasab and O. Damen. Partial lattice reduction and subspace detection of large-scale MIMO systems. In *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, pages 1–6, 2019.

[49] M. K. Izadinasab and O. Damen. Reduced complexity ordering in subspace MIMO detection algorithms. In *2019 16th Canadian Workshop on Information Theory (CWIT)*, pages 1–5, 2019.

[50] M. K. Izadinasab and O. Damen. Bridging the gap between MMSE-DFE and optimal detection of MIMO systems. *IEEE Transactions on Communications*, 68(1):220–231, 2020.

[51] M. K. Izadinasab, O. Damen, and H. Najafi. Efficient lattice-reduction-aided conditional detection for MIMO systems. In *2017 15th Canadian Workshop on Information Theory (CWIT)*, pages 1–5, 2017.

[52] J. Jaldén, L. G. Barbero, B. Ottersten, and J. S. Thompson. The error probability of the fixed-complexity sphere decoder. *IEEE Transactions on Signal Processing*, 57:2711–2720, 2009.

[53] J. Jaldén and P. Elia. DMT optimality of LR-aided linear decoders for a general class of channels, lattice designs, and system models. *IEEE Transactions on Information Theory*, 56(10):4765–4780, 2010.

[54] J. Jaldén and B. Ottersten. On the maximal diversity order of spatial multiplexing with transmit antenna selection. *IEEE Transactions on Information Theory*, 53(11):4273–4276, Nov 2007.

[55] C. Jeon, R. Ghods, A. Maleki, and C. Studer. Optimality of large MIMO detection via approximate message passing. In *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pages 1227–1231, 2015.

[56] J. Jin, C. Xiao, W. Chen, and Y. Wu. Channel-statistics-based hybrid precoding for millimeter-wave MIMO systems with dynamic subarrays. *IEEE Transactions on Communications*, 67(6):3991–4003, Jun. 2019.

[57] A. Kammoun, A. Müller, E. Björnson, and M. Debbah. Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO systems. *IEEE J. Sel. Topics Signal Process.*, 8(5):861–875, 2014.

[58] D. N. C Tse L. Zheng. Diversity and multiplexing : A fundamental tradeoff in multiple-antenna channels. *IEEE Transactions on Information Theory*, 49(5):1073–1096, 2003.

[59] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta. Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.*, 52(2):186–195, 2014.

[60] E. G. Larsson and J. Jaldén. Fixed-complexity soft MIMO detection via partial marginalization. *IEEE Transactions on Signal Processing*, 56(8):3397–3407, Aug. 2008.

[61] A. K. Lenstra, H. W. Lenstra, and L. Lovász. Factoring polynomials with rational coefficients. *Mathematische Annalen", year="1982*, 261(4):515–534.

[62] P. Li and R. D. Murch. Multiple output selection-LAS algorithm in large MIMO systems. *IEEE Communications Letters*, 14(5):399–401, May 2010.

[63] Y. Li and Z. Luo. Parallel detection for V-BLAST system. In *IEEE International Conference on Communications*, volume 1, pages 340–344, 2002.

[64] C. Ling, W. H. Mow, and L. Gan. Dual-lattice ordering and partial lattice reduction for SIC-based MIMO detection. *IEEE Journal of Selected Topics in Signal Processing*, 3(6):975–985, 2009.

[65] L. Liu, G. Peng, and S. Wei. *Massive MIMO Detection Algorithm and VLSI Architecture*. Springer Singapore, 2019.

[66] X. Liu, W. Zou, and S. Chen. Joint design of analog and digital codebooks for hybrid precoding in millimeter wave massive MIMO systems. *IEEE Access*, 6:69818–69825, 2018.

[67] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang. An overview of massive MIMO: Benefits and challenges. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):742–758, 2014.

[68] A. Maaref and S. Aïssa. Eigenvalue distributions of Wishart-type random matrices with application to the performance analysis of MIMO MRC systems. *IEEE Transactions on Wireless Communications*, 6(7):2678–2689, 2007.

[69] R. K. Mallik. The pseudo-Wishart distribution and its application to MIMO systems. *IEEE Transactions on Information Theory*, 49(10):2761–2769, 2003.

[70] A. O. Martínez, E. De Carvalho, and J. Nielsen. Towards very large aperture massive MIMO: A measurement based study. In *2014 IEEE Globecom Workshops*, pages 281–286, 2014.

[71] T. L. Marzetta. Non-cooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Transactions on Wireless Communications*, 9(11):3590–3600, 2010.

[72] A. H. Mehana and A. Nosratinia. Diversity of MMSE MIMO receivers. *IEEE Transaction on Information Theory*, 58(11):6788–6805, 2012.

[73] A. Müller, A. Kammoun, E. Björnson, and M. Debbah. Efficient linear precoding for massive MIMO systems using truncated polynomial expansion. In *Proc. IEEE Sensor Array and Multichannel Signal Process. Workshop (SAM)*, pages 273–276, 2014.

[74] H. Najafi and M. O. Damen. Lattice-reduction-aided conditional detection for MIMO systems. *IEEE Transactions on Communications*, 62(11):3864–3873, 2014.

[75] H. Najafi, M. E. D. Jafari, and M. O. Damen. Adaptive soft-output detection in MIMO systems. In *Allerton Conference on Communication, Control, and Computing*, Sept 2008.

[76] H. Najafi, M. E. D. Jafari, and M. O. Damen. On adaptive lattice reduction over correlated fading channels. *IEEE Transactions on Communications*, 59(5):1224–1227, 2011.

[77] L. G. Ordóñez, D. P. Palomar, A. Pagès-zamora, and J. R. Fonollosa. High-SNR analytical performance of spatial multiplexing MIMO systems with CSI. *IEEE Transactions on Signal Processing*, 55(11):5447–5463, 2007.

[78] D. Persson and E. G. Larsson. Partial marginalization soft MIMO detection with higher order constellations. *IEEE Transactions on Signal Processing*, 59(1):453–458, Jan 2011.

[79] C. Qian, J. Wu, Y. R. Zheng, and Z. Wang. Low complexity detection algorithm for under-determined MIMO systems. *IEEE International Conference on Communications*, pages 5580–5585, 2014.

[80] R. Rajashekar and L. Hanzo. Iterative matrix decomposition aided block diagonalization for mm-wave multiuser MIMO systems. *IEEE Transactions on Wireless Communications*, 16(3):1372–1384, Mar. 2017.

[81] 3GPP TS 38.211 V16.0.0. New radio; physical channels and modulation. Dec. 2019.

[82] P. Robertson, E. Villebrun, and P. Hoeher. A comparison of optimal and sub-optimal MAP decoding algorithms operating in the log domain. In *IEEE International Conference on Communications*, June 1995.

[83] T. H. Robey and D. L. Sulsky. Row ordering for a sparse QR decomposition. *SIAM Journal on Matrix Analysis and Applications*, 15(4):1208–1225, Oct. 1994.

[84] E. Juárez Ruiz, R. Cortes Maldonado, and F. Pérez Rodríguez. Relationship between the inverses of a matrix and a submatrix. *Computación y Sistemas*, 20, 2016.

[85] 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.*, 30(1):40–60, 2013.

[86] H. Sarieddeen, M. M. Mansour, and A. Chehab. Large MIMO detection schemes based on channel puncturing: Performance and complexity analysis. *IEEE Transactions on Communications*, 66(6):2421–2436, June 2018.

[87] A. Scaglione, P. Stoica, S. Barbarossa, G. B. Giannakis, and H. Sampath. Optimal designs for space-time linear precoders and decoders. *IEEE Transactions on Signal Processing*, 50(5):1051–1064, May 2002.

[88] G. M. A. Sessler and F. K. Jondral. Rapidly converging polynomial expansion multiuser detector with low complexity for CDMA systems. *Electronics Letters*, 38(17):997–998, 2002.

[89] G. M. A. Sessler and F. K. Jondral. Low-complexity polynomial expansion multiuser detector for CDMA systems. *IEEE Transactions on Vehicular Technology*, 54(4):1379–1391, 2005.

[90] S. Sirianunpiboon, Y. Wu, A. R. Calderbank, and S. D. Howard. Fast optimal decoding of multiplexed orthogonal designs by conditional optimization. *IEEE Transactions on Information Theory*, 56(3):1106–1113, 2010.

[91] M. Siti and M. P. Fitz. A novel soft-output layered orthogonal lattice detector for multiple antenna communications. In *IEEE International Conference on Communications*, volume 4, pages 1686–1691, June 2006.

[92] W. Song, X. Chen, L. Wang, and X. Lu. Joint conjugate gradient and Jacobi iteration based low complexity precoding for massive MIMO systems. In *IEEE International Conference on Communications in China*, pages 1–5, 2016.

[93] N. Srinidhi, S. K. Mohammed, A. Chockalingam, and B. Sundar Rajan. Low-complexity near-ML decoding of large non-orthogonal STBCs using reactive tabu search. In *IEEE International Symposium on Information Theory*, pages 1993–1997, June 2009.

[94] M. Taherzadeh, A. Mobasher, and A. K. Khandani. LLL reduction achieves the receive diversity in MIMO decoding. *IEEE Transactions on Information Theory*, 53(12):4801–4805, Dec 2007.

[95] V. Tarokh, N. Seshadri, and A. R. Calderbank. Space-time codes for high data rate wireless communication: performance criterion and code construction. *IEEE Transactions on Information Theory*, 44(2):744–765, 1998.

[96] M. L. Walker, J. Tao, J. Wu, and Y. R. Zheng. Low complexity turbo detection of coded under-determined MIMO systems. *IEEE International Conference on Communications*, 2011.

[97] D. W. Waters and J. R. Barry. The Chase family of detection algorithms for multiple-input multiple-output channels. *IEEE Transactions on Signal Processing*, 56(2):739–747, Feb 2008.

[98] C. Wen, C. Wang, S. Jin, K. Wong, and P. Ting. Bayes-optimal joint channel-and-data estimation for massive MIMO with low-precision ADCs. *IEEE Trans. Signal Process.*, 64(10):2541–2556, 2016.

[99] C. Windpassinger and R. F. H. Fischer. Low-complexity near-maximum-likelihood detection and precoding for MIMO systems using lattice reduction. In *IEEE Information Theory Workshop*, pages 345–348, 2003.

[100] M. Wu, C. Dick, J. R. Cavallaro, and C. Studer. High-throughput data detection for massive MU-MIMO-OFDM using coordinate descent. *IEEE Trans. Circuits Syst. I, Reg. Papers*, 63(12):2357–2367, 2016.

[101] M. Wu, B. Yin, A. Vosoughi, C. Studer, J. R. Cavallaro, and C. Dick. Approximate matrix inversion for high-throughput data detection in the large-scale MIMO uplink. In *IEEE International Symposium on Circuits and Systems*, pages 2155–2158, 2013.

[102] M. Wu, B. Yin, G. Wang, C. Dick, J. R. Cavallaro, and C. Studer. Large-scale MIMO detection for 3GPP LTE: Algorithms and FPGA implementations. *IEEE Journal of Selected Topics in Signal Processing*, 8(5):916–929, 2014.

[103] S. Wu, L. Kuang, Z. Ni, J. Lu, D. Huang, and Q. Guo. Low-complexity iterative detection for large-scale multiuser MIMO-OFDM systems using approximate message passing. *IEEE J. Sel. Topics Signal Process.*, 8(5):902–915, 2014.

[104] D. Wübben, R. Böhnke, V. Kühn, and K. D. Kammeyer. MMSE extension of V-BLAST based on sorted QR decomposition. In *IEEE Vehicular Technology Conference*, volume 1, pages 508–512, Oct. 2003.

[105] X. Yang, F. Cao, M. Matthaiou, and S. Jin. On the uplink transmission of multi-user extra-large scale massive MIMO systems, 2019.

[106] H. Yao and G. W. Wornell. Lattice-reduction-aided detectors for MIMO communication systems. In *IEEE Global Communications Conference*, volume 1, pages 424–428, 2002.

[107] B. Yin, M. Wu, J. R. Cavallaro, and C. Studer. Conjugate gradient-based soft-output detection and precoding in massive MIMO systems. In *2014 IEEE Global Communications Conference*, pages 3696–3701, 2014.

[108] B. Yin, M. Wu, C. Studer, J. R. Cavallaro, and C. Dick. Implementation trade-offs for linear detection in large-scale MIMO systems. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pages 2679–2683, 2013.

[109] X. Yu, J. Shen, J. Zhang, and K. B. Letaief. Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems. *IEEE Journal of Selected Topics in Signal Processing*, 10(3):485–500, Apr. 2016.

[110] S. Zarei, W. Gerstacker, and R. Schober. Low-complexity widely-linear precoding for downlink large-scale MU-MISO systems. *IEEE Commun. Lett.*, 19(4):665–668, 2015.

[111] H. Zhang, H. Dai, Q. Zhou, and B. L. Hughes. On the diversity order of spatial multiplexing systems with transmit antenna selection: A geometrical approach. *IEEE Transactions on Information Theory*, 52(12):5297–5311, 2006.

[112] T. Zhang, C. Wen, S. Jin, and T. Jiang. Mixed-ADC massive MIMO detectors: Performance analysis and design optimization. *IEEE Trans. Wireless Commun.*, 15(11):7738–7752, 2016.

[113] J. Zhou, Y. Ye, and J. Hu. Biased MMSE soft-output detection based on jacobi method in massive MIMO. In *IEEE International Conference on Computational Problem-Solving*, pages 442–445, 2014.

[114] H. Zhu, Z. Lei, and F. P. S. Chin. An improved square-root algorithm for BLAST. *IEEE Signal Processing Letters*, 11(9):772–775, 2004.

APPENDICES

Appendix A

The LLL Algorithm

The following table contains the LLL algorithm [17,61] for a real-valued basis $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_M]$. Two sub-routines, i.e., SIZE_REDUCE and SWAP, are called during the execution of the algorithm. The resultant reduced basis $\mathbf{B}$ from this algorithm satisfy the following conditions

$$|p_{i,j}| \leq 1/2 \quad \text{for} \quad 1 \leq j < i \leq M, \tag{A.1}$$

$$\| \hat{\mathbf{h}}_i + p_{i,i-1}\hat{\mathbf{h}}_{i-1} \|^2 \geq \kappa \| \hat{\mathbf{h}}_{i-1} \|^2 \quad 1 < i \leq M, \tag{A.2}$$

where $1/4 < \kappa \leq 1$ for real-valued LLL reduction and also $\mathbf{H} = \hat{\mathbf{H}}\mathbf{P}^T$ is the Gram-Schmidt orthogonalization where $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, ..., \hat{\mathbf{h}}_M]$ and $\mathbf{P} = [p_{i,j}]$ is a lower triangular matrix. In is worth noting that a larger κ results in a stronger but slower reduction. This algorithm has been extended for complex-valued lattices in [29] where the following condition

$$|\Re\{(p_{i,j}\}| \leq 1/2, \quad |\Im\{p_{i,j}\}| \leq 1/2, \quad \text{for} \quad 1 \leq j < i \leq M, \tag{A.3}$$

$$\| \hat{\mathbf{h}}_i + p_{i,i-1}\hat{\mathbf{h}}_{i-1} \|^2 \geq \kappa \| \hat{\mathbf{h}}_{i-1} \|^2 \quad 1 < i \leq M, \tag{A.4}$$

are satisfied by the LLL-reduced complex-valued basis where $1/2 < \kappa \leq 1$.

Algorithm A.1 The LLL Algorithm

Input: basis $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_M]$, $1/4 < \kappa \leq 1$ (commonly $\kappa = 3/4$)
Output: LLL-reduced basis $\mathbf{B}$, transformation matrix $\mathbf{U}$

Step 1: Initialization

$k \leftarrow 2$, $k_{max} \leftarrow 1$, $\hat{\mathbf{h}}_1 \leftarrow \mathbf{h}_1$, $\mathcal{H}_1 \leftarrow \langle \mathbf{h}_1 \mathbf{h}_1 \rangle$, and $\mathbf{U} \leftarrow \mathbf{I}_M$

Step 2: Incremental Gram-Schmidt

if $k \leq k_{max}$ **then**
    go to Step 3
else
    $k_{max} \leftarrow k$
    $\hat{\mathbf{h}}_k \leftarrow \mathbf{h}_k$
    for $j = 1, \cdots, k-1$ **do**
        $p_{k,j} \leftarrow \langle \mathbf{h}_k, \hat{\mathbf{h}}_j \rangle / \mathcal{H}_j$
        $\hat{\mathbf{h}}_k \leftarrow \hat{\mathbf{h}}_k - p_{k,j} \hat{\mathbf{h}}_j$
    end for
    $\mathcal{H}_k \leftarrow \langle \hat{\mathbf{h}}_k, \hat{\mathbf{h}}_k \rangle$
end if

Step 3: Test LLL Condition

$\mathbf{H}$, $\mathbf{U}$, $\mathcal{H}$, $[p]_{i,j} \leftarrow$ SIZE_REDUCE($\mathbf{H}, \mathbf{U}, [p]_{i,j}, k, k-1$)
if $\mathcal{H}_k < (\kappa - p_{k,k-1}^2)\mathcal{H}_{k-1}$ **then**
    $\mathbf{H}$, $\mathbf{U}$, $\hat{\mathbf{H}}$, $[p]_{i,j}$, $\mathcal{H}_k$, $\mathcal{H}_{k-1} \leftarrow$ SWAP($\mathbf{H}$, $\mathbf{U}$, $\hat{\mathbf{H}}$, $[p]_{i,j}$, $\mathcal{H}_k$, $\mathcal{H}_{k-1}$, k)
    $k \leftarrow max(2, k-1)$
    go to Step 3
else
    for $l = k-2, k-3, \cdots, 1$ **do**
        $\mathbf{H}$, $\mathbf{U}$, $\mathcal{H}$, $[p]_{i,j} \leftarrow$ SIZE_REDUCE($\mathbf{H}, \mathbf{U}, [p]_{i,j}, k, l$)
    end for
    $k \leftarrow k + 1$
end if

Step 4: Termination

if $k \leq M$ **then**
    go to Step 2
else
    return $\mathbf{B} = \mathbf{H}$, $\mathbf{U}$
end if

Algorithm A.2 Sub-routine SIZE_REDUCE

Input: $\mathbf{H}$, $\mathbf{U}$, $[p]_{i,j}$, indices k, l
Output: $\mathbf{H}$, $\mathbf{U}$, $[p]_{i,j}$

if $|p_{k,l}| > 1/2$ **then**
  $q \leftarrow \lceil p_{k,l} \rfloor$
  $\mathbf{h}_k \leftarrow \mathbf{h}_k - q\mathbf{h}_l$
  $\mathbf{u}_k \leftarrow \mathbf{u}_k - q\mathbf{u}_l$
  $p_{k,l} \leftarrow p_{k,l} - q$
  for $i = 1, \cdots, l-1$ **do**
    $p_{k,i} \leftarrow p_{k,i} - qp_{l,i}$
  end for
end if

Algorithm A.3 Sub-routine SWAP

Input: $\mathbf{H}$, $\mathbf{U}$, $\hat{\mathbf{H}}$, $[p]_{i,j}$, $\mathcal{H}_k$, $\mathcal{H}_{k-1}$, index k
Output: $\mathbf{H}$, $\mathbf{U}$, $\hat{\mathbf{H}}$, $[p]_{i,j}$, $\mathcal{H}_k$, $\mathcal{H}_{k-1}$

exchange the vectors $\mathbf{h}_k$ and $\mathbf{h}_{k-1}$, $\mathbf{u}_k$ and $\mathbf{u}_{k-1}$
if $k > 2$ **then**
  for $j = 1, \cdots, k-2$ **do**
    exchange $p_{k,j}$ with $p_{k-1,j}$
  end for
end if
$p \leftarrow p_{k,k-1}$
$\mathcal{H} \leftarrow \mathcal{H}_k + p^2\mathcal{H}_{k-1}$
$p_{k,k-1} \leftarrow p.\mathcal{H}_{k-1}/\mathcal{H}$
$\mathbf{h} \leftarrow \hat{\mathbf{h}}_{k-1}$
$\hat{\mathbf{h}}_{k-1} \leftarrow \hat{\mathbf{h}}_k + p.\mathbf{h}$
$\hat{\mathbf{h}}_k \leftarrow -p_{k,k-1}\hat{\mathbf{h}}_k + \mathcal{H}_k/\mathcal{H}\mathbf{h}$
$\mathcal{H}_k \leftarrow \mathcal{H}_{k-1}\mathcal{H}_k/\mathcal{H}$
$\mathcal{H}_{k-1} \leftarrow \mathcal{H}$
for $i = k+1, \cdots, k_{max}$ **do**
  $t \leftarrow p_{i,k}$,
  $p_{i,k} \leftarrow p_{i,k-1} - pt$
  $p_{i,k-1} \leftarrow t + p_{k,k-1}p_{i,k}$
end for

Appendix B

Proof of Theorem 1

In Algorithm 2.1, since all submatrices are LLL reduced following the approach in [74] the maximum receiver diversity is achieved (more details below). We consider two cases for Algorithm 2.2:

- *Case 1:* In the first part of the detection, the submatrix with the best performance metric is the one that has been reduced using the LLL algorithm and one symbol is detected by exhaustive search in the second part. This case happens when $\mathbf{B}_M$ is chosen as the first submatrix. We denote the probability of error corresponding to this case as $P_{e|case1}$.

- *Case 2:* In some cases the best submatrix is not actually LLL reduced. Since in the algorithm the submatrix with the best performance metric is chosen, its error probability is less than or equal to the case when the LLL reduced submatrix is chosen. The error probability of this case is denoted by $P_{e|case2}$.

  If we denote the probability of cases 1 and 2 as p_{case1} and p_{case2}, respectively, the error

probability of Algorithm 2.2 is defined as

$$P_{e,Algorithm2.2} = P_e(\hat{\mathbf{x}}_{Algorithm2.2} \neq \mathbf{x}) \tag{B.1}$$

$$= P_{e|case1}.p_{case1} + P_{e|case2}.p_{case2}$$

$$\leq P_{e|case1}.p_{case1} + P_{e|case1}.p_{case2}$$

$$= P_{e|case1}, \tag{B.2}$$

where $\hat{\mathbf{x}}_{Algorithm2.2}$ is the detected signal vector using Algorithm 2.2. Note that $P_{e|case1}$ relates to the error probability of LLL reduced submatrices. Based on Theorem 3 of [94], because the submatrix is LLL reduced, the error probability of the first part of the detection, P_{e_1} is bounded by

$$P_{e_1} \leq \alpha_1 snr^{-N}, \tag{B.3}$$

where α_1 is a constant independent of $\mathbf{H_1}$. Moreover, because of the exhaustive search over all possible symbols, the probability of error of the second part given that the first part is detected correctly is bounded by the error probability of the ML detection for the whole signal, i.e.

$$P_{e_2|c_1} \leq \alpha_2 snr^{-N}, \tag{B.4}$$

where α_2 is a constant independent of the channel. As a result the error probability of the detection of both parts is bounded by

$$P_e = P_{e|case1} \leq P_{e_1} + P_{e_2|c_1} \leq (\alpha_1 + \alpha_2) snr^{-N}, \tag{B.5}$$

where P_{e_1} is the error probability of the first subvector and $P_{e_2|c_1}$ is the error probability of the second subvector of the conditional optimization detection assuming that the first subvector is detected correctly. Note that since all submatrices of Algorithm 2.1 are LLL

reduced, such an upper bound can be obtained for it. Consequently, for Algorithm 2.1 and Algorithm 2.2 the receive diversity order is

$$d = \lim_{snr \to \infty} -\frac{\log P_e}{\log snr} = N. \tag{B.6}$$

Appendix C

Proof of Theorem 3.7

For the proof, we follow the footsteps of [52] and apply the required modifications. In [111] it is shown that the maximum diversity order of an antenna selection systems with L selected transmit antennas is

$$d_{AS,opt} = (N - L + 1)(M - L + 1), \tag{C.1}$$

where the diversity order d is defined as

$$d \triangleq \lim_{snr \to \infty} -\frac{\log P_e}{\log snr}, \tag{C.2}$$

where snr is the average SNR at a given receive antenna.

It is obvious that $p \geq (M - N)$ because otherwise the second step of the FSD scheme which is equal to ZF or DFE path is still under-determined and the diversity is zero. If such p full expansion levels are performed, (C.1) provides an upper-bound on the d_{SE}^{UD} with $L = M - p$, i.e.

$$d_{SE}^{UD} \leq d_{AS,opt} = (N - M)(p + 1) + (p + 1)^2 \tag{C.3}$$

$$= (M - N)(p_{eq} + 1) + (p_{eq} + 1)^2. \tag{C.4}$$

Note that $p_{eq} \in \{0, 1, ..., N\}$. Suppose that $\mathbf{H}_o = [\mathbf{H}_{o1},\ \mathbf{H}_{o2}]$ is the ordered channel matrix. In the SE stage, we consider ZF or ZF-DFE detector.

The diversity order of the ZF and ZF-DFE receivers is lower-bounded as

$$d_{SE}^{UD} \geq -\lim_{snr\to\infty} \frac{\log P(\rho_{min}^{(ZF)} \leq \gamma_{th})}{\log snr}, \tag{C.5}$$

where γ_{th} is a given threshold SNR at the receiver and $\rho_{min}^{(ZF)}$ is the minimum post-processing SNR of ZF detector which is given by

$$\rho_{min}^{(ZF)} \triangleq \min_j \frac{snr}{L.[(\mathbf{H}_{o1}^H \mathbf{H}_{o1})^{-1}]_{j,j}}. \tag{C.6}$$

On the other hand, ρ_{min}^{ZF} is lower-bounded by the minimum eigenvalue[1] of $\mathbf{H}_{o1}^H \mathbf{H}_{o1}$ as

$$\rho_{min}^{(ZF)} \geq \lambda_1(\mathbf{H}_{o1}^H \mathbf{H}_{o1}) \frac{snr}{L}. \tag{C.7}$$

Note that $\mathbf{H}_{o1}$ is the remained submatrix after removing $p \geq M - N$ columns. From (C.5) and (C.7) the diversity order corresponding to this submatrix is written as

$$d_{SE}^{UD} \geq -\lim_{snr\to\infty} \frac{\log P(\lambda_1(\mathbf{H}_{o1}^H \mathbf{H}_{o1}).\frac{snr}{L} \leq \gamma_{th})}{\log snr}. \tag{C.8}$$

In order to provide a lower-bound on d_{SE}^{UD} we need to find an expression for $\lambda_1(\mathbf{H}_{o1}^H \mathbf{H}_{o1})$. From block matrices multiplication we have:

$$\tilde{\mathbf{H}}_o^H \tilde{\mathbf{H}}_o = \mathbf{H}_o^H \mathbf{H}_o + \frac{1}{snr} \mathbf{I}_M \tag{C.9}$$

$$= \left[\begin{array}{c|c} \mathbf{H}_{o1}^H \mathbf{H}_{o1} & \mathbf{H}_{o1}^H \mathbf{H}_{o2} \\ \hline \mathbf{H}_{o2}^H \mathbf{H}_{o1} & \mathbf{H}_{o2}^H \mathbf{H}_{o2} \end{array}\right] + \frac{1}{snr} \mathbf{I}_M. \tag{C.10}$$

[1] $\lambda_1(\mathbf{Q}) \leq ... \leq \lambda_n(\mathbf{Q})$ denote the ordered eigenvalues of the $n \times n$ matrix $\mathbf{Q}$.

If we denote $\tilde{\mathbf{Q}}_{o1} = \mathbf{H}_{o1}^H \mathbf{H}_{o1} + \frac{1}{snr}\mathbf{I}_{M-p}$, we have $\lambda_1(\tilde{\mathbf{Q}}_{o1}) = \lambda_1(\mathbf{H}_{o1}^H \mathbf{H}_{o1}) + \frac{1}{snr}$. Therefore,

$$\lambda_1(\mathbf{H}_{o1}^H \mathbf{H}_{o1}) = \lambda_1(\tilde{\mathbf{Q}}_{o1}) - \frac{1}{snr}. \tag{C.11}$$

Moreover, by performing the ordering on $\tilde{\mathbf{Q}}$, according to [52, corollary 3] we have

$$\lambda_1(\tilde{\mathbf{Q}}_{o1}) \sim \lambda_{p+1}(\tilde{\mathbf{Q}}), \tag{C.12}$$

which implies that the minimum eigenvalue of $\tilde{\mathbf{Q}}_{o1}$ behaves similarly to $p+1$-th smallest eigenvalue of $\tilde{\mathbf{Q}}$. The eigenvalues of the channel Gram matrix relate to the eigenvalues of the augmented channel Gram matrix by the following equation

$$\lambda_{k+1}(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}}) = \lambda_{k+1}(\mathbf{H}^H \mathbf{H}) + \frac{1}{snr}; k = 0, 1, ..., M. \tag{C.13}$$

Hence, from (C.12) and (C.13), we have

$$\lambda_1(\tilde{\mathbf{Q}}_{o1}) \sim \lambda_{p+1}(\tilde{\mathbf{Q}}) = \lambda_{p+1}(\mathbf{H}^H \mathbf{H}) + \frac{1}{snr}. \tag{C.14}$$

Moreover, the N non-zero eigenvalues of $\mathbf{H}^H\mathbf{H}$ have the same distribution as the N eigenvalues of the full-rank matrix $\mathbf{H}\mathbf{H}^H$ [69]. Hence, one may write

$$\lambda_{p+1}(\mathbf{H}^H \mathbf{H}) = \lambda_{p_{eq}+1}(\mathbf{H}\mathbf{H}^H); p_{eq} = p - (M - N). \tag{C.15}$$

Note that the $\{\lambda_{M-N}(\mathbf{H}^H\mathbf{H}), ..., \lambda_2(\mathbf{H}^H\mathbf{H}), \lambda_1(\mathbf{H}^H\mathbf{H})\}$ are zero. From (C.11), (C.14) and (C.15) we have

$$\lambda_1(\mathbf{H}_{o1}^H \mathbf{H}_{o1}) \sim \lambda_{p_{eq}+1}(\mathbf{H}\mathbf{H}^H); p_{eq} = p - (M - N). \tag{C.16}$$

By inserting (C.16) into (C.8) and some manipulations we have

$$d_{SE}^{UD} \geq \lim_{x \to 0} \frac{\log P(\lambda_1(\mathbf{H}_{o1}^H \mathbf{H}_{o1}) \leq x)}{\log x} \tag{C.17}$$

$$= \lim_{x \to 0} \frac{\log P(\lambda_{p_{eq}+1}(\mathbf{H}\mathbf{H}^H) \leq x)}{\log x} \tag{C.18}$$

$$= (M - N)(p_{eq} + 1) + (p_{eq} + 1)^2, \tag{C.19}$$

where the last equality is obtained from the pdf of the eigenvalues of the Wishart matrix $\mathbf{H}\mathbf{H}^H$ given in [58, 77]. Since the lower-bound and the upper-bound coincide, the proof is completed.

Appendix D

Proof of Proposition 1

We prove here Proposition 1 for a more general scenario where conditional optimization detection is done on all selections of p columns ($p = 1$ is used in the proposed MMSE-DFE detector). The proof follows in the footsteps of [7, 74] and is included here for completeness (the difference with [7, 74] is the use of the regularized channel matrix and the round robin selection of the different columns for exhaustive search). The decoding error probability of the conditional optimization detectors [74] is upper bounded by

$$P_e \leq P_{e_1} + P_{e_2|c_1}, \tag{D.1}$$

where P_{e_1} is the error probability of the first subvector and $P_{e_2|c_1}$ is the error probability of the second subvector of the conditional optimization detection assuming that the first subvector is detected correctly. Assuming a correct detection for the first part, the error probability of the second part is [74, 95]

$$P_{e_2|c_1} \leq \alpha_1 snr^{-N}, \tag{D.2}$$

and,

$$P_{e_1} \leq \alpha_2 snr^{-d_1}, \tag{D.3}$$

where, α_1 and α_2 are two constants which are independent of the channel. Moreover, d_1 is the diversity order corresponding to the first subvector of the detection.

At high SNRs, using (D.2) and (D.3) the diversity order is lower bounded as

$$d \geq min(N, d_1). \tag{D.4}$$

Similar to [52], this implies

$$d = min(N, d_1). \tag{D.5}$$

Since in the proposed detector, a search is performed over all possible submatrices, hence, the detector outperforms the FSD [7] or RFSD scheme with p full search levels along with the FSD ordering and RFSD ordering, respectively. Therefore, we have

$$d_1 \geq d_{FSD} = (N - M)(p + 1) + (p + 1)^2, \tag{D.6}$$

where the right hand side of inequality (D.6) is the diversity of the FSD scheme or RFSD scheme with p full search levels. On the other hand, based on the concept of antenna selection [54], by choosing l antennas out of M transmit antennas, the maximum diversity order is

$$d_1 \leq d_{1,max} \triangleq (M - l + 1)(N - l + 1). \tag{D.7}$$

Because of the p full search levels, our scheme is equivalent to the problem of antenna selection with $l = M - p$. Hence, from (D.7) we have

$$d_1 \leq (N - M)(p + 1) + (p + 1)^2. \tag{D.8}$$

Since the lower bound and the upper bound in (D.6) and (D.8) coincide, we can write

$$d = min(N, (N - M)(p + 1) + (p + 1)^2). \tag{D.9}$$

The diversity order of the proposed detector is obtained when $p = 1$, which completes the proof.