POWER ALLOCATION AND PRECODING FOR DEVICE-TO-DEVICE COMMUNICATIONS IN MASSIVE MIMO NETWORKS

by

Belal Salama Amin Korany

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
ELECTRONICS AND COMMUNICATIONS ENGINEERING

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2015

POWER ALLOCATION AND PRECODING FOR DEVICE-TO-DEVICE COMMUNICATIONS IN MASSIVE MIMO NETWORKS

by

Belal Salama Amin Korany

A Thesis Submitted to the
Faculty of Engineering at Cairo University
in Partial Fulfillment of the
Requirements for the Degree of
MASTER OF SCIENCE
in
ELECTRONICS AND COMMUNICATIONS ENGINEERING

Under the Supervision of

Prof. Hebat-Allah M. Mourad
Professor
Electronics and Communications Eng. Dept.
Faculty of Engineering, Cairo University

Dr. Mahmoud H. Ismail
Associate Professor
Electronics and Communications Eng. Dept.
Faculty of Engineering, Cairo University

Dr. Ahmed S. Ibrahim
Assistant Professor
Electronics and Communications Eng. Dept.
Faculty of Engineering, Cairo University

FACULTY OF ENGINEERING, CAIRO UNIVERSITY
GIZA, EGYPT
2015

Acknowledgment

In the name of Allah the most merciful the most gracious; all thanks to Allah and peace be upon his messenger Mohamed and his companions. Thence, I would like to thank all those who helped and supported me during the course of my research.

I would like to thank Prof. Hebat-Allah Mourad for her valuable suggestions during the course of this work.

I would also like to thank Dr. Mahmoud H. Ismail, and Dr. Ahmed S. Ibrahim for their continuous support during and before this work. Their guidance and help have had the most remarkable effect on my life choices.

Thanks shall go to Eng. Yahia Ramadan and Eng. Omar Abdallah in Cairo University for their technical suggestions and ideas which helped a lot in the existence of this work.

Last but not least, I am most grateful to my parents, family, and my fiancée for their patience and love. For without them, this work would have never been brought into light.

Belal S. Amin

Contents

List of Tables

List of Figures

List of Symbols and Abbreviations

Symbols

ε	Accuracy of the solution of the bisection algorithm.
$R_d^{\mathbf{D2D}}$	Achievable rate at the d_{th} D2D receiver.
$R_k^{\mathbf{CUE}}$	Achievable rate at the k_{th} CUE.
n_k	Additive white gaussian noise at the k^{th} CUE.
z_d	Additive white gaussian noise at the d^{th} D2D receiver.
θ	Angle step size in the surface-projected gradient descent algorithm.
$\mathbf{e}_l^{(m)}$	column vector of m zeros with the l^{th} element set to 1.
$\mathscr{C}(\mu, \sigma^2)$	Complex Gaussian random vector with mean μ and variance σ^2.
$\preceq$	Component-wise inequality.
∇	Gradient operator.
I_d	Identity matrix with the $d^t h$ diagonal element set to zero..
Λ_o	Initial power allocation vector for the OptSum solution.
F	Matrix of all the channel vectors between the BS and the D2D receivers.
$P_{\mathbf{BS}}$	Maximum transmit power for the base station.
P_D	Maximum transmit power for D2D transmitters and BS.

γ_{th}	Minimum acceptable Signal to Interference Plus Noise Ratio at the CUEs.
$\hat{\mathbf{f}}_d$	$N \times 1$ channel vector from the BS to the d^{th} D2D receiver.
$\hat{\mathbf{h}}_k$	$N \times 1$ channel vector from the BS to the k^{th} cellular UE.
$\mathbf{w}$	$N \times 1$ precoding vector at the BS.
N_o	Noise power spectral density.
χ	Normalization constant of the zeroforcing precoder.
$\mathbf{Null}\{.\}$	Null space of a matrix.
D	Number of D2D pairs.
K	Number of Cellular User Equipments.
N	Number of BS Antennas.
$N_{\mathbf{IPM}}$	Number of iterations of the interior point method.
$N_{\mathbf{OptSum}}$	Number of iterations of the OptSum power allocation solution.
$\mathbf{Q}$	Number of iterations of the gradient descent algorithm.
C_d	Parameter representing the strength of the channel of the d^{th} D2D pair..
J_d	Parameter representing the share of the d^{th} D2D transmitter in the interference on a CUE.
W_d	Parameter representing the share of the d^{th} D2D transmitter in the interference drop on a CUE.
α	Path Loss Exponent.
$\beta_{d,k}^{(g)}$	Path loss of the link between the d^th D2D transmitter and the k^{th} CUE..
$\beta_d^{(f)}$	Path loss of the link between the BS and the d^{th} D2D receiver..
$\beta_k^{(h)}$	Path loss of the link between the BS and the k^{th} CUE..

$\beta_{l,m}^{(\rho)}$	Path loss of the link between the l^{th} D2D transmitter and the m^{th} D2D receiver..
λ_d	Power adjustment factor for the d^{th} D2D transmitter.
$\lambda_{\mathbf{BS}}$	Power adjustment factor for the BS.
rank(.)	Rank of a matrix.
$\mathbf{f}_d$	Rayleigh distributed small scale fading $N \times 1$ channel vector from the BS to the d^{th} D2D receiver.
$\mathbf{h}_k$	Rayleigh distributed small scale fading $N \times 1$ channel vector from the BS to the k^{th} cellular UE.
$\rho_{l,m}$	Rayleigh distributed small scale fading channel coefficient from the l^{th} D2D transmitter to the m^{th} D2D receiver.
$g_{d,k}$	Rayleigh distributed small scale fading channel coefficient from the d^{th} D2D transmitter to the k^{th} cellular UE.
a_k	Received power (from the BS) at the k^{th} CUE.
$y_d^{\mathbf{D2D}}$	Received symbol at the d^{th} D2D receiver.
$y_k^{\mathbf{CUE}}$	Received symbol at the k^{th} CUE.
$\gamma_d^{\mathbf{D2D}}$	Signal to Interference Plus Noise ratio at the d^{th} D2D receiver.
μ	Step size in linear gradient algorithms.
$\mathbf{r}$	Surface projection vector.
$\hat{\rho}_{l,m}$	The channel from the l^{th} D2D transmitter to the m^{th} D2D receiver.
$\hat{g}_{d,k}$	The channel from the d^{th} D2D transmitter to the k^{th} cellular UE.
$E\{.\}$	The Expectation operator.
δ_{th}	The maximum allowable interference on a cellular user.
$\mathbb{R}$	The set of real numbers.
δ_k	The total interference on the k^{th} CUE.

ξ_k The weight of the k^{th} CUE in power distribution of the conventional beamforming precoder.

$Tr(.)$ Trace of a matrix operator.

s_B Transmitted symbol at the base station.

s_d Transmitted symbol at the d^{th} D2D transmitter.

Λ Vector with the power adjustment factors of all D2D transmitters and the BS.

$\mathbf{b}_k$ Vector of received interference powers (from the D D2D transmitters) at the $k^t h$ CUE.

$\mathbf{c}_d$ Vector of received powers (from all D D2D transmitters and the BS) at the $d^t h$ D2D receiver.

ζ Weighting factor of the interference drop factors in the PA heuristic.

Abbreviations

2-D Two-dimensional.

BF Beamforming.

BS Base Station.

CA Carrier Aggregation.

CCCH Common Control Channel.

CDF Cumulative Distribution Function.

CoMP Coordinated Multipoint.

CP Conic Programming.

CSI Channel State Information.

CUE Cellular User Equipment.

D2D Device-to-Device.

DCP Difference of Convex Program.

eNB Evolved Node B (a Base Station in LTE).

GDA Gradient Descent Algorithms.

GLFP Generalized Linear Fractional Program.

IA Interference Alignment.

IC Interference Cancellation.

ILA Interference-Limited Area.

IMT-A International Mobile Telecommunications - Advanced.

ITU-R International Telecommunications Union - Radio communications sector.

JFI Jain's Fairness Index.

LAN Local Area Network.

LP Linear Programming.

LTE Long Term Evolution.

MANET Mobile Ad-hoc Network.

MIMO Multiple-Input-Multiple-Output.

MTD Machine-type devices.

MU-MIMO Multi-User Multiple Input Multiple Output.

OFDM Orthogonal Frequency Division Multiplexing.

PA Power Allocation.

PMC Perfect magnetic conductor.

PSO Particle Swarm Optimization.

QoS Quality of Service.

RA Resource Allocation.

RB	Resource Block.
SDP	Semi-Definite Programming.
SINR	Signal to Interference plus Noise Ratio.
SLNR	Signal-to-Leakage-plus-Noise Ratio.
SNR	Signal to Noise Ratio.
SOCP	Second Order Cone Programming.
TDMA	Time Division Multiple Access.
ZF	Zeroforcing.

Abstract

Device-to-device communication (D2D) is an emerging technology that is proposed by researchers to enhance the performance of next generation wireless communication systems. D2D communication enables two mobile stations to communicate directly without traversing the Base Station (BS). In this work, we look into the problem of deploying D2D communications in a single-cell network that uses very large number of antennas at the BS. Usage of large number of antenna elements is called Massive Multiple-Input-Multiple-Output (MIMO), and is also considered one of the important enhancements for the next generation wireless systems, and is proven to have great impact on the achievable rates of the communication links. We investigate the problems of designing the BS precoder and allocating power values for the D2D transmitter in a way that will maximize the achievable rates of the D2D pairs while maintaining Quality of Service (QoS) constraints on the communication links of the Cellular User Equipments (CUEs). We propose two algorithms for the power allocation problem. The first is an optimal DC-programming-based solution. Due to its complexity, we propose another sub-optimal, less complex heuristic algorithm. We also propose two solutions for the problem of precoder design at the BS. The first one is based on a subfield of convex optimization called Semi-Definite Programming. Also, due to its high complexity, we propose another solution based on the gradient descent algorithm. Finally, we propose to merge the solutions of both problems to solve the joint optimization problem. Simulations show that our proposed schemes have better performance than the ones proposed previously in the literature.

Chapter 1

Introduction

In the past few years, the mobile users demand on data rates in wireless communication systems has been drastically increasing. The new wireless technologies such as WiMax and LTE, despite their high data rates compared to previous wireless technologies, have not been able to satisfy such growing demands. Researchers have been proposing new techniques and paradigms that would allow the existing technologies to meet the International Mobile Telecommunications-Advanced (IMT-A) requirements set by the International Telecommunications Union - Radio communications sector (ITU-R). These technologies include, but are not limited to, Coordinated Multipoint (CoMP), Carrier Aggregation (CA), Massive MIMO, Device-to-device communication (D2D), and others.

In this work, we focus our attention to D2D communications and Massive MIMO. Device-to-device communication is a scenario in which two mobile nodes communicate directly without traversing the Base Station (BS) or the core network [1]. This is most beneficial when those two mobile nodes are in proximity, a situation which is abundant nowadays due to the proximity services like gaming and media sharing. Direct communication would then take advantage of the short-range to construct a good communication link with high Signal-to-Noise Ratio (SNR) and achievable rate, increasing the capacity of the network. If this communication channel takes place in the same frequency resources used by the core network, this would lead to enhanced spectral efficiency (more bits/s/Hz), but will induce interference on the cellular links. Other utilization could be lowering the transmit power of the device since its intended receiver is in proximity, extending the battery life of the mobile device.

On the other hand, Massive MIMO systems are those that use antenna arrays with large number of antennas, in the order of a hundred antennas or more, serving a much smaller number of mobile stations [2]. Massive MIMO is an emerging technology that has proven to be beneficial in so many ways. First, the capacity of the network can be considerably increased due to the fact that the more antennas we have, the more independent data streams that can be sent. Second, energy efficiency is improved because of the ability to concentrate all the energy into the direction of the receiver. Third, reliability of the communication link is enhanced due to the fact that there exist more distinct paths that the signal can propagate over, and hence the channel is flattened. Other advantages are available in the literature [3].

In this work, we invistigate a scenario that deploys both D2D communications and Massive MIMO technologies, where the central BS is equipped with large number of antennas. We assume that the D2D pairs reuse the multicasting frequency bands of the BS. We target the problem of optimizing the power allocation of the BS and the D2D transmitters, and the choice of the BS precoder, in a way that will increase the acheivable rates of the D2D pairs, while maintaining some QoS constraints on the CUEs.

1.1 Contributions

The main contributions of this thesis could be summarized as follows:

1.1.1 Power Allocation Algorithms for D2D communications

As mentioned before, if D2D communication takes place in the same frequency resource as the core network, it induces interference on the cellular links. Good interference management/coordination algorithms should be developed to overcome this issue. As will be thouroughly mentioned in chapter 2, most of the previous attempts were directed towards not to make the D2D link operate at all if its full power transmission will induce higher-than-acceptable interference on the cellular links, and making it operate on other frequency resources, leading to a resource allocation problem. But what if the system was a single-carrier wideband system with only one frequency resource? Why should the D2D transmitter operate only with full power or not operate at all? In this work, we come up with Power Allocation (PA) algorithms to assign transmission power values for the D2D transmitters that will maximize the sum rates of the D2D pairs while maintaining some QoS constraints on cellular links. Those techniques can then be used in single-carrier systems, or applied to each frequency resource -called Resource Block (RB)- in multicarrier systems.

1.1.2 Precoding Algorithms for BS in D2D underlaid networks

Since multiple-antenna BSs have been considered for wireless communication systems, and have already been implemented in the fourth-generation Long-Term Evolution (LTE) networks, D2D underlaid networks with multiple antennas at the BS shall be considered. As will also be mentioned in chapter 2, those networks were given less attention in the literature. Good precoding at the BS may greatly help alleviate the problem of mutual interference between the cellular links and the D2D links. Most of the works in the literature concentrate on the performance of evaluation of the traditional precoders, such as the Beamforming precoder (BF) or the Interference Cancellation (IC) or ZeroForcing precoder (ZF). What if we do not have to stick to those conventional precoders? What if we can design new precoders that will yield better performance results? In this work, we come up with precoding techniques at the BS that will balance the performance of cellular links and D2D links, e.g. maximize the rates of D2D links while preserving QoS requirements at the cellular links. We make use of the recent technology of deploying huge number of antennas at the BS -Massive MIMO-, which is considered for next generation wireless systems, to get the solutions of these optimization problems.

1.2 Thesis Outline

This thesis is organized as follows:

- In chapter 2, we present a survey for the algorithms proposed in the literature for interference coordination between the D2D pairs and the cellular users.

- In chapter 3, we present the system model of the network under investigation and formulate the problem that we will address afterwards.

- In 4, we propose the algorithms that can be used to optimize the power allocation for the BS and D2D pairs, as well as their simulation results.

- While in chapter 5, we propose the algorithms that can be used to optimize the design of the precoder of the BS.
- In chapter 6, we address the joint problem of designing both the PA and precoder.
- Chapter 7 concludes our work and presents future extensions of this work.

Chapter 2

D2D Communications: Literature Survey

In this chapter, we present a detailed review on the previous work in the literature for interference coordination, avoidance, and/or mitigation techniques between the underlaying D2D communicationg devices and the cellular users. A lot of attention has been drawn to enhance the performance of the network in terms of spectral efficiency and/or power efficiency, with QoS/power constraints on the network or the D2D pairs.

2.1 D2D communications

Device-to-device communication is a scenario in which two mobile nodes communicate directly without traversing the Base Station (BS) or the core network [1]. Advantages of D2D communications include enhancing the spectral efficiency of the network [4], extending the battery life of the mobile stations, and mobile data offloading.

In literature, D2D has been proposed to be used in so many contexts and scenarios, such as:

1. Multicasting [5], where the BS transmits the data to a subset of cellular users. These users, in return, share the data with geographically close devices, making use of the good short-range communication links.

2. Cellular offloading [6], where the network makes use of the data available on some mobile device, and instructs it to serve the requests of the near mobile devices locally.

3. Machine-to-machine communication [7], where D2D communication is used as a means of attaching massive number of low-power Machine-Type Devices (MTDs) to the network without overloading the base station.

4. Others.

One thing that would come in mind is that D2D communications are like cognitive radios or Mobile Ad-hoc Networks (MANETs), but the key difference between them is that D2D is controlled by the BS [4]. Two main types of D2D communications can take place:

1. Inband D2D: devices communicate through the same frequency band used by the network.

2. Outband D2D: D2D communicating devices use other frequency bands (mostly the ISM band).

Inband D2D is further divided into two categories:

1. Underlay Inband D2D: D2D devices communicate using the same frequency resources used by cellular users,

2. Overlay Inband D2D: D2D devices use dedicated frequency resources that cellular users are not allowed to use.

Advantages of underlay inband D2D include increasing the spectral efficiency of the network, since D2D devices reuse the same frequency resources of CUEs, which allows for more bits per second to be squeezed into the same available frequency resources. An obvious disadvantage of the underlay inband D2D is the interference between D2D pairs and CUEs. On the other hand, outband D2D avoids the problem of interference between the D2D pairs and the CUEs, but is not generally as spectrally efficient as the inband D2D. This classification is shown in fig 2.1.

The first paper to propose the idea of D2D communications was [8] in 2000, where the authors suggested that a mobile device could relay the data received by the BS to another mobile device, hence, introducing the concept of multihop communications in cellular networks. Then, starting 2007, researchers have started

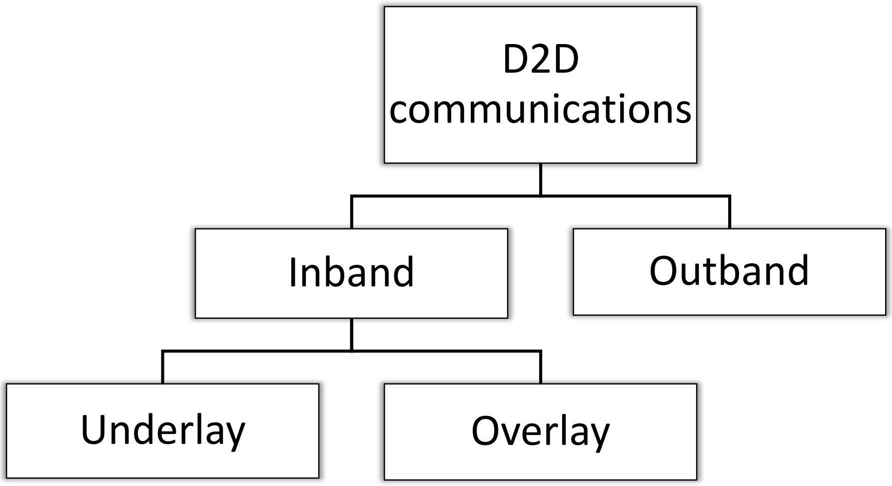

Figure 2.1: Categorization of D2D

to consider D2D for cellular communications and investigated its advantages of increasing the spectral efficiency of the network.

2.2 Implementation-related aspects

While many researchers in academia investigated the issues of deploying D2D in cellular networks, others were concerned with more sophisticated details of deploying it in existing fourth generation LTE networks. Doppler et al. [4] [9] considered the tight integration between D2D and the existing LTE standard. They proposed ways to setup and manage a D2D session and how to incorporate that in the control plane of the LTE system architecture. They proposed the allocation of a dedicated control channel for D2D sessions setup and management. Their simulations show that the D2D throughput could go up to 15 Mb/s compared to a 2.5 Mb/s throughput in case the devices communicated through the LTE Evolved NodeB (eNB). In [10], the authors proposed a scheme to deploy D2D communications in LTE networks. First, an interference range is identified by letting the CUEs monitor the received signal from the D2D transmitters on the Common Control Channel (CCCH), and reporting the results to the eNB. Then, the eNB classifies the RBs into suitable RBs and bad physical RBs, where the bad RBs are defined as the RBs in which there will be high interference on the CUE (if in downlink channel) or high interference on the D2D receiver (if in uplink channel). CUEs are scheduled on the suitable RBs and D2D pairs are scheduled on the bad RBs.

Wu et al. [11] presented FlashLinQ, a complete network architecture for device-to-device communications with channel aware distributed scheduling techniques. The performance of their proposed architecture was compared to the existing 802.11g architecture for Wireless-LANs. The performance of the two schemes was tested using an implemented FPGA platform for FlashLinQ. Their experiments show a 590% increase on the achieved rates of FlashLinQ over 802.11g in indoor environments and 540% increase in the outdoor environments.

2.3 Schemes for Resource and/or Power Allocation

Apart from the implementation-related issues, most of the available literature on D2D concentrate on underlay inband D2D and its underlying problem of interference, proposing techniques for interference coordination, avoidance, and/or mitigation between the D2D pais and the CUEs. Many researchers proposed schemes and algorithms for resource allocation and/or power allocation to control such mutual interference. In the next subsection, we will get a glimpse of these works.

2.3.1 Spectral efficiency enhancing schemes (QoS constrained)

In [12], the authors proposed two mechanisms for interference avoidance between the D2D pairs and the cellular users in an uplink scenario. The first mechanism avoids high interference from the CUEs to the D2D receivers, by allowing the D2D receivers to read the RB information and avoid communication on the resources which are used by CUEs nearby. The second mechanism avoids high interference from the D2D transmitters to the BS, by sending information to the D2D transmitters about the expected interference from them on all RBs, thus, they can apply some power control and RB selection to operate on RBs in which they will not produce high interference on the BS. Those mechanisms have proven to be efficient by improving the system throughput by ~40%, but the underlying assumption of a multicarrier system makes it difficult to apply the algorithms in single carrier wideband systems. If it was, D2D pairs may not operate at all in the single carrier system if there is a CUE nearby. In [13], a similar power control scheme was used in an uplink scenario. The BS suffers interference from the D2D transmitters. So, the D2D transmitter must adjust their power in order for the SINR at the BS not to fall below a certain required threshold. In order to do so, the D2D transmitters measure the received power from the BS in the downlink slots. They then measure the path loss between them and the BS and adjust their power accordingly. If the minimum transmit power at the D2D transmitter for acceptable performace at the receiver will induce high interference on the BS, the D2D pair is turned off.

In [14], the authors proposed to use a recent research proposal that would enhance the cellular capacity, Interference Alignment (IA) [15] to deploy D2D in a cellular network. The authors proposed to divide the D2D pairs into groups. Each group contains 3 D2D pairs. Then, the pairs use IA to precode transmission such that they would not introduce interference on the cellular network.

More sophisticated mathematical techniques were used in other papers. In [16], the authors used graph theory to propose an interference aware graph based solution to the resource allocation problem. In their solution, each vertex in the graph represents a communication link (cellular or D2D) and each edge represents the mutual interference between the two links represented by the vertices the edge is connected. Their simulation results show that their solution performs near the optimal exhaustive search resource allocation algorithm. Papers [17–19] applied game theory concepts to solve the resource allocation problem, where all the D2D pairs (and possibly the CUEs) are driven into an auction game and the auction is resolved such that a certain performance metric is optimized.

2.3.2 Location-based schemes

Multiple location-based schemes were proposed in the literature. In these schemes, CUEs and D2D pairs are prohibited to operate simultaneously if the distance between them is smaller than a certain threshold. In [20], the authors proposed a framework in which an Interference-Limited Area (ILA) around each D2D receiver is defined. The ILA is defined as the area in which the interference from a CUE -if it existed within- on the D2D receiver is greater then a certain predefined threshold. If a CUE existed within this area, it is prohibited to operate on that RB and is scheduled to operate on another RB which is assigned to another

D2D receiver whose ILA does not cover the CUE. While the scheme proved to enhance the performance by more than 100% over the conventional interference management schemes, it certainly reduces the multiuser diversity in the network, since it reduces the scheduling alternatives for the BS on each RB. A similar approach was considered in [21], where ILAs are defined in the same way, but before an RB is assigned to any D2D receiver. After the ILA is defined, resources are allocated in a way where CUEs and D2D receivers in the same ILA are assigned different RBs.

A similar location-based scheme is depicted in [22], but the ILA is calculated in real time inherently by the measurements done by the CUE on a dedicated D2D control channel. If a CUE senses a higher-than-threshold D2D activity, indicating a near D2D receiver, it reports this measurement back to the BS, which in turn takes the decision to prohibit the D2D from operating on the CUE assigned RB. Simulations show that the scheme improves the system throughput with more than 300% of its value with no scheme applied. The authors in [23] adopted the same strategy but with taking the interference between the D2D pairs and each other into consideration.

2.3.3 Constrained optimization problem schemes

Researchers have also formulated optimization problems with QoS and/or power constraints. In [24], the authors proposed a resource allocation method that guarantees QoS requirements for both the CUEs and the D2D pairs. Their optimization problem is non-linear, non-convex. They proposed to solve the problem in two steps. First, the BS decides whether a D2D pair can be admitted operation or not based on SINR requirements. Second, a bipartite matching based scheme is used for the resource allocation of both the CUEs and the D2D pairs. In [25] and [26], the authors formulate a problem of maximizing the system throughput subject to maintaining minimum data rate requirements on the CUEs and D2D pairs, then they use Particle Swarm Optimization (PSO) techniques to solve the optimization problem. The simulation results of [25] show ~15% gain in system throughput over the orthogonal resource allocation schemes (overlay D2D). In [27], the authors assume a time slotted network and formulate a QoS-constrained optimization problem to maximize the sum rate of the system by proper scheduling. They then use a stochastic sub-gradient algorithm solution that was proven to improve the sum rate of the system by 500%. The authors of [28] found a closed form solution for the problem of maximizing the sum rate of a cellular network with one CUE and 2 D2D pairs subject to power constraints on the BS and the devices. They also proposed a choose-out-of-set solution for the same problem yielding a 45% improvement on the network sum rate.

In [29] the authors formulate a problem of the total downlink transmit power minimization of a single cell Orthogonal Frequency Division Multiplexing (OFDM) network. They proposed a heuristic algorithm to solve their optimization problem. In their heuristic, they first apply the schemes presented in [30, 31] for the resource allocation (subcarrier and bit allocation) for the cellular users. Then, they perform mode selection and power allocation for the D2D pairs. If the required transmit power of a certain D2D transmitter is higher than a certain threshold, this D2D pair becomes two normal CUEs and communicate through the BS. Their simulation results show 20% less power consumption compared to existing schemes used for D2D-free OFDM networks. The authors of [32] also formulate a QoS-and-power constrained optimization problem in which they want to maximize a utility function that they proposed. Their function is related to the concept of the power efficiency which is directly proportional to the achievable rates of the links and inversely proportional to the power consumed. They proposed a heuristic scheme in which the power efficiency is calculated for all the mobile devices if they operate as CUEs and if they operate as D2D pairs. Then each mobile devices is admitted operation with the mode that maximizes its power efficiency. The allocation of power in each mode is done by maximizing a lower bound for the nonconvex utility function.The algorithm is complex since it performs exhaustive search for the optimum solution. A very similar approach is used in [33], where the authors proposed to calculate the achievable rates of the mobile users if they either operate in cellular mode (communicating with the BS), or in D2D mode (communicating directly with each other), and then choosing the mode that will achieve a higher rate.

In [34], the authors proposed an iterative algorithm to allocate power values to the BS and D2D transmitters in a single cell network with only one CUE and two D2D pairs. 10% improvement in the system capacity was achieved over the full power case. The authors of [35] investigate a similar scenario with one CUE and one D2D pair and formulate a nonconvex optimization problem of mazimizing the total sum rate of the network subject to preserving power constraints. They propose an iterative solution whose iterations are based on fast barrier methods to solve convex subproblems.

2.3.4 D2D underlaying MIMO networks

While BSs with multiple antennas are considered for future wireless networks and even already implemented in the existing wireless standards, the problem of deploying D2D communication in a MIMO network, i.e. a network whose BS has multiple number of antennas, has drawn less attention in the literature. One of the attempts to exploit this scenario was done in [36], where the authors presented an algorithm to allocate power values to the D2D pairs underlaying a Multi-usero MIMO (MU-MIMO) network, with both Beamforming (BF) and Zeroforcing (ZF) precoders considered. They used a binary power allocation scheme where D2D pairs are -again- either turned on with full power or turned off completely.

In [37], the authors investigated the design of precoders at both the BS and the D2D transmitter (assuming multiple antennas at each), by choosing them from predefined set called a codebook. The problem was decoupled to the choice of each individually aiming at maximization of the Signal-to-Leakage-plus-Noise-Ratio (SLNR) and Signal-to-Interference-plus-Noise-Ratio (SINR). The authors were able to achieve remarkable gains in the throughput of the network and approach the performance of the optimum exhaustive search.

2.3.5 Other deployment schemes

There has also been some effort in the literature to deploy D2D communications in non-traditional ways. In [38], the authors proposed that a D2D receiver, while receiving its intended data from the D2D transmitter, can simultaneously act as a relay node to the cellular transmission. Simulations show that the capacity region of the entire network can be enlarged by as much as 60% over the traditional separate CUE and D2D links.

2.3.6 Performance Evaluation

Other researchers were interested in the performance evaluation of D2D-underlaid networks. In [39], the authors evaluate the performance of a D2D underlaid network under fading circumstances. They found that the network exhibits rate loss and throughput outage when channel fades are present. The losses increase with increasing the distances between each D2D transmitter and its receiver. They also concluded that these losses can be compensated by a strategy that favors the performance of the CUEs over D2D pairs. The authors of [40] provide an analytical framework for the calculation of the SINR outage probability and the spectrum efficiency of a D2D underlaid uplink cellular network. They study the effect of the power control cutoff threshold (which is defined as the minimum acceptable SINR threshold of a communication link) and the D2D biasing factor (a factor which expresses the tendency of a mobile node to operate as a D2D node) on the aforementioned performance metrics.

In MIMO networks, researchers presented performance evaluations for both the conventional BF and ZF precoders [41,42]. It has been found, using both analytical and numerical results, that ZF outperforms BF in terms of the sum capacity of the network in the high SNR regime due to its ability to suppress interference on the D2D pairs and, hence, enhance their performance. BF, on the other hand, outperforms ZF in the low SNR regime due to its ability to direct the available low power to the intended CUEs. The authors of [41] also presented some possible enhancements for the precoders by using closed loop techniques.

One recent strong paper [43] addressed the performance evaluation of D2D underlaid networks with huge number of BS antennas (Massive MIMO) and presented closed form formulas for the spectral efficiency of the network under both perfect and imperfect Channel State Information (CSI) at the BS.

2.4 Literature voids

There are two underlying assumptions of most of the mentioned work: a) the availability of multiple frequency resources, which makes most of the attempts focus on resource allocation, where D2D pairs are allowed to operate with full power on some RBs and not operate at all on the others, but not to work with an optimized transmit power on any, and b) neglecting the exploitation of the advantage of having multiple antennas at the BS, by either assuming one antenna at the BS, or assuming multiple antennas with fixed conventional precoding schemes (BF, ZF, or codebook-based).

To the best of our knowledge, algorithms for precoding and non-binary power allocation for D2D communications underlaying MIMO (or Massive MIMO) networks have been absent in the literature, and this work is our attempt to fill that void.

Table 2.1: Quick summary of the literature survey

Paper	Main Contribution	Technique	D2D PA	Precoding
[12]	Resource Allocation (RA) scheme for D2D deployment in LTE networks.	Heuristic	Binary	N/A
[13]	Per-D2D RA and PA scheme in an uplink frame for SINR threshold at the BS.	Analytical	Minimum allowable for QoS threshold	N/A
[16]	RA scheme to maximize the total sum rate of the network.	Graph theory	Binary	N/A
[14]	D2D grouping heuristic to use IA in managing the interference on the CUEs	IA - Heuristic	N/A	N/A
[20]	Location-based scheme that prohibits a D2D pair and a CUE to operate in the same calculated geographical area.	Analytical	Binary	Traditional
[21]	Location-based scheme to make a smart RA to CUEs and D2D pairs in the same geographical area.	Analytical / Heuristic	Binary	N/A
[22]	Distributed location-based RA scheme in LTE networks.	Heuristic	Binary	N/A
[24]	Solving a QoS and power constrained network sum rate optimization problem by solving two subproblems of mode selection and power control.	Analytical	Continuous	N/A
[25], [26]	Solving a QoS and power constrained network sum rate optimization problem using particle swarm optimization techniques.	PSO	Continuous	N/A
[27]	Solving a scheduling problem in a time-slotted network to maximize the total throughput of the network.	Analytical	Binary	N/A
[28]	Closed form solution for transmit power optimization in a network with one BS, one CUE and one D2D pair.	Analytical	Continuous	N/A
[29]	BS transmission power minimization scheme for QoS guaranteed operation.	Heuristic	Binary	N/A

Paper	Main Contribution	Technique	D2D PA	Precoding
[34]	Solving an optimization problem to maximize the total throughput of a network with one CUE and two D2D pairs.	Heuristic	Continuous	N/A
[35]	Solving an optimization problem to maximize the total throughput of a network with one CUE and two D2D pairs.	Analytical / Heuristic	Continuous	N/A
[4]	Proposed session setup and management schemes for deploying D2D in LTE networks	Heuristic	Constant-step control	N/A
[32]	Solving an QoS and power constrained optimization problem to maximize the energy efficiency of the network.	Heuristic	Continuous	N/A
[40]	Performance evaluation for the SINR outage probability of D2D underlaid networks	Analytical	Binary	N/A
[36]	Solving a joint optimization problem for optimal PA and precoding by decoupling the problem.	Heuristic	Binary	Traditional
[37]	Precoder optimization for SLNR or SINR maximization.	Heuristic	Binary	Codebook based
[41]	Performance evaluation of different MIMO schemes with the assumption of multiple antennas at both the BS and the mobile stations.	Analytical	Binary	Traditional
[43]	Performance evaluation of Massive MIMO underlaid networks with both perfect and imperfect CSI.	Analytical	Binary	Massive / Traditional
[38]	Proposal of the usage of D2D mobiles as relays to the CUEs	Analytical	Binary	N/A
Thesis	Maximizing the sum rate of the D2D pairs with QoS constraints on the CUEs	Heuristic	Continuous	Optimized

Chapter 3

System Model

In this chapter, we present the system model of th network under investigation and introduce all notations that will be used in our work. Then, we will formulate the optimization problem that we will solve in the course of our work.

We consider a single-cell downlink cellular network with a BS which is equipped with N antennas. There are K CUEs uniformly distributed in the cell. There exists D D2D pairs. The D2D transmitters are uniformly distributed in the cell, whereas each D2D receiver is located in a circle centered at its transmitter. Fig. 3.1 summarizes the previous information. Each of the CUEs and the D2D devices is equipped with a single omnidirectional antenna. We will consider a scenario where the number of antennas N is much larger than K ($N \gg K$) to make use of the benefits of Massive MIMO. We assume that the D2D pairs operate in the multicasting channel of the BS. In a multicasting setting, the BS transmits the same data symbol s_B to all the CUEs. This assumption will dispense the need to consider inter-CUE interference and focus on the mutual interference between the CUEs and D2D pairs. The d^{th} D2D transmitter sends a symbol s_d to its corresponding receiver. We assume that all the data symbols have an average unit power, i.e., $E\{|s_B|^2\} = E\{|s_d|^2\} = 1$, where $E\{.\}$ is the expectation operator.

The received signals at the k^{th} CUE and the d^{th} D2D receiver are given by

$$y_k^{\text{CUE}} = \sqrt{\lambda_{\text{BS}} P_{\text{BS}}}\, \hat{\mathbf{h}}_k^H \mathbf{w} s_B + \sum_{d=1}^{D} \sqrt{\lambda_d P_D}\, \hat{g}_{d,k} s_d + n_k, \tag{3.1}$$

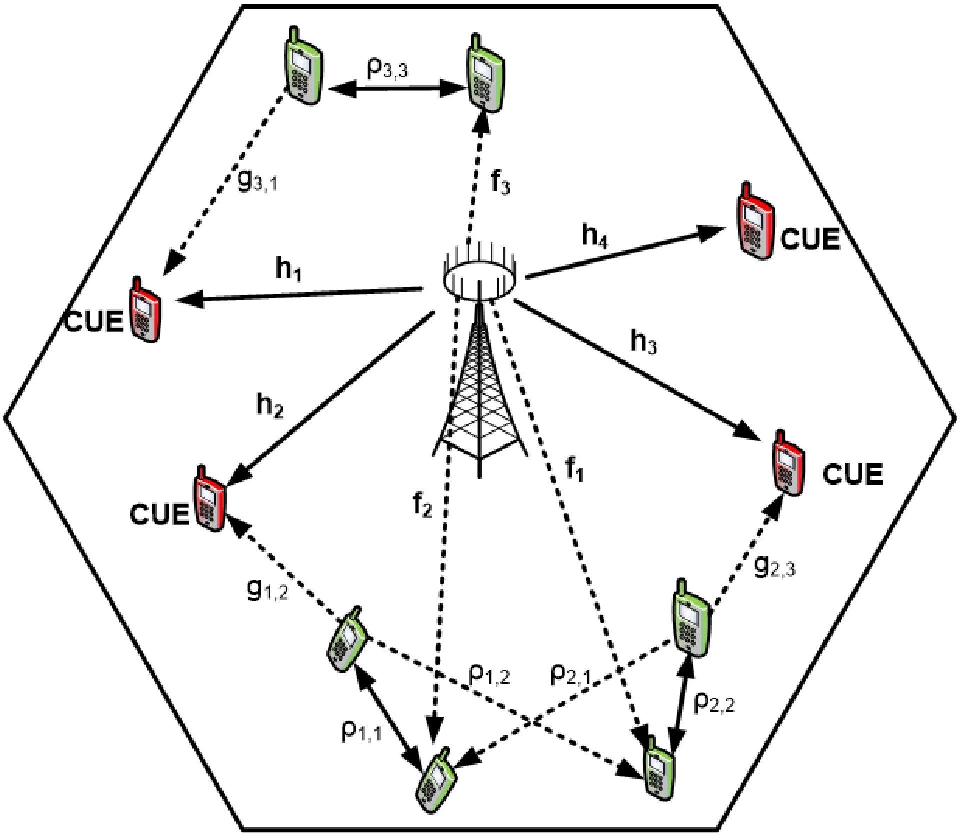

Figure 3.1: Single cell multicasting network with Massive MIMO, $K = 4$ cellular users (red terminals), $D = 3$ D2D pairs (green terminals). Dashed lines represent interference.

and

$$y_d^{\text{D2D}} = \sqrt{\lambda_d P_D}\,\hat{\rho}_{d,d} s_d + \sum_{\substack{d'=1 \\ d' \neq d}}^{D} \sqrt{\lambda_{d'} P_D}\,\hat{\rho}_{d',d} s_{d'} + \sqrt{\lambda_{\text{BS}} P_{\text{BS}}}\,\hat{\mathbf{f}}_d^H \mathbf{w} s_B + z_d, \quad (3.2)$$

respectively, where

- P_{BS} and P_D are the maximum powers of the BS and a D2D transmitter, respectively,
- λ_{BS} and λ_d are power adjustment factors for the BS and the d^{th} D2D transmitter, where $\lambda_{\text{BS}}, \lambda_d \in [0,1]\ \forall d \in \{1,\ldots,D\}$,
- $\mathbf{w}$ is an $N \times 1$ precoding vector at the BS,
- n_k and z_d are additive white Gaussian noise$\sim \mathscr{C}(0,N_o)$ at the k^{th} CUE and the d^{th} D2D receiver, respectively, where N_o is the noise power.

The channel gain between any two nodes is modeled as a Rayleigh fading channel with path loss as follows

- $\hat{\mathbf{h}}_k = \sqrt{\beta_k^{(h)}}\mathbf{h}_k$ is the $N \times 1$ channel vector from the BS to the k^{th}CUE, $\beta_k^{(h)}$ denotes the path loss effect and $\mathbf{h}_k$ is the small scale fading complex Gaussian $\mathscr{C}(0,1)$ $N \times 1$ channel coefficients,
- $\hat{g}_{d,k} = \sqrt{\beta_{d,k}^{(g)}}g_{d,k}$ is the channel from the d^{th} D2D transmitter to the k^{th}CUE, $\beta_{d,k}^{(g)}$ denotes the path loss effect and $g_{d,k}$ is the small scale fading complex Gaussian $\mathscr{C}(0,1)$ channel coefficient,
- $\hat{\rho}_{l,m} = \sqrt{\beta_{l,m}^{(\rho)}}\rho_{l,m}$ is the channel from the l^{th} D2D transmitter to the m^{th} D2D receiver, $\beta_{l,m}^{(\rho)}$ denotes the path loss effect and $\rho_{l,m}$ is the small scale fading complex Gaussian $\mathscr{C}(0,1)$ channel coefficient,
- $\hat{\mathbf{f}}_d = \sqrt{\beta_d^{(f)}}\mathbf{f}_d$ is $N \times 1$ channel vector from the BS to the d^{th} D2D receiver, $\beta_d^{(f)}$ denotes the path loss effect and $\mathbf{f}_d$ is the small scale fading complex Gaussian $\mathscr{C}(0,1)$ $N \times 1$ channel coefficients.

The path loss coefficient of the channel of any link is calculated as $\beta = c^{-\alpha}$ where c is the distance between the two nodes of the link and α is the path loss exponent

In (3.1), the first term represents the intended data to be sent to the CUE (all CUEs are intended to receive the same data symbol s_B, while the second term represents the interference from the D D2D transmitters on the k^{th} CUE. The Signal to Interference plus Noise Ratio (SINR) at the k^{th} CUE can be calculated as

$$\gamma_k^{\text{CUE}} = \frac{\lambda_{\text{BS}} P_{\text{BS}} \, |\hat{\mathbf{h}}_k^H \mathbf{w}|^{\mathbf{2}}}{\sum_{d=1}^{D} \lambda_d P_D \, |\hat{g}_{d,k}|^2 + N_o}. \tag{3.3}$$

In (3.2), the first term represents the intended signal from the d^{th} D2D transmitter to its receiver, while the second term represents the interference from the other D2D transmitters, and the third one represents the interference from the BS. The total interference power can be calculated as the sum of the powers of these interference sources. This is due to the fact that the interference from these sources is independent. Thus, the SINR of the d^{th} D2D pair is calculated as

$$\gamma_d^{\text{D2D}} = \frac{\lambda_d P_D \, |\hat{\rho}_{d,d}|^2}{\sum_{\substack{d'=1 \\ d' \neq d}}^{D} \lambda_{d'} P_D \, |\hat{\rho}_{d',d}|^2 + \lambda_{\text{BS}} P_{\text{BS}} |\hat{\mathbf{f}}_d^H \mathbf{w}|^2 + N_o}. \tag{3.4}$$

Using Shannon's formula for the achievable rate of a communication link, let

$$R_k^{\text{CUE}} = \log\left(1 + \frac{\lambda_{\text{BS}} P_{\text{BS}} \, |\hat{\mathbf{h}}_k^H \mathbf{w}|^{\mathbf{2}}}{\sum_{d=1}^{D} \lambda_d P_D \, |\hat{g}_{d,k}|^2 + N_o}\right) \tag{3.5}$$

and

$$R_d^{\text{D2D}} = \log\left(1 + \frac{\lambda_d P_D \, |\hat{\rho}_{d,d}|^2}{\sum_{\substack{d'=1 \\ d' \neq d}}^{D} \lambda_{d'} P_D \, |\hat{\rho}_{d',d}|^2 + \lambda_{\text{BS}} P_{\text{BS}} |\hat{\mathbf{f}}_d^H \mathbf{w}|^2 + N_o}\right) \tag{3.6}$$

be the achievable rates of the k^{th} CUE link, and d^{th} D2D pair link, respectively, where the log here and throughout the thesis denotes the base-2 logarithm.

Chapter 4

Power Allocation

In this chapter, we will formulate the problem of optimizing the PA factors $\lambda_1, \lambda_2, \ldots, \lambda_D, \lambda_{\text{BS}}$ for a certain determined choice of $\mathbf{w}$. We will propose two solutions for this subproblem and compare between their performance. We will also propose an alternative formulation for the problem that will try to maximize the minimum D2D rate instead of the sum rates of the D2D pairs, for fairness issues. We will use the Beamforming Precoder (BF) and Zeroforcing Precoder (ZF) as conventional choices for $\mathbf{w}$.

4.1 Problem Formulation

We want to determine the optimal values of the PA factors $\lambda_1, \lambda_2, \ldots, \lambda_D, \lambda_{\text{BS}}$ in a way that will optimize a certain performance metric in the network. Possible objective functions include

1. maximizing the sum rate of the D2D pairs,
2. maximizing the sum rate of the CUEs,
3. maximizing the total sum rate of the network (D2D pairs + CUEs),
4. maximizing the minimum rate of the D2D pairs,
5. maximizing the minimum rate of the CUEs,
6. maximizing the minimum rate of the network, or others.

We note that the large number of antennas at the BS insures good communication links, with very high SNRs, between the BS and the CUEs [44]. Hence, there is no need to consider maximizing their rates in the optimization problem, and guaranteeing good communication link suffices. So, we will focus our attention on maximizing the sum rate of the D2D pairs while preserving QoS constraints on the CUEs. In subsequent sections, we will consider maximizing the minimum rate of the D2D pairs with the same QoS constraints.

The sum rate optimization problem can be formulated as

$$\max_{\lambda_1,\lambda_2,\ldots,\lambda_D,\lambda_{\mathrm{BS}}} \quad \sum_{d=1}^{D} R_d^{\mathrm{D2D}}, \tag{4.1a}$$

$$\text{subject to} \quad \gamma_k^{\mathrm{CUE}} \geq \gamma_{th}, \quad \forall k \in \{1,\ldots,K\}, \tag{4.1b}$$

$$0 \leq \lambda_i \leq 1, \quad \forall i \in \{1,2,\ldots,D,\mathrm{BS}\}, \tag{4.1c}$$

where γ_{th} is the minimum required SINR at the CUEs for satisfactory performance. Constraint (4.1b) represents the QoS constraint, while (4.1c) represents the power constraint.

The optimization problem (4.1a)–(4.1c) is nonlinear, non convex, and hence, not directly solvable (See appendix A). Before going into the details of the problem solution, we will first choose a specific precoder $\mathbf{w}$ from one of two traditional options: Beamforming or Zeroforcing precoders.

Beamforming Precoder (BF)

In [45], it was shown that the optimum precoder for a multicasting massive MIMO system, without the existence of D2D users, is a weighted sum of the channels of the CUEs, i.e., $\mathbf{w} = \sum_{k=1}^{K} \xi_k \hat{\mathbf{h}}_k$ where the weights ξ_k act as power distribution factors. For simplicity, we will consider equal power distribution among the CUEs. In this case, the precoder coefficients are given by

$$\mathbf{w}_{\mathrm{BF}} = \frac{\sum_{k=1}^{K} \hat{\mathbf{h}}_k}{\sqrt{N \sum_{k=1}^{K} \beta_k^{(h)}}}, \tag{4.2}$$

where $\beta_k^{(h)}$ is the path loss between the BS and the k^{th} CUE, and the factor $\sqrt{N \sum_{k=1}^{K} \beta_k^{(h)}}$ is a normalization factor for $|\mathbf{w}_{\mathrm{BF}}^H \mathbf{w}_{\mathrm{BF}}|$ to be 1. Due to the law of large numbers, and the fact that $\hat{\mathbf{f}}_d$ and $\mathbf{w}_{\mathrm{BF}}$ are independent, $E\left\{|\mathbf{f}_d^H \mathbf{w}_{\mathrm{BF}}|^2\right\} \xrightarrow[N\to\infty]{\text{a.s.}} 1$, where a.s. denotes asymptotic convergence [46]. Hence, the interference from the BS on the D2D users in (3.4) is, on average, constant.

Zero Forcing (ZF)

The second variant is the ZF precoder where the BS will transmit its data in the projection of the channels of the CUEs on the null space of the channels between the BS and the D2D receivers, thus cancelling out the interference made by the BS on them. Since $\mathbf{w}$ is wanted to satisfy $\left[\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \ldots, \hat{\mathbf{f}}_D\right]^H \mathbf{w}_{\text{ZF}} = 0$, $\mathbf{w}_{\text{ZF}}$ can be calculated as

$$\mathbf{w}_{\text{ZF}} = \frac{1}{\chi} \left(\text{Null}\left(F^H\right)\right) \left(\text{Null}\left(F^H\right)\right)^H \sum_{k=1}^{K} \hat{\mathbf{h}}_k, \tag{4.3}$$

where $F = \left[\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \ldots, \hat{\mathbf{f}}_D\right]$, Null(.) is the null space of a matrix, and χ is a normalization factor to ensure that $|\mathbf{w}_{\text{ZF}}^H \mathbf{w}_{\text{ZF}}| = 1$.

4.2 OptSum: PA problem solution 1

After choosing a specific precoder $\mathbf{w}$, let

$$\begin{aligned}
\Lambda &\triangleq [\lambda_1, \lambda_2, \ldots, \lambda_D, \lambda_{\text{BS}}]^T, && (4.4)\\
a_k &\triangleq P_{\text{BS}} |\hat{\mathbf{h}}_k^H \mathbf{w}|^{\mathbf{2}}, && (4.5)\\
\mathbf{b}_k &\triangleq [P_D |\hat{g}_{1,k}|^2, P_D |\hat{g}_{2,k}|^2, \ldots, P_D |\hat{g}_{D,k}|^2, 0]^T, && (4.6)\\
\mathbf{c}_d &\triangleq [P_D |\hat{\rho}_{1,d}|^2, P_D |\hat{\rho}_{2,d}|^2, \ldots, P_D |\hat{\rho}_{D,d}|^2, P_{BS} |\hat{\mathbf{f}}_d^H \mathbf{w}|^2]^T && (4.7)
\end{aligned}$$

Also let $\mathbf{e}_l^{(m)}$ be a column vector of m zeros with the l^{th} element set to 1, and I_d be the identity matrix with the d^{th} diagonal element set to zero. The optimization problem (4.1a)–(4.1c) can now be rewritten in matrix form as

$$\begin{aligned}
\Lambda^* = \underset{\Lambda}{\text{argmax}} \quad & \sum_{d=1}^{D} \log \left(\frac{\mathbf{c}_d^T \Lambda + N_o}{\mathbf{c}_d^T I_d \Lambda + N_o} \right), && (4.8a)\\
\text{subject to} \quad & \mathbf{m}_k^T \Lambda - \gamma_{th} N_o \geq 0, \quad \forall k \in \{1, \ldots, K\}, && (4.8b)\\
& 0 \preceq \Lambda \preceq 1, && (4.8c)
\end{aligned}$$

where $\mathbf{m}_k = a_k \mathbf{e}_{D+1}^{(D+1)} - \gamma_{th} \mathbf{b}_k$, and $\preceq$ denotes component-wise inequality. The optimization problem (4.8a)–(4.8c) is nonconvex, difference of convex programming (DCP) problem [47].

To solve the DCP program, we propose an iterative solution. By approximating the logarithm $\log\left(\mathbf{c}_d^T I_d \Lambda + N_o\right)$ with its first order Taylor expansion, the optimization problem (4.8a)–(4.8c) is approximated to

$$\Lambda^* = \underset{\Lambda}{\operatorname{argmax}} \sum_{d=1}^{D} \left\{ \log\left(\mathbf{c}_d^T \Lambda + N_o\right) - \log\left(\mathbf{c}_d^T I_d \Lambda_o + N_o\right) - \frac{I_d^T \mathbf{c}_d}{\mathbf{c}_d^T I_d \Lambda_o + N_o} \left(\Lambda - \Lambda_o\right) \right\}, \tag{4.9a}$$

$$\text{subject to} \quad \mathbf{m}_k^T \Lambda - \gamma_{th} N_o \geq 0, \quad \forall k \in \{1, ..., K\}, \tag{4.9b}$$

$$0 \preceq \Lambda \preceq 1, \tag{4.9c}$$

where Λ_o is an initial value for the PA vector Λ. The optimization problem (4.9a)–(4.9c) is concave, and can be solved using interior point methods to get the optimum value Λ^*(see appendix A). Since (4.9a) is an approximate value of the original objective function, the obtained Λ^* is not the optimum solution to (4.8a)–(4.8c). We propose to set $\Lambda_o = \Lambda^*$ and solve the optimization problem again. The new optimum value is then used as an initial value for another iteration, and so on. Eventually, the optimum value Λ^* will converge to the optimum solution of the problem (4.8a)–(4.8c) as will be shown later. Interior point methods solve the problem in typically 10 to 100 steps, with each step requiring an order of $\max\left\{(D+1)^3, (D+1)^2(K+2D), K+DN+N^2\right\}$ operations [48], where $K+DN+N^2$ is the cost of evaluating the first and second derivatives of (4.9a) and (4.9b). Thus, if we denote the number of iterations of the algorithm as N_{OptSum}, and the number of steps of the interior point method as N_{IPM}, the complexity of the OptSum solution is

$$\mathscr{O}\left(N_{\text{OptSum}}.N_{\text{IPM}}.\max\left\{D^2K + 2D^3, K + DN + N^2\right\}\right). \tag{4.10}$$

4.3 Heuristic: PA problem solution 2

Due to the high complexity of the previous solution, we propose a suboptimal heuristic that tries to maximize the sum rate of the D2D pairs while maintaining the QoS constraints on the CUEs. The details of the heuristic are shown in Algorithm 4.1. The steps and intuition behind the heuristic are summarized in the following points:

1. The BS as well as the D2D transmitters are initialized greedily to operate with full power.

2. CUEs are ordered ascendingly according to their SINRs. If the lowest SINR is above than the SINR threshold, then the QoS constraint is satisfied for all the CUEs and the algorithm terminates. Otherwise, the CUEs are chosen one by one according to their ascending order of the SINRs.

3. Once a CUE is chosen, if its SINR is below the SINR threshold, the total interference power on it, which is equal to

$$\delta_k = \sum_{d=1}^{D} \lambda_d P_D |\hat{g}_{d,k}|^2 \tag{4.11}$$

, should be reduced to

$$\delta_{th} = \frac{P_{\text{BS}} |\hat{\mathbf{h}}_k^H \mathbf{w}|^2}{\gamma_{th}} - N_o \tag{4.12}$$

such that this CUE operates on the SINR threshold.

4. To achieve this the D2D transmitters powers should be adjusted (reduced). Two factors are taken into consideration in this power reduction:

   (a) a D2D pair with high channel gain between its transmitter and the CUE is preferred to have its power reduced because it induces higher interference on the CUE. To express this, each D2D pair is given a parameter
   $$J_d = \frac{\lambda_d P_D |\hat{g}_{d,k}|^2}{\delta_k} \tag{4.13}$$
   , and

   (b) a D2D pair with higher channel gain between its transmitter and receiver is preferred to retain its high power, due to the possibility to achieve higher rate. To express this, each D2D pair is given a parameter
   $$C_d = \frac{|\hat{\rho}_{d,d}|^2}{\sum_{d=1}^{D} |\hat{\rho}_{d,d}|^2} \tag{4.14}$$
   .

5. A D2D pair with high J_d and low C_d is more preferable to have a higher share in the interference drop. Hence, each D2D pair is given a score

$$W_d = (1-\zeta) J_d + \zeta (1 - C_d) \tag{4.15}$$

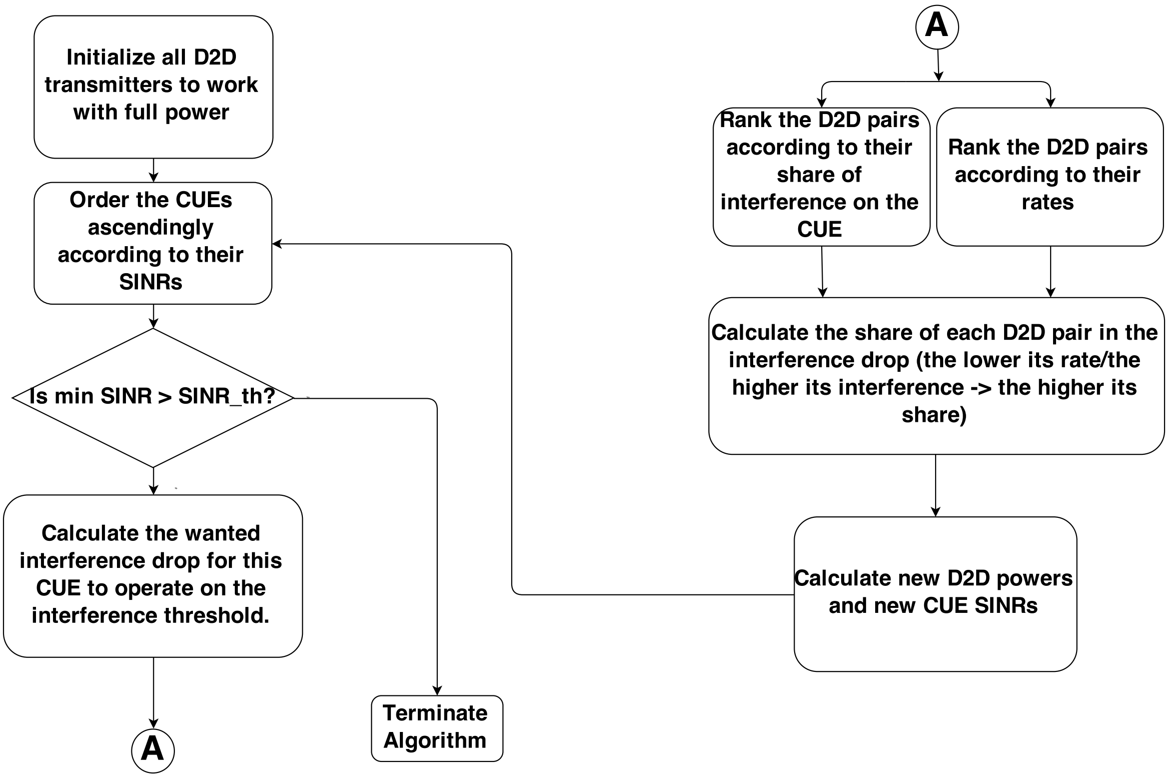

Figure 4.1: Flowchart of the Heuristic PA algorithm

which is directly proportional to J_d and inversely proportional to C_d. The higher the score W_d of a certain D2D pair is, the lower its power gets.

6. The weighting factor $\zeta \in [0,1]$ defines the weights of the parameters J_d and C_d in the calculation of W_d. If $\zeta = 1$, pairs with lower possible rates are the ones that contribute to the power adjustment, which is preferable to maintain higher sum rate, but this may violate the constraint that $\Lambda^{\text{new}} \succeq 0$. In the other extreme $\zeta = 0$, the possible rates of the D2D pairs are not taken into consideration and the pairs adjust their powers according to their interference on the CUE. A simple linear search for a suitable choice of ζ is used.

The algorithm has a computational complexity of order $\mathscr{O}\left(KD^2\right)$.

Algorithm 4.1 D2D power allocation heuristic

Step 1: Initialize: $\Lambda = \mathbf{1}_{1\times(D+1)}$

Step 2: Calculate the interference on CUEs

$$\delta_k = \sum_{d=1}^{D} \lambda_d P_D |\hat{g}_{d,k}|^2, \quad k \in \{1,\ldots,K\}$$

Sort the CUEs in a descending order according to the interference power on them.

Step 3:

for $k = 1$ to K **do**

  if $\gamma_k^{\mathrm{CUE}} < \gamma_{th}$ **then**

    Step 4: Calculate the wanted drop in interference on the k^{th} CUE as $\delta_k^{drop} = \delta_k - \delta_{th}$ where

$$\delta_{th} = \frac{P_{\mathrm{BS}}|\hat{\mathbf{h}}_k^H \mathbf{w}|^2}{\gamma_{th}} - N_o$$

    Step 5: Calculate the share of each D2D pair in the interference drop

    for $d = 1$ to D **do**

      - Calculate the share of the d^{th} D2D transmitter in the current interference as $J_d = \lambda_d P_D |\hat{g}_{d,k}|^2 / \delta_k$

      - Calculate an indicator for the strength of the d^{th} pair channel as $C_d = |\hat{\rho}_{d,d}|^2 / \sum_{d=1}^{D} |\hat{\rho}_{d,d}|^2$

    end for

    for $\zeta = 1 : -0.1 : 0$ **do**

      - Calculate the shares of the D2D transmitters of the interference drop as

$$W_d = (1-\zeta)J_d + \zeta(1-C_d)$$

      - Calculate the new PA factors as

$$\lambda_d^{\mathrm{new}} = \lambda_d - \frac{W_d I_{drop}}{P_D |\hat{g}_{d,k}|^2 \sum_{d=1}^{D} W_d}$$

      if Λ^{new} has an all positive values **then**

        $\Lambda = \Lambda^{\mathrm{new}}$

        break;

      end if

    end for

  end if

end for

4.4 MaxMin Problem Formulation and Solution

Maximizing the sum rate of the D2D pairs does not guarantee fairness between the pairs. Another objective function that guarantees some kind of fairness is to maximize the minimum rate of the D2D pairs. Maximization of the minimum D2D rate corresponds to the maximization of the minimum D2D SINR. The optimization problem is then formulated as

$$\begin{aligned} \Lambda^* = \underset{\Lambda}{\operatorname{argmax}} \quad & \min_{d} \left\{ \frac{\mathbf{c}_d^T \Lambda + N_o}{\mathbf{c}_d^T I_d \Lambda + N_o} \right\}_{d=1}^{D}, && (4.16a) \\ \text{subject to} \quad & M\Lambda - \gamma_{th} N_o \succeq 0, && (4.16b) \\ & 0 \preceq \Lambda \preceq 1, && (4.16c) \end{aligned}$$

where M is a $K \times (D+1)$ matrix with $\mathbf{m}_k^T$ as its rows. The optimization problem (4.16a)–(4.16c) is a Generalized Linear Fractional Program (GLFP) that can be solved using the bisection algorithm for quasiconvex optimization problems [48].

Algorithm 4.2 Bisection algorithm for generalized linear fractional programs

Initialize: $L = 10 * \min_{d} \left\{ \frac{\mathbf{c}_d^T \mathbf{1} + N_o}{\mathbf{c}_d^T I_d \mathbf{1} + N_o} \right\}$, $R = 0$, $Counter = 1$, $FeasibilityCounter = 0$

while stopping condition not satisfied **do**

  Take $S = (L+R)/2$ and solve the feasibility problem

$$\begin{aligned} \text{find} \quad & \Lambda, && (4.17a) \\ \text{subject to} \quad & C\Lambda + N_o \geq S(C'\Lambda + N_o), && (4.17b) \\ & F\Lambda - \gamma_{th} N_o \succeq 0, && (4.17c) \\ & 0 \preceq \Lambda \preceq 1 && (4.17d) \end{aligned}$$

  if infeasible **then**

    $L = S$

  else

    $R = S, FeasibilityCounter = FeasibilityCounter + 1$

  end if

  $Counter = Counter + 1$

end while

In the algorithm, L and R denotes the left and right initial limits for the search space.C is a matrix with $\mathbf{c}_d^T$ as its rows, and C' is a matrix with $\mathbf{c}_d^T I_d$ as its rows.

The stopping condition is

$$(R - L < \varepsilon) \,|\, (Counter > C_{th} \& FeasibilityCounter \geq 1) \,|\, Counter > C_{max} \quad (4.18)$$

The algorithm terminates whenever a specific accuracy is reached. If the number of iteration exceeded some value C_{th}, the algorithm terminates whenever it finds a feasible solution. If the algorithm does not find any feasible solution for a certain number of iterations C_{max}, it terminates and declares infeasible problem. Since the bisection algorithm halves the search space in each iteration, the total number of iterations is $\lceil \log_2 (\varepsilon_o/\varepsilon) \rceil$ where ε_o is the accuracy of the initial solution, and ε is the required final accuracy. We use arbitrary values of $\varepsilon_o = 10^{-4}$, $C_{th} = 40$, and $C_{max} = 200$. Each iteration is a Linear Programming (LP) feasibility problem that can be solved in polynomial time.

4.5 Performance Evaluation

In this section, we use Monte-Carlo simulations for a downlink single-cell network to compare the performance of the different PA and precoding schemes. We provide the simulation results of our proposed algorithms, comparing their performance with a conventional QoS-constrained Time Division Multiple Access (TDMA) scheme. In the TDMA scheme, for each channel instance, e.g., coherence time, only one D2D pair is chosen for operation in a round-robin fashion. The power adjustment factor λ for the operating D2D transmitter is calculated such that it does not violate the QoS constraint on all the CUEs.

4.5.1 Simulation Parameters

The simulation parameters are summarized in Table 4.1.

4.5.2 Simulation Results

4.5.2.1 Convergence of the OptSum solution

Fig. 4.2 shows the convergence of the iterative algorithm used in the OptSum solution for the MaxSum optimization problem. The rate achieved by the optimum PA obtained from the convex optimization solver after each iteration is plotted. The power allocation vector Λ_o is initiated with random numbers uniformly distibuted

Table 4.1: PA techniques - Simulation Parameters

Parameter	Value
Cell radius (R)	200 m
Number of base station antennas (N)	100
Number of CUEs (K)	5
Base station maximum power (P_{BS})	30 dBm
Mobile terminal maximum power (P_D)	13 dBm
Path loss exponent (α)	3
Noise power	10^{-7}mW
Users distribution inside the cell	uniform
Average inter-D2D distance	12 m
Threshold SINR for CUEs (γ_{th})	20.96 dB
Number of iterations of OptSum $\left(N_{\mathrm{OptSum}}\right)$	8

in the range$\sim U(0,1)$. It is shown that after each iteration of the algorithm, the objective sum rate increases. The value of the global maximum of the sum rate obtained by the complex exhaustive search algorithm is plotted for reference. As shown, the algorithm converges to the global optimum value obtained by exhaustive search.

4.5.2.2 Sum rates of the D2D pairs

To study the performance of the different PA algorithms on the sum rate of the D2D pairs, Fig. 4.3 shows this sum rate as a function of the number of D2D pairs D. The precoder is chosen as ZF. It can be seen that among the PA schemes, OptSum algorithm outperforms the MaxMin and TDMA algorithms. All the schemes have an increasing sum rate with increasing D, except for the TDMA scheme, due to the fact that, regardless of the number of the D2D pairs, only one pair is active at a time. The PA heuristic has inferior performance to the OptSum Algorithm and achieves about 75% of its optimum rate. The reason why this happens is the greedy full power initialization of the heuristic does not take the inter-D2D interference into consideration, and hence, high interference between the D2D pairs is present and degrades the performance.

4.5.2.3 Minimum rate of the D2D pairs

As for the minimum of the rates of the D2D pairs, Fig. 4.4 shows the performance of the different PA algorithms as a function of the number of D2D pairs D. MaxMin algorithm outperforms OptSum and TDMA algorithms. Once D is greater than 1, the TDMA algorithm achieves a zero minimum since $D-1$ pairs

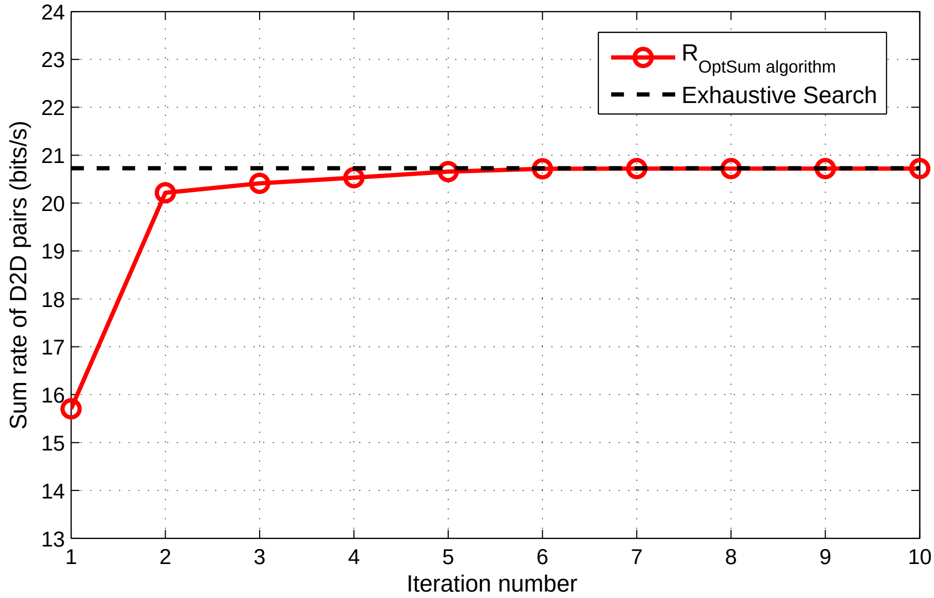

Figure 4.2: Convergence of the iterative algorithm used in the OptSum solution with uniform random initialization. $D = 3$.

Table 4.2: Comparison between the PA schemes. $D = 3$.

	OptSum	Heuristic	MaxMin
Sum rate (bits/s/Hz)	20.723	16.98	15.78
Min rate (bits/s/Hz)	2.4	3.81	5.25
JFI	0.66	0.77	0.98

are always turned off in every channel instance. It is also seen that the heuristic algorithm achieves the same performance as both OptSum and MaxMin for a single D2D pair in the cell.

To have an indication about the fairness of the PA schemes, Jain's Fairness Index (JFI) [49] is used, where

$$JFI = \frac{\left(\sum_{d=1}^{D} R_d\right)^2}{D.\left(\sum_{d=1}^{D} R_d^2\right)} \tag{4.19}$$

and has a maximum value of 1 and minimum value of $1/D$. Table 4.2 compares between the different PA schemes in terms of the D2D performance. It can be seen that the MaxMin algorithm is the best in terms of fairness.

4.5.2.4 Distribution of the D2D rates

To study the effect of the different precoding techniques on the performance of the network, Fig. 4.5 shows the Cumulative Distribution Function (CDF) of the D2D

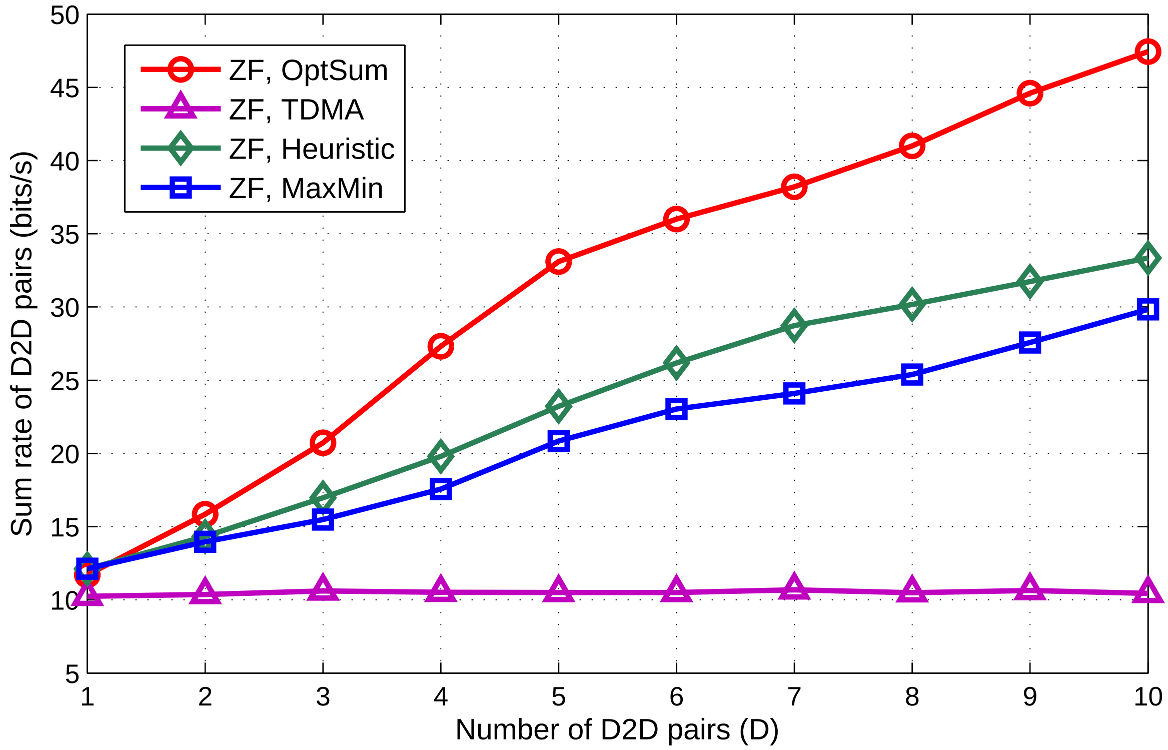

Figure 4.3: Effect of number of D2D pairs on the achievable sum of D2D rates for different PA techniques.

achievable rates under different PA schemes with using BF and ZF precoders. It is clear that the ZF precoder is better than the BF precoder due to its ability to suppress interference on the D2D receivers. The figure also shows that the CDF of the D2D rates under MaxMin PA exhibit higher minimum rate and lower maximum rate than those achieved by OptSum PA. The PA heuristic algorithm acts as a tradeoff between the two schemes, by achieving lower minimum rate than MaxMin and lower sum rate than OptSum.

4.5.2.5 Distribution of the CUEs SINRs

As for the performance of the CUEs, Fig. 4.6 shows the effect of the different PA and precoding schemes on the CDF of the CUEs SINRs. All the algorithms have similar performance in terms of preserving the QoS constraint. BF precoder performs better than ZF precoder, since it has the ability to direct the available power at the BS to the CUEs. For the PA heuristic, one of the K CUEs in each channel instance operates on the SINR threshold. This CUE is the one that is most affected by the operation of the D2D pairs. For $K = 5$, it is shown in the figure that $1/K = 20\%$ of all the CUEs have an SINR equal to the SINR threshold.

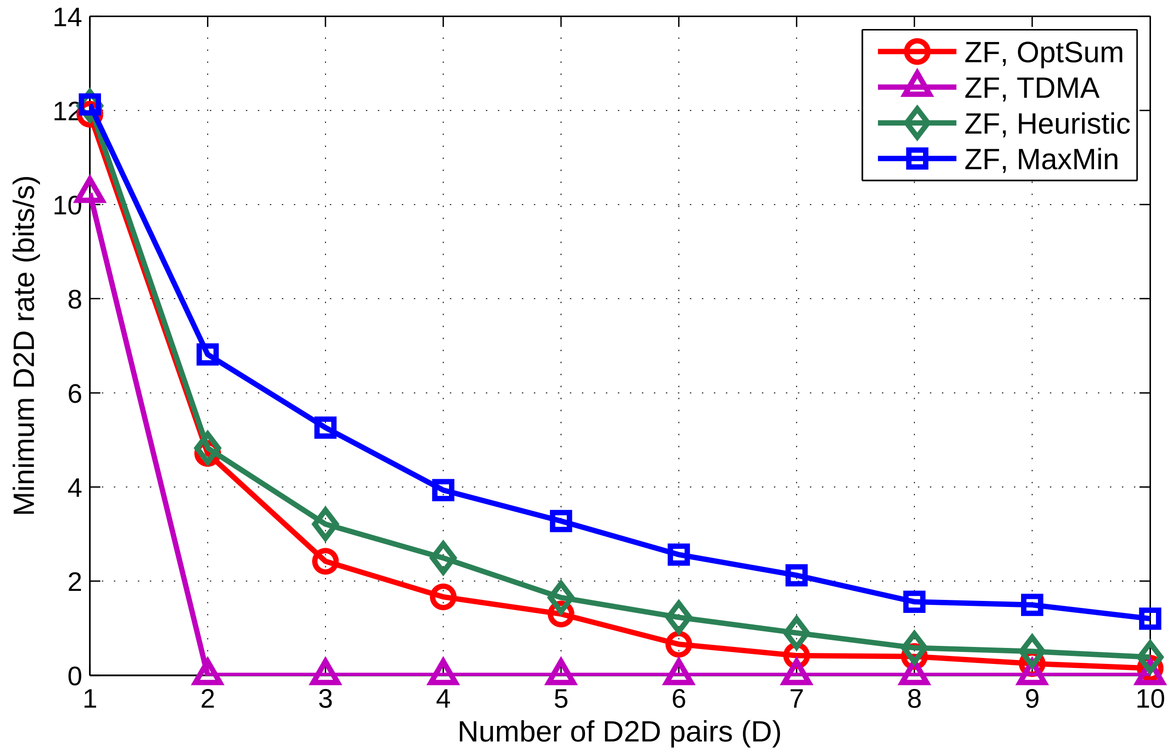

Figure 4.4: Effect of number of D2D pairs on the minimum of D2D rates for different PA techniques.

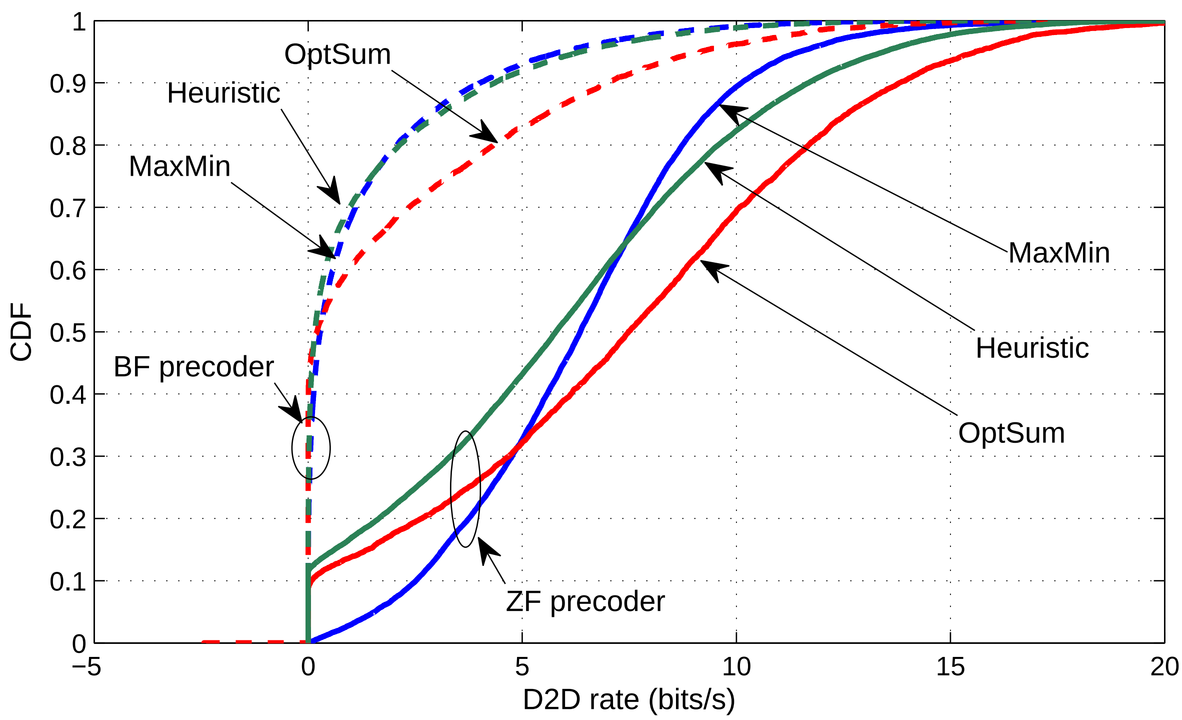

Figure 4.5: CDF of the D2D achievable rates under different PA and precoding schemes. $D = 3$.

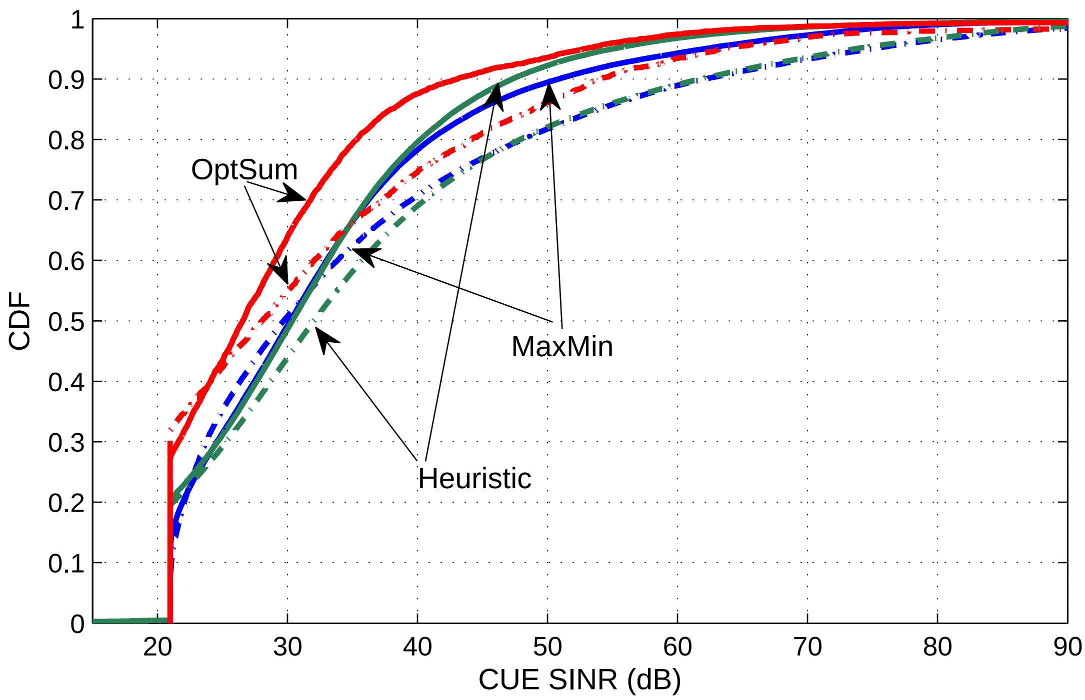

Figure 4.6: CDF of the SINR of the CUEs under different PA and precoding techniques. Solid lines represent ZF precoder. Dashed lines represent BF precoder. $D = 3$.

Chapter 5

BS Precoding

In this chapter, we will formulate the problem of optimizing the precoder $\mathbf{w}$ for a certain determined choice of the PA factors $\lambda_1, \lambda_2, \ldots, \lambda_D, \lambda_{\text{BS}}$. In this problem, we will assume that all the D2D transmitters are working with the same transmit power, i.e. $\lambda_1 = \lambda_2 = \ldots = \lambda_D = \lambda$. The precoder is optimized to maximize the rates of the D2D pairs with maintaining the QoS constraint on the cellular links. From the point of view of the D2D receivers, the obvious optimum precoder choice is the ZeroForcing (ZF) precoder, with their transmitters working with full power. Since this will most likely violate the QoS constraint, the D2D transmitters will back-off with their powers, eventually reaching a suboptimal solution. We will propose two solutions for this problem and compare between their performance.

5.1 Problem Formulation

Assuming known equal PA factors at the D2D transmitters, the optimization problem can be formulated as

$$\mathbf{w}^* = \underset{\mathbf{w}}{\operatorname{argmax}} \sum_{d=1}^{D} \log\left(1 + \frac{\lambda_d P_D \, |\hat{\rho}_{d,d}|^2}{\sum_{\substack{d'=1 \\ d' \neq d}}^{D} \lambda_{d'} P_D \, |\hat{\rho}_{d',d}|^2 + \lambda_{\text{BS}} P_{\text{BS}} |\hat{\mathbf{f}}_d^H \mathbf{w}|^2 + N_o}\right) \tag{5.1a}$$

$$\text{subject to} \qquad |\hat{\mathbf{h}}_k^H \mathbf{w}|^{\mathbf{2}} \geq \tilde{\gamma}_k, \quad \forall k \in \{1, ..., K\}, \tag{5.1b}$$

$$|\mathbf{w}^H \mathbf{w}| = 1, \tag{5.1c}$$

.

Summing up the constant terms, the optimization problem can be rewritten as

$$\mathbf{w}^* = \underset{\mathbf{w}}{\text{argmax}} \sum_{d=1}^{D} \log \left(\frac{|\hat{\mathbf{f}}_d^H \mathbf{w}|^2 + A_d}{|\hat{\mathbf{f}}_d^H \mathbf{w}|^2 + B_d} \right), \tag{5.2a}$$

$$\text{subject to} \qquad |\hat{\mathbf{h}}_k^H \mathbf{w}|^{\mathbf{2}} \geq \tilde{\gamma}_k, \quad \forall k \in \{1, ..., K\}, \tag{5.2b}$$

$$|\mathbf{w}^H \mathbf{w}| = 1, \tag{5.2c}$$

where

$$A_d = \frac{1}{\lambda_{\text{BS}} P_{\text{BS}}} \left(\sum_{d'=1}^{D} \lambda_{d'} P_D \, |\hat{\rho}_{d',d}|^2 + N_o \right), \tag{5.3}$$

$$B_d = \frac{1}{\lambda_{\text{BS}} P_{\text{BS}}} \left(\sum_{\substack{d'=1 \\ d' \neq d}}^{D} \lambda_{d'} P_D \, |\hat{\rho}_{d',d}|^2 + N_o \right), \text{ and} \tag{5.4}$$

$$\tilde{\gamma}_k = \frac{\gamma_{th}}{\lambda_{\text{BS}} P_{\text{BS}}} \left(\sum_{d=1}^{D} \lambda_d P_D \, |\hat{g}_{d,k}|^2 + N_o \right). \tag{5.5}$$

5.2 Solution 1: Semi-Definite Programming (SDP)

If we define $W = \mathbf{w}\mathbf{w}^H$, $F_d = \hat{\mathbf{f}}_d \hat{\mathbf{f}}_d^H$, and $H_k = \hat{\mathbf{h}}_k \hat{\mathbf{h}}_k^H$, the optimization problem in (5.2a)–(5.2c) can be rewritten as

$$W^* = \underset{W}{\text{argmax}} \sum_{d=1}^{D} \log \left(\frac{\text{Tr}(F_d W) + A_d}{\text{Tr}(F_d W) + B_d} \right), \tag{5.6a}$$

$$\text{subject to} \qquad \text{Tr}(H_k W) \geq \tilde{\gamma}_k, \quad \forall k \in \{1, ..., K\}, \tag{5.6b}$$

$$\text{Tr}(W) = 1, \tag{5.6c}$$

$$W \succeq 0, \tag{5.6d}$$

$$\text{rank}(W) = 1, \tag{5.6e}$$

where we used the fact that

$$|\hat{\mathbf{f}}_d^H \mathbf{w}|^2 = \hat{\mathbf{f}}_d^H \mathbf{w}\mathbf{w}^H \hat{\mathbf{f}}_d = \text{Tr}(\hat{\mathbf{f}}_d^H \mathbf{w}\mathbf{w}^H \hat{\mathbf{f}}_d) = \text{Tr}(\hat{\mathbf{f}}_d \hat{\mathbf{f}}_d^H \mathbf{w}\mathbf{w}^H) = \text{Tr}(F_d W) \tag{5.7}$$

. Constraint (5.6b) represents the QoS constraint, while (5.6c) represents the power constraint, and (5.6d)–(5.6e) emerged naturally from the definition of W.

The non-convex objective function and the rank-one contraint prevent the convexity of the optimization problem. The rank constraint can be dealt with by dropping it to relax the optimization problem.

The objective function can be rewritten as

$$\begin{aligned}
W^* = \; & \underset{W}{\text{argmax}} \; \sum_{d=1}^{D} \log\left(\frac{\text{Tr}(F_d W) + A_d}{\text{Tr}(F_d W) + B_d}\right) \\
= \; & \underset{W}{\text{argmax}} \; \sum_{d=1}^{D} \log\left(\frac{A_d\left(1 + \frac{\text{Tr}(F_d W)}{A_d}\right)}{B_d\left(1 + \frac{\text{Tr}(F_d W)}{B_d}\right)}\right) \\
\approx \; & \underset{W}{\text{argmax}} \; \sum_{d=1}^{D} \left\{\log\left(\frac{A_d}{B_d}\right) + \frac{\text{Tr}(F_d W)}{A_d} - \frac{\text{Tr}(F_d W)}{B_d}\right\} \\
= \; & \underset{W}{\text{argmin}} \; \sum_{d=1}^{D} \left\{\left(\frac{A_d - B_d}{A_d B_d}\right) \text{Tr}(F_d W)\right\}
\end{aligned} \tag{5.8}$$

where we used the fact that, for the optimum W, $\text{Tr}(F_d W)$ should be very small as it represents the interference from the BS onto the D2D receivers, and $\log(1+x) \approx x$ for very small x. The term $\log(A_d/B_d)$ has been neglected because it is constant with respect to the optimization variable W. This assumption is valid due to the fact that we are dealing with a Massive MIMO system, hence, the BS has a lot of degrees of freedom to null its transmitted power in the direction of the D2D receivers, without too much effect on the links of the cellular UEs. Substituting the approximate objective function (5.8) and relaxing the problem, (5.6a)–(5.6e) becomes

$$W^* = \underset{W}{\text{argmin}} \sum_{d=1}^{D} \left(\frac{A_d - B_d}{A_d B_d}\right) \text{Tr}(F_d W), \tag{5.9a}$$

$$\text{subject to} \quad \text{Tr}(H_k W) \geq \tilde{\gamma}_k, \forall k \in \{1, ..., K\}, \tag{5.9b}$$

$$\text{Tr}(W) = 1, \tag{5.9c}$$

$$W \succeq 0 \tag{5.9d}$$

The optimization problem in (5.11a)–(5.11d) is a Semi-Definite Program (SDP) and can be solved using interior point methods (see appendix A).

One important parameter in the solution is the choice of λ, the unified power adjustment factor for all the D2D transmitters. Initially, λ is initialized to 1 for full power operation. If the optimization problem is infeasible, i.e. there is no way the BS could precode its data such that all the QoS constraints are met, the D2D transmitters back-off with a constant step of $-\Delta\lambda$ until the optimization problem becomes feasible. If λ becomes equal to zero, and the optimization problem is yet infeasible, that means that the constraints cannot be met even in the absence of the D2D pairs, hence, the BS shuts down in such channel instance and waits for another channel instance in time (or operates on another coherence bandwidth if in a multicarrier system).

If, for a certain λ, the optimization problem becomes feasible, it is solved and the optimum W^* , which doesn't necessarily has a rank of 1, is obtained. If W^*has a rank of 1, its principal component is chosen as the optimum precoder $\mathbf{w}^*$ and the solution terminates. Otherwise, a randomization technique such as the one used in [50] can be used. First, the eigen decomposition of $W_{opt} = U\Sigma U^H$ is calculated and a number of M precoding vectors $\mathbf{w}_m$, $m = 1:M$ are calculated as

$$\mathbf{w}_m = U\Sigma^{0.5}\mathbf{e}_m \tag{5.10}$$

, where $[\mathbf{e}_m]_n = e^{j\theta_{m,n}}$,$n = 1:N$, where $\theta_{m,n}$ are independent and uniformly distributed on $[0,2\pi)$. This way ensures that $\mathbf{w}_m\mathbf{w}_m^H = 1$ regardless the value of $\mathbf{e}_m$. From those M precoding vectors, the one that maximizes the sum rate of the D2D pairs is chosen as the optimum precoder $\mathbf{w}^*$. The overall algorithm is summarized in 5.1.

Such an SDP problem can be solved using interior point methods in a worst-case complexity of $\mathcal{O}\left(\frac{1}{\Delta\lambda}\left(K+D+N^2\right)^{3.5}\right)$. [50]

5.3 Solution 2: Gradient Descent Algorithm (GD)

Due to the high complexity of the SDP algorithm, a gradient descent algorithm is proposed to find the optimum precoder $\mathbf{w}$ which minimizes the function $-\sum_{d=1}^{D} R_{d2d}$. Gradient descent algorithms are based on starting from an initial solution $\mathbf{w}_o$ and then updating this solution by moving in the opposite direction of the gradient of the objective function w.r.t the variable $\mathbf{w}$. GD algorithms are

Algorithm 5.1 SDP algorithm for precoding

1: Initialize: $\lambda = \lambda_1 = \lambda_2 = \ldots = \lambda_D = 1$
2: Calculate a_d, b_d, c_k
3: Solve the SDP optimization problem

$$W^* = \underset{W}{\operatorname{argmin}} \sum_{d=1}^{D} \left(\frac{A_d - B_d}{A_d B_d} \right) \operatorname{Tr}(F_d W), \tag{5.11a}$$

$$\text{subject to} \quad \operatorname{Tr}(H_k W) \geq \tilde{\gamma}_k, \forall k \in \{1, ..., K\}, \tag{5.11b}$$

$$\operatorname{Tr}(W) = 1, \tag{5.11c}$$

$$W \succeq 0 \tag{5.11d}$$

4: **if** infeasible **then**
5: $\quad \lambda = \lambda - \Delta\lambda$
6: $\quad$ Go to 2.
7: **end if**
8: **if** rank(W^*) = 1 **then**
9: $\quad$ **w** = first eigenvector of W^*
10: **else**
11: $\quad$ Calculate the eigen decomposition of $W^* = U \Sigma U^H$
12: $\quad$ **for** $m = 1 : M$ **do**
13: $\quad\quad$ Calculate $\mathbf{w}_m = U \Sigma^{0.5} \mathbf{e}_m$, where $[\mathbf{e}_m]_n = e^{j\theta_{m,n}}$, where $\theta_{m,n}$ are independent and uniformly distributed on $[0, 2\pi)$
14: $\quad\quad$ Calculate $R_m = \sum_{d=1}^{D} \log \left(\frac{|\hat{\mathbf{f}}_d^H \mathbf{w}_m|^2 + A_d}{|\hat{\mathbf{f}}_d^H \mathbf{w}_m|^2 + B_d} \right)$
15: $\quad$ **end for**
16: $\quad$ Choose $\mathbf{w} = \mathbf{w}_{m^*}$, where $m^* = \underset{m}{\operatorname{argmax}} R_m$
17: **end if**

used for non-constrained optimization. In order to account for the constraints of the problem at hand, the following modifications are done:

CUEs QoS constraint

To account for the CUEs QoS constraint, a log-barrier function

$$\phi(\mathbf{w}) = -\sum_{k=1}^{K} \log\left(|\hat{\mathbf{h}}_k^H \mathbf{w}|^{\mathbf{2}} - \tilde{\gamma}_k\right) \tag{5.12}$$

is added to the objective function. The added function turns to ∞ when the constraint is about to be violated. Hence, the descent algorithm turns away from that point.

Precoder Unit Power Constraint

The conventional update equation for a GD algorithm is a straightforward equation, $\mathbf{w}_{new} = \mathbf{w}_{old} - \mu\nabla_{\mathbf{w}}$, where $\nabla_{\mathbf{w}}$ is the gradient of the objective function w.r.t $\mathbf{w}$ and μ is an arbitrary step size where $\mu \in \mathbb{R}$. It can be seen that $||\mathbf{w}_{old}||$ being equal to 1 does not necessarily mean that $||\mathbf{w}_{new}|| = 1$, where $||.||$ denotes the norm of a vector. One solution is to always normalize the new precoder by dividing it by its norm. Another solution was proposed in literature that proved to be better in terms of convergence. The gradient is projected onto the sphere on which all the points satisfy that $||\mathbf{w}|| = \mathbf{1}$. Hence, "moving along the gradient direction" can be thought of as "rotating along the unit sphere in the gradient direction". $\mathbf{r}$ is the projection of $\nabla_{\mathbf{w}}$ on the sphere, and is calculated by subtracting the projection of $\nabla_{\mathbf{w}}$ on $\mathbf{w}$ which is always normal to the sphere surface

$$\mathbf{r} = \frac{\nabla_{\mathbf{w}} - \mathbf{w}^H \nabla_{\mathbf{w}} \mathbf{w}}{||\nabla_{\mathbf{w}} - \mathbf{w}^H \nabla_{\mathbf{w}} \mathbf{w}||} \tag{5.13}$$

. The update equation is then a linear combination of the unity-norm vectors $\mathbf{r}$ and $\mathbf{w}$. Another advantage of using the new update equation is the limited search space for the "step size" $\theta \in [0, 2\pi[$ instead of $\mathbb{R}$ in the traditional linear step update equation.

It is worth noting that for the algorithm to work, the initial solution should be a feasible solution, i.e. satisfying the constraints. Hence, as a best effort to find an initial feasible solution, the beamforming precoder is used and then the powers of the D2D transmitters are adjusted such that none of the QoS constraints is violated. Once the initial solution is obtained, the GD agorithm is performed and $\mathbf{w}$ is updated to a new value. Before this value is further updated, the PA factor λ is revisited and recalculated to be as maximum as possible for a feasible solution with the new $\mathbf{w}$.

The complexity of the GD algorithm is

$$\mathscr{O}\left(Q(K+D)N^2\right) \tag{5.14}$$

.

Algorithm 5.2 Gradient descent algorithm for precoding

1: Define an objective function with log barrier:

$$f(\mathbf{w}) = \sum_{d=1}^{D} \log\left(\frac{|\hat{\mathbf{f}}_d^H\mathbf{w}|^2+B_d}{|\hat{\mathbf{f}}_d^H\mathbf{w}|^2+A_d}\right) - \sum_{k=1}^{K} \log\left(|\hat{\mathbf{h}}_k^H\mathbf{w}|^2-\tilde{\gamma}_k\right) \tag{5.15}$$

2: Initialize: $\lambda = \lambda_1 = \lambda_2 = \ldots = \lambda_D = 1$
3: Initialize: $\mathbf{w} = \mathbf{w}_{BF}$
4: **while** $|\hat{\mathbf{h}}_k^H\mathbf{w}|^2 < \tilde{\gamma}_k, \quad$ for any $\quad k \in \{1,...,K\}$ **do**
5: $\quad \lambda = \lambda - \Delta\lambda$
6: **end while**
7: **for** $q = 1$ to Q **do**
8: $\quad$ Calculate the gradient w.r.t $\mathbf{w}$ as

$$\nabla_{\mathbf{w}^{(q)}} = \sum_{d=1}^{D} \frac{2\hat{\mathbf{f}}_d\hat{\mathbf{f}}_d^H\mathbf{w}^{(q)}(A_d - B_d)}{(|\hat{\mathbf{f}}_d^H\mathbf{w}^{(q)}|^2+A_d)(|\hat{\mathbf{f}}_d^H\mathbf{w}^{(q)}|^2+B_d)} - \sum_{k=1}^{K} \frac{2\hat{\mathbf{h}}_k\hat{\mathbf{h}}_k^H\mathbf{w}^{(q)}}{|\hat{\mathbf{h}}_k^H\mathbf{w}^{(q)}|^2-\tilde{\gamma}_k} \tag{5.16}$$

9: $\quad$ Calculate the surface projection vector $\mathbf{r}$

$$\mathbf{r}_{(q)} = \frac{\nabla_{\mathbf{w}^{(q)}} - {\mathbf{w}^{(q)}}^H\nabla_{\mathbf{w}^{(q)}}\mathbf{w}^{(q)}}{\left\|\nabla_{\mathbf{w}^{(q)}} - {\mathbf{w}^{(q)}}^H\nabla_{\mathbf{w}^{(q)}}\mathbf{w}^{(q)}\right\|} \tag{5.17}$$

10: $\quad$ Calculate α where

$$\theta = \underset{\theta}{\operatorname{argmin}}\ f\left(\cos(\theta)\mathbf{w}^{(q)} + \sin(\theta)\mathbf{r}^{(q)}\right) \tag{5.18}$$

11: $\quad$ Calculate $\mathbf{w}^{(q+1)} = \cos(\theta)\mathbf{w}^{(q)} + \sin(\theta)\mathbf{r}^{(q)}$
12: **end for**

Table 5.1: Precoding techniques - Simulation Parameters

Parameter	Value
Cell radius (R)	200 m
Number of CUEs (K)	5
Base station maximum power (P_{BS})	30 dBm
Mobile terminal maximum power (P_D)	13 dBm
Path loss exponent (α)	3
Noise power	10^{-7}mW
Users distribution inside the cell	uniform
Average inter-D2D distance	12 m
Threshold SINR for CUEs (γ_{th})	10.96 dB
Power backoff factor of SDP ($\Delta\lambda$)	0.05
Randomization iterations of SDP (M)	100
Number of Gradient descent iterations (Q)	20

5.4 Performance Evaluation

In this section, we use Monte-Carlo simulations for a downlink single-cell network to compare the performance of the different precoding schemes. We provide the simulation results of our proposed algorithms, comparing their performance with the conventional Beamforming (BF) and Zeroforcing (ZF) precoders, where

$$\mathbf{w}_{\text{BF}} = \frac{\sum_{k=1}^{K} \hat{\mathbf{h}}_k}{\sqrt{N \sum_{k=1}^{K} \beta_k^{(h)}}}, \tag{5.19}$$

and $\mathbf{w}_{\text{ZF}}$ is chosen such that

$$\mathbf{w}_{\text{ZF}} = \frac{1}{\chi} \left(\text{Null}\left(F^H\right)\right) \left(\text{Null}\left(F^H\right)\right)^H \sum_{k=1}^{K} \hat{\mathbf{h}}_k, \tag{5.20}$$

where $F = \left[\hat{\mathbf{f}}_1, \hat{\mathbf{f}}_2, \ldots, \hat{\mathbf{f}}_D\right]$, Null(.) is the null space of a matrix, and χ is a normalization factor to ensure that$|\mathbf{w}_{\text{ZF}}^H \mathbf{w}_{\text{ZF}}| = 1$. For fair comparison, in both the ZF and BF precoders, the D2D transmitters operate initially with full power, and then back-off linearly until the satisfaction of the QoS constraint.

5.4.1 Simulation Parameters

The simulation parameters are shown in Table 5.1.

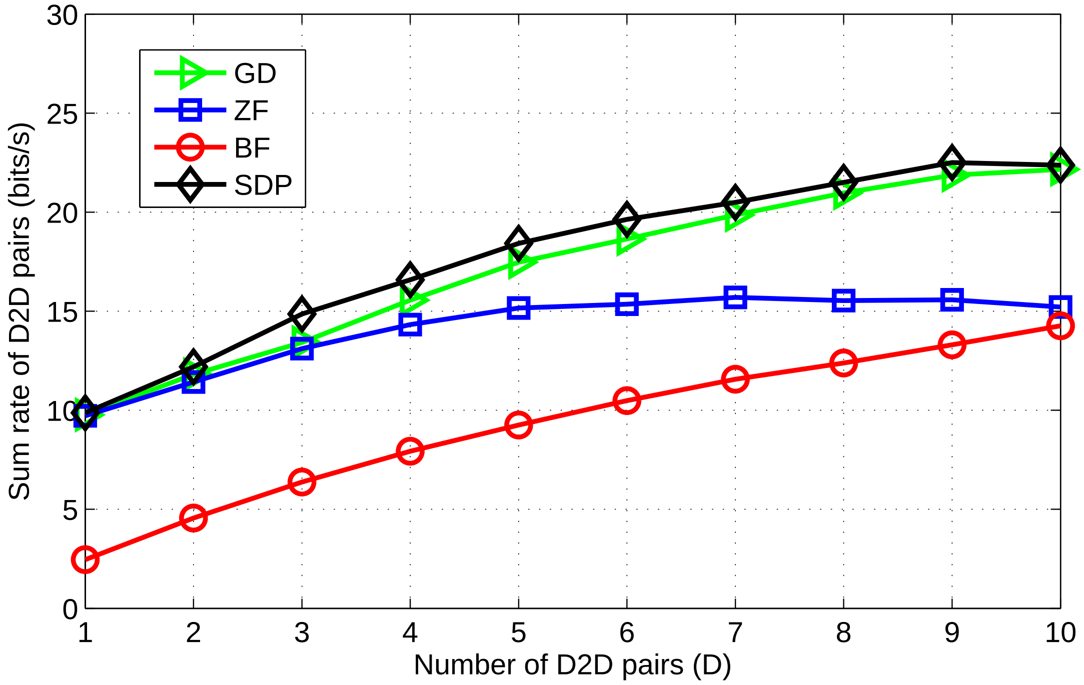

Figure 5.1: Effect of the number of D2D pairs on the performance of different precoding schemes. $N = 100$.

5.4.2 Simulation results

5.4.2.1 Sum rate of D2D pairs vs. D

Fig. 5.1 shows the effect of the number of D2D pairs on the performance of the different precoding schemes in terms of the D2D pairs sum rate. It can be seen the SDP algorithm outperforms all the other algorithms since it finds the optimal precoder. The GD algortihm can be seen to achieve more than 90% of the rates achieved by the SDP . The ZF precoder is supposed to be optimal in terms of the D2D rates since it suppresses all the interference on the D2D receivers, but due to the fact that when the number of D2D pairs increases, the total interference on a CUE gets higher, so the D2D transmitters have to lower their powers more in order to satisfy the QoS contraint. That is why the performance of the ZF precoder is not optimal and does not increase much with increasing D2D pairs. The BF precoder has the worst performance since its interference on the D2D receivers is not managed, but the power back-off step has lower effect since the SINRs at the CUEs are generally better than the case with ZF.

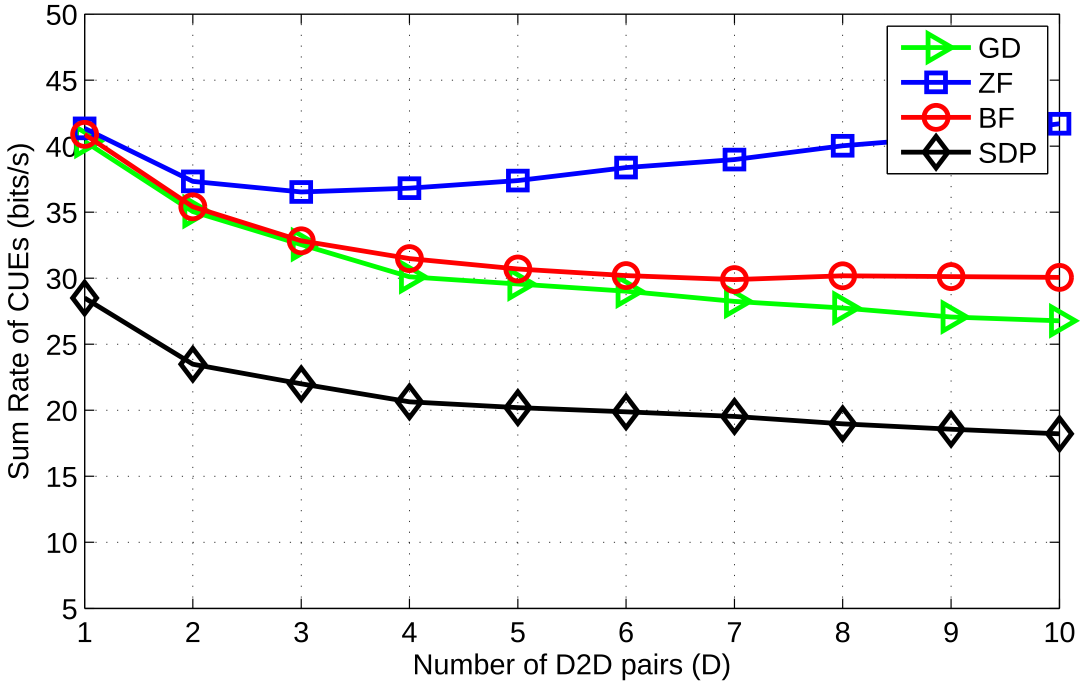

Figure 5.2: Effect of number of D2D pairs on the achievable sum of CUE rates for different precoding techniques. $N = 100$

5.4.2.2 Sum rate of the CUEs vs. D

To study the performance of the different precoding algorithms on the sum rate of the CUEs, Fig. 5.2 shows this sum rate as a function of the number of D2D pairs D. The performance of the ZF precoder is unintuitevly better than all the other precoders, but this is expected due to the huge power back-off step after precoding. The GD algorithm has better performance than the SDP algorithm and achieves 1.5 times more CUE sum rate due to the log-barrier function used in the GD algorithm. This barrier prohibits the CUEs from operating on or near the SINR threshold, and pushes the CUEs away from it. While in SDP, there is no hindrance that a CUE operates on the SINR threshold. We can say that the GD algorithm favors the perforamance of the CUEs on account of performance loss in terms of the D2D rates as shown previously. The performance of the precoders expectedly gets worse with increasing the number of D2D pairs due to the high interference on the CUEs.

5.4.2.3 Distribution of the CUEs SINRs

The same arguments presented above can be further proven by plotting the Cumulative Distribution Function (CDF) of the SINRs of the CUEs under different

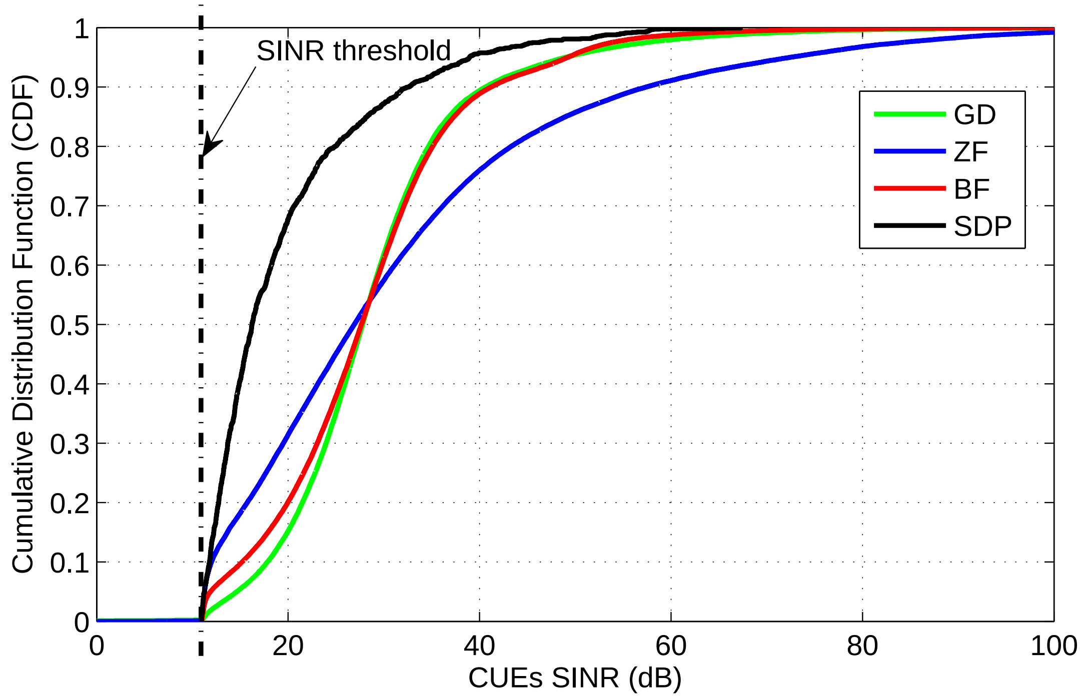

Figure 5.3: CDF of the SINR of the CUEs under different precoding techniques. $N = 100, D = 3$.

precoders. This is shown in 5.3. The SINRs achieved by the GD algorithm are further away from the SINR threshold than those achieved by the SDP.

5.4.2.4 Effect of the number of BS antennas N on performance

Fig. 5.4 shows the effect of the number of the BS antennas (N) on the sum rate of the D2D pairs. For the sake of comparison, the peformance of a scheme in which ZF precoding is used at the BS and the D2D transmitters operate with full power is evaluated and plotted. The achievable rate of that scheme is constant and does not depend on the number of BS antennas since the D2D pairs always operate in a BS-interference-free environment. We can see that the performance of the SDP algorithm approaches that optimum algorithm with increasing number of BS antennas since the approximation made in the SDP algorithm 5.8 becomes more valid as explained in tha algorithm.

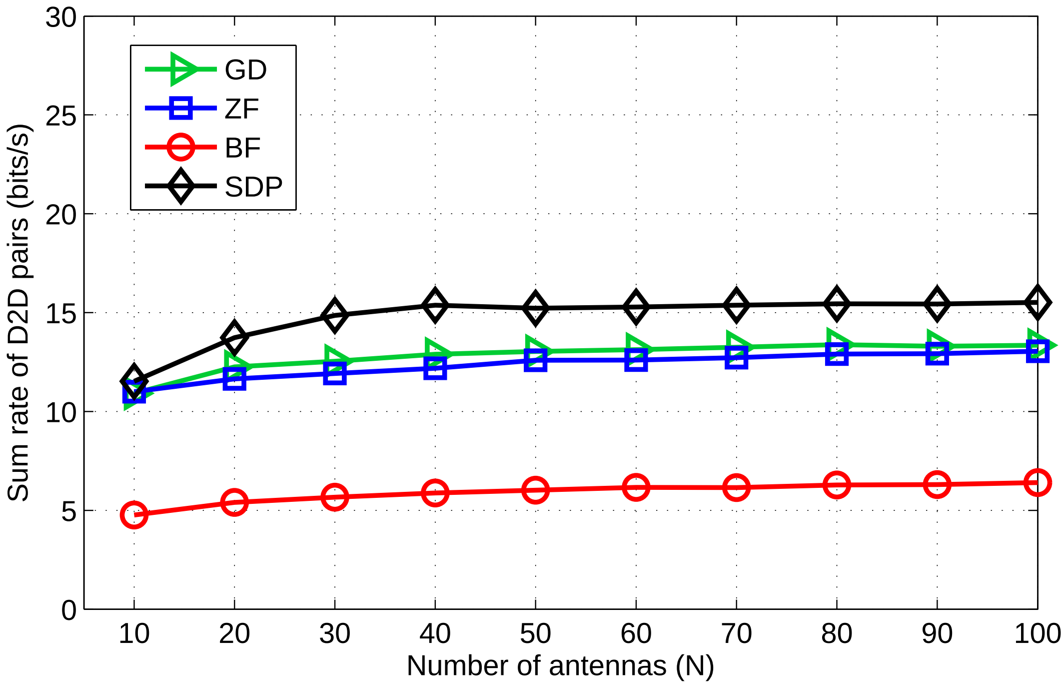

Figure 5.4: Effect of the number of BS antennas on D2D performance. $D = 3$.

Chapter 6

Joint Solution

In this section we target the more general problem of optimizing both the precoder and the power allocation factors to deploy D2D communications in a Massive MIMO network. We will also compare that general solution to some attempts of underlaying D2D communications in cellular networks that are presented in the literature. It is worth mentioning that, to the best of our knowledge, the combination of the system model that we consider and the objective function that we target has not been considered as is in any of the previous papers in that area, and that should be taken into consideration when comparing our results to the previous ones.

6.1 Problem Formulation

To optimize both the precoder and the power allocation factors, the optimization problem is stated as

$$\max_{\lambda_1,\lambda_2,\ldots,\lambda_D,\lambda_{\mathrm{BS}},\mathbf{w}} \quad \sum_{d=1}^{D} log\left(1+\frac{\lambda_d P_D\,|\hat{\rho}_{d,d}|^2}{\sum_{\substack{d'=1\\ d'\neq d}}^{D}\lambda_{d'}P_D\,|\hat{\rho}_{d',d}|^2+\lambda_{\mathrm{BS}}P_{\mathrm{BS}}|\hat{\mathbf{f}}_d^H\mathbf{w}|^2+N_o}\right), \tag{6.1a}$$

$$\text{subject to} \quad \gamma_k^{\mathrm{CUE}} \geq \gamma_{th}, \quad \forall k\in\{1,\ldots,K\}, \tag{6.1b}$$

$$0\leq\lambda_i\leq 1, \quad \forall i\in\{1,2,\ldots,D,\mathrm{BS}\}, \tag{6.1c}$$

$$|\mathbf{w}^H\mathbf{w}|=1, \tag{6.1d}$$

Table 6.1: Joint PA and Precoding techniques - Simulation Parameters

Parameter	**Value**
Cell radius (R)	200 m
Number of CUEs (K)	5
Base station maximum power (P_{BS})	30 dBm
Mobile terminal maximum power (P_D)	13 dBm
Path loss exponent (α)	3
Noise power	10^{-7}mW
Users distribution inside the cell	uniform
Average inter-D2D distance	12 m
Threshold SINR for CUEs (γ_{th})	10.96 dB
Number of iterations of OptSum (N_{OptSum})	8
Power backoff factor of SDP ($\Delta\lambda$)	0.05
Randomization iterations of SDP (M)	100
Number of Gradient descent iterations (Q)	20

The problem in (6.1a)–(6.1d) is an NP hard, nonconvex optimization problem. To solve the problem, we propose to adopt the solutions we presented in the previous chapters for the subproblems of optimizing the power allocation factors and the precoding vector separately in a global iterative solution.

Algorithm 6.1 Combination of the PA and precoding scheme

```
1: Initialize: λ = λ_1 = λ_2 = ... = λ_D = 0, λ_BS = 1
2: Initialize: w = w_BF
3: if QoS constraints not met then
4:     Break the algorithm
5: end if
6: for iterations = 1 : 10 do
7:     Find Λ* from OptSum or Heuristic Algorithm
8:     Set Λ = Λ*
9:     Find w* using algorithm 5.1 without steps 4 to 7, or using steps 7 to 12 in
       algorithm 5.2.
10:    Set w = w*
11: end for
```

6.2 Performance Evaluation

6.2.1 Simulation Parameters

The simulation parameters are shown in Table 6.1.

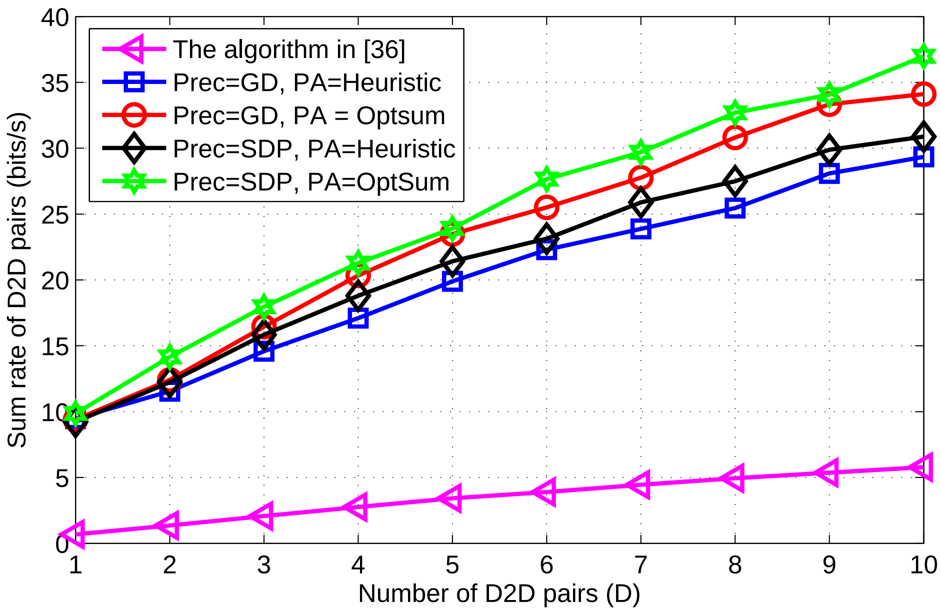

Figure 6.1: Effect of number of D2D pairs on the achievable sum of D2D rates for different combinations of precoding and PA techniques.

6.2.2 Sum rates of the D2D pairs

Fig. 6.1 shows the performance of the different combinations of the precoding and power allocation schemes in terms of the sum rate of the D2D pairs (the objective function). The performance of the algorithm in [36] is plotted for comparison. It should be noted that the objective function of [36] was the total sum rate of the network and the authors did not force QoS constraints in the optimization problem. It can be seen that the higher the complexities of the algorithms involved are, the higher the achievable rates are. For instance, the combination of the two optimal schemes (OptSum for PA and SDP for precoding) achieves around 2.5 times higher sum rate than the combination of the less-complex schemes (Heuristic for PA and GD for precoding). It can also be seen that the performance gap between the algorithms involving the Heuristic scheme as PA and the ones involving OptSum is generally higher than the performance gap between schemes with the same PA and difference precoders. For instance, at $D = 5$, the sum rate achieved by the combination GD/Heuristic is ~22.3 bits/s, while that achieved by the combination GD/OptSum is ~25.5 bits/s. On the other hand, the sum rate achieved by the combination GD/Heuristic is ~22.3 bits/s, while that achieved by the combination SDP/Heuristic is ~23 bits/s.

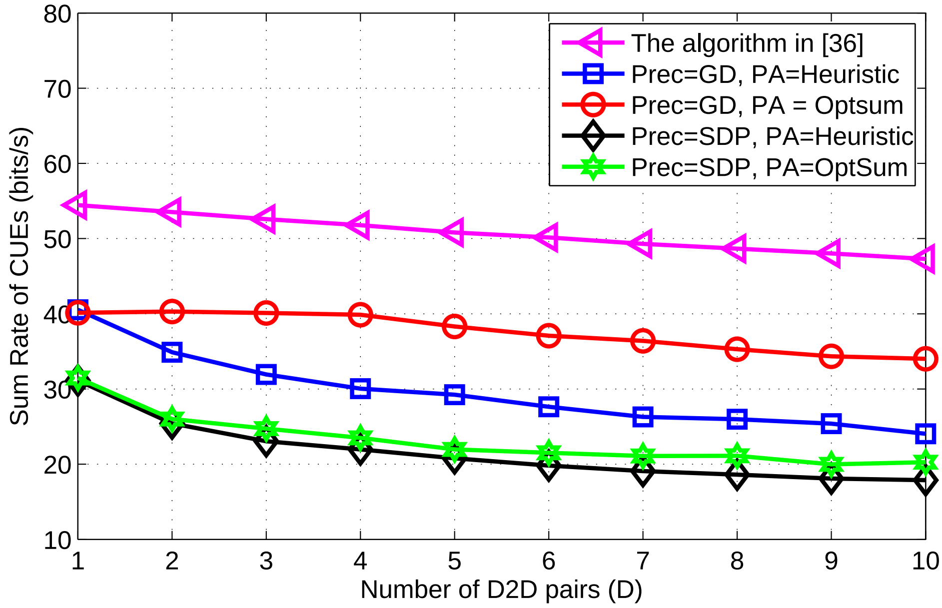

Figure 6.2: Effect of number of D2D pairs on the achievable sum of D2D rates for different combinations of precoding and PA techniques.

6.2.3 Sum rates of the CUEs

Fig. 6.2 shows the performance of the different combinations of the precoding and power allocation schemes in terms of the sum rate of the CUEs. The performance of the algorithm in [36] is also plotted for comparison. With both figures 6.1 and 6.2 in mind, it can be seen that the combination of the GD algorithm for precoding and the OptSum algorithm for PA has an acceptable performance with moderate complexity, since it has the best performance in terms of the CUEs (around 1.8 times better than the combinations including SDP as a precoding technique) with a little loss in the performance of the D2D pairs as shown previously.

6.2.4 SINRs of the CUEs

Fig. 6.3 shows the CDF of the SINRs of the CUEs under different combinations of precoding and PA techniques. It is also seen that combinations involving the GD algorithm have an average median SINR of ~30dB while those involving SDP have that of ~17dB. This confirms the results obtained before that GD algorithm pushes the performance of the CUEs away from the SINR threshold due to the barrier function.

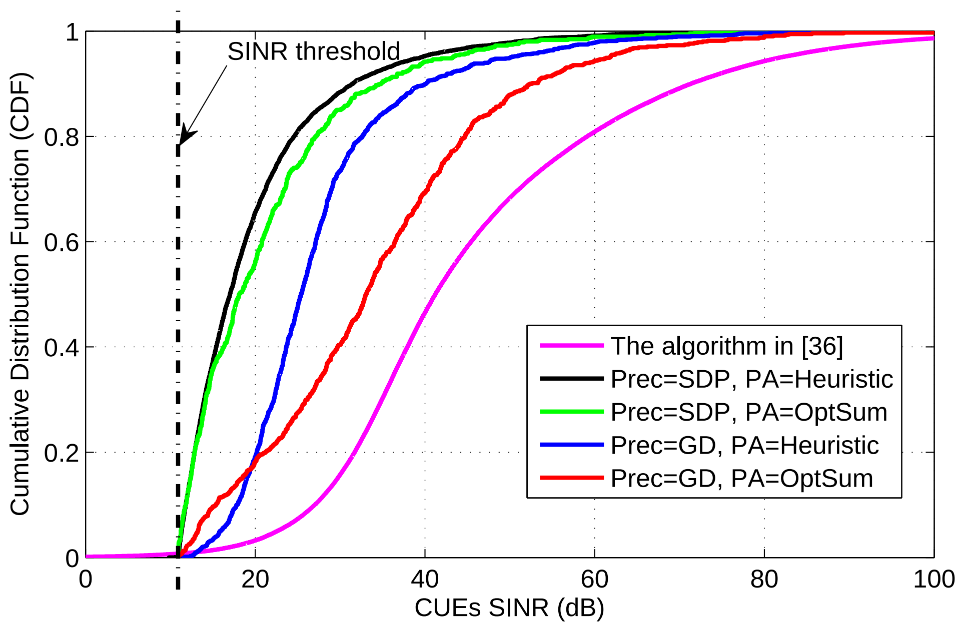

Figure 6.3: CDF of SINRs of the CUEs under different combinations of the precoding and PA schemes $D = 3$.

Chapter 7

Conclusions and Suggested Future Work

7.1 Conclusion

In this thesis, we proposed solutions to deploy D2D communications in a single-cell network with very large number of antennas at the BS. Concerning the PA for the D2D transmitters, two solutions were proposed to maximize the sum rate of the D2D pairs while preserving QoS requirements on the CUEs. OptSum solution was shown to converge to the global optimum in few number of iterations. A heuristic algorithm was proposed to reduce the complexity, on the account of worse performance. We also formulated a problem of maximizing the minimum rate of the D2D pairs instead of their sum, to improve the fairness between the D2D pairs, we proposed a solution for that problem based on the bisection algorithm for quasiconvex problems.

Concerning the precoder design, we also proposed two solutions for the same problem mentioned above. The first one -based on SDP- was shown to achieve above 95% of the sum rate of the D2D pairs in a BS-interference-free precoding, the ZF precoder. Another suboptimal, but far less complex, solution was proposed based on the gradient descent algorithm. The GD algorithm was shown to achieve above 90% of the sum rates achieved by the SDP algorithm. Barrier function and surface projection techniques were used to account for the constraints of our problem.

Finally, we joined both solutions of the precoding problem and the power allocation problem to solve the joint optimization problem. We showed that the performance of the network is greatly enhanced by the proposed schemes (~around 6 times greater D2D sum rate) in comparison with the conventional precoders (BF, ZF, ...) and the conventional PA schemes (uniform, random, ...) and the schemes presented in the literature.

7.2 Suggestions for future work

- Investigate the problem of underlaying D2D communications in a downlink shared (non-multicasting) channel, where different data symbols are forwarded to different CUEs. This would greatly affect the design of the precoder.
- Investigate a scenario where partial CSI is available at the BS.
- Design distributed algorithms for PA problem to reduce the computational overhead at the BS.

List of Publications

1. B. S. Amin, Y. R. Ramadan, A. S. Ibrahim, and M. H. Ismail, “Power Allocation for Device-to-Device Communication Underlaying Massive MIMO Multicasting Networks," in IEEE Wireless Communications and Networking Conference (WCNC), March 2015, pp. 1237–1242.

References

[1] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys Tutorials*, vol. 16, no. 4, pp. 1801–1819, 2014.

[2] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, pp. 186–195, February 2014.

[3] L. Lu, G. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 742–758, Oct 2014.

[4] K. Doppler, M. Rinne, C. Wijting, C. Ribeiro, and K. Hugl, "Device-to-device communication as an underlay to LTE-advanced networks," *IEEE Communications Magazine*, vol. 47, pp. 42–49, Dec 2009.

[5] J. Du, W. Zhu, J. Xu, Z. Li, and H. Wang, "A compressed HARQ feedback for device-to-device multicast communications," in *IEEE Vehicular Technology Conference (VTC Fall)*, pp. 1–5, Sept 2012.

[6] X. Bao, U. Lee, I. Rimac, and R. R. Choudhury, "Dataspotting: Offloading cellular traffic via managed device-to-device data transfer at data spots," *SIGMOBILE Mob. Comput. Commun. Rev.*, vol. 14, pp. 37–39, Dec. 2010.

[7] N. K. Pratas and P. Popovski, "Low-rate machine-type communication via wireless device-to-device (D2D) links," *CoRR*, vol. abs/1305.6783, 2013.

[8] Y.-D. Lin and Y.-C. Hsu, "Multihop cellular: a new architecture for wireless communications," in *IEEE INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies.*, vol. 3, pp. 1273–1282 vol.3, Mar 2000.

[9] K. Doppler, M. Rinne, P. Janis, C. Ribeiro, and K. Hugl, "Device-to-device communications; functional prospects for LTE-advanced networks," in *IEEE International Conference on Communications Workshops*, pp. 1–6, June 2009.

[10] S. Xu and K. S. Kwak, "Effective interference coordination for D2D underlaying LTE networks," in *2014 IEEE 79th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, May 2014.

[11] X. Wu, S. Tavildar, S. Shakkottai, T. Richardson, J. Li, R. Laroia, and A. Jovicic, "Flashlinq: A synchronous distributed scheduler for peer-to-peer ad hoc networks," in *2010 48th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 514–521, Sept 2010.

[12] T. Peng, Q. Lu, H. Wang, S. Xu, and W. Wang, "Interference avoidance mechanisms in the hybrid cellular and device-to-device systems," in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 617–621, Sept 2009.

[13] B. Kaufman and B. Aazhang, "Cellular networks with an overlaid device to device network," in *2008 42nd Asilomar Conference on Signals, Systems and Computer,*, pp. 1537–1541, Oct 2008.

[14] H. Elkotby, K. Elsayed, and M. Ismail, "Exploiting interference alignment for sum rate enhancement in D2D-enabled cellular networks," in *2012 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1624–1629, April 2012.

[15] V. Cadambe and S. Jafar, "Interference alignment and spatial degrees of freedom for the k user interference channel," in *IEEE International Conference on Communication (ICC)*, pp. 971–975, May 2008.

[16] R. Zhang, X. Cheng, L. Yang, and B. Jiao, "Interference-aware graph based resource sharing for device-to-device communications underlaying cellular networks," in *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 140–145, April 2013.

[17] C. Xu, L. Song, Z. Han, Q. Zhao, X. Wang, and B. Jiao, "Interference-aware resource allocation for device-to-device communications as an underlay using sequential second price auction," in *2012 IEEE International Conference on Communications (ICC)*, pp. 445–449, June 2012.

[18] C. Xu, L. Song, Z. Han, D. Li, and B. Jiao, "Resource allocation using a reverse iterative combinatorial auction for device-to-device underlay cellular networks," in *IEEE Global Communications Conference (GLOBECOM)*, pp. 4542–4547, Dec 2012.

[19] K. Akkarajitsakul, P. Phunchongharn, E. Hossain, and V. Bhargava, "Mode selection for energy-efficient D2D communications in LTE-advanced networks: A coalitional game approach," in *2012 IEEE International Conference on Communication Systems (ICCS)*, pp. 488–492, Nov 2012.

[20] H. Min, J. Lee, S. Park, and D. Hong, "Capacity enhancement using an interference limited area for device-to-device uplink underlaying cellular networks," *IEEE Transactions on Wireless Communications*, vol. 10, pp. 3995–4000, December 2011.

[21] X. Chen, L. Chen, M. Zeng, X. Zhang, and D. Yang, "Downlink resource allocation for device-to-device communication underlaying cellular networks," in *IEEE 23rd International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 232–237, Sept 2012.

[22] S. Xu, H. Wang, T. Chen, Q. Huang, and T. Peng, "Effective interference cancellation scheme for device-to-device communication underlaying cellular networks," in *2010 IEEE 72nd Vehicular Technology Conference (VTC)*, pp. 1–5, Sept 2010.

[23] P. Janis, V. Koivunen, C. Ribeiro, J. Korhonen, K. Doppler, and K. Hugl, "Interference-aware resource allocation for device-to-device radio underlaying cellular networks," in *IEEE 69th Vehicular Technology Conference (VTC)*, pp. 1–5, April 2009.

[24] D. Feng, L. Lu, Y. Yuan-Wu, G. Li, G. Feng, and S. Li, "Device-to-device communications underlaying cellular networks," *IEEE Transactions on Communications*, vol. 61, pp. 3541–3551, August 2013.

[25] L. Su, Y. Ji, P. Wang, and F. Liu, "Resource allocation using particle swarm optimization for D2D communication underlay of cellular networks," in *IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 129–133, April 2013.

[26] T. Liang, T. Zhang, J. Cao, and C. Feng, "Joint resource allocation and power control scheme for device-to-device communication underlaying cellular networks," in *2014 International Symposium on Wireless Personal Multimedia Communications (WPMC)*, pp. 568–572, Sept 2014.

[27] M.-H. Han, B.-G. Kim, and J.-W. Lee, "Subchannel and transmission mode scheduling for D2D communication in OFDMA networks," in *2012 IEEE Vehicular Technology Conference (VTC Fall)*, pp. 1–5, Sept 2012.

[28] C.-H. Yu, O. Tirkkonen, K. Doppler, and C. Ribeiro, "Power optimization of device-to-device communication underlaying cellular communication," in *IEEE International Conference on Communications (ICC)*, pp. 1–5, June 2009.

[29] X. Xiao, X. Tao, and J. Lu, "A QoS-aware power optimization scheme in ofdma systems with integrated device-to-device (d2d) communications," in *IEEE Vehicular Technology Conference (VTC Fall)*, pp. 1–5, Sept 2011.

[30] G. Zhang, "Subcarrier and bit allocation for real-time services in multiuser OFDM systems," in *2004 IEEE International Conference on Communications*, vol. 5, pp. 2985–2989 Vol.5, June 2004.

[31] G. Youjun, X. HaiBo, T. Hui, and Z. Ping, "A QoS-guaranteed adaptive resource allocation algorithm with low complexity in OFDMA system," in *International Conference on Wireless Communications, Networking and Mobile Computing (WiCOM)*, pp. 1–4, Sept 2006.

[32] M. Jung, K. Hwang, and S. Choi, "Joint mode selection and power allocation scheme for power-efficient device-to-device (D2D) communication," in *2012 IEEE 75th Vehicular Technology Conference (VTC Spring)*, pp. 1–5, May 2012.

[33] K. Doppler, C.-H. Yu, C. Ribeiro, and P. Janis, "Mode selection for device-to-device communication underlaying an LTE-advanced network," in *2010 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1–6, April 2010.

[34] B. Guo, S. Sun, and Q. Gao, "Power optimization of D2D communications underlying cellular networks in multi-user scenario," in *2014 IEEE International Conference on Communication Systems (ICCS)*, pp. 212–216, Nov 2014.

[35] W. Zhao and S. Wang, "Low complexity power allocation for device-to-device communication underlaying cellular networks," in *2014 IEEE International Conference on Communications (ICC)*, pp. 5532–5537, June 2014.

[36] L. Wei, R. Hu, T. He, and Y. Qian, "Device-to-device (D2D) communications underlaying MU-MIMO cellular networks," in *IEEE Globecom Workshops (GC Wkshps)*, pp. 4902–4907, Dec 2013.

[37] H. Tang, C. Zhu, and Z. Ding, "Cooperative MIMO precoding for D2D underlay in cellular networks," in *2013 IEEE International Conference on Communications (ICC)*, pp. 5517–5521, June 2013.

[38] Y. Pei and Y.-C. Liang, "Resource allocation for device-to-device communication overlaying two-way cellular networks," in *2013 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 3346–3351, April 2013.

[39] C.-H. Yu, K. Doppler, C. Ribeiro, and O. Tirkkonen, "Performance impact of fading interference to device-to-device communication underlaying cellular networks," in *2009 IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 858–862, Sept 2009.

[40] H. ElSawy, E. Hossain, and M.-S. Alouini, "Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks," *IEEE Transactions on Communications*, vol. 62, pp. 4147–4161, Nov 2014.

[41] P. Janis, V. Koivunen, C. Ribeiro, K. Doppler, and K. Hugl, "Interference-avoiding MIMO schemes for device-to-device radio underlaying cellular networks," in *IEEE 20th International Symposium on Personal, Indoor and Mobile Radio Communications*, pp. 2385–2389, Sept 2009.

[42] W. Xu, L. Liang, H. Zhang, S. Jin, J. Li, and M. Lei, "Performance enhanced transmission in device-to-device communications: Beamforming or interference cancellation?," in *2012 IEEE Global Communications Conference (GLOBECOM)*, pp. 4296–4301, Dec 2012.

[43] X. Lin, R. W. H. Jr., and J. G. Andrews, "The interplay between massive MIMO and underlaid D2D networking," *IEEE Transactions on Wireless Communications, submitted*, vol. abs/1409.2792, 2014.

[44] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3590–3600, November 2010.

[45] Z. Xiang, M. Tao, and X. Wang, "Massive MIMO multicasting in noncooperative cellular networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, pp. 1180–1193, June 2014.

[46] J. Hoydis, "Massive MIMO and HetNets: Benefits and challenges," *Newcom Summer School on Interference Management for Tomorrow's Wireless Networks*, 2013.

[47] R. Horst and N. Thoai, "DC programming: Overview," *Journal of Optimization Theory and Applications*, vol. 103, no. 1, pp. 1–43, 1999.

[48] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York, NY, USA: Cambridge University Press, 2004.

[49] R. Jain, D. Chiu, and W. Hawe, "A quantitative measure of fairness and discrimination for resource allocation in shared computer systems," *CoRR*, vol. cs.NI/9809099, 1998.

[50] N. Sidiropoulos, T. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Transactions on Signal Processing*, vol. 54, pp. 2239–2251, June 2006.

[51] "Requirements related to technical performance for imt-advanced radio interface(s)," Tech. Rep. M.2134-0, ITU-R, Dec. 2008.

[52] B. S. Amin, Y. R. Ramadan, A. S. Ibrahim, and M. H. Ismail, "Power allocation for device-to-device communication underlaying massive mimo multicasting networks," in *2015 IEEE Wireless Communications and Networking Conference (WCNC)*, pp. 1237–1242, March 2015.

Appendix A

Convex Optimization

A mathematical optimization problem has the form

$$\begin{array}{lcc} \text{minimize} & f_o(x) & \\ \text{subject to} & f_i(x) \leq b_i & i = 1, \ldots, M \end{array}$$

The vector $x = [x_1, \ldots, x_n]$ is called the optimization variable, the function f_o : $\mathbb{R}^n \rightarrow \mathbb{R}$ is called the objective function, and the functions $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$ are called the constraint functions, and the constants b_i are the boundaries of the constraints. A vector x is called optimal if it has the smallest objective function value among all the vectors that satisfy the constraints.

Optimization problems are classified according to the type of the objective and constraint functions. For example, if f_i, $i = 0, \ldots, M$ are all linear functions of x, the optimization problem is called a linear program, i.e. they all satisfy the condition that

$$f_i(\alpha x + \beta y) = \alpha f_i(x) + \beta f_i(y), \qquad \forall x, y \in \mathbb{R}^n \text{ and } \alpha, \beta \in \mathbb{R} \tag{A.1}$$

If only one of the objective or the constraint functions is not linear, the optimization problem is called nonlinear program. Convex optimization is a type of nonlinear optimization problems in which all the objective function and the constraints functions satisfy the inequality

$$f_i(\alpha x + \beta y) \leq \alpha f_i(x) + \beta f_i(y) \tag{A.2}$$

for all $x, y \in \mathbb{R}^n$ and $\alpha, \beta \in \mathbb{R}$ and $\alpha, \beta \geq 0$ and $\alpha + \beta = 1$. It is obvious that linear and convex problems are not the only classes of optimization, they are, however, two of the most important and famoud due to the availability of effective, fast, and efficient algorithms that solve those two types of problems and find their optimal solution.

Convex optimization problems do not have a closed-form analytical solution. Rather, iterative methods and algorithms were developped to find the optimal solution. The most reliable and well-known methods are called interior point methods. They solve the convex problems in typically 10 to 100 iterations. Research is still done in the area of interior point methods to increase their efficiency and decrease their complexity.

Nonconvex optimization problems do not satisfy A.1 or A.2, hence, convex solutions do not apply on them, but they still play an important role. One of the uses is finding an approximate convex formulation for the nonconvex problem and solving the approximate problem.

Convex optimization is subcategorized into various subfields:

1. Linear Programs (LP) : the ones in which A.2 is satisfied with equality.

2. Second-Order Cone Programs (SOCP): is a type of convex problems that takes the form

$$\begin{array}{ll} \text{minimize} & a^T x \\ \text{subject to} & \|A_i x + b_i\|_2 \leq c_i^T x + d_i, \quad i = 1, \ldots, M \\ & Fx = g \end{array}$$

   where $a, b_i, c_i \in \mathbb{R}^n$, $A_i \in \mathbb{R}^{n_i \times n}$, $d_i \in \mathbb{R}$, $F \in \mathbb{R}^{k \times n}$, and $g \in \mathbb{R}^k$. SOCP can be thought of as a quadratically constrained linear program. It should be noted that if all A_i in an SOCP is equal to zero, the SOCP simplifies to a LP.

3. Conic Programming (CP): a type of convex programs in which the objective function is $f_o : \mathbb{C} \to \mathbb{R}$, where $\mathbb{C}$ is a cone (a vector space that is closed under linear combinations with positive coefficients), and the constraint functions f_i define an affince subspace $\mathscr{H}$.

4. Semi-Definite Programming (SDP): a type of conic programs in which the cone $\mathbb{C}$ is all the positive semidefinite $n \times n$ matrices $\mathbb{S}_+^n$. SDP deals with

optimization problems whose optimization variable is a positive semidefinite matrix and has the general form of

$$\begin{array}{ll} \text{minimize} & \operatorname{Tr}(CX) \\ \text{subject to} & \operatorname{Tr}(A_iX) = b_i, \quad i = 1, \ldots, M \\ & X \succeq 0 \end{array}$$

where Tr(.) is the trace operator, $C, A_1, \ldots, A_M \in \mathbb{S}^n$.