Deterministic and Stochastic Optimization Techniques for Massive MIMO Systems

A Thesis Submitted
in Partial Fulfilment of the Requirements
for the Degree of
Master of Technology

by
Reithick A

to the
DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY
KANPUR

July, 2022

<u>CERTIFICATE</u>

It is certified that the work contained in the thesis entitled ***"Deterministic and Stochastic Optimization Techniques for Massive MIMO Systems"*** by ***"Reithick A"***, has been carried out under my supervision and this work has not been submitted elsewhere for a degree.

Prof. Rohit Budhiraja
Department of Electrical Engineering
Indian Institute of Technology Kanpur

July, 2022

DECLARATION

This is to certify that the thesis titled ***"Deterministic and Stochastic Optimization Techniques for Massive MIMO Systems"***, has been authored by me. It presents the research conducted by me under the supervision of Prof. Rohit Budhiraja.
To the best of my knowledge, it is an original work, both in terms of research content and narrative, and has not been submitted elsewhere, in part or in full, for a degree. Further, due credit has been attributed to the relevant state-of-the-art and collaborations with appropriate citations and acknowledgments, in line with established norms and practices.

Reithick A

17807563

Department of Electrical Engineering
Indian Institute of Technology Kanpur

July, 2022

Synopsis

Name:	**Reithick A**
Roll No.:	**17807563**
Degree for which submitted:	**Master of Technology**
Department:	**Electrical Engineering**
Thesis Title:	**Deterministic and Stochastic Optimization Techniques for Massive MIMO Systems**
Thesis Supervisor:	**Prof. Rohit Budhiraja**
Month and year of submission:	**July, 2022**

This thesis is a two-part work, both with different system model succeeds in framing optimization techniques for Deterministic and Stochastic objective functions. First part, considers a device-to-device (D2D) communication enabled multi-cell massive multi-input multi-output (MIMO) system both operating under short blocklength regime for Ultra Reliable Low Latency Communications and algorithms are approached by Quadratic Transform (QT) and Langrange Dual Transform (LDT) and also stochastic Majorization-Minimization approach for stochastic functions. Second part, considers a multi-relay multi -cell mMIMO NOMA downlink system and are optimized by Weighted Minimum Man Square Error (WMMSE) approach and also stochastic sequential upper-bound minimization technique (SSUM) approach stochastic optimization. To show the efficacy of the proposed solution, numerical simulation are conducted to the show the convergence and inexpensive, low-complexity nature of the optimization frameworks.

"Don't be afraid to give your best to what seemingly are small jobs.
Every time you conquer one it makes you that much stronger.
If you do the little jobs well,
the big ones will tend to take care of themselves."

Dale Carnegie

To my family and friends

ACKNOWLEDGEMENTS

First and foremost, I would like to wholeheartedly thank the Institute, my Supervisor Prof. Rohit Budhiraja and my Phd guide Mr. Dheeraj N Amudala for providing the support throughout my research even in the times of COVID crisis. Their invaluable guidance and their consistent effort have encouraged me to go all the way, in to finishing my thesis under them. They gave me ample amounts of opportunities and a degree of freedom beyond the expectations of a student. In a way, I'm very opportune to have them as my guidance.

I also acknowledge my department, my professors and my institute for teaching classroom as well life lessons. I would also like to thank my parents and my sister, for their everlasting love and support throughout my life. Last but not the least I thank the thickest of my friends for the time and moral support during my stay at the campus. I also thank my lab-mates and my seniors for their constant support throughout.

July, 2022 **Reithick A**

Contents

List of Figures

Chapter 1

Introduction

In the recent decade Massive MIMO(mMIMO) communications are widely popular due to their capability to provide substantial spectral efficiency (SE) and energy efficiency (EE) gains, higher reliability and lower end-to-end latency. Massive antenna arrays are installed at base stations (BSs) in cellular mMIMO systems, which maximises spatial diversity to multiplex tens of users sharing a single spectrum resource while increasing system SE and reliability. In order to limit the amount of inter-user interference in a multi-user mMIMO system, orthogonal multiple access (OMA) techniques are employed which reduce the amount of inter-user interference that occurs. These approaches distribute orthogonal time-frequency resources to each of the system's users (IUI). Even while OMA is very efficient at suppressing IUI, it lowers system SE when users observe channels with low quality. [1, 2].

The paper works with the optimization techniques for two mMIMO systems. The paper approaches both the system models with a varied approach inorder to find a very-low time complexity based algorithms and it provides optimization framework for both Deterministic and Stochastic Models for the both the systems with a iterative closed-form solutions. The papers has also extended its work in optimizing deterministic precoders for the downlink systems using stochastic optimizing frameworks. The system models are acquired from [3] and [4].

The first part of the paper proposes framework to optimize the GEE of D2D underlaid multi-cell Uplink mMIMO systems with URLLC enabled. URLLC is on of the use-case of fifth (5G) generation wireless systems which considers error probability by adding a dispersion term in the Shannon capacity. The framework solves deterministic and stochastic GEE objective function with Quadratic Transform (QT), Lagrange Dual Transform (LDT) and other classic optimization techniques. Chapters 2-5 form the first part of the paper with i)*Introduction*, ii)*System Model and SE*, iii)*GEE maximization* and iv)*Simulation Results and Conclusion* respectively.

The second part of the paper proposes a framework to optimize the sum SE of a Multi-Cell, Multi-Relay Downlink mMIMO NOMA system, which solves both stochastic and deterministic sum SE non-convex problem with iterative closed form solutions and a very-low time complexity based on the WMMSE approach. The paper also extends its work to obtain optimal deterministic precoder allocation based on sum SE maximization which reduces the time and memory complexity of precoder formation and sum SE optimization. Chapters 6-9 form the second part of the paper with i)*Introduction*, ii)*System Model and SE*, iii)*sum SE maximization* and iv)*Simulation Results and Conclusion* respectively.

Chapter 2

D2D underlaid multi-cell mMIMO URLLC system : Introduction

Massive multi-input multi-output (mMIMO) and device-device (D2D) communications are widely popular due to their capability to provide substantial spectral efficiency (SE) and energy efficiency (EE) gains, higher reliability and lower end-to-end latency [5]. In a cellular mMIMO system, the base station (BS) equipped with massive antenna arrays, exploit spatial diversity to multiplex tens of users on the same spectral resource improving the system SE, EE and reliability [5, 6]. D2D communication, on the other hand, enables two users within proximity to communicate with each other without BS's intervention, thereby reducing the end-to-end latency and improving the user EE [7, 8]. Underlaying D2D with cellular mMIMO has recently gained significant interest due to their aforementioned benefits [7–11].

In a D2D underlaid mMIMO system, cellular users (CUs) and D2D user-pairs coexist in a geographical area and share the same time-frequency resources to communicate, thereby causing D2D-to-cellular and cellular-to-D2D interferences at the CUs and the D2D users, respectively [11–13]. He *et al.* in [11] considers a D2D underlaid multi-cell mMIMO system and proposed various power allocation schemes to mitigate the mutual interference between the CUs and D2D users. Authors in [12] derived asymptotic SE expressions for the downlink of multi-cell mMIMO system with D2D users. Ghazanfari *et al.* in [13] derived closed-form SE expressions for an uplink multi-cell mMIMO system with D2D users and uncorrelated Rayleigh faded channels. All these works derived asymptotic /closed-form SE expressions using Shannon's capacity limit, which is assumes very long code blocklengths [14, 15].

Ultra-reliable and low-latency communication (URLLC) is one of the use-case in fifth generation (5G) wireless systems [14–16]. URLLC requires latency in fraction of milliseconds ($\leq$ 1ms) and a very high reliability (error probability $\leq 1e^{-5}$) [14–16]. To satisfy these requirements, URLLC system transmit short packets of data over short channel blocklengths, which are prone to decoding errors [16]. Shannon's capacity limit, which ignores these errors, greatly overestimates the performance of URLLC systems [16, 17]. Authors in [17] derived an SE bound that captures the effect of limited channel blocklength and reliability of URLLC. Using this bound, following works analyzed the SE of URLLC enabled mMIMO systems [18–22]. Östman *et al.* in [18] analyzed the error probability of a single-cell mMIMO URLLC system with maximal ratio (MR) combining scheme. Ren *et al.* in [19] derived closed-form SE expression for an uncorrelated single-cell mMIMO URLLC system with maximal ratio (MR) and zero forcing (ZF) combining schemes. Nasir *et al.* in [20] considered URLLC enabled cell-free mMIMO system and derived closed-form SE expression assuming Rayleigh fading channels. References [18–20] considered URLLC enabled mMIMO system, *but ignored D2D users.* Yang *et al.* in [22] derived closed-form SE expressions and optimized the pilot and data transmit powers for a URLLC enabled D2D underlaid single-cell mMIMO system. All these URLLC works, except [20], considered a single-cell mMIMO system.

The aforementioned works on D2D underlaid mMIMO with URLLC traffic [21, 22] or non-URLLC traffic [11–13], derived closed-form/asymptotic SE expressions assuming uncorrelated Rayleigh faded channels. Practical mMIMO systems, due to the proximity among BS, CUs and D2D users, have a deterministic line-of-sight component along with the non-LoS components in their channels [23]. The BS, further, to maintain smaller design form-factor, consists of closely packed antenna arrays, which observe spatially-correlated channel entries [5]. Such a channel is practically modelled as spatially-correlated Rician-faded [23]. It is crucial to consider spatially-correlated Rician-fading channels as they can change the key findings of the existing URLLC systems. The current works fill this gap by considering a URLLC enabled D2D underlaid multi-cell mMIMO system with spatially-correlated Rician-faded channels.

Design of resource allocation techniques to enhance the reliability and latency of a URLLC system is an active research area [19,21,22]. Ren *et al.* in [19] maximized the weighted sum SE of a single-cell mMIMO URLLC system. Authors in [21] solved the multi-objective problem of transmit power minimization and sum SE maximization, for a D2D underlaid mMIMO URLLC system. Yang *et al.* in [22] maximized the minimum SE of the D2D users under quality of service (QoS) constraints. All these works optimized variants of SE metric i.e., sum SE, weighted sum SE and max-min SE. Exponential growth in URLLC use-cases inevitably led to energy-efficient design of URLLC systems [15, 24]. The global energy efficiency (GEE) metric, defined as a ratio of system throughput to its power consumption, signifies the information (in bits) transmitted per unit of energy consumed (in J) [25, 26]. Sun *et al.* in [15] optimized the EE of a URLLC massive MIMO system under QoS constraints. Keshav *et al.* in [24] optimized beamforming vectors and decoding error probabilities by maximizing the EE of an URLLC system. Both these works, however, considered *single-cell mMIMO system and without D2D users.* To the best of our knowledge, none of the existing works optimized the GEE of a multi-cell D2D underlaid mMIMO system with URLLC.

The SE and GEE metrics in the aforementioned optimization works [15, 19, 21, 22, 24], are stochastic functions of random fading channels and receiver noise. These works adopted a statistically approximated average (SAA) optimization approach, wherein an approximate closed-form expression of SE/GEE is first obtained by statistically averaging out the randomness due to channel fading, and the resultant deterministic expression is then optimized using an iterative algorithm. As a result SAA approach requires more memory and has higher computational and time complexity [27]. Stochastic optimization, in contrast to SAA, constructs an approximate sample objective function that captures the randomness in real-time and then solves it using iterative algorithm [27]. The stochastic optimization approach, therefore, has much lesser time complexity than SAA method [27, 28]. In this work, we propose a generic framework to optimize the GEE of a D2D underlaid mMIMO system, which can be used in both stochastic and SAA settings.

Table 2.1: Summary of mMIMO literature focusing on URLLC, D2D users and Rician fading.

Ref.	Cell	D2D	URLLC	Channel	Corr.	BS combining	Metric
[11]	Multi-cell	✓	✗	Rayleigh	✗	ZF	✗
[13]	Multi-cell	✓	✗	Rayleigh	✗	MR and ZF	max-min SE
[15]	Multi-cell	✗	✓	Rayleigh	✗	MR	EE
[19]	Single-cell	✗	✓	Rayleigh	✗	MR and ZF	WSEE
[22]	Single-cell	✓	✓	Rayleigh	✗	MR	max-min SE
Pr	**Multi-cell**	✓	✓	**Rician**	✓	**MR, ZF, MMSE**	**SE and GEE**

The existing literature on D2D underlaid mMIMO system with/without URLLC is summarized in Table 2.1. We infer from Table 2.1, that none of the existing works derived closed-form SE expression with spatially correlated Rician fading channels and optimized the SE/GEE of a URLLC enabled D2D underlaid multi-cell mMIMO system. The current work bridges these gaps with its main contributions listed below

.

Chapter 3

D2D underlaid multi-cell mMIMO URLLC system : System Model and SE Expression

3.1 System and Signal Model

We consider a D2D underlaid multi-cell mMIMO URLLC system, wherein L cells and D D2D user-pairs coexist in the same geographic area. Each cell comprises of a mMIMO BS that serve K URLLC capable cellular users (CU), while each D2D user-pair consists of a D2D transmitter which communicates its low-latency information with a D2D receiver without the intervention of BSs. The CUs are uniform randomly located within each cell-area, while the D2D user-pairs are randomly located in the whole network and do not belong to any specific cell. We assume that each mMIMO BS is equipped with M antennas, whereas both CU and D2D users have single-antenna to transmit/receive its data. We consider a realistic scenario, wherein the BS and the D2D receivers do not have the knowledge of channels and acquire the same with the help of pilot symbols transmitted by the CUs and D2D transmitters. The communication takes place over a coherence block of τ_c samples, which is split into two phases: channel training phase of duration τ_p samples and data transmission phase of duration $\tau_c - \tau_p$ samples.

3.1.1 Channel Model

Due to insufficient antenna spacing at the BS and the presence of line-of-sight (LoS) links between the communicating nodes, we adopt a spatially correlated Rician fading channel model, which captures both these aspects.

CU to BS and D2D user to BS channels

The spatially correlated Rician-faded channels from kth CU in lth cell $U_{l,k}^{c}$ to jth BS and from d' th D2D transmitter $U_{d',t}^{d}$ to jth BS, denoted respectively as $\mathbf{h}_{lk}^{j} \in \mathbb{C}^{M\times 1}$ and $\mathbf{h}_{d'}^{j} \in \mathbb{C}^{M\times 1}$, are modelled as follows:

$$\mathbf{h}_{z}^{j} = \sqrt{\frac{\beta_z^j K_z^j}{K_z^j + 1}}\overline{\mathbf{h}}_{m_z}^{j} + \sqrt{\frac{\beta_z^j}{K_z^j + 1}}\mathbf{\Sigma}_z^{j\,\frac{1}{2}}\mathbf{h}_{w_z}^{j} \triangleq \overline{\mathbf{h}}_{z}^{j} + \mathbf{R}_z^{j\,\frac{1}{2}}\mathbf{h}_{w_z}^{j}, \text{ where } z = \begin{cases} lk, & \text{for CU,} \\ d', & \text{for D2D.} \end{cases} \tag{3.1}$$

Here the scalars β_z^j and K_z^j denotes the large-scale fading coefficient and rician K factor corresponding to channel $\mathbf{h}_z^j$, respectively. The large-scale fading coefficient captures the effect of distance dependent pathloss and shadowing, while the rician K factor determines the relative strength of the LoS component over the NLoS part. The vector $\overline{\mathbf{h}}_{m_z}^{j}$ represents the deterministic LoS component and is given as $\overline{\mathbf{h}}_{z}^{j} = [1, e^{j2\pi d_h \sin\left(\theta_z^j\right)}, \cdots, e^{j2\pi(M-1)d_h \sin\left(\theta_z^j\right)}]^T$. Here d_h is the antenna spacing at the BS and θ_z^j is the nominal angle between the arbitrary user U_z^α and jth BS, where $(z,\alpha) = (lk, c)$ for CU and; $(z,\alpha) = (d', d)$ for D2D user. The matrix $\mathbf{\Sigma}_z^j$ denotes the spatial correlation matrix and the vector $\mathbf{h}_{w_z}^{j}$ models the small-scale fading whose entries are independent and identically distributed (i.i.d.) with probability distribution function (pdf) $\mathcal{CN}(0,1)$.

CU to D2D user and inter-D2D user channels

The channel from CU $U^c_{l,k}$ to the dth D2D receiver $U^d_{d,r}$ and from d' th D2D transmitter $U^d_{d',t}$ to dth D2D receiver $U^d_{d,r}$ are modelled as

$$g^d_z = \bar{g}^d_z + \sqrt{\frac{\beta^d_z}{K^d_z + 1}} g^d_{w_z}, \text{ with } \bar{g}^d_z = \sqrt{\frac{K^d_z \beta^d_z}{K^d_z + 1}}. \tag{3.2}$$

Here $\bar{g}^d_z$ is the deterministic LoS component and the scalar $g^d_{w_z} \sim \mathcal{CN}(0, 1)$ represents the small scale fading. The scalars K^d_z and β^d_z are the rician K factors and large scale fading coefficients corresponding to channels g^d_z respectively. Here $z = lk$ for CU and; $z = d'$ for D2D user.

3.1.2 Channel Training

In this phase, CUs and D2D transmitters transmit pilot sequences of length τ_p samples, using which the BS and the D2D receivers estimate the channels $\{\mathbf{h}^j_{lk}, \mathbf{h}^j_d, g^d_{lk}, g^d_{d'}\}$[1]. Due to finite coherence block length, it is difficult to assign orthogonal pilots to all the users in the system. The users therefore reuse the same pilots thereby causing pilot contamination. We consider a set of $\tau_p = N + K$, with $N < D$, orthogonal pilots $\mathbf{\Phi} \in \mathbb{C}^{\tau_p \times \tau_p}$, which we partition into $\mathbf{\Phi} = [\mathbf{\Phi}^c \ \mathbf{\Phi}^d]$. Here $\mathbf{\Phi}^c = [\boldsymbol{\phi}^c_1, \cdots, \boldsymbol{\phi}^c_K] \in \mathbb{C}^{\tau_p \times K}$ and $\mathbf{\Phi}^d = [\boldsymbol{\phi}^d_1, \cdots, \boldsymbol{\phi}^d_N] \in \mathbb{C}^{\tau_p \times N}$, denote the pilot sequences assigned to CUs and D2D users, respectively. We similar to [23], assume orthogonal pilots to CUs in each cell and the set $\mathcal{C}_{lk}$ denotes the set of CUs that are using the same pilot as that of $U^c_{l,k}$. We also partition the D2D user-pairs into N sets $\mathcal{D}_1, \cdots, \mathcal{D}_N$, such that users in set $\mathcal{D}_n$ share the pilot sequence $\boldsymbol{\phi}^d_n$. Let $\boldsymbol{\phi}^c_{lk}$ and $\boldsymbol{\phi}^d_d$ be the pilot sequences assigned to CU $U^c_{l,k}$ and the D2D transmitter $U^d_{d,t}$, respectively. The signal received by the jth BS and dth D2D receiver over the duration of τ_p symbols are respectively

[1] We assume $k = 1, \cdots, K$, $l = 1, \cdots, K$, $j = 1, \cdots, L$, $d = 1, \cdots, D$ and $d' = 1, \cdots, D$, throughout this paper.

given as

$$\mathbf{Y}^j = \sum_{l'=1}^{L}\sum_{k'=1}^{K}\sqrt{p^{\rho}_{l'k'}}\mathbf{h}^j_{l'k'}(\boldsymbol{\phi}^c_{l'k'})^T + \sum_{d'=1}^{D}\sqrt{p^{\rho}_{d'}}\mathbf{h}^j_{d'}(\boldsymbol{\phi}^d_{d'})^T + \mathbf{W}^j \text{ and} \tag{3.3}$$

$$\mathbf{y}^d = \sum_{d'=1}^{D}\sqrt{p^{\rho}_{d'}}g^d_{d'}(\boldsymbol{\phi}^d_{d'})^T + \sum_{l'=1}^{L}\sum_{k'=1}^{K}\sqrt{p^{\rho}_{l'k'}}g^d_{l'k'}(\boldsymbol{\phi}^c_{l'k'})^T + \mathbf{w}^d. \tag{3.4}$$

Here p^{ρ}_{lk} (resp. $p^{\rho}_{d'}$) represents the pilot power of CU $U^c_{l,k}$ (resp. $U^d_{d'}$). The vector $\mathbf{w}^d \in \mathbb{C}^{1\times\tau_p}$ and the matrix $\mathbf{W}^j \in \mathbb{C}^{M\times\tau_p}$ denote the additive white Gaussian noise at the jth BS and dth D2D receiver, respectively.

Channel estimation at BS

The jth BS, to estimate channels $\mathbf{h}^j_{lk}$ and $\mathbf{h}^j_{d}$, correlates its received signal matrix $\mathbf{Y}^j$ with respective pilot sequences i.e $(\boldsymbol{\phi}^c_{lk})^*$ or $(\boldsymbol{\phi}^c_{d})^*$ to obtain,

$$\mathbf{y}^{\rho}_{j,lk} = \mathbf{Y}^j(\boldsymbol{\phi}^c_{lk})^* \overset{(a)}{=} \sum_{(l',k')\in\mathcal{C}_{lk}}\sqrt{p^{\rho}_{l'k'}}\tau_p\mathbf{h}^j_{l'k'} + \mathbf{W}^j(\boldsymbol{\phi}^c_{lk})^* \text{ and} \tag{3.5}$$

$$\mathbf{y}^{\rho}_{j,d} = \mathbf{Y}^j(\boldsymbol{\phi}^c_{d})^* \overset{(a)}{=} \sum_{d'\in\mathcal{D}_d}\sqrt{p^{p}_{d'}}\tau_p\mathbf{h}^j_{d'} + \mathbf{W}^j(\boldsymbol{\phi}^d_{d})^*, \tag{3.6}$$

Here equality (a) is because of orthogonality of pilot sequences only CUs and D2D users which share same pilot sequences remains. The MMSE estimates of channels $\mathbf{h}^j_{lk}$ and $\mathbf{h}^j_{d}$, treating (3.5) and (3.6) as sufficient statistics, are given as [23]

$$\hat{\mathbf{h}}^j_{lk} = \overline{\mathbf{h}}^j_{lk} + \sqrt{p^{\rho}_{lk}}\mathbf{R}^j_{lk}\boldsymbol{\Psi}^{-1}_{lk}(\mathbf{y}^{\rho}_{j,lk} - \overline{\mathbf{y}}^{\rho}_{j,lk}) \quad \text{and} \quad \hat{\mathbf{h}}^j_{d} = \overline{\mathbf{h}}^j_{d} + \sqrt{p^{\rho}_{d}}\mathbf{R}^j_{d}\boldsymbol{\Psi}^{-1}_{d}(\mathbf{y}^{\rho}_{j,d} - \overline{\mathbf{y}}^{\rho}_{j,d}). \tag{3.7}$$

where $\overline{\mathbf{y}}^{\rho}_{j,lk} = \sum_{(l',k')\in\mathcal{C}_{lk}}\sqrt{p^{\rho}_{l'k'}}\tau_p\overline{\mathbf{h}}^j_{l'k'}$, $\overline{\mathbf{y}}^{\rho}_{j,d} = \sum_{d'\in\mathcal{D}_d}\sqrt{p^{\rho}_{d'}}\tau_p\overline{\mathbf{h}}^j_{d'}$,

$$\boldsymbol{\Psi}^j_{lk} = \frac{\mathrm{Cov}\{\mathbf{y}^{\rho}_{j,lk}\}}{\tau_p} = \sum_{(l',k')\in\mathcal{C}_{lk}}\tau_p p^{\rho}_{l'k'}\mathbf{R}^j_{l'k'} + \sigma^2\mathbf{I}_M \text{ and } \boldsymbol{\Psi}^j_{d} = \frac{\mathrm{Cov}\{\mathbf{y}^{\rho}_{j,d}\}}{\tau_p} = \sum_{d\in\mathcal{D}_d}p^{\rho}_{d}\tau_p\mathbf{R}^j_{d} + \sigma^2\mathbf{I}_M.$$

Channel estimation at the D2D receiver

The dth D2D receiver, on similar lines, correlates its received signal $\mathbf{y}^d$ with pilot signals $\boldsymbol{\phi}^c_{lk}$ and $\boldsymbol{\phi}^d_{d'}$ to estimate the channels g^d_{lk} and $g^d_{d'}$, respectively. The processed receive signals are respectively given as

$$y^{\rho}_{d,lk} = \sum_{(l',k')\in\mathcal{C}_{l,k}} \tau_p \sqrt{p^{\rho}_{l'k'}} g^d_{l'k'} + w_m \text{ and } y^{\rho}_{d,d'} = \sum_{d''\in\mathcal{D}_d} \tau_p \sqrt{p^{\rho}_{d''}} g^d_{d''} + w_m \tag{3.8}$$

The MMSE channel estimates of channels g^d_{lk} and $g^d_{d'}$ are respectively given as,

$$\hat{g}^d_{lk} = \overline{g}^d_{lk} + \frac{\tau_p \sqrt{p^{\rho}_{lk}} \beta^d_{lk}}{1 + \sum\limits_{(l',k')\in\mathcal{C}_{lk}} p^{\rho}_{l'k'} \tau_p \beta^d_{l'k'}} (y^{\rho}_{d,lk} - \overline{y}^{\rho}_{d,lk}) \text{ and } \hat{g}^d_{d'} = \overline{g}^d_{d'} + \frac{\tau_p \sqrt{p^{\rho}_{d'}} \beta^d_{d'}}{1 + \sum\limits_{d''\in\mathcal{D}_{d'}} p^{\rho}_{d''} \tau_p \beta^d_{d''}} (y^{\rho}_{d,d'} - \overline{y}^{\rho}_{d,d'}). \tag{3.9}$$

Here $\overline{y}^{\rho}_{d,lk} = \sum\limits_{(l',k')\in\mathcal{C}_{lk}} \tau_p \sqrt{p^{\rho}_{lk}} \overline{g}^d_{l'k'}$ and $\overline{y}^{p}_{d,d'} = \sum\limits_{d''\in\mathcal{D}_{d'}} \tau_p \sqrt{p^{p}_{d''}} \overline{g}^d_{d''}$.

3.1.3 Data Transmission

The kth CU in lth cell $U^c_{l,k}$ transmits its signal $\sqrt{p_{lk}} s_{lk}$ to its respective BS and simultaneously the dth D2D transmitter $U^d_{d,t}$ transmits its signal $\sqrt{q_d} s_d$ to its D2D receiver $U^d_{d,r}$. Here $s_{lk} \in \mathbb{C}$ and $s_d \in \mathbb{C}$ denotes the transmit symbol of $U^c_{l,k}$ and $U^d_{d,t}$ respectively, which have zero mean and unit variance. The scalars p_{lk} and q_d represents the data transmit powers of $U^c_{l,k}$ and $U^d_{d,t}$ respectively. The BSs and the D2D receivers, consequently, receive a combination of CU and D2D transmit signals. The receive signal of jth BS is given as

$$\mathbf{y}_j = \sum_{l'=1}^{L} \sum_{k'=1}^{K} \sqrt{p_{l'k'}} \mathbf{h}^j_{l'k'} s_{l'k'} + \sum_{d=1}^{D} \sqrt{q_d} \mathbf{h}^j_d s_d + \mathbf{w}_j. \tag{3.10}$$

Here $\mathbf{w}_j \in \mathbb{C}^{M\times 1}$, with i.i.d. $\mathcal{CN}(0, \sigma^2)$ entries, denote the AWGN at the jth BS. Prior to decoding, the jth BS correlates its receive signal with a combining vector

$\mathbf{v}_{jk} \in \mathbb{C}^{M\times 1}$ designed using the estimated channel state information (CSI). The processed signal after combining is rewritten to show different interference terms as follows:

$$\hat{s}_{jk} = \mathbf{v}_{jk}^H \mathbf{y}_j = \underbrace{\sqrt{p_{jk}}\mathbf{v}_{jk}^H \mathbf{h}_{jk}^j s_{jk}}_{\text{desired signal}} + \underbrace{\sum_{k'\neq k}^{K} \sqrt{p_{jk'}}\mathbf{v}_{jk}^H \mathbf{h}_{jk'}^j s_{jk'}}_{\text{intra-cell interference}} + \underbrace{\sum_{l'\neq j}^{L}\sum_{k'=1}^{K} \sqrt{p_{l'k'}}\mathbf{v}_{jk}^H \mathbf{h}_{l'k'}^j s_{l'k'}}_{\text{inter-cell interference}}$$
$$+ \underbrace{\sum_{d=1}^{D} \sqrt{q_d}\mathbf{v}_{l,k}^H \mathbf{h}_d^j s_d}_{\text{D2D interference}} + \underbrace{\mathbf{v}_{jk}^H \mathbf{w}_j}_{\text{AWGN at BS}} \tag{3.11}$$

Here the intra- and inter-cell inteference is caused by the transmit signals of CUs served by same BS and different BS respectively. The D2D interference is caused by the signal transmitted by the D2D transmitters. The dth D2D receiver, on similar lines, correlates its receive signal with the channel estimate $\hat{g}_d^d$ to obtain

$$\begin{aligned} y_d &= \sum_{d'=1}^{D} \sqrt{q_{d'}}\hat{g}_d^{d^*} g_{d'}^d s_{d'} + \sum_{l'=1}^{L}\sum_{k'=1}^{K} \sqrt{p_{l'k'}}\hat{g}_d^{d^*} g_{l'k'}^d s_{l'k'} + \hat{g}_d^{d^*} w_d \\ &= \underbrace{\sqrt{q_d}\hat{g}_d^{d^*} g_d^d s_d}_{\text{Desired Signal}} + \underbrace{\sum_{d'\neq d}^{D} \sqrt{q_{d'}}\hat{g}_d^{d^*} g_{d'}^d s_{d'}}_{\text{inter-D2D interference}} + \underbrace{\sum_{l'=1}^{L}\sum_{k'=1}^{K} \sqrt{p_{l'k'}}\hat{g}_d^{d^*} g_{l'k'}^d s_{l'k'}}_{\text{CU Interference}} + \underbrace{\hat{g}_d^{d^*} w_d}_{\substack{\text{AWGN at} \\ \text{D2D receiver}}}, \end{aligned} \tag{3.12}$$

Here $w_d \sim \mathcal{CN}(0, \sigma^2)$ is the AWGN at the dth D2D receiver. In (3.12), the desired signal is the signal transmitted by the dth D2D transmitter, inter-D2D interference - caused by the transmit signals of other D2D pairs, CU interference is caused by the transmit signal of CUs.

BS combiners

In this work, we analyze the system performance for three combining schemes namely MRC, interference-aware ZF (Ia-ZF) and interference-aware MMSE (Ia-MMSE). Conventional MMSE and ZF combiners are designed to cancel the inter- and intra-cell interference experienced by the user. These techniques, however, cannot cancel the D2D interference, which is crucial for an D2D underlaid multi-cell system []. The Ia-ZF and Ia-MMSE schemes, by utilizing the estimates of both CU-BS and D2D transmitter-BS channels, mitigates the intra-cell, inter-cell and D2D interference. The Ia-ZF and Ia-MMSE combiners are given as

$$\mathbf{V}^j = \begin{cases} \widehat{\mathbf{H}}_j^c, & \text{for MRC} \\ \widehat{\mathbf{H}}_j\left[\widehat{\mathbf{H}}_j^H\widehat{\mathbf{H}}_j\right]^{-1}, & \text{for Ia-ZF} \\ \left[\sum\limits_{l'=1}^{L}\widehat{\mathbf{H}}_{l'}\overline{\mathbf{P}}_{l'}^{cd}\widehat{\mathbf{H}}_{l'}^H + \sum\limits_{l'=1}^{L}\left(\sum\limits_{k'=1}^{K}p_{l'k'}\mathbf{C}_{l'k'}^j + \sum\limits_{d=1}^{D}q_d\mathbf{C}_d^j\right) + \mathbf{I}_M\right]^{-1}\widehat{\mathbf{H}}_j\overline{\mathbf{P}}_j^{cd}, & \text{for Ia-MMSE.} \end{cases} \tag{3.13}$$

Here $\mathbf{V}^j = [\mathbf{v}_{j1}, \cdots, \mathbf{v}_{jK}] \in \mathbb{C}^{M\times K}$ denotes the set of combiners used by jth BS. The matrices $\widehat{\mathbf{H}}_j^c = [\hat{\mathbf{h}}_{j1}^j, \cdots, \hat{\mathbf{h}}_{jK}^j] \in \mathbb{C}^{M\times K}$ and $\widehat{\mathbf{H}}_j^d = [\hat{\mathbf{h}}_1^j, \cdots, \hat{\mathbf{h}}_D^j] \in \mathbb{C}^{M\times D}$ denote the set of channels from K CUs in jth cell and D D2D transmitters to the jth BS respectively, which we concatenate as $\widehat{\mathbf{H}}_j = [\widehat{\mathbf{H}}_j^c\ ;\ \widehat{\mathbf{H}}_j^d] \in \mathbb{C}^{M\times(K+D)}$. Further $\mathbf{P}_{l'} = \mathrm{diag}(p_{l'1}, \cdots, p_{l'K}) \in \mathbb{R}_+^{K\times K}$, $\mathbf{Q} = \mathrm{diag}(q_1, \cdots, q_D) \in \mathbb{R}_+^{D\times D}$ contain the transmit powers of CUs in jth cell and D D2D transmitters, respectively and $\overline{\mathbf{P}}_j^{cd} = \mathrm{blkdiag}(\mathbf{P}_j,\ \mathbf{Q})$.

3.2 Achievable Spectral Efficiency Analysis

In this section, we first explain the achievable SE of a URLLC enabled mMIMO system. We later analyze the achievable SE of CUs and D2D users for various BS combining schemes. A URLLC system, due to its low-latency requirements, have shorter blocklength and support short packet communications []. Due to finite block lengths, decoding errors are more prominent in a URLLC system. Shannon capacity, defined as $R(\beta) = \log_2(1+\beta)$ with β being signal to noise ratio, greatly underestimates the SE of a URLLC system as it fails to capture these errors. The finite blocklength theory, an extension to Shannon's theorem, captures the effect of errors by approximating the rate expression as

$$R(\beta) = \log_2(1+\beta) - Q^{-1}(\epsilon)\log(2)\sqrt{\frac{V}{LB}} \text{ where } V = \frac{2\beta}{1+\beta}. \tag{3.14}$$

Here V and ϵ denotes the channel dispersion and decoding error probability, respectively []. The scalars B and L denote coherence bandwidth and transmission blocklength, respectively. The function $Q^{-1}(\cdot)$ is the inverse of Gaussian Q-function and ϵ is the decoding error probability.

3.2.1 Spectral Efficiency of Cellular Users

We first give the ergodic sum SE of the CUs and then derive a closed-form SE expression using use-and-then-forget (UaTF) technique. The ergodic sum SE of CUs is obtained by treating the signal received over estimated channel as desired signal in (3.11) as

$$R^c_{\text{sum}} = \left(\frac{\tau_c - \tau_p}{\tau_c}\right)\sum_{l=1}^{L}\sum_{k=1}^{K}\mathbb{E}\left[\log_2(1+\text{SINR}^c_{lk}) - \frac{Q^{-1}(\epsilon)\log_2(e)}{\sqrt{\tau_c-\tau_p}}\left(\frac{2\,\text{SINR}^c_{lk}}{1+\text{SINR}^c_{lk}}\right)^{\frac{1}{2}}\right]. \tag{3.15}$$

Here SINR_{lk} is the signal-to-interference-plus-noise ratio of CU U^c_{lk} obtained using (3.11) as:

$$\text{SINR}^c_{lk} = \frac{p_{jk}\left|\mathbf{v}^H_{jk}\hat{\mathbf{h}}^j_{jk}\right|^2}{\left\{\begin{matrix}\sum\limits_{k'\neq k}^{K} p_{jk'}\left|\mathbf{v}^H_{jk}\hat{\mathbf{h}}^j_{jk'}\right|^2 + \sum\limits_{l'=1}^{L}\sum\limits_{k'\neq j}^{K} p_{l'k'}\left|\mathbf{v}^H_{jk}\hat{\mathbf{h}}^j_{l'k'}\right|^2 + \sum\limits_{d=1}^{D} q_d\left|\mathbf{v}^H_{jk}\hat{\mathbf{h}}^j_{d}\right|^2 \\ +\mathbf{v}^H_{jk}\left(\sum\limits_{l'=1}^{L}\sum\limits_{k'=1}^{K} p_{l'k'}\mathbf{C}^j_{l'k'} + \sum\limits_{d=1}^{D} q_d\mathbf{C}^d_d + \sigma^2\mathbf{I}_M\right)\mathbf{v}_{jk}\end{matrix}\right\}}. \tag{3.16}$$

The expectation in (8.25) is over the channel estimates $\{\hat{\mathbf{h}}^j_{lk}, \hat{\mathbf{h}}^j_d\}$. The ergodic sum SE expression in (8.25) is difficult to analyze due to the expectation outside the logarithm. We use this expression later to perform stochastic optimization. We now derive a tractable lower-bound on the sum SE using UaTF technique. We rewrite the signal received at the jth BS in (3.11) as

$$\hat{s}_{jk} = \underbrace{\sqrt{p_{jk}}\mathbb{E}\left[\mathbf{v}_{jk}^H\mathbf{h}_{jk}^j\right] s_{jk}}_{\substack{\text{signal received over}\\ \text{hardened channel}}} + z_{jk}^{\text{eff}}, \text{ where}$$

$$z_{jk}^{\text{eff}} = \sqrt{p_{jk}}\left[\mathbf{v}_{jk}^H\mathbf{h}_{jk}^j - \mathbb{E}(\mathbf{v}_{jk}^H\mathbf{h}_{jk}^j)\right] s_{jk} + \sum_{k'\neq k}^{K}\sqrt{p_{jk'}}\mathbf{v}_{jk}^H\mathbf{h}_{jk'}^j s_{jk'} + \sum_{l'\neq j}^{L}\sum_{k'=1}^{K}\sqrt{p_{l'k'}}\mathbf{v}_{jk}^H\mathbf{h}_{l'k'}^j s_{l'k'}$$
$$+ \sum_{d=1}^{D}\sqrt{q_d}\mathbf{v}_{jk}^H\mathbf{h}_d^j s_d + \mathbf{v}_{jk}^H\mathbf{w}_j. \tag{3.17}$$

We see that the signal received over hardened channel is uncorrelated with the effective noise z_{jk}^{eff}. Though the exact distribution of the effective noise is unknown, therefore we invoke central limit theorem and treat z_{jk}^{eff} as worst case Gaussian noise. The statistical lower bound on the SE of CU $U_{j,k}^c$ is, therefore, obtained using (3.17) as

$$\overline{R}_{jk}^c = \left(1 - \frac{\tau_p}{\tau_c}\right)\log_2(1 + \overline{\text{SINR}}_{jk}^c) - \frac{Q^{-1}(\epsilon)\log_2(e)}{\sqrt{\tau_c - \tau_p}}\left[\frac{2\,\overline{\text{SINR}}_{jk}^c}{1 + \overline{\text{SINR}}_{jk}^c}\right]^{\frac{1}{2}}, \text{ where} \tag{3.18}$$

$$\overline{\text{SINR}}_{jk}^c = \frac{p_{jk}\left|\Upsilon_{0_{jk}}\right|^2}{\sum\limits_{l'=1}^{L}\sum\limits_{k'=1}^{K} p_{l'k'}\Upsilon_{1_{jk}}^{l'k'} - p_{jk}\left|\Upsilon_{0_{jk}}\right|^2 + \sum\limits_{d=1}^{D} q_d\Upsilon_{2_{jk}}^d + \sigma^2\Upsilon_{3_{jk}}} \triangleq \frac{\Delta_{jk}^c}{\Lambda_{jk}^c}. \tag{3.19}$$

Here $\Upsilon_{0_{jk}} = \mathbb{E}\left\{\mathbf{v}_{jk}^H\mathbf{h}_{jk}^j\right\}$, $\Upsilon_{1_{jk}}^{l'k'} = \mathbb{E}\left\{\left|\mathbf{v}_{jk}^H\mathbf{h}_{l'k'}^j\right|^2\right\}$, $\Upsilon_{2_{jk}}^d = \mathbb{E}\left\{\left|\mathbf{v}_{jk}^H\mathbf{h}_d^j\right|^2\right\}$ and $\Upsilon_{3_{jk}} = \mathbb{E}\left\{\mathbf{v}_{jk}^H\mathbf{v}_{jk}\right\}$. The SE lower bound in (3.19) is valid for any BS combining scheme. Due to presence of spatially correlated rician faded channels and complicated design of Ia-ZF and Ia-MMSE combiners, theoretical computation of these expectations is a non-trivial task. We therefore analyze the SE of Ia-ZF and Ia-MMSE combiners by numerically evaluating the expectations. However, for MRC we can compute these expectations in closed-form. This closed-form SE expression is calculated in the paper [3] and hence the proofs are referenced to it. We now provide the closed-form SE expression for MRC case in the following theorem.

Theorem 1. *The closed-form expression of SINR of CU $U_{j,k}^c$ $\overline{SINR}_{jk}$ with MR* *combining and spatially-correlated Rician-faded channels is given as*

$$\overline{SINR}^{c,MR}_{jk} = \frac{p_{jk}a^j_{jk}}{\sum\limits_{l'=1}^{L}\sum\limits_{k'=1}^{K} p_{l'k'}b^{jk}_{l'k'} + \sum\limits_{(l',k')\in\mathcal{P}_{jk}\setminus(j,k)} p_{l'k'}c^{jk}_{l'k'} - p_{jk}d^j_{jk} + \sum\limits_{d=1}^{D} q_d e^d_{jk} + \sigma^2} \quad \textit{where} \tag{3.20}$$

$$\begin{aligned}
a^j_{jk} &= \tau_p p^\rho_{jk}\,\mathrm{Tr}\left(\mathbf{R}^j_{jk}\boldsymbol{\Psi}^j_{jk}\mathbf{R}^j_{jk}\right) + \left\|\overline{\mathbf{h}}^j_{jk}\right\|^2, \quad d^j_{jk} = \frac{\left\|\overline{\mathbf{h}}^j_{jk}\right\|^4}{a^j_{jk}},\\
b^{jk}_{l'k'} &= \frac{1}{a^j_{jk}}\left[p^\rho_{jk}\tau_p\,\mathrm{Tr}\left(\mathbf{R}^j_{l'k'}\mathbf{R}^j_{jk}\boldsymbol{\Psi}^j_{jk}\mathbf{R}^j_{jk}\right) + \overline{\mathbf{h}}^{j^H}_{jk}\mathbf{R}^j_{l'k'}\overline{\mathbf{h}}^j_{jk} + p^\rho_{jk}\tau_p\overline{\mathbf{h}}^{j^H}_{l'k'}\mathbf{R}^j_{jk}\boldsymbol{\Psi}^j_{jk}\mathbf{R}^j_{jk}\overline{\mathbf{h}}^j_{l'k'} + \left|\overline{\mathbf{h}}^{j^H}_{jk}\overline{\mathbf{h}}^j_{l'k'}\right|^2\right]\\
c^{jk}_{l'k'} &= \frac{1}{a^j_{jk}}\left[p^\rho_{jk}p^\rho_{l,i}\tau_p^2|\,\mathrm{Tr}\left(\mathbf{R}^j_{jk}\boldsymbol{\Psi}^j_{jk}\mathbf{R}^j_{jk}\right)|^2 + 2\sqrt{p^\rho_{jk}p^\rho_{l'k'}}\tau_p Re\Big\{\mathrm{Tr}\left(\mathbf{R}^j_{l'k'}\boldsymbol{\Psi}^j_{jk}\mathbf{R}^j_{jk}\right)\overline{\mathbf{h}}^{j^H}_{l'k'}\overline{\mathbf{h}}^j_{jk}\Big\}\right] \quad \textit{and}\\
e^d_{jk} &= \frac{1}{a^j_{jk}}\left[p^\rho_{jk}\tau_p\,\mathrm{Tr}\left(\mathbf{R}^j_d\mathbf{R}^j_{jk}\boldsymbol{\Psi}^j_{jk}\mathbf{R}^j_{jk}\right) + \overline{\mathbf{h}}^{j^H}_{jk}\mathbf{R}^j_d\overline{\mathbf{h}}^j_{jk} + p^\rho_{jk}\tau_p\overline{\mathbf{h}}^{j^H}_d\mathbf{R}^j_{jk}\boldsymbol{\Psi}^j_{jk}\mathbf{R}^j_{jk}\overline{\mathbf{h}}^j_d + \left|\overline{\mathbf{h}}^{j^H}_{jk}\overline{\mathbf{h}}^j_d\right|^2\right].
\end{aligned} \tag{3.21}$$

Proof. Please refer to Appendix 4.1 of [3]. □

3.2.2 Spectral Efficiency of D2D Users

To detect the desired data from dth transmitter, the D2D receiver $U^d_{d,t}$ treats the channels it estimates i.e, $\hat{g}^d_{d'}$, $\hat{g}^d_{jk}$ as side information. Using the side-information bound in [], a lower bound on the sum SE of D2D users can be obtained using (3.12) as

$$\begin{aligned}
R^d_{\mathrm{sum}} &= \sum_{d=1}^{D}\left(\frac{\tau_c-\tau_p}{\tau_c}\right)\mathbb{E}\left[\log_2(1+\mathrm{SINR}^d_d) - \frac{Q^{-1}(\epsilon)\log_2(e)}{\sqrt{\tau_c-\tau_p}}\left(\frac{2\,\mathrm{SINR}^d_d}{1+\mathrm{SINR}^d_d}\right)^{\frac{1}{2}}\right] \quad \text{with}\\
\mathrm{SINR}^d_d &= \frac{q_d\left|\mathbb{E}\left[\hat{g}^{d^*}_d g^d_d\middle|\mathcal{G}^d\right]\right|^2}{q_d\mathbb{V}\mathrm{ar}\left\{\hat{g}^{d^*}_d g^d_d\middle|\mathcal{G}^d\right\} + \mathbb{V}\mathrm{ar}\left\{\sum\limits_{l=1}^{L}\sum\limits_{k=1}^{K}\sqrt{p_{lk}}\hat{g}^{d^*}_d g^d_{l,k}s_{lk} + \sum\limits_{d'\neq d}^{D}\sqrt{q_{d'}}\hat{g}^{d^*}_d g^d_{d'}s_{d'} + \hat{g}^{d^*}_d w_d\middle|\mathcal{G}^d\right\}}.
\end{aligned} \tag{3.22}$$

Here the set $\mathcal{G}^d = \left\{\hat{g}^d_{d'}, \hat{g}^d_{l,k}\right\}$ for $l = 1, \cdots, K$, $k = 1, \cdots, K$ and $d' = 1, \cdots, D$, denotes the set of side-information available at dth D2D receiver. We note from (3.22) that the lower-bound consists of conditional expectations, which are difficult to calculate numerically. We similar to [13], use the approximation [, lemma ??] to obtain a closed-form SE expression as shown in the following theorem.

Theorem 2. *The closed-form SE expression of the dth D2D receiver is given as*

$$\overline{R}^d_d = \left(\frac{\tau_c - \tau_p}{\tau_c}\right)\left[\log_2(1 + SINR^d_d) - \frac{Q^{-1}(\epsilon)\log_2(e)}{\sqrt{\tau_c - \tau_p}}\left(\frac{2\,\overline{SINR}^d_d}{1 + \overline{SINR}^d_d}\right)^{\frac{1}{2}}\right]. \quad (3.23)$$

Here $\overline{SINR}^d_d = \dfrac{q_d a^d_d}{q_d b^d_d + \sum\limits_{l=1}^{L}\sum\limits_{k=1}^{K} p_{lk} c^d_{lk} + \sum\limits_{d' \neq d}^{D} q_{d'} e^d_{d'} + \sigma^2} \triangleq \dfrac{\Delta^d_d}{\Lambda^d_d}$, $a^d_d = \left(\delta^d_d + |\bar{g}^d_d|^2\right)$, $b^d_d = (\beta^d_d - \delta^d_d)$, $c^d_{lk} = \left(|\bar{g}^d_{lk}|^2 + \beta^d_{lk}\right)$, $e^d_{d'} = \left(|\bar{g}^d_{d'}| + \beta^d_{d'}\right)$ *and* $\delta^d_d = \frac{q^\rho_d \tau^2_p (\beta^d_d)^2}{1 + \sum\limits_{d' \in \mathcal{D}_d} q^\rho_{d'} \tau_p \beta^d_{d'}}$.

Proof. Please refer to Appendix 4.3 of [3] $\square$

Chapter 4

D2D underlaid multi-cell mMIMO URLLC system : GEE maximization

4.1 GEE Formulation

In this section, we optimize the global energy efficiency (GEE) of the considered system. *Previous multi-cell mMIMO works with/without D2D users did not consider URLLC.* The current work fills this gap by optimizing the network centric GEE metric of a URLLC enabled D2D underlaid mMIMO system. The GEE of the system is defined using the SE expressions in (3.18) and (3.23) is defined as

$$f_{GEE} = B \frac{\sum_{d'=1}^{D} \overline{R}_{d'}^{d} + \sum_{l'=1}^{L} \sum_{k'=1}^{K} \overline{R}_{l'k'}^{c}}{\left(1 - \frac{\tau_p}{\tau_c}\right) \left[\sum_{d=1}^{D} \mu_d q_d + \sum_{l'=1}^{L} \sum_{k'=1}^{K} \mu_{l'k'} p_{l'k'}\right] + P_\rho + P_{net}}. \tag{4.1}$$

Here μ_d and $\mu_{l'k'}$ denotes the power amplifier inefficiency of D2D transmitter $U_{d',t}^{d}$ and CU $U_{l',k'}^{c}$, respectively. The scalar $P_\rho = \frac{\tau_p}{\tau_c} \left[\sum_{d=1}^{D} \mu_d q_d + \sum_{l'=1}^{L} \sum_{k'=1}^{K} \mu_{l'k'} p_{l'k'}\right]$ is sum of transmit pilot powers of CUs and D2D users and P_{net} is the circuit power consumption of the entire network. We now briefly discuss the circuit power consumption model.

4.1.1 Power Consumption Model

: We adopt a realistic circuit power consumption model that captures the power consumed by transceiver chains, signal processing and MMSE channel estimation. The term P_{net} is therefore modelled as [29]

$$P_{net} = \sum_{l=1}^{L}[P_{\mathrm{FIX},l} + P_{\mathrm{TC},l} + P_{\mathrm{CE},l} + P_{\mathrm{SP},l}] + \sum_{d=1}^{D}[P_{\mathrm{TC},d} + P_{\mathrm{CE},d} + P_{\mathrm{SP},d}] \tag{4.2}$$

Here $P_{\mathrm{FIX},l}$ are the fixed circuit power for lth cell which is required for site-cooling and load independent back-haul. $P_{\mathrm{TC},l}$ and $P_{\mathrm{TC},d}$ accounts for power consumed by the RF transmit/receive chains of CUs and BS in lth cell and dth D2D transmitter-receiver pair, respectively. Here

$$P_{\mathrm{TC},l} = MP_{\mathrm{BS},j} + P_{\mathrm{LO},j} + KP_{\mathrm{CU},j}, \quad P_{\mathrm{TC},d} = 2P_{\mathrm{D2D}},$$

where $P_{\mathrm{BS},j}$, $P_{\mathrm{CU},j}$ and $P_{\mathrm{D2D},j}$ is the per-antenna power consumed by the jth BS, the CUs in jth cell and the dth D2D user-pair, respectively. The term $P_{\mathrm{LO},j}$ is the power required for the local oscillator of jth BS. The terms $P_{\mathrm{CE},l}$ (resp. $P_{\mathrm{CE},d}$) in (4.2) is the circuit power consumed to perform MMSE channel estimation at jth BS (resp. dth D2D receiver), and is modelled as

$$P_{\mathrm{CE},l} = \frac{3B}{\tau_d L_{\mathrm{BS}}}(K+D)\left(M\tau_p + M^2\right) \text{ and } P_{\mathrm{CE},d} = \frac{3B}{\tau_d L_{\mathrm{D}}}(K+D)\left(\tau_p + 1\right).$$

Here L_{BS} and L_{D}, in flops/W, is the computational efficiency of BS and D2D receiver respectively. The terms $P_{\mathrm{SP},d}$ and $P_{\mathrm{SP},l}$ are the power consumed for the signal processing at dth D2D receiver and lth BS, respectively and are given as

$$P_{\mathrm{SP},d} = \frac{3B}{\tau_d L_{\mathrm{D}}} D(\tau_c - \tau_p) \text{ and } P_{\mathrm{SP},l} = \frac{3B}{\tau_d L_{\mathrm{BS}}} KM(\tau_c - \tau_p) + \Delta. \tag{4.3}$$

Here Δ is the power consumed in computing the BS combining vectors, which is different for different combining schemes. The values of Δ for MR, Ia-ZF and Ia-MMSE combiners, are chosen from Table ?? of [].

4.2 Deterministic GEE Optimization:

The GEE optimization problem of the URLLC system, by using (4.1) and ignoring the constant $B(1-\frac{\tau_p}{\tau_c})$, is formulated as

$$\mathbf{P1}: \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}} \; \frac{\sum\limits_{(l',k')\in\mathcal{C}} \log\left(1+\overline{\text{SINR}}^{c}_{l'k'}\right) - \alpha\sqrt{\frac{\overline{\text{SINR}}^{c}_{l'k'}}{1+\overline{\text{SINR}}^{c}_{l'k'}}} + \sum\limits_{d'=1}^{D} \log\left(1+\overline{\text{SINR}}^{d}_{d'}\right) - \alpha\sqrt{\frac{\overline{\text{SINR}}^{d}_{d'}}{1+\overline{\text{SINR}}^{d}_{d'}}}}{\left(1-\frac{\tau_p}{\tau_c}\right)\left[\sum\limits_{(l',k')\in\mathcal{C}} \mu_{l'k'}p_{l'k'} + \sum\limits_{d'=1}^{D} \mu_{d'}q_{d'}\right] + P_\rho + P_{net}}$$

$$\text{s.t.} \qquad 0 \le p_{lk} \le P_{max} \text{ and } 0 \le q_d \le q_{\max}.$$

Here the set $\mathcal{C} = \{(l',k') : l' = 1,\cdots,L \text{ and } k' = 1,\cdots,K\}$ denotes the set of CUs and the constant $\alpha = \frac{Q^{-1}(\epsilon)\sqrt{2}}{\sqrt{\tau_c-\tau_p}}$. The CU transmit power matrix $\mathbf{P} = \text{blkdiag}(\mathbf{P}_1,\cdots,\mathbf{P}_L) \in \mathbb{R}^{LK\times LK}$. We make the following observations from the structure of problem **P1**:

- R1: The functions $\log(x)$ and $\sqrt{x}$ are both concave and non-decreasing for $x \in \mathbb{R}_{++}$. The numerator of GEE objective due to URLLC, therefore, consists of difference of concave functions in terms of $\overline{\text{SINR}}^{c}_{l'k'}$ and $\overline{\text{SINR}}^{d}_{d'}$.
- R2: From (3.18) and (3.23), we see that each SINR term is a fractional function of transmit powers $\{p_{jk}, q_d\}$.

Problem **P1**, due to aforementioned reasons, is a challenging non-convex problem with difference of concave functions and nested fractional functions of optimization variables. To the best of our knowledge, this is the first work to optimize the GEE of a URLLC enabled system with coexisting CUs and D2D users. We now leverage the fractional programming (FP) tools: LDT and QT, and develop an iterative closed-form solution for the non-convex GEE problem. The proposed approach first transforms the non-convex problem into a linear FP problem and then use LDT and QT to decouple various fractional terms. The resultant problem is then solved by updating the transmit powers in closed-form. We now introduce epigraph variable and recast the problem **P1** as

$$\mathbf{P2}: \underset{\mathbf{P},\mathbf{Q},t}{\text{Maximize}} \ \frac{t}{\mathcal{A}\Big[\sum\limits_{(l',k')\in\mathcal{C}} \mu_{l'k'} p_{l'k'} + \sum\limits_{d'=1}^{D} \mu_{d'} q_{d'}\Big] + \widetilde{P}_{net}} \tag{4.5a}$$

$$\text{s.t.} \quad \sum_{(l',k')\in\mathcal{C}} \log\big(1+\overline{\text{SINR}}^{c}_{l'k'}\big) - \alpha\sqrt{V^{c}_{l'k'}} + \sum_{d'=1}^{D} \log\Big(1+\overline{\text{SINR}}^{d}_{d'}\Big) - \alpha\sqrt{V^{d}_{d'}} \geq t \tag{4.5b}$$

$$0 \leq p_{lk} \leq P_{max} \text{ and } 0 \leq q_d \leq q_{\max}. \tag{4.5c}$$

Here $V^{c}_{l'k'} = \frac{\overline{\text{SINR}}^{c}_{l'k'}}{1+\overline{\text{SINR}}^{c}_{l'k'}}$, $V^{d}_{d'} = \frac{\overline{\text{SINR}}^{d}_{d'}}{1+\overline{\text{SINR}}^{d}_{d'}}$. The constants $\mathcal{A} = \Big(1 - \frac{\tau_p}{\tau_c}\Big)$ and $\widetilde{P}_{net} = P_\rho + P_{net}$. We note that the objective of **P2** consists of single ratio with both numerator and denominator being affine functions in $(\mathbf{P}, \mathbf{Q}, t)$. We now use QT [**?**, Corollary 2] to decouple the ratio in the objective function. The problem **P2** is recast as

$$\mathbf{P3}: \underset{\mathbf{P},\mathbf{Q},t,z}{\text{Maximize}} \ 2z\sqrt{t} - z^2 \left\{ \mathcal{A}\left[\sum_{(l',k')\in\mathcal{C}} \mu_{l'k'} p_{l'k'} + \sum_{d'=1}^{D} \mu_{d'} q_{d'}\right] + \widetilde{P}_{net} \right\} \tag{4.6a}$$

$$\text{s.t.} \quad \sum_{(l',k')\in\mathcal{C}} \log\big(1+\overline{\text{SINR}}^{c}_{l'k'}\big) - \alpha\sqrt{V^{c}_{l'k'}} + \sum_{d'=1}^{D} \log\Big(1+\overline{\text{SINR}}^{d}_{d'}\Big) - \alpha\sqrt{V^{d}_{d'}} \geq t \tag{4.6b}$$

$$0 \leq p_{lk} \leq P_{max} \text{ and } 0 \leq q_d \leq q_{\max}. \tag{4.6c}$$

The auxilary variable z decouples the ratio in the objective function. Problem **P3**, for a given $(\mathbf{P}, \mathbf{Q}, t)$ is concave in z, and its optimal value is computed as

$$z^* = \frac{\sqrt{t}}{\mathcal{A}\Big[\sum_{(l',k')\in\mathcal{C}} \mu_{l'k'} p_{l'k'} + \sum_{d'=1}^{D} \mu_{d'} q_{d'}\Big] + \widetilde{P}_{net}}. \tag{4.7}$$

Problem **P3**, for a fixed $z = z^*$, can be rewritten as outer optimization over $(\mathbf{P}, \mathbf{Q})$ and inner maximization over t for a fixed $(\mathbf{P}, \mathbf{Q})$ i.e.,

$$\mathbf{P3} : \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}} \quad \underset{t}{\text{Maximize}} \quad f_3(\mathbf{P}, \mathbf{Q}, , t, z^*) \tag{4.8a}$$

$$\text{s.t.} \;\; f_{3,1}(\mathbf{P}, \mathbf{Q}) \geq t, \tag{4.8b}$$

$$\text{s.t.} \qquad 0 \leq p_{lk} \leq P_{max} \text{ and } 0 \leq q_d \leq q_{\max}. \tag{4.8c}$$

Here $f_3(\mathbf{P}, \mathbf{Q}, z^*, t) = 2z^*\sqrt{t} - (z^*)^2\Big\{\mathcal{A}\Big[\sum_{(l',k')\in\mathcal{C}} \mu_{l'k'} p_{l'k'} + \sum_{d'=1}^{D} \mu_{d'} q_{d'}\Big] + \widetilde{P}_{net}\Big\}$ and $f_{3,1}(\mathbf{P}, \mathbf{Q}) = \sum_{(l',k')\in\mathcal{C}} \log\big(1 + \overline{\text{SINR}}^c_{l'k'}\big) - \alpha\sqrt{V^c_{l'k'}} + \sum_{d'=1}^{D} \log\Big(1 + \overline{\text{SINR}}^d_{d'}\Big) - \alpha\sqrt{V^d_{d'}}$. The lagrangian function for the inner maximization problem in **P3** is given as

$$\begin{aligned} &L_3(t, \lambda, \mathbf{P}, \mathbf{Q}, z^*) \\ &= 2z^*\sqrt{t} - (z^*)^2\Big\{\mathcal{A}\Big[\sum_{(l',k')\in\mathcal{C}} \mu_{l'k'} p_{l'k'} + \sum_{d'=1}^{D} \mu_{d'} q_{d'}\Big] + \widetilde{P}_{net}\Big\} - \lambda\big[t - f_{3,1}(\mathbf{P}, \mathbf{Q})\big]. \end{aligned} \tag{4.9}$$

Here λ is the lagrangian dual variable associated to constraint (4.8b). The inner maximization problem in **P3** can be rewritten using the lagrangian function as

$$\mathbf{P3}_{\text{inner}} : \underset{t}{\text{Maximize}} \;\; \underset{\lambda\geq 0}{\text{Minimize}} \; L_3(t, \lambda, \mathbf{P}, \mathbf{Q}, z^*). \tag{4.10}$$

We see from (4.9) that $\mathbf{P3}_{\text{inner}}$ is convex (resp. concave) in λ (resp. t). Due to strong duality, the constraint (4.8b) holds i.e., $t^* = f_{3,1}(\mathbf{P}, \mathbf{Q})$ and the optimal values of λ

and t are calculated as

$$\frac{\partial L_3(t,\lambda,\mathbf{P}_j,\mathbf{Q},z^*)}{\partial t}\Big|_{\lambda=\lambda^*,t=t^*}=0 \implies \frac{z^*}{\sqrt{t^*}}-\lambda^*=0 \implies \lambda^*=\frac{z^*}{\sqrt{t^*}} \text{ and}$$

$$t^*=\sum_{(l',k')\in\mathcal{C}}\log\big(1+\overline{\mathrm{SINR}}^c_{l'k'}\big)-\alpha\sqrt{\frac{\overline{\mathrm{SINR}}^c_{l'k'}}{1+\overline{\mathrm{SINR}}^c_{l'k'}}}+\sum_{d'=1}^{D}\log\Big(1+\overline{\mathrm{SINR}}^d_{d'}\Big)-\alpha\sqrt{\frac{\overline{\mathrm{SINR}}^d_{d'}}{1+\overline{\mathrm{SINR}}^d_{d'}}}. \tag{4.11}$$

The optimal value of the inner maximization problem is therefore

$$L_3(t^*,\lambda^*,\mathbf{P},\mathbf{Q},z^*)=z^*\sqrt{t^*}-(z^*)^2\Big\{\mathcal{A}\Big[\sum_{(l',k')\in\mathcal{C}}\mu_{l'k'}p_{l'k'}+\sum_{d'=1}^{D}\mu_{d'}q_{d'}\Big]+\widetilde{P}_{net}\Big\}-\frac{z^*}{\sqrt{t^*}}f_{3,1}(\mathbf{P},\mathbf{Q}). \tag{4.12}$$

The outer maximization problem in **P3** is therefore recast using (4.12) as

$$\mathbf{P4}:\underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}}\ \sum_{(l',k')\in\mathcal{C}}\log\left(1+\frac{\Delta^c_{l'k'}}{\Lambda^c_{l'k'}}\right)-\alpha\sqrt{\frac{\Delta^c_{l'k'}}{\Delta^c_{l'k'}+\Lambda^c_{l'k'}}}+\sum_{d'=1}^{D}\log\left(1+\frac{\Delta^d_{d'}}{\Lambda^d_{d'}}\right)-\alpha\sqrt{\frac{\Delta^d_{d'}}{\Delta^d_{d'}+\Lambda^d_{d'}}}$$

$$-z^*\sqrt{t^*}\Big\{\mathcal{A}\Big[\sum_{(l',k')\in\mathcal{C}}\mu_{l'k'}p_{l'k'}+\sum_{d'=1}^{D}\mu_{d'}q_{d'}\Big]+\widetilde{P}_{net}\Big\}+t^* \tag{4.13a}$$

$$\text{s.t.}\qquad 0\le p_{lk}\le P_{max} \text{ and } 0\le q_d\le q_{\max}. \tag{4.13b}$$

Here the objective function is obtained by i) expanding the function $f_{3,1}(\mathbf{P},\mathbf{Q})$ in (4.12) by substituting $\overline{\mathrm{SINR}}^c_{l'k'}=\frac{\Delta^c_{l'k'}}{\Lambda^c_{l'k'}}$, $\overline{\mathrm{SINR}}^d_{l'k'}=\frac{\Delta^d_{d'}}{\Lambda^d_{d'}}$, $\mathbf{V}^c_{l'k'}=\frac{\Delta^c_{l'k'}}{\Delta^c_{l'k'}+\Lambda^c_{l'k'}}$, $\mathbf{V}^d_{d'}=\frac{\Delta^d_{d'}}{\Delta^d_{d'}+\Lambda^d_{d'}}$ and; ii) scaling down the objective by constant $\frac{\sqrt{t^*}}{z^*}$. Problem **P4** is non-convex in $(\mathbf{P},\mathbf{Q})$ due to its difference-of-concave-functions form (i.e., difference of $\log_2(\cdot)$ and $\sqrt{\cdot}$ functions in (4.13a)), and the presence of fractional terms $\frac{\Delta^\xi_z}{\Lambda^\xi_z}$ and $\frac{\Delta^\xi_z}{\Delta^\xi_z+\Lambda^\xi_z}$ with $(z,\xi)=(l'k',c)$ for CUs; and $(z,\xi)=(d',d)$ for D2D users. To decouple the fractions and to eliminate the difference-of-concave-functions form, we do the following changes

- rewrite the term $\sqrt{\frac{\Delta^\xi_z}{\Delta^\xi_z+\Lambda^\xi_z}}$ as $\left[\frac{\Delta^\xi_z+\Lambda^\xi_z}{\Delta^\xi_z}\right]^{\frac{-1}{2}}$ $\forall(z,\xi)$, which effectively converts the difference-of-concave-functions to concave - convex form which is concave using composition principles [].

- introduce the epigraph variables to move the fractional functions out of $\log_2(\cdot)$ and $[\cdot]^{\frac{-1}{2}}$ functions.

Problem **P4**, by doing the aforementioned changes is recast as

$$\begin{aligned}
\mathbf{P5}: \underset{\substack{\mathbf{P},\mathbf{Q}\\ \boldsymbol{\nu}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^c,\boldsymbol{\theta}^d}}{\text{Maximize}} \quad & \sum_{(l',k')\in\mathcal{C}} \log\left(1+\nu^c_{l'k'}\right)-\alpha(\theta^c_{l'k'})^{\frac{-1}{2}}+\sum_{d'=1}^{D}\log\left(1+\nu^d_{d'}\right)-\alpha(\theta^d_{d'})^{\frac{-1}{2}} \\
& - z^*\sqrt{t^*}\left\{\mathcal{A}\left[\sum_{(l',k')\in\mathcal{C}}\mu_{l'k'}p_{l'k'}+\sum_{d'=1}^{D}\mu_{d'}q_{d'}\right]+\widetilde{P}_{net}\right\}+t^* \quad (4.14a)\\
\text{s.t.} \quad & \nu^c_{l'k'} \le \frac{\Delta^c_{l'k'}}{\Lambda^c_{l'k'}},\ \frac{1}{\theta^c_{l'k'}} \le \frac{\Delta^c_{l'k'}}{\Delta^c_{l'k'}+\Lambda^c_{l'k'}},\ \nu^d_{d'} \le \frac{\Delta^d_{d'}}{\Lambda^d_{d'}},\ \frac{1}{\theta^d_{d'}} \le \frac{\Delta^d_{d'}}{\Delta^d_{d'}+\Lambda^d_{d'}}, \quad (4.14b)\\
& 0 \le p_{lk} \le P_{max} \text{ and } 0 \le q_d \le q_{\max}. \quad (4.14c)
\end{aligned}$$

Here $\boldsymbol{\nu}^c=[\nu_{1,1},\cdots,\nu_{LK}]$, $\boldsymbol{\theta}^c=[\theta_{1,1},\cdots,\theta_{LK}]$, $\boldsymbol{\nu}^d=[\nu_1,\cdots,\nu_D]$ and $\boldsymbol{\theta}^d=[\theta_1,\cdots,\theta_D]$ are the epigraph variables. For fixed $\mathbf{P},\mathbf{Q}$, we again formulate the lagrangian function of the inner problem of **P5** over variable set $(\boldsymbol{\nu}^c,\boldsymbol{\theta}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^d)$ using (4.23a) and (4.14b) is given as

$$\begin{aligned}
& L_5(\boldsymbol{\nu}^c,\boldsymbol{\theta}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^d,\boldsymbol{\lambda}^c_\nu,\boldsymbol{\lambda}^c_\theta,\boldsymbol{\lambda}^d_\nu,\boldsymbol{\lambda}^d_\theta,\mathbf{P},\mathbf{Q})\\
&= f_5(\boldsymbol{\nu}^c,\boldsymbol{\theta}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^d,\mathbf{P},\mathbf{Q},t^*,z^*) - \sum_{(l'k')\in\mathcal{C}}\left\{\lambda^c_{\nu,l'k'}\left[\nu^c_{l'k'}-\frac{\Delta^c_{l'k'}}{\Lambda^c_{l'k'}}\right]+\lambda^c_{\theta,l'k'}\left[\frac{1}{\theta^c_{l'k'}}-\frac{\Delta^c_{l'k'}}{\Delta^c_{l'k'}+\Lambda^c_{l'k'}}\right]\right\}\\
&\quad -\sum_{d'=1}^{D}\left\{\lambda^d_{\nu,d'}\left[\nu^d_{d'}-\frac{\Delta^d_{d'}}{\Lambda^d_{d'}}\right]+\lambda^d_{\theta,d'}\left[\frac{1}{\theta^d_{d'}}-\frac{\Delta^d_{d'}}{\Delta^d_{d'}+\Lambda^d_{d'}}\right]\right\}. \quad (4.15)
\end{aligned}$$

Here $f_5(\boldsymbol{\nu}^c,\boldsymbol{\theta}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^d,\mathbf{P},\mathbf{Q},t^*,z^\blacktriangle)$ is the objective function of problem **P5**. The vectors $\boldsymbol{\lambda}^c_\nu=[\lambda^c_{\nu,11},\cdots,\lambda^c_{\nu,LK}]$, $\boldsymbol{\lambda}^c_\theta=[\lambda^c_{\theta,11},\cdots,\lambda^c_{\theta,LK}]$, $\boldsymbol{\lambda}^d_\nu=[\lambda^d_{\nu,1},\cdots,\lambda^d_{\nu,D}]$ and $\boldsymbol{\lambda}^d_\theta=[\lambda^d_{\theta,1},\cdots,\lambda^d_{\theta,D}]$ are the lagrangian dual variables corresponding to the constraints in (4.14b). Problem **P5** for a fixed $(\mathbf{P},\mathbf{Q})$, can be recast using the Lagrangian function in (4.15) as

$$\mathbf{P6}: \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}} \quad \underset{\boldsymbol{\nu}^c,\boldsymbol{\theta}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^d}{\text{Maximize}} \quad \underset{\boldsymbol{\lambda}_\nu^c,\boldsymbol{\lambda}_\theta^c,\boldsymbol{\lambda}_\nu^d,\boldsymbol{\lambda}_\theta^d \succeq \mathbf{0}}{\text{Minimize}} \quad L_5(\boldsymbol{\nu}^c,\boldsymbol{\theta}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^d,\boldsymbol{\lambda}_\nu^c,\boldsymbol{\lambda}_\theta^c,\boldsymbol{\lambda}_\nu^d,\boldsymbol{\lambda}_\theta^d,\mathbf{P},\mathbf{Q})$$
$$\text{s.t.} \qquad 0 \le p_{l'k'} \le P_{\max} \text{ and } 0 \le q_d \le q_{\max}. \tag{4.16}$$

We see that problem **P6**, for a fixed $(\mathbf{P},\mathbf{Q})$, is individually concave (resp. convex) in each of the epigraph variables $(\boldsymbol{\nu}^c,\boldsymbol{\theta}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^d)$ (resp. Lagrangian parameters $(\boldsymbol{\lambda}_\nu^c,\boldsymbol{\lambda}_\theta^c,\boldsymbol{\lambda}_\nu^d,\boldsymbol{\lambda}_\theta^d)$). Using KKT conditions, the optimal values of lagrangian parameters, for fixed $(\mathbf{P},\mathbf{Q},\boldsymbol{\nu}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^c,\boldsymbol{\theta}^d)$, are obtained as

$$\lambda_{\nu,l'k'}^{c,*} = \frac{1}{(1+\nu_{l'k'}^c)}, \quad \lambda_{\nu,d'}^{d,*} = \frac{1}{(1+\nu_{d'}^d)}, \quad \lambda_{\theta,l'k'}^{c,*} = \frac{\alpha\sqrt{\theta_{l'k'}^c}}{2} \quad \text{and} \quad \lambda_{\theta,d'}^{d,*} = \frac{\alpha\sqrt{\theta_{d'}^{d,*}}}{2}. \tag{4.17}$$

The lagrangian function $L_5(\cdot)$ by substituting the optimal values of $\boldsymbol{\lambda}_\nu^c$,$\boldsymbol{\lambda}_\nu^d$, $\boldsymbol{\lambda}_\theta^c$ and $\boldsymbol{\lambda}_\theta^d$ is given as

$$\begin{aligned}
&L_5(\boldsymbol{\nu}^c,\boldsymbol{\theta}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^d,\boldsymbol{\lambda}_\nu^{c,*},\boldsymbol{\lambda}_\theta^{c,*},\boldsymbol{\lambda}_\nu^{d,*},\boldsymbol{\lambda}_\theta^{d,*},\mathbf{P},\mathbf{Q}) \\
&= f_5(\boldsymbol{\nu}^c,\boldsymbol{\theta}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^d,\mathbf{P},\mathbf{Q},t^*,z^*) - \sum_{(l'k')\in\mathcal{C}} \left\{ \lambda_{\nu,l'k'}^{c,*}\left[\nu_{l'k'}^c - \frac{\Delta_{l'k'}^c}{\Lambda_{l'k'}^{c,*}}\right] + \lambda_{\theta,l'k'}^{c,*}\left[\frac{1}{\theta_{l'k'}^c} - \frac{\Delta_{l'k'}^c}{\Delta_{l'k'}^c+\Lambda_{l'k'}^c}\right]\right\} \\
&\qquad - \sum_{d'=1}^{D}\left\{ \lambda_{\nu,d'}^{d,*}\left[\nu_{d'}^d - \frac{\Delta_{d'}^d}{\Lambda_{d'}^d}\right] + \lambda_{\theta,d'}^{d,*}\left[\frac{1}{\theta_{d'}^d} - \frac{\Delta_{d'}^d}{\Delta_{d'}^d+\Lambda_{d'}^d}\right]\right\}.
\end{aligned} \tag{4.18}$$

The inner maximization problem in **P6** eventually reduces to optimization over the epigraph variables $(\boldsymbol{\nu}^c,\boldsymbol{\nu}^d,\boldsymbol{\theta}^c,\boldsymbol{\theta}^d)$, whose optimal values are given as

$$\nu_{l'k'}^{c,*} = \frac{\Delta_{l'k'}^c}{\Lambda_{l'k'}^c}, \quad \nu_{d'}^{d,*} = \frac{\Delta_{d'}^d}{\Lambda_{d'}^d}, \quad \theta_{l'k'}^{c,*} = \frac{\Delta_{l'k'}^c}{\Delta_{l'k'}^c+\Lambda_{l'k'}^c} \quad \text{and} \quad \theta_{d'}^{d,*} = \frac{\Delta_{d'}^d}{\Delta_{d'}^d+\Lambda_{d'}^d}. \tag{4.19}$$

We now substitute the optimal values of epigraph variables in the lagrangian function. The resultant problem is a function of optimization variables $(\mathbf{P},\mathbf{Q})$ given as

$$\mathbf{P7}: \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}}\ f_7(\mathbf{P},\mathbf{Q}) \text{ s.t. } 0 \leq p_{lk} \leq P_{max} \text{ and } 0 \leq q_d \leq q_{\max}. \tag{4.20}$$

Here

$$\begin{aligned} f_7(\mathbf{P},\mathbf{Q}) = & \sum_{(l',k')\in\mathcal{C}} \Bigg[\log\left(1+\nu^{c,*}_{l'k'}\right) - \nu^{c,*}_{l'k'} + \frac{(1+\nu^{c,*}_{l'k'})\Delta^c_{l'k'}}{(\Delta^c_{l'k'}+\Lambda^c_{l'k'})} - \frac{\alpha}{2}(\theta^{c,*}_{l'k'})^{-\frac{1}{2}} - \frac{\alpha\sqrt{\theta^{c,*}_{l'k'}}\Delta^c_{l'k'}}{2(\Delta^c_{l'k'}+\Lambda^c_{l'k'})} \Bigg] \\ & + \sum_{d'=1}^{D} \Bigg[\log\left(1+\nu^{d,*}_{d'}\right) - \nu^{d,*}_{d'} + \frac{(1+\nu^{d,*}_{d'})\Delta^d_{d'}}{\Delta^d_{d'}+\Lambda^d_{d'}} - \frac{\alpha}{2}(\theta^{d,*}_{d'})^{-\frac{1}{2}} - \frac{\alpha\sqrt{\theta^{d,*}_{d'}}\Delta^d_{d'}}{\Delta^d_{d'}+\Lambda^d_{d'}} \Bigg] \\ & - z^*\sqrt{t^*}\Bigg\{ \mathcal{A}\Bigg[\sum_{(l',k')\in\mathcal{C}} \mu_{l'k'}p_{l'k'} + \sum_{d'=1}^{D}\mu_{d'}q_{d'} \Bigg] + \widetilde{P}_{net} \Bigg\} + t^*. \end{aligned} \tag{4.21}$$

The resultant problem using the above function $f_7(\mathbf{P},\mathbf{Q})$ is recast as

$$\begin{aligned} \mathbf{P8}: \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}} & \sum_{(l'k')\in\mathcal{C}} \frac{C_1(\nu^{c,*}_{l'k'},\theta^{c,*}_{l'k'})\Delta^c_{l'k'}}{(\Delta^c_{l'k'}+\Lambda^c_{l'k'})} + \sum_{d'=1}^{D}\frac{C_1(\nu^{d,*}_{d'},\theta^{d,*}_{d'})\Delta^d_{d'}}{(\Delta^d_{d'}+\Lambda^d_{d'})} + C_2(\boldsymbol{\nu}^{c,*},\boldsymbol{\nu}^{d,*},\boldsymbol{\theta}^{c,*},\boldsymbol{\theta}^{d,*},t^*) \\ & - z^*\sqrt{t^*}\Bigg\{ \mathcal{A}\Bigg[\sum_{(l'k')\in\mathcal{C}} \mu_{l'k'}p_{l'k'} + \sum_{d'=1}^{D}\mu_{d'}q_{d'} \Bigg] + \widetilde{P}_{net} \Bigg\} \end{aligned} \tag{4.22a}$$

$$\text{s.t.} \qquad 0 \leq p_{lk} \leq P_{max} \text{ and } 0 \leq q_d \leq q_{\max}. \tag{4.22b}$$

Here the constant $C_1(\nu^{\xi,*}_z,\theta^{\xi,*}_z) = \left(1+\nu^{\xi,*}_z - \frac{\alpha\sqrt{\theta^{\xi,*}_z}}{2}\right)$ for $(z,\xi) = \{(l'k',c),(d',d)\}$ and $C_2(\boldsymbol{\nu}^{c,*},\boldsymbol{\nu}^{d,*},\boldsymbol{\theta}^{c,*},\boldsymbol{\theta}^{d,*},t^*) = \sum_{(l'k')\in\mathcal{C}}\left[\log(1+\nu^{c,*}_{l'k'}) - \nu^{c,*}_{l'k'} - \frac{\alpha}{2}(\theta^{c,*}_{l'k'})^{\frac{-1}{2}}\right]$ $+ \sum_{d'=1}^{D}\left[\log\left(1+\nu^{d,*}_{d'}\right) - \nu^{d,*}_{d'} - \frac{\alpha}{2}(\theta^{d,*}_{d'})^{\frac{-1}{2}}\right] + t^*$. In problem **P8**, the objective function consists of scalar fractional terms $\frac{\Delta^c_{l'k'}}{(\Delta^c_{l'k'}+\Lambda^c_{l'k'})}$ and $\frac{\Delta^d_{d'}}{(\Delta^d_{d'}+\Lambda^d_{d'})}$, which makes the objective non-convex. We now use quadratic transform (QT) to decouple each of these scalar ratios. The equivalent problem is given as

$$
\begin{aligned}
\mathbf{P9}: \underset{\mathbf{P},\mathbf{Q},\mathbf{y}^c,\mathbf{y}^d}{\text{Maximize}} \quad & \sum_{(l'k')\in\mathcal{C}} 2y^c_{l'k'}\sqrt{C_1(\nu^{c,*}_{l'k'},\theta^{c,*}_{l'k'})\Delta^c_{l'k'}} - (y^c_{l'k'})^2(\Delta^c_{l'k'} \\
& + \Lambda^c_{l'k'}) + \sum_{d'=1}^{D} 2y^d_{d'}\sqrt{C_1(\nu^{d,*}_{d'},\theta^{d,*}_{d'})\Delta^d_{d'}} - (y^d_{d'})^2(\Delta^d_{d'} + \Lambda^d_{d'}) \\
& + C_2(\boldsymbol{\nu}^{c,*},\boldsymbol{\nu}^{d,*},\boldsymbol{\theta}^{c,*},\boldsymbol{\theta}^{d,*},t^*) \\
& - z^*\sqrt{t^*}\left\{\mathcal{A}\left[\sum_{(l'k')\in\mathcal{C}} \mu_{l'k'}p_{l'k'} + \sum_{d'=1}^{D}\mu_{d'}q_{d'}\right] + \widetilde{P}_{net}\right\} \qquad (4.23a) \\
\text{s.t.} \quad & 0 \le p_{lk} \le P_{max} \text{ and } 0 \le q_d \le q_{\max}. \qquad (4.23b)
\end{aligned}
$$

Here $\mathbf{y}^c = [y^c_{11},\cdots,y^c_{LK}]^T \in \mathbb{R}^{LK\times 1}$ and $\mathbf{y}^d = [y^d_1,\cdots,y^d_D]^T \in \mathbf{R}^{D\times 1}$ are the set of auxilary variables that decouple the scalar ratios. For a given $\mathbf{P},\mathbf{Q}$, the optimal values of $\mathbf{y}^c$ and $\mathbf{y}^d$ are calculated in closed form as

$$
y^c_{l'k'} = \frac{\sqrt{C_1(\nu^{c,*}_{l'k'},\theta^{c,*}_{l'k'})\Delta^c_{l'k'}}}{(\Delta^c_{l'k'} + \Lambda^c_{l'k'})} \text{ and } y^d_{d'} = \frac{\sqrt{C_1(\nu^{d,*}_{d'},\theta^{d,*}_{d'})\Delta^d_{d'}}}{(\Delta^d_{d'} + \Lambda^d_{d'})}. \qquad (4.24)
$$

Problem **P9**, for a fixed values of $\mathbf{y}^c,\mathbf{y}^d$, is a concave maximization problem in $(\mathbf{P},\mathbf{Q})$. To solve the concave maximization problem, we now differentiate the objective function of **P9** with respect to the optimization variables $p_{l'k'}$ and $q^d_{d'}$ in the following theorem.

Theorem 3. *For fixed auxiliary variables t, $y_{l'k'}$,$y^d_{d'}$ and z, the optimal $p_{l'k'}, q^d_{d'}$, using the first order condition, are obtained as follows*

$$
p^*_{l'k'} = \min\left\{ \frac{y^2_{l'k'}C_1(\nu^{c,*}_{l'k'},\theta^{c,*}_{l'k'})a^{l'}_{l'k'}}{\left(\sum_{l=1}^{L}\sum_{k=1}^{K} y^2_{l,k}b^{l,k}_{l'k'} + \sum_{(l,k)\in\mathcal{P}_{l'k'}} y^2_{l,k}c^{l,k}_{l'k'} + \sum_{d=1}^{D} y^2_d c^d_{l'k'} + \mathcal{A}\mu_{l'k'}z^*\sqrt{t^*}\right)^2}, P_{max} \right\} \qquad (4.25)
$$

$$
q^*_d = \min\left\{ \frac{y^2_d C_1(\nu^{d,*}_d,\theta^{d,*}_z)a^d_d}{\left(\sum_{l=1}^{L}\sum_{k=1}^{K} y^2_{l,k}e^d_{jk} + \sum_{d'=1} y^2_{d'}e^{d'}_d + \mathcal{A}\mu_d z^*\sqrt{t^*}\right)^2}, q_{max} \right\} \qquad (4.26)
$$

Proof. Please refer to Appendix A. □

We observe that (4.25) and (4.26) are fixed-point equations, where the RHS expression depend themselves on $p_{l'k'}$ and $q_{d'}$. We therefore develop an iterative algorithm to solve the problem **P9** by starting from a feasible transmit power vector and iteratively updating the transmit powers as the solution of (4.25), (4.26) and rest of the auxiliary variables $z, t, y^c_{l'k'}, y^d_{d'}, \nu^c_{l'k'}, \nu^d_{d'}, \theta^c_{l'k'}, \theta^d_{d'}$. The resulting formal procedure to solve **P9** in (4.23) is provided in Algorithm 1.

Algorithm 1: GEE maximization algorithm Closed Form Approach

Input: Given a tolerance $\epsilon > 0$, the maximum number of iterations N and maximum power constraint P_{max} for UE U_{jk} and D2D U_d. Calculate the initial values $p_{l'k'}, q_{d'}$ with random power allocation for all UEs and D2D users i.e., $\mathbf{P}^{(1)} \sim \mathbb{U}[0, P_{max}]$ and $\mathbf{Q}^{(1)} \sim \mathbb{U}[0, q_{max}]$.

Output: $p^*_{l'k'}$ and q^*_d.

for $n \leftarrow 1$ **to** N **do**

- Given a feasible $p^{(n)}_{l'k'}$ and $p^{(n)}_d$, update auxiliary variables $z^{(n)}$, $(y^{c(n)}_{l'k'}, y^{d(n)}_{d'})$ using (4.7) and (4.24)
- Update the epigraph variables t and $(\nu^c_{l'k'}, \nu^d_{d'}, \theta^c_{l'k'}, \theta^d_{d'})$ using (4.11) and (4.19), respectively
- Compute $p^{(n+1)}_{l'k'}$ and $q^{(n+1)}_{d'}$ using (4.25) and (4.26).
- Do until convergence If$\sum\limits_{l=1}^{L}\sum\limits_{l=1}^{K}\left(q^{(n+1)}_{l'k'} - p^{(n)}_{l'k'}\right) + \sum\limits_{d=1}^{D}\left(q^{(n+1)}_d - q^{(n)}_d\right) < \epsilon$ break.

return $\mathbf{p}^*, \mathbf{q}^*$.;

4.3 Stochastic MM-based GEE optimization

In the previous section, we formulated a deterministic GEE problem using SE lower-bound expressions of CUs D2D users. The optimal power allocation scheme in Theorem 3, computes the optimal transmit powers as a function of expectations over random channels (see Theorem 3). It is easy to compute these expectations in closed-form for MRC, but computing them is a non-trivial task for Ia-ZF and Ia-MMSE combining schemes. As a result, for Ia-ZF and Ia-MMSE schemes, the optimal power allocation scheme in Theorem 3 uses statistical averages, which requires the collection of a large number of random channel realizations before updating the transmit powers. The GEE optimization in Theorem 3, therefore, requires more memory to store the samples and has higher computational complexity. To reduce the memory requirement and the computational complexity, we now recast the GEE problem in **P1** as a stochastic optimization problem and then optimize it using a low-complexity stochastic MM framework. The stochastic GEE problem is formulated as

$$\begin{aligned} \mathbf{P1}_{\text{sto}} : &\underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}}\ \mathbb{E}\left[g(\mathbf{P},\mathbf{Q},\mathcal{F})\right] \triangleq \overline{f}_{\text{GEE}}(\mathbf{P},\mathbf{Q}) \\ &\text{s.t.} \qquad 0 \le p_{lk} \le P_{max} \text{ and } 0 \le q_d \le q_{\max}. \end{aligned}$$

Here $g(\mathbf{P},\mathbf{Q},\mathcal{F})$ denotes the instantaneous GEE of the system, with $\mathcal{F} = \{\mathbf{h}_{lk}^j, \mathbf{h}_{d'}^j, g_d^{d'}, g_{lk}^{d'}\}, \forall l,k,j,d'$. It is defined as

$$g(\mathbf{P},\mathbf{Q},\mathcal{F}) = \frac{\sum\limits_{(l',k')\in\mathcal{C}} \left[\log(1+\text{SINR}_{l'k'}^c) - \alpha\sqrt{\overline{V_{l'k'}^c}}\right] + \sum\limits_{d'=1}^{D} \left[\log\left(1+\text{SINR}_{d'}^d\right) - \alpha\sqrt{\overline{V_{d'}^d}}\right]}{\mathcal{A}\left[\sum\limits_{(l',k')\in\mathcal{C}} \mu_{l'k'}p_{l'k'} + \sum\limits_{d'=1}^{D} \mu_{d'}q_{d'}\right] + \widetilde{P}_{net}}, \tag{4.28}$$

where $\overline{V_{l'k'}^c} = \frac{\text{SINR}_{l'k'}^c}{1+\text{SINR}_{l'k'}^c}$ and $\overline{V_{d'}^d} = \frac{\text{SINR}_{d'}^d}{1+\text{SINR}_{d'}^d}$, with $\text{SINR}_{l'k'}^c$ and $\text{SINR}_{d'}^d$ being the instantaneous SINR expressions of CU and D2D user as given in (3.16) and (3.22), respectively. The outer expectations in the objective function are over the random channels $\mathcal{F}$. Problem $\mathbf{P1}_{\text{sto}}$, due to fractional instantaneous GEE function and stochastic objective, is a non-convex optimization problem. We now solve it using stochastic minorization maximization (SMM) which is a stochastic variant of MM. We now briefly explain the SMM framework in the following proposition.

Proposition 1. *Consider the following stochastic maximization problem with expectation over the random parameter ζ:*

$$\mathbf{P}: \quad \underset{\mathcal{P}}{\mathit{Maximize}} \quad f(\mathcal{P}) = \mathbb{E}[g(\mathcal{P},\zeta)] \quad \textit{subject to} \quad \mathcal{P} \in \mathcal{P}. \tag{4.29}$$

The SMM framework solves problem $\mathbf{P}$ *by generating a sequence of points* $\overline{\mathcal{P}}^t$ *solving the following problem*

$$\mathbf{P}_1: \ \overline{\mathcal{P}}^t = \underset{\mathcal{P}}{argmax} \quad \tilde{f}(\mathcal{P}) \quad subject\ to \quad \mathcal{P} \in \mathcal{P}. \tag{4.30}$$

Here $\tilde{f}(\mathcal{P})$ *is the recursive surrogate function which is updated at every new realization of random parameter* $\zeta = \zeta^t$ *and* $\mathcal{P} = \mathcal{P}^t$ *as*

$$\tilde{f}(\mathcal{P}) = (1-\rho^t) f^{t-1} + \rho^t \hat{g}(\mathcal{P}; \mathcal{P}^t, \zeta^t), \tag{4.31}$$

The scalar f^t *is an approximation of* $\mathbb{E}[g(\mathcal{P}, \zeta)]$*, which is updated recursively as* $f^t = (1-\rho^t) f^{t-1} + \rho^t g(\mathcal{P}^t, \mathbf{h}^t)$ *and* $f^{-1} = 0$*. Here* $\hat{g}(\mathcal{P}; \mathcal{P}^t, \zeta^t)$ *is the sample surrogate function constructed at the point* $\mathcal{P} = \mathcal{P}^t$ *and for a realization* $\zeta = \zeta^t$*. It satisfies the following properties*

$$g(\mathcal{P}^t, \zeta^t) = \hat{g}(\mathcal{P}^t; \mathcal{P}^t, \zeta^t) \ and \ \nabla g(\mathcal{P}, \zeta^t)\big|_{\mathcal{P}=\mathcal{P}^t} = \nabla \hat{g}(\mathcal{P}; \mathcal{P}^t, \zeta^t)\big|_{\mathcal{P}=\mathcal{P}^t}. \tag{4.32}$$

For a given $\overline{\mathcal{P}}^t$*, the optimal value of* $\mathcal{P}$ *is calculated as*

$$\mathcal{P}^{t+1} = (1-\gamma^t)\mathcal{P}^t + \gamma^t \overline{\mathcal{P}}^t. \tag{4.33}$$

The constants $(\gamma^t, \rho^t) \in (0,\ 1]$ *denote the step sizes, which satisfies* $\vartheta^t \to 0$, $\sum_t \vartheta^t = \infty$ *and* $\sum_t (\vartheta^t)^2 < \infty$*, where* $\vartheta = \{\gamma, \rho\}$*. By iterating between i) solving the problem* $\mathbf{P1}$ *in* (4.30)*; ii) updating the optimal transmit power* $\mathcal{P}$ *using the*

We now construct a valid surrogate function on the instantaneous GEE expression and later use it to solve the ergodic GEE problem using SMM technique. The instantaneous GEE problem, for a fixed $\mathcal{F} = \mathcal{F}^t$, is given as

$$\mathbf{P2}_{\text{sto}}: \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}} \quad g(\mathbf{P}, \mathbf{Q}, \mathcal{F}^t) \quad \text{s.t.} \quad 0 \le p_{lk} \le P_{max} \text{ and } 0 \le q_d \le q_{\max}.$$

We see that for a fixed set of random channels, the function $g(\mathbf{P}, \mathbf{Q}, \mathcal{F}^t)$ looks similar to that of the deterministic GEE objective in problem **P1**. We therefore, similar to the deterministic GEE optimization in Section **??**, apply LDT and QT iteratively to decouple the scalar ratios. The final surrogate function on the instantaneous GEE expression calculated at a feasible point $(\mathbf{P}^r, \mathbf{Q}^r)$ is given as

$$\begin{aligned}
&\hat{g}(\mathbf{P}, \mathbf{Q}; \mathbf{P}^r, \mathbf{Q}^r, \mathcal{F}^t) \\
&= \sum_{(l'k')\in\mathcal{C}} 2\bar{y}^c_{l'k'} \sqrt{C_1(\bar{\nu}^{c,*}_{l'k'}, \bar{\theta}^{c,*}_{l'k'})\bar{\Delta}^c_{l'k'}} - (\bar{y}^c_{l'k'})^2(\bar{\Delta}^c_{l'k'} + \bar{\Lambda}^c_{l'k'}) + \sum_{d'=1}^{D} 2\bar{y}^d_{d'} \sqrt{C_1(\bar{\nu}^{d,*}_{d'}, \bar{\theta}^{d,*}_{d'})\bar{\Delta}^d_{d'}} \\
&- (\bar{y}^d_{d'})^2(\bar{\Delta}^d_{d'} + \bar{\Lambda}^d_{d'}) + C_2(\bar{\boldsymbol{\nu}}^{c,*}, \bar{\boldsymbol{\nu}}^{d,*}, \bar{\boldsymbol{\theta}}^{c,*}, \bar{\boldsymbol{\theta}}^{d,*}, \bar{t}^*) - \bar{z}^*\sqrt{\bar{t}^*}\left\{\mathcal{A}\left[\sum_{(l'k')\in\mathcal{C}} \mu_{l'k'}p_{l'k'} + \sum_{d'=1}^{D}\mu_{d'}q_{d'}\right] + \widetilde{P}_{net}\right\}
\end{aligned} \tag{4.35}$$

Here $C_1(\bar{\nu}^{\xi,*}_z, \bar{\theta}^{\xi,*}_z) = \left(1 + \bar{\nu}^{\xi,*}_z - \frac{\alpha\sqrt{\bar{\theta}^{\xi,*}_z}}{2}\right)$ for $(z,\xi) = \{(l'k', c), (d', d)\}$ and $C_2(\bar{\boldsymbol{\nu}}^{c,*}, \bar{\boldsymbol{\nu}}^{d,*}, \bar{\boldsymbol{\theta}}^{c,*}, \bar{\boldsymbol{\theta}}^{d,*}, \bar{t}^*) = \sum_{(l'k')\in\mathcal{C}} \left[\log(1+\bar{\nu}^{c,*}_{l'k'}) - \bar{\nu}^{c,*}_{l'k'} - \frac{\alpha}{2}(\bar{\theta}^{c,*}_{l'k'})^{\frac{-1}{2}}\right]$ $+ \sum_{d'=1}^{D}\left[\log\left(1+\bar{\nu}^{d,*}_{d'}\right) - \bar{\nu}^{d,*}_{d'} - \frac{\alpha}{2}(\bar{\theta}^{d,*}_{d'})^{\frac{-1}{2}}\right] + \bar{t}^*$ for $(z,\xi) = \{(l'k', c), (d', d)\}$. The scalars $\nu^{\xi,*}_z, \theta^{\xi,*}_z, y^{\xi}_z, \bar{z}^*$ and $\bar{t}^*$ are calculated using the feasible values $(\mathbf{P}^r, \mathbf{Q}^r)$ as

$$\bar{\nu}^{\xi,*}_z = \frac{\overline{\Delta}^{\xi}_z}{\overline{\Lambda}^{\xi}_z}, \quad \bar{\theta}^{\xi,*}_z = \frac{\overline{\Delta}^{\xi}_z}{\overline{\Delta}^{\xi}_z + \overline{\Lambda}^{\xi}_z}, \quad y^{\xi}_z = \frac{C_1(\nu^{\xi,*}_z, \theta^{\xi,*}_z)\overline{\Lambda}^{\xi}_z}{\overline{\Delta}^{\xi}_z + \overline{\Lambda}^{\xi}_z} \tag{4.36}$$

$$\begin{aligned}
\bar{t}^* &= \sum_{(l',k')\in\mathcal{C}} \log(1+\mathrm{SINR}^c_{l'k'}) - \alpha\sqrt{\frac{\mathrm{SINR}^c_{l'k'}}{1+\mathrm{SINR}^c_{l'k'}}} + \sum_{d'=1}^{D}\log\left(1+\mathrm{SINR}^d_{d'}\right) - \alpha\sqrt{\frac{\mathrm{SINR}^d_{d'}}{1+\mathrm{SINR}^d_{d'}}} \\
\bar{z}^* &= \frac{\sqrt{\bar{t}^*}}{\mathcal{A}\left[\sum_{(l',k')\in\mathcal{C}} \mu_{l'k'}p_{l'k'} + \sum_{d'=1}^{D}\mu_{d'}q_{d'}\right] + \widetilde{P}_{net}}.
\end{aligned} \tag{4.37}$$

Here $\mathrm{SINR}^c_{l'k'}$ and $\mathrm{SINR}^d_{d'}$ are computed at the feasible point $(\mathbf{P}^r, \mathbf{Q}^r)$ and the channel realization $\mathcal{F} = \mathcal{F}^t$. It is easy to show that the function $g(\mathbf{P}, \mathbf{Q}; \mathbf{P}^r, \mathbf{Q}^r, \mathcal{F}^t)$ satisfies the properties in (4.32), and is a valid surrogate function on $g(\mathbf{P}, \mathbf{Q}, \mathcal{F}^t)$. For any feasible point

The stochastic problem $\mathbf{P1}_{\mathrm{sto}}$ is therefore recast using the surrogate function in (4.35) as

$$\mathbf{P2}_{\text{sto}}: \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}}\ \tilde{f}(\mathbf{P},\mathbf{Q}) = (1-\rho^t)f^{t-1} + \rho^t \hat{g}(\mathbf{P},\mathbf{Q};\mathbf{P}^r,\mathbf{Q}^r,\mathcal{F}^t)$$
$$\text{s.t.} \qquad 0 \le p_{lk} \le P_{max} \text{ and } 0 \le q_d \le q_{\max}.$$

We see that the above problem is concave in $(\mathbf{P},\mathbf{Q})$. Inorder to find the optimal values of the function, we need to find the function of surrogate function in a simplistic way. There expanding the function $g(\mathbf{P},\mathbf{Q};\mathbf{P}^r,\mathbf{Q}^r,\mathcal{F}^t)$ by expanding the Denominator an Numerator of the SINR and grouping together transmit power-wise($Pmat$) and it given as:

$$\begin{aligned}&\hat{g}(\mathbf{P},\mathbf{Q};\mathbf{P}^r,\mathbf{Q}^r,\mathcal{F}^t)\\ &= \sum_{(l'k')\in\mathcal{C}} \alpha^t_{l'k'}p_{l'k'} + 2\beta^t_{l'k'}\sqrt{p_{l'k'}} + \sum_{d'=1}^{D}\chi^t_{d'}q_{d'} + 2\omega^t_{d'}\sqrt{q_{d'}} + Const(z,t,P_{net}). \end{aligned} \tag{4.39}$$

where the constants are $\alpha^t_{l'k'}, \beta^t_{l'k'}, \chi^t_{d'}, \omega_{d'}$ at the t th realization are expressed as:

$$\alpha^t_{l'k'} = -\left(\sum_{l=1}^{L}\sum_{k=1}^{K} y^2_{l,k}b^{l,k}_{l'k'} + \sum_{(l,k)\in\mathcal{P}_{l'k'}} y^2_{l,k}c^{l,k}_{l'k'} + \sum_{d=1}^{D} y^2_d c^d_{l'k'} + \mathcal{A}\mu_{l'k'}z^*\sqrt{t^*}\right) \tag{4.40}$$

$$\beta^t_{l'k'} = \sqrt{y^2_{l'k'}C_1(\nu^{c,*}_{l'k'},\theta^{c,*}_{l'k'})a^{l'}_{l'k'}} \tag{4.41}$$

$$\chi_{d'} = -\left(\sum_{l=1}^{L}\sum_{k=1}^{K} y^2_{l,k}e^d_{jk} + \sum_{d'=1} y^2_{d'}e^{d'}_d + \mathcal{A}\mu_d z^*\sqrt{t^*}\right) \tag{4.42}$$

$$\omega^t_{d'} = \sqrt{y^2_d C_1(\nu^{d,*}_d,\theta^{d,*}_z)a^d_d} \tag{4.43}$$

The grouped convex approximate function as in (4.39), provides a detail that the convex approximate remains the same for any t realizations. Therefore, by mathematical induction the surrogate can be constructed as:

$$\tilde{f}(\mathbf{P},\mathbf{Q}) = \sum_{(l'k')\in\mathcal{C}} \boldsymbol{\alpha}^{1:t}_{l'k'}p_{l'k'} + 2\boldsymbol{\beta}^{1:t}_{l'k'}\sqrt{p_{l'k'}} + \sum_{d'=1}^{D}\boldsymbol{\chi}^{1:t}_{d'}q_{d'} + 2\boldsymbol{\omega}^{1:t}_{d'}\sqrt{q_{d'}} + Const(z,t,P_{net}) \tag{4.44}$$

where the surrogate constants $\boldsymbol{\alpha}_{l'k'}^{1:t}, \boldsymbol{\beta}_{l'k'}^{1:t}, \boldsymbol{\chi}_{d'}^{1:t}, \boldsymbol{\omega}_{d'}^{1:t}$ are updated as:

$$\boldsymbol{\alpha}_{l'k'}^{1:t} = (1-\rho^t)\boldsymbol{\alpha}_{l'k'}^{1:t-1} + \rho^t \alpha_{l'k'}^t \tag{4.45}$$

$$\boldsymbol{\beta}_{l'k'}^{1:t} = (1-\rho^t)\boldsymbol{\beta}_{l'k'}^{1:t-1} + \rho^t \beta_{l'k'}^t \tag{4.46}$$

$$\boldsymbol{\chi}_{d'}^{1:t} = (1-\rho^t)\boldsymbol{\chi}_{d'}^{1:t-1} + \rho^t \chi_{d'}^t \tag{4.47}$$

$$\boldsymbol{\omega}_{d'}^{1:t} = (1-\rho^t)\boldsymbol{\omega}_{d'}^{1:t-1} + \rho^t \chi_{d'}^t \tag{4.48}$$

With this surrogate function to find the optimal values of power variables, the problem formulation $\mathbf{P2}_{sto}$ is reconstructed as:

$$\mathbf{P3}_{\text{sto}} : \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}} \sum_{(l'k')\in\mathcal{C}} \boldsymbol{\alpha}_{l'k'}^{1:t} p_{l'k'} + \boldsymbol{\beta}_{l'k'}^{1:t}\sqrt{p_{l'k'}} + \sum_{d'=1}^{D} \boldsymbol{\chi}_{d'}^{1:t} q_{d'} + \boldsymbol{\omega}_{d'}^{1:t}\sqrt{q_{d'}} + Const(z,t,P_{net})$$

$$\text{s.t.} \qquad 0 \le p_{lk} \le P_{max} \text{ and } 0 \le q_d \le q_{\max}.$$

The optimal values of the transmit powers $p_{l'k'}$ and $q_{d'}$ for the t th realization is are therefore calculated in closed-form as

$$p_{l'k'}^{t+1} = \min\left\{\left(\frac{\boldsymbol{\alpha}_{l'k'}^{1:t}}{\boldsymbol{\beta}_{l'k'}^{1:t}}\right)^2, P_{max}\right\} \tag{4.50}$$

$$q_{d'}^{t+1} = \min\left\{\left(\frac{\boldsymbol{\chi}_{d'}^{1:t}}{\boldsymbol{\omega}_{d'}^{1:t}}\right)^2, q_{max}\right\} \tag{4.51}$$

So, we have developed an iterative algorithm to solve the problem $\mathbf{P1}_{sto}$ with transmit power and by receiving the channel realization at each iteration updating the surrogate function $\tilde{f}(\mathbf{P},\mathbf{Q})$ with convex approximate, the transmit powers as the solution of (4.50), (4.51) and rest of the auxiliary variables $z, t, y_{l'k'}^c, y_{d'}^d, \nu_{l'k'}^c, \nu_{d'}^d, \theta_{l'k'}^c, \theta_{d'}^d$.The resulting formal procedure to solve is provided in Algorithm 2.

Algorithm 2: Stochastic GEE maximization, Closed Form Approach

---

Input: Given a tolerance $\epsilon > 0$, the maximum number of iterations N and maximum power constraint P_{max} for UE U_{jk} and D2D U_d. Calculate the initial values $p_{l'k'}, q_{d'}$ with random power allocation for all UEs and D2D users i.e., $\mathbf{P}^{(1)} \sim \mathbb{U}[0, P_{max}]$ and $\mathbf{Q}^{(1)} \sim \mathbb{U}[0, q_{max}]$.

Output: $p^*_{l'k'}$ and q^*_d.

for $n \leftarrow 1$ **to** N **do**

- Obtain the n^{th} channel realization.
- Update the surrogate function $\tilde{f}(\mathbf{P}, \mathbf{Q})$ by updating $\boldsymbol{\alpha}^{1:n}_{l'k'}, \boldsymbol{\beta}^{1:n}_{l'k'}, \boldsymbol{\chi}^{1:n}_{d'}, \boldsymbol{\omega}^{1:n}_{d'} using (4.40)$
- Given a feasible $p^{(n)}_{l'k'}$ and $p^{(n)}_d$, update auxiliary variables $z^{(n)}$, $(y^{c(n)}_{l'k'}, y^{d(n)}_{d'})$ $(\nu^c_{l'k'}, \nu^d_{d'}, \theta^c_{l'k'}, \theta^d_{d'})$ using (4.36) and (4.37)
- Compute $p^{(n+1)}_{l'k'}$ and $q^{(n+1)}_{d'}$ using (4.50) and (4.51).
- Do until convergence $\sum_{l'=1}^{L} \sum_{k'=1}^{K} \left(p^{(n+1)}_{l'k'} - p^{(n)}_{l'k'} \right) + \sum_{d=1}^{D} \left(q^{(n+q)}_d - q^{(n-1)}_d \right) < \epsilon$
  break.

return $\mathbf{p}^*$,$\mathbf{q}^*$;

Chapter 5

D2D underlaid multi-cell mMIMO URLLC system : Simulation Results and Conclusion

5.1 Simulation Results

In this section, we evaluate the optimizations of the URLLC-enabled D2D overlayed mMIMO uplink system and study the optimizations' convergence as well as the low complexity of it. We run a simulation of a four-cell system in which each cell is made up of ten CUs and ten D2D pairings are spread out evenly throughout the whole region. It is envisaged that the four BSs will be spread out across an area of 1 km2, and the wrap-around technique will be used in order to prevent any edge effects from occurring. The D2D transmitter and receiver are separated by a distance of 10 metres from one another. The carrier frequency is set at 2 GHz, the bandwidth is set at 20 MHz, and there are a total of 50 symbols included inside the coherence interval. The total number of pilots that are allocated to D2D pairs is 5, and each D2D pair chooses one of them at random to be their pilot. In a similar vein, there are a total of ten pilots allocated to CUs in a cell, with pilot reuse accounting for one. In addition, a value of -94 dBm has been assigned to the noise variance.

We now model the large-scale fading coefficients, Rician factor and correlation matrices of different channels as follows:

- *Large-scale fading coefficient and Rician factor model for BSs:* We model the large-scale fading coefficient β^j_{jk}, β^j_d and the Rician factor K^j_{lk}, K^j_d using the 3GPP specifications [30] as

$$\beta^j_{lk} = -30.18 - 26\log_{10}(d^j_{lk}) + F^j_{lk} \text{ dB and } K^j_{lk} = 13 - 0.03 d^j_{lk} \text{ dB}, \tag{5.1}$$

$$\beta^j_d = -30.18 - 26\log_{10}(d^j_d) + F^j_d \text{ dB and } K^j_d = 13 - 0.03 d^j_d \text{ dB}, \tag{5.2}$$

where d^j_{lk}, d^j_d are the distance between the UE U_{lk} and the j-th BS and d^{th} D2D transmitter and the j-th BS respectively, and F^j_{lk}, F^j_d are the shadow fading for UE U_{lk} in lth cell and d^{th} D2D transmitter respectively, which are distributed as $\mathcal{N}(0, \sigma^2_{sf})$. We set $\sigma^2_{sf} = 4$ dB. The large scale fading coefficients $\beta^j_{l,k}, \beta^j_d$ are normalized with respect to noise variance σ^2.

- *Large-scale fading coefficient and Rician Factor model for D2D receivers:* We model the large-scale fading coefficient $g^{d'}_{jk}, g^{d'}_d$ and the Rician factor $K^{d'}_{lk}, K^{d'}_d$ using the 3GPP specifications [30] as

$$g^{d'}_{lk} = -30.18 - 26\log_{10}(d^{d'}_d) + F^{d'}_{lk} \text{ dB and } K^{d'}_{lk} = 13 - 0.03 d^{d'}_d \text{ dB}, \tag{5.3}$$

$$g^{d'}_d = -30.18 - 26\log_{10}(d^{d'}_d) + F^{d'}_d \text{ dB and } K^{d'}_{R_d} = 13 - 0.03 d^{d'}_d \text{ dB}, \tag{5.4}$$

where $d^{d'}_{lk}, d^{d'}_d$ are the distance between the UE U_{lk} and the d'-th D2D receiver and d^{th} D2D transmitter and the d'-th D2D receiver respectively, and $F^{d'}_{lk}, F^{d'}_d$ are the shadow fading for UE U_{lk} in lth cell and d^{th} D2D transmitter respectively, which are distributed as $\mathcal{N}(0, \sigma^2_{sf})$. We set $\sigma^2_{sf} = 4$ dB. The large scale fading coefficients $g^{d'}_{l,k}, g^{d'}_d$ are normalized with respect to noise variance σ^2.

- *Spatial correlation model:* We assume, similar to [31], that the BS is equipped with uniform linear array and model the spatial correlation matrices $\{\mathbf{R}_{lk}^{j}\}_{\forall l,k,j}$, $\{\mathbf{R}_{d}^{j}\}_{\forall d,j}$ as follows

$$\left[\mathbf{R}_{lk}^{j}\right]_{m,n} = \frac{\beta_{lk}^{j,\text{NLoS}}}{N} \sum_{n=1}^{N} e^{j\pi(n-m)\sin\left(\phi_{lk,n}^{j}\right)} e^{-\frac{\sigma_{\phi}^{2}}{2}\left(\pi(n-m)\cos\left(\phi_{lk,n}^{j}\right)\right)^{2}} \tag{5.5}$$

$$\left[\mathbf{R}_{d}^{j}\right]_{m,n} = \frac{\beta_{d}^{j,\text{NLoS}}}{N} \sum_{n=1}^{N} e^{j\pi(n-m)\sin\left(\phi_{d,n}^{j}\right)} e^{-\frac{\sigma_{\phi}^{2}}{2}\left(\pi(n-m)\cos\left(\phi_{d,n}^{j}\right)\right)^{2}} \tag{5.6}$$

Here $\left[\mathbf{R}_{lk}^{j}\right]_{m,n}$ denotes the (m,n)th entry of the matrix $\mathbf{R}_{lk}^{j}$, $\rho \in [0,1)$ is the correlation coefficient, and ϕ_{lk}^{j} is the nominal angle between the kth UE in lth cell U_{lk} and BS j and the notations are extended to d^{th} D2D transmitter and BS j. The cluster have a Gaussian scattered distributed Angle of Arrival(AoA) with Angular Standard deviation(ASD) $\sigma_{\phi} = 5^{o}$.

- *Power Parameters:* The power parameters are given by $P_{\text{FIX},tot} = 5W$, $P_{\text{BS},l} = 0.1W$, $P_{\text{LO},l} = 0.1W$, $P_{\text{LO},l} = 0.1W$, $P_{\text{CU},l} = 0.1W$, $L_{\text{BS}} = 750$ Gflops/W, $L_{\text{D}} = 750$ Gflops/W.

- *Other parameters:* The decode error probability present in the URLLC term of D2D users $\epsilon = 10^{-5}$.

5.1.1 Validation of Optimization Technique

As the work of the paper is of optimization techniques, it is necessary to understand that the optimal value has been achieved. The ϵ which is the tolerance is set to be as 0.001 and the framework converges smoothly. In the Fig.5.1, we plot the GEE (with UATF and Ergodic SE) of users vs the maximum transmit power allocation to both CUs and D2Ds. The plots are compared among i)Optimized Power Allocation (OPA) through optimization of GEE (GEE-MAX) ii) OPA through maximization of

SE (SE-MAX) iii) Full Power Allocation (FPA) where the users (CU and D2D) are allocated power equally to the maximum. The difference between the SE-MAX and GEE-MAX is that the auxiliary variable z is tuned down to zero, which means it does not carry the weight of the GEE Denominator. The Fig.5.1a, provides the OPA plot of the converged value of deterministic GEE optimization. Similarly, the Figs.9.1b-d provides the OPA plot of the converged value of stochastic GEE optimization of the Ergodic GEE with the decoders of MR, IaZF, IaMMSE respectively. At higher powers, you can clearly see that MR can maximize the sum SE much better that the MMSE, which should be true because the MMSE is built based on maximizing the SINR at FPA. Thus strengthening the work. The paper has taken 10 User Setups and have averaged it for the uniformity between optimization techniques.

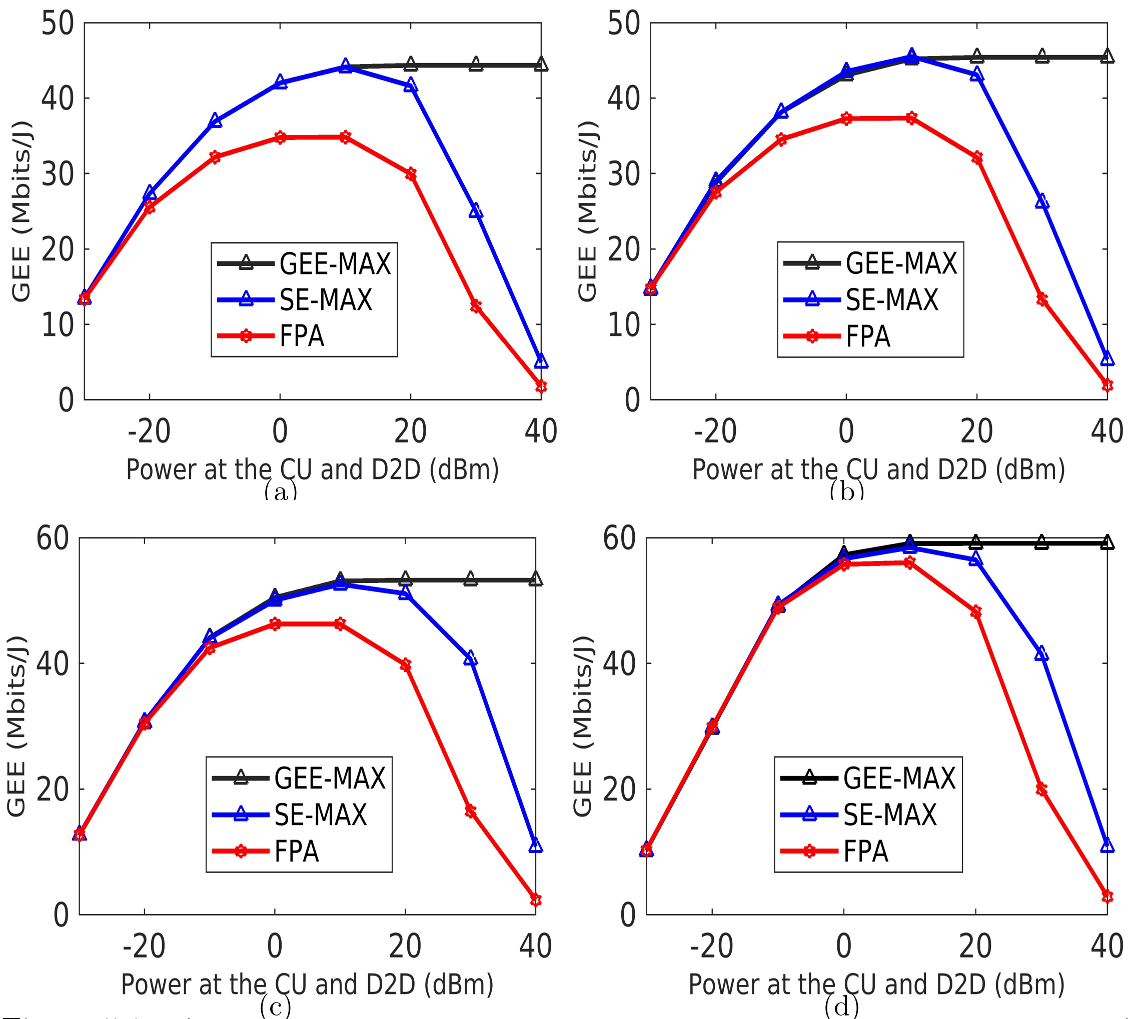

Figure 5.1: a) Deterministic GEE versus transmit power P_T, Stochastic GEE with b) MR; c) IaZF and; d) IaMMSE

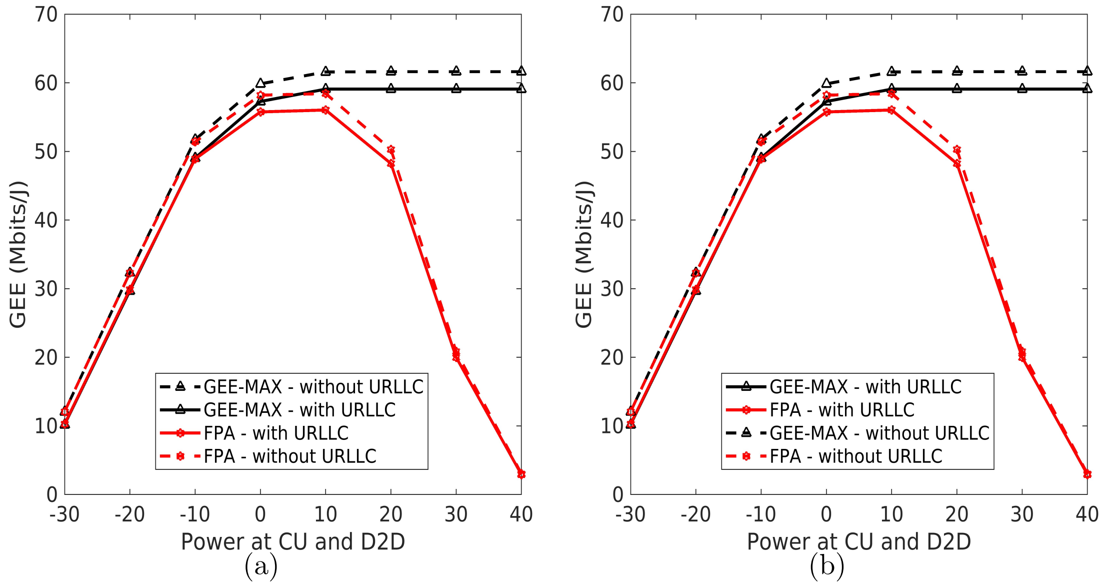

Figure 5.2: Dispersion Term inference by GEE versus BS transmit power with a) UATF b) IaMMSE.

5.1.2 Impact of Dispersion channel Term

In the Fig.5.2 we plot GEE vs transmit power to understand the effect of URLLC in the optimization. The important work of this optimization framework lies on the dispersion term and how it is handled. Therefore, the plots in the Fig.5.2 a-b, are compared with the GEE optimization with and without URLLC (constant of dispersion term to be 0) on UATF GEE and Ergodic GEE with IaMMSE combiner scheme. As you can see the Dip between the URLLC and non-URLLC in both FPA and GEE-MAX which indicates that the URLLC is handled in the optimization. For better understanding, one can also say that the optimal powers between the non-URLLC and URLLC plots are reshuffled such that the dispersion term is handle in the optimization, thus the dip. If the optimal power were same, then dispersion term would have more effect and would die down.

5.1.3 Algorithm Time Complexity vs Convergence

In the Fig.5.3 we plot the CPU time while optimizing the GEE at each iterations with maximum and equal power allocation to the CUs and D2Ds of $20dBm$. The

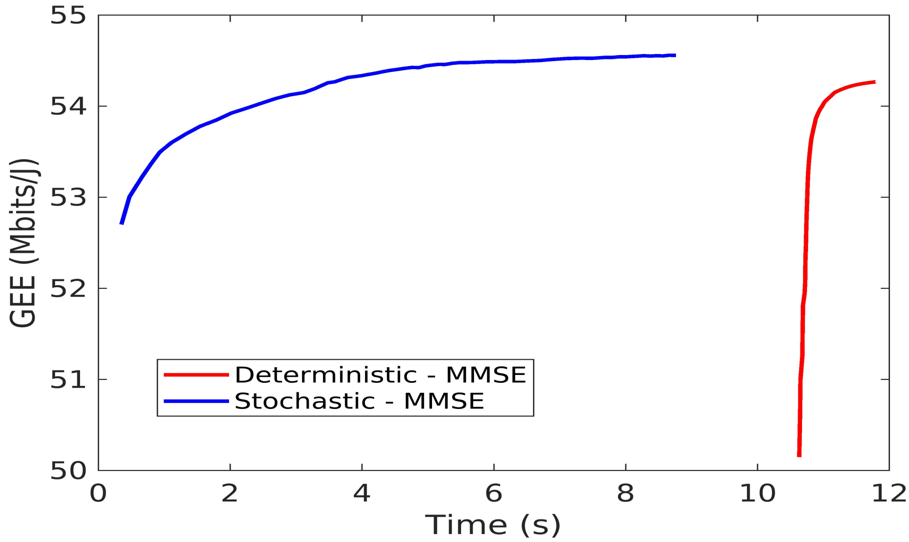

Figure 5.3: GEE Maximization vs Time.

setup is compared in a Intel® Core™ i5-8250U CPU @ 1.60GHz ×8 with a 8GB Ram. In this plot, we compare the time taken to converge, between Deterministic GEE framework with SAA approach and the current framework (Stochastic GEE maximization), both with IaMMSE combiner scheme. The no. of iterations both are around 500. As you can see, the Deterministic GEE requires statistical averages to solve get the objective function, whereas the later uses the channel realizations while iterating. It stochastic GEE maximization can also pipe-lined effectively increasing the latency of the system. The plot also proves that the stochastic GEE optimization reaches higher optimal value as it does not follow the Channel Hardening Laws for closed-form expressions.

5.2 Conclusion

The paper has worked on the D2D-enabled underlay multi-cell massive MIMO network with spatially correlated Rician fading channels with URLLC. The paper has formulated two optimization frameworks at a very low cost, and observed with the time complexity plot. The paper has provided Deterministic GEE maximization approach, Stochastic GEE maximization approach and handle the URLLC term in the optimization.

Chapter 6

Multi-Relay Multi-Cell mMIMO NOMA system : Introduction

It has been shown that non-orthogonal multiple access, also known as NOMA, has a greater SE than its orthogonal counterpart, OMA. NOMA does this by using overlap coding to multiplex the signals of the many users in the power domain. [1, 2, 32]. At the receiver, users of NOMA employ a technique known as successive interference cancellation (SIC), which helps to buffer the IUI impact.Integrating NOMA in mMIMO systems has lately been looked into in [2, 4, 32–36] due to the fact that doing so may greatly increase the system SE. Within the single cell mMIMO NOMA system, channel estimation and straightforward linear BS precoding techniques are used to efficiently minimise IUI [2]. When there are several cells involved, and there is a finite coherence interval, the pilot sequences are replayed in the cells that are close to each other. As a consequence, the channel estimates of users who share pilots get tainted, which brings to a decline in the quality of the SIC and, ultimately, the SE. [34, 36]. The paper [4] which is the base for this paper, works on closed form MRC SE of the Multi-Cell Multi-Relay NOMA Systems.

In the recent decade Massive MIMO(mMIMO) communications are widely popular due to their capability to provide substantial spectral efficiency (SE) and energy efficiency (EE) gains, higher reliability and lower end-to-end latency. Massive antenna arrays are installed at base stations (BSs) in cellular mMIMO systems, which maximises spatial diversity to multiplex tens of users sharing a single spectrum resource while increasing system SE and reliability. In order to limit the amount of inter-user interference in a multi-user mMIMO system, orthogonal multiple access (OMA) techniques are employed which reduce the amount of inter-user interference that occurs. These approaches distribute orthogonal time-frequency resources to each of the system's users (IUI). Even while OMA is very efficient at suppressing IUI, it lowers system SE when users observe channels with low quality. [1, 2].

The system coverage area, as well as its SE in [37], may be improved by the use of cooperative relaying, which involves the BS serving several customers by way of relays. a summary of the uses of NOMA in the relay-aided, single hop mMIMO systems and cooperative relaying was presented in Vaezi *et al.* in [37]. Recent research conducted in has focused on analysing multi-relay-aided uni-cell mMIMO NOMA systems [38–43].

The WMMSE algorithm was first proposed in [44] to optimize the weighted sum rate to design a linear transmit filter. The fundamental concept behind the approach is to transform the objective problem into a WMMSE maximisation, in which the weights are modified in an iterative fashion. The algorithm covers a wide range of issues, including sum rate use.

The sum SE and GEE metrics in the aforementioned optimization works [15, 19, 21, 22, 24], are stochastic functions of random fading channels and receiver noise. These works adopted a statistically approximated average (SAA) optimization approach, wherein an approximate closed-form expression of SE/GEE is first obtained by statistically averaging out the randomness due to channel fading, and the resultant deterministic expression is then optimized using an iterative algorithm. As a result SAA approach requires more memory and has higher computational and time complexity [27]. Stochastic optimization, in contrast to SAA, constructs an approximate sample objective function that captures the randomness in real-time and then solves it using iterative algorithm [27]. The stochastic optimization approach, therefore, has much lesser time complexity than SAA method [27, 28].

The base paper for this part of the thesis [4] uses the low-complexity inexpensive algorithm to solve the *deterministic* GEE and sum SE non-convex problems. The algorithm of [4] tries to convert the hard non-convex *deterministic* problems into a convex approximate using Quadratic Transform (QT) and Lagrangian Dual transform (LDT) which is solved through approach based on modified low-complexity(AMM).

In this work, we propose a generic framework to optimize the sum SE of a Multi-Cell, Multi-Relay Downlink mMIMO NOMA system, which solves both **stochastic** and **deterministic** sum SE non-convex problem with iterative closed form solutions and a very-low time complexity based on the WMMSE approach. This framework is then compared with the framework of the base paper [4]. The paper also extends its work to obtain **optimal deterministic precoder allocation** based on sum SE maximization which reduces the time and memory complexity of precoder formation and sum SE optimization.

Chapter 7

Multi-Relay Multi-Cell mMIMO NOMA system : System Model and SE Expression

7.1 System Model

The system model is acquired from [4]. and this section follows the order of the [4]. This is a briefed section and for any explanation revert back to the [4]. The paper investigates the downlink communication of an multi-cell (L) mMIMO NOMA system where multi-antenna(N) BS supply data to the single-antenna cellular users which forms clusters by deploying NOMA through half-duplex single-antenna AF relays which acts as the intermediate between BS and Users. These relays are installed such that the cluster of users forms around them and are positioned so that both the BS-relay and the user-relay channels include LoS and NLoS components. The cluster $\mathcal{U}_{lk}$ is the user from lth cell and R_{lk} relay .

The communication runs in the TDD mode, and its coherence interval of τ_c symbols is separated into the channel estimation (CE) phase, which consists of the transmission of pilot symbols, and the data transmission (DT) phase, which consists of τ and (τ_c - τ) symbols respectively. During the CE phase, the pilots are sent from the relays to the base stations (BS) as well as the Users for the purposes of transmit beamforming and receive forming, respectively. During the DT phase, the NOMA architecture is put into place, and users are serviced through relays.

The communication operates in TDD mode, with τ_c symbols coherence interval is divided into channel estimation (CE) by transmitting pilots and data transmission (DT) phases of τ and (τ_c - τ) symbols respectively. In the CE phase, the pilots are transmitted from the relays to both BS and the Users for transmit beamforming and receive forming respectively. In DT phase, NOMA is deployed and users are served via relays.

7.1.1 Channel Modeling

As there are two hops, one from BS-relay and the other from relay-BS which channel modeling. As it based on the paper [4], We just provide the mathematical expression for the channel.

For the modelling of BS-relay channel, as mentioned earlier that channel contain LoS and NLos components and which forms Rician channel. $\mathbf{h}_{lk}^{j}$ is the label given for the channel that runs from the l'th BS to R_{lk}.In light of this, the representation of it is as follows:

$$\mathbf{h}_{lk}^{l'} = \overline{\mathbf{h}}_{lk}^{l'} + (\mathbf{R}_{lk}^{l'})^{\frac{1}{2}} \mathbf{h}_{lk}^{l',\mathrm{NLoS}} \tag{7.1}$$

In order to define the spatial correlation of the NLoS component, the matrix $\overline{\mathbf{R}}_{lk}^{k'}$ is used. The $\mathbf{h}_{lk}^{l',\mathrm{NLoS}}$ is random variable which follows complex normal pdf. And the vector $\overline{\mathbf{h}}_{lk}^{l'}$ is the LoS component of the channel. The second-order statistics and its generation are explained in [4].

The relay-user channel is a scalar as both are single antennas. The channel pdf is a rician pdf as it contains Los and NLoS components. $g_{lk,n}^{l'k'}$ is the lable given for the channel scalar to the nth user $\mathcal{U}_{lk}$ from $R_{l'k'}$.In light of this, the representation of it is as follows:

$$g_{lk,n}^{l'k'} = \overline{g}_{lk,n}^{l'k'} + (\upsilon_{lk,n}^{l'k'})^{\frac{1}{2}} g_{lk,n}^{l'k',\text{NLoS}}. \tag{7.2}$$

The scalar $g_{lk,n}^{jk',\text{NLoS}}$ is random variable with complex normal distribution. In order to define the spatial correlation of the NLoS component, $\upsilon_{lk,n}^{l'k'}$ is used.

7.1.2 The CE phase:

In this system model which follows [4], there is no direct link between users and BS as we considering they are far apart and causes huge shadowing and path loss and which leads to not having to estimate the BSs and user channel end-to-end. Therefore the relay helps to estimate the channel between user-relay and BS-relay, the local CSI. To estimate the local CSI, the K relays in the cluster R_{lk} transmits $K = \tau_p$ pilots (ψ_k) which are mutually orthogonal and with magnitude K to both lth BS and cluster of users $\mathcal{U}_{lk}$. Each cell's relays share the pilots, causing pilot contamination.

BS-Relay Estimate: The channel estimate of $\mathbf{h}_{lk}^{l}$ can be considered as a simple MMSE channel estimation between BS and relay. From [4], without using math, the MMSE estimate of channel $\mathbf{h}_{lk}^{l}$ is:

$$\widehat{\mathbf{h}}_{lk}^{l} = \overline{\mathbf{h}}_{lk}^{l} + \sqrt{p_p}\mathbf{R}_{lk}^{l}\mathbf{\Psi}_{lk}\left[\tilde{\mathbf{y}}_k^l - \overline{\mathbf{y}}_k^l\right], \tag{7.3}$$

where $\mathbf{\Psi}_{lk} = \left(\mathbf{I}_N + \tau p_p \sum_{l'=1}^{L} \mathbf{R}_{l'k}^{l}\right)^{-1}$, $\overline{\mathbf{y}}_k^l = \sum_{l'=1}^{L} \sqrt{p_p}\tau\overline{\mathbf{h}}_{l'k}^{l}$ and $\tilde{\mathbf{y}}_k^l = \sum_{l'=1}^{L} \sqrt{p_p}\tau\mathbf{h}_{l'k}^{l} + \mathbf{N}_k^{l}{}^*$ and p_p is the pilot power and $\boldsymbol{N}_k^{l*}$ is AWGN of the lth BS from kth relay.

Relay-User Estimate: The channel estimate of $g_{lk,n}^{lk}$, is can be formed similar to the MMSE channel estimate of D2D section mentioned in the previous chapter, The MMSE estimate of $g_{lk,n}^{lk}$ is obtained from [4] is expressed as:

$$\widehat{g}_{lk,n}^{lk} = \overline{g}_{lk,n}^{lk} + \frac{\sqrt{p_p}\gamma_{lk,n}^{l'k}}{1 + \sum\limits_{l'=1}^{L} \tau p_p \gamma_{lk,n}^{l'k}} \left[\tilde{y}_{lk,n}^{p} - \overline{y}_{lk,n}^{p} \right], \quad \text{where } \overline{y}_{lk,n}^{p} = \sum_{l'=1}^{L} \sqrt{p_p}\tau \overline{g}_{lk,n}^{l'k}. \quad (7.4)$$

where $\tilde{y}_{lk,n}^{p} = \sum\limits_{l'=1}^{L} \sqrt{p_p}\tau g_{lk,n}^{l'k} + \mathbf{n}_{lk,n}^{p*}$

7.1.3 The DT phase:

Two time slots are necessary for the transfer of data from the base station to the users through relay. The operation of these time slots is outlined as follows:

The 1^{st} time slot

The First slot is used for BS-to-Relay transmission, and NOMA superposes user transmit signals in its cell, then precodes and sends them to relays. The equation for calculating NOMA precoded signal that is transmitted by the lth BS, with the $s_{lk,n}$ denoting nth users signal in U_{lk} obtained from [4] and is expressed as:

$$\mathbf{x}^{l} = \sum_{k=1}^{K} \mathbf{w}_{lk} \sum_{n=1}^{\mathcal{U}_{lk}} \sqrt{p_{lk,n}} s_{lk,n} \triangleq \sum_{k=1}^{K} \mathbf{w}_{lk} x_{lk}. \quad (7.5)$$

Here x_{lk} is the signal transmitted for the cluster $\mathcal{U}_{jk}$ to relay R_{lk} and the precoder $\mathbf{w}_{lk} \in \mathbb{C}^{N\times 1}$ is based on the channel estimate of $\mathbf{H}$ with unit norm due to power constraints. The received signal from the BS l to the kth relay in lth cell is given as

$$y_{R_{lk}} = \underbrace{\sum_{(l',k')} (\mathbf{h}_{lk}^{l'})^{T} \mathbf{w}_{l'k'} x_{l'k'}}_{\tilde{y}_{R_{lk}}} + z_{R_{lk}}. \quad (7.6)$$

The scalar $z_{R_{lk}}$ is the AWGN at the relay R_{lk}.

The 2^{nd} time slot

The second slot is used for Relay-BS transmission after the signal received to Relay k. The received signal is transmitted again after the amplification of the signal by AF relay to the users of the cluster $\mathcal{U}_{lk}$. The transmit relay power is limited to q_{lk}, by $\overline{\mu}_{lk}$ which is the amplification factor. The $\overline{\mu}_{lk}$ cited from [4] is given as:

$$\overline{\mu}_{lk}^2 \mathbb{E}(|y_{R_{lk}}|^2) = q_{lk} \implies \overline{\mu}_{lk} = \sqrt{\frac{q_{lk}}{\left(\sum_{(l',k')} \kappa_{l'k',lk} p_{l'k'} + \sum_{l'=1}^{L} \rho_{l'k,lk} p_{l'k} + 1\right)}}. \tag{7.7}$$

where $p_{l'k'}$ is sum of the user transmit power from the cluster $\mathcal{U}_{l'k'}$. And the variables $\kappa_{l'k',lk}, \rho_{l'k,lk}$ are provided in the appendix A of [4]. The nth user's received NOMA signal is given by equation (11) in [4] is:

$$\begin{aligned}
y_{lk,n} &= f_{lk,n} g_{lk,n}^{lk} \mu_{lk} \mathbf{h}_{lk}^{l^T} \mathbf{w}_{lk} \sqrt{p_{lk,n}} s_{lk,n} + f_{lk,n} g_{lk,n}^{lk} \mu_{lk} \mathbf{h}_{lk}^{l^T} \mathbf{w}_{lk} \sum_{n' \neq n}^{\mathcal{U}_{lk}} \sqrt{p_{lk,n'}} s_{lk,n'} \\
&+ \sum_{l' \neq l}^{L} f_{lk,n} g_{lk,n}^{lk} \mu_{lk} \mathbf{h}_{lk}^{l'^T} \mathbf{w}_{l'k} x_{l',k'} + \sum_{l'=1}^{L} \sum_{k' \neq k}^{K} f_{lk,n} g_{lk,n}^{lk} \mu_{lk} \mathbf{h}_{lk}^{l'^T} \mathbf{w}_{l'k'} x_{l'k'} + \sum_{l' \neq l}^{L} f_{k,n} g_{lk,n}^{l'k} \mu_{l'k} \tilde{y}_{R_{l'k}} \\
&+ \sum_{l'=1}^{L} \sum_{k' \neq k}^{K} f_{lk,n} g_{lk,n}^{l'k'} \mu_{l'k'} \tilde{y}_{R_{l'k'}} + \sum_{l'=1}^{L} \sum_{k'=1}^{K} f_{lk,n} g_{lk,n}^{l'k'} \mu_{l'k'} z_{R_{l'k'}} + f_{k,n} z_{k,n}.
\end{aligned} \tag{7.8}$$

The explanation of the equation is provided in the system model section of the paper [4]. Relay-associated users execute SIC to minimise inter-relay interference. We suppose that these users are sorted by decreasing path loss. After cancelling IRI from users connected with the nth relay, using the SIC method [45], the first n-1 users' signals are treated as intrinsic intra-cluster interference [45]. $\mathbb{E}\left[\mathbf{h}_k^T \mathbf{w}_k\right]$ and $\hat{g}_{k,k,n}$ are used by the user to execute SIC which is expresse in the below equation:

$$f_{lk,n} g_{lk,n}^{lk} \mu_{lk} \mathbf{h}_{lk}^{l^T} \mathbf{w}_{lk} \sum_{n'=1}^{n-1} \sqrt{p_{lk,n'}} s_{lk,n'} + \sum_{n'=n+1}^{\mathcal{U}_{lk}} \mu_{lk} \left[f_{lk,n} g_{lk,n}^{lk} \mathbf{h}_{lk}^{l^T} \mathbf{w}_{lk} - f_{lk,n} \widehat{g}_{lk,n}^{lk} \mathbb{E}\left[\mathbf{h}_{lk}^{l^T} \mathbf{w}_{lk}\right]\right] \sqrt{p_{lk,n'}} s_{lk,n'}.$$

Equation (12) from [4] is used to calculate the received NOMA signal at the n user of $\mathcal{U}_{lk}$ after the SIC and is stated as:

$$\bar{y}_{lk,n} = f_{lk,n} g_{lk,n}^{lk} \mu_{lk} \mathbf{h}_{lk}^{lT} \mathbf{w}_{lk} \sqrt{p_{lk,n}} s_{lk,n} + f_{lk,n} g_{lk,n}^{lk} \mu_{lk} \mathbf{h}_{lk}^{lT} \mathbf{w}_{lk} \sum_{n'=1}^{n-1} \sqrt{p_{lk,n'}} s_{lk,n'} \tag{7.9}$$
$$+ \sum_{n'=n+1}^{\mathcal{U}_{lk}} \mu_{lk} \left[f_{lk,n} g_{lk,n}^{lk} \mathbf{h}_{lk}^{lT} \mathbf{w}_{lk} - f_{lk,n} \widehat{g}_{lk,n}^{lk} \mathbb{E}\left[\mathbf{h}_{lk}^{lT} \mathbf{w}_{lk} \right] \right] \sqrt{p_{lk,n'}} s_{lk,n'}$$
$$+ \sum_{l' \neq l}^{L} f_{lk,n} g_{lk,n}^{lk} \mu_{lk} \mathbf{h}_{lk}^{l'T} \mathbf{w}_{l'k} x_{l',k'} + \sum_{l'=1}^{L} \sum_{k' \neq k}^{K} f_{lk,n} g_{lk,n}^{lk} \mu_{lk} \mathbf{h}_{lk}^{l'T} \mathbf{w}_{l'k'} x_{l'k'} + \sum_{l' \neq l}^{L} f_{k,n} g_{lk,n}^{l'k} \mu_{l'k} \tilde{y}_{R_{l'k}}$$
$$+ \sum_{l'=1}^{L} \sum_{k' \neq k}^{K} f_{lk,n} g_{lk,n}^{l'k'} \mu_{l'k'} \tilde{y}_{R_{l'k'}} + \sum_{l'=1}^{L} \sum_{k'=1}^{K} f_{lk,n} g_{lk,n}^{l'k'} \mu_{l'k'} z_{R_{l'k'}} + f_{k,n} z_{k,n}.$$

BS combiners

In this work, we analyze the system performance for three combining schemes namely MRC, ZF and MMSE. The MMSE and ZF combiners are designed to cancel the inter- and intra-cell interference experienced by the user. These techniques, however, cannot cancel the Relay to User interference, of the system. The ZF and MMSE schemes, by utilizing the statistics and the channel realizations of the relay transmitters, mitigates the intra-cell, inter-cell of the relay interference. The ZF and MMSE combiners are given as

$$\overline{\mathbf{W}}_l = \begin{cases} \mathbf{H}_l, & \text{for MRC} \\ \mathbf{H}_l \left[\mathbf{H}_l^H \mathbf{H}_j \right]^{-1}, & \text{for ZF} \\ \left[\sum\limits_{l'=1}^{L} \mathbf{H}_{l'} \overline{\mathbf{P}}_{l'}^{cd} \mathbf{H}_{l'}^H + \mathbf{I}_M \right]^{-1} \mathbf{H}_j \overline{\mathbf{P}}_j, & \text{for MMSE.} \end{cases} \tag{7.10}$$

Here $\overline{\mathbf{W}^l} = [\mathbf{w}_{l1}, \cdots, \mathbf{w}_{lK}] \in \mathbb{C}^{N \times K}$ denotes the set of combiners used by lth BS for the relays in the lth cell. The matrices $\mathbf{H}_l = [\mathbf{h}_{l1}^l, \cdots, \mathbf{h}_{lK}^l] \in \mathbb{C}^{N \times K}$ denote the set of channels from K relays in lth cell to the lth BS respectively. Further $\mathbf{P}_{l'} = \text{diag}(p_{l'1}, \cdots, p_{l'K}) \in \mathbb{R}_+^{K \times K}$ where $p_{l'k} = \sum\limits_{n'=1}^{\mathcal{U}_{l'k}} p_{l'k,n'}$contains the sum of transmit powers of the users of cluster $\mathcal{U}_{l'k}$.

7.2 Achievable Spectral Efficiency Expression

Ergodic Spectral Efficiency

From the equation (12) of the base paper [4], The Ergodic sum SE of the system using Genie-bound with finite BS antennas of the system model, is given as

$$R_{\text{sum}}^{\text{e}} = \sum_{(l,k)} \sum_{n=1}^{\mathcal{U}_{lk}} \mathbb{E}\left[\frac{1}{2}\left(1-\frac{\tau}{\tau_c}\right)\log_2\left(1+\frac{\overline{\Delta}_{lk,n}^{i}}{\overline{\Omega}_{lk,n}^{i}}\right)\right], \text{ where } \overline{\Omega}_{lk,n}^{i} = \sum_{m=1}^{5} \hat{I}_{lk,n}^{(m)} + 1, \tag{7.11}$$

$$\hat{\Delta}_{lk,n} = \hat{A}_{lk,n}\overline{\mu}_{lk}^{2} p_{lk,n}, \ \hat{I}_{lk,n}^{(1)} = \sum_{n'=1}^{n-1} \hat{C}_{lk,n}^{(1)} \overline{\mu}_{lk}^{2} p_{lk,n'}, \ \hat{I}_{lk,n}^{(2)} = \sum_{n'=n+1}^{\mathcal{U}_{lk}} \hat{C}_{lk,n}^{(2)} p_{lk,n'} \overline{\mu}_{lk}^{2},$$

$$\hat{I}_{lk,n}^{(3)} = \sum_{(l',k')\neq(l,k)} \sum_{n'=1}^{\mathcal{U}_{l'k'}} \hat{C}_{l'k',lk,n}^{(3)} p_{l'k',n'} \overline{\mu}_{l'k'}^{2} \ \hat{I}_{lk,n}^{(4)} = \sum_{(l',k')} \sum_{(l'',k'')\neq(l',k')} \sum_{n'=1}^{\mathcal{U}_{l''k''}} C_{\substack{l''k'',l'k',\\ lk,n}}^{(4)} p_{l''k'',n'} \overline{\mu}_{l'k'}^{2}$$

$$\hat{I}_{k,n}^{(5)} = \sum_{l'=1}^{L} \sum_{k'=1}^{K} \hat{C}_{l'k',lk,n}^{(5)} \overline{\mu}_{l'k'}^{2}, \text{ and } \overline{\mu}_{lk} = \sqrt{\frac{q_{lk}}{\left(\sum\limits_{(l',k')} \kappa_{l'k',lk} p_{l'k'} + \sum\limits_{l'=1}^{L} {}_{l'k,lk} p_{l'k} + 1\right)}}.$$

The terms $\hat{A}_{lk,n}$, $\hat{C}_{lk,n}^{(1)}$, $\hat{C}_{lk,n}^{(2)}$, $\hat{C}_{l'k',lk,n}^{(3)}$, $\hat{C}_{\substack{l''k'',l'k',\\ lk,n}}^{(4)}$ and $\hat{C}_{l'k',lk,n}^{(5)}$ are functions of instantaneous channel realizations which are given as

$$\hat{A}_{lk,n} = |f_{lk,n} g_{lk,n}^{lk} \mathbf{h}_{lk}^{l^H} \mathbf{w}_{lk}|^2, \ \hat{C}_{lk,n}^{(1)} = |f_{lk,n} g_{lk,n}^{lk} \mathbf{h}_{lk}^{l^H} \mathbf{w}_{lk}|^2, \ \hat{C}_{lk,n}^{(2)} = |f_{lk,n} g_{lk,n}^{lk} \mathbf{h}_{lk}^{l^H} \mathbf{w}_{lk} - f_{lk,n} \hat{g}_{lk,n}^{lk} \mathbb{E}[\mathbf{h}_{lk}^{l^H} \mathbf{w}_{lk}$$

$$\hat{C}_{l'k',lk,n}^{(3)} = |f_{lk,n} g_{lk,n}^{l'k'} \mathbf{h}_{l'k'}^{l'^H} \mathbf{w}_{l'k'}|^2, \ \hat{C}_{\substack{l''k'',l'k',\\ lk,n}}^{(4)} = |f_{lk,n} g_{lk,n}^{l'k'} \mathbf{h}_{l'k'}^{l''^H} \mathbf{w}_{l''k''}|^2, \ \hat{C}_{l'k',lk,n}^{(5)} = |f_{lk,n} g_{lk,n}^{l'k'}|^2 \tag{7.12}$$

Closed-form Spectral Efficiency

From the derivation of the base paper [4], The UATF sum SE of the system with MRC precoder using Hard-bound technique with finite BS antennas of the system model, is given as:

$$R^c_{\text{sum}} = \sum_{(l,k)} \sum_{n=1}^{\mathcal{U}_{lk}} \frac{1}{2}\left(1 - \frac{\tau}{\tau_c}\right) \log_2\left(1 + \frac{\overline{\Delta}^c_{lk,n}}{\overline{\Omega}^c_{lk,n}}\right), \text{ where } \overline{\Omega}_{lk,n} = \sum_{m=0}^{7} \overline{I}^{(m)}_{lk,n} + 1, \quad (7.13)$$

$$\overline{\Delta}_{lk,n} = A_{lk,n}\overline{\mu}^2_{lk} p_{lk,n},\ \overline{I}^{(0)}_{lk,n} = C^{(0)}_{lk,n} p_{lk,n}\overline{\mu}^2_{lk},\ \overline{I}^{(1)}_{lk,n} = C^{(1)}_{lk,n}\sum_{n'=1}^{n-1}\overline{\mu}^2_{lk} p_{lk,n'},\ \overline{I}^{(2)}_{lk,n} = C^{(2)}_{lk,n}\sum_{n'=n+1}^{\mathcal{U}_{lk}} p_{lk,n'}\overline{\mu}^2_{lk},$$

$$\sum_{m=3}^{6}\overline{I}^{(m)}_{lk,n} = \sum_{(l',k')\neq(l,k)} \sum_{n'=1}^{\mathcal{U}_{l'k'}} C^{(3)}_{l'k',lk,n} p_{l'k',n'}\mu^2_{l'k'} + \sum_{(l',k')} \sum_{(l'',k'')\neq(l',k')} \sum_{n'=1}^{\mathcal{U}_{l''k''}} C^{(4)}_{\substack{l''k'',l'k',\\ lk,n}} p_{l''k'',n'}\overline{\mu}^2_{l'k'},$$

$$\overline{I}^{(7)}_{k,n} = \sum_{l'=1}^{L}\sum_{k'=1}^{K} C^{(5)}_{l'k',lk,n}\overline{\mu}^2_{l'k'}, \text{ and } \overline{\mu}_{lk} = \sqrt{\frac{q_{lk}}{\left(\sum\limits_{(l',k')} \kappa_{l'k',lk} p_{l'k'} + \sum\limits_{l'=1}^{L} \rho_{l'k,lk} + 1\right)}}. \quad (7.14)$$

Here $p_{lk} = \sum\limits_{n=1}^{\mathcal{U}_{lk}} p_{lk,n}$, $A_{lk,n} = \frac{\pi \upsilon^{lk}_{lk,n}\delta_{lk}}{4}\left[L_{1/2}\left(-\left|\overline{g}^{lk}_{lk,n}\right|^2 / \upsilon^{lk}_{lk,n}\right)\right]^2$, with $L_{\frac{1}{2}}(\cdot)$ being Laguerre polynomial [46]. The terms $C^{(0)}_{lk,n}$, $C^{(1)}_{lk,n}$, $C^{(2)}_{lk,n}$, $C^{(3)}_{l'k',lk,n}$, $C^{(4)}_{\substack{l''k'',l'k',\\ lk,n}}$, and $C^{(5)}_{l'k',lk,n}$ are functions are provided in the base paper's [4] Appendix A.

Chapter 8

Multi-Relay Multi-Cell mMIMO NOMA system : Sum SE maximization

8.1 Preface to WMMSE Algorithm

The WMMSE algorithm was first proposed in [44] to optimize the weighted sum rate to design a linear transmit filter. The fundamental concept behind the approach is to transform the objective problem into a WMMSE maximisation, in which the weights are modified in an iterative fashion. The algorithm covers a wide range of issues, including sum rate use. The following fundamental equation serves as the foundation for the WMMSE method.

$$\text{SINR} = \max_{u} \gamma = \max_{u} \frac{1}{e} - 1 \tag{8.1}$$

where u can be considered as the receiving beamformer to the signal received and also acts as an auxiliary variable in the algorithm, and γ as the corresponding SINR with respect to the decoded signal and e is the *Mean Square Error*(MSE) of the decoded signal. This vital equation can be best explained by a special case of uplink single cell multi user SISO system. The signal received by the only BS is:

$$y_k = h_k \sqrt{p_k} s_k + \sum_{\substack{i=1 \\ i \neq k}}^{N} h_i \sqrt{p_i} s_i + n \tag{8.2}$$

where h is the channel matrix from Cellular User(CU) to the BS, s and p are the transmit symbol and power from the CU and n is the white Gaussian noise. The SE of the kth CU is:

$$\begin{aligned} \mathrm{SE}_k^{CU} &= C \log\left(1 + \mathrm{SINR}_k^{CU}\right) = \frac{\Delta_k}{\Omega_k}, \text{ where} \\ \mathrm{SINR}_k^{CU} &= \frac{|h_k|^2 p_k}{\sum\limits_{\substack{i=1 \\ i \neq k}}^{N} |h_i|^2 p_i + 1}. \end{aligned} \tag{8.3}$$

The received signal y_k at the BS of the CU k is decoded with a received beamformer u_k. The decoded signal is given as;

$$\hat{s_k} = u_k y_k = u_k h_k \sqrt{p_k} s_k + \sum_{\substack{i=1 \\ i \neq k}}^{N} u_k h_i \sqrt{p_i} s_i + u_k n, \tag{8.4}$$

the SINR of CU k of the decoded signal is:

$$\gamma_k = \frac{|u_k h_k|^2 p_k}{\sum_{\substack{i=1 \\ i \neq k}}^{N} |u_k h_i|^2 p_i + |u_k|^2}, \tag{8.5}$$

and the MSE of the decoded signal with the transmit symbol is

$$e_k = E(|\hat{s_k} - s_k|^2) = |1 - u_k h_k \sqrt{p_k}|^2 + \sum_{\substack{i=1 \\ i \neq k}}^{N} |u_k h_i|^2 p_i + |u_k|^2 \tag{8.6}$$

Upon minimizing the MSE e_k with respect to u_k. The equivalence one can observe is that:

$$\mathrm{SINR}_k^{CU} = \frac{1}{\min\limits_{u_k} e_k} - 1 = \max_{u_k} \frac{1}{e_k} - 1 \tag{8.7}$$

which follows the equation mentioned earlier (8.1). In this part of the thesis, this equation serves as to manipulate the non-convex sum-rate utilization into pseudo concave functions for optimizations.

As the model becomes more complex and different, this relationship between MSE and SINR does not hold. Instead, it helps in understanding and providing relationship at a more general version which is provided in the below proposition.

Proposition 1: For any SINR with Δ and Ω as its numerator and denominator, can be reconstructed as,

$$\begin{aligned} \text{SINR} &= \max_{u} \frac{1}{e} - 1 \text{ where,} && (8.8)\\ e &= |1 - u\left(\sqrt{\Delta}\right)|^2 + |u|^2\,(\Omega) && (8.9) \end{aligned}$$

Reason: If we look carefully the e_k variable in (8.6), it can be interpreted as,

$$e_k = |1 - u_k \underbrace{h_k\sqrt{p_k}}_{\sqrt{\Delta_k}}|^2 + |u_k|^2 \underbrace{\left(\sum_{\substack{i=1\\ i\neq k}}^{N} |h_i|^2 p_i + 1\right)}_{\Omega_k} \tag{8.10}$$

For the sum-rate maximization, the WMMSE algorithm uses the auxiliary variables for optimization which is presented in the following proposition.

Proposition 2: For any SINR with Δ and Ω as its numerator and denominator, the SE can be reconstructed as:

$$\begin{aligned} \text{SE} &= \log\left(1 + \frac{\Delta}{\Omega}\right) = \underset{w,u}{\text{Maximize}}\ \log(w) - we \text{ where,} && (8.11)\\ e &= |1 - u\left(\sqrt{\Delta}\right)|^2 + |u|^2\,(\Omega)\ \text{ and} && (8.12)\\ w &= \frac{1}{e}. && (8.13) \end{aligned}$$

Proof : Using the proposition 1, the SE can be rewritten as:

$$\begin{aligned} \mathrm{SE} &= \log\left(1+\frac{\Delta}{\Omega}\right) = \log\left(\max_u \frac{1}{e}\right) \text{ where,} && (8.14) \\ e &= |1-u\left(\sqrt{\Delta}\right)|^2 + |u|^2\left(\Omega\right) && (8.15) \end{aligned}$$

As log is a monotonically increasing function, the log function can be brought inside the max function. Also using a epigrapgh trick on the e variable, the above SE is reconstructed as:

$$\begin{aligned} \mathrm{SE} &= \log\left(1+\frac{\Delta}{\Omega}\right) = \max_u \log\left(w\right) \text{ where,} && (8.16) \\ w &\leq \frac{1}{e}. && (8.17) \end{aligned}$$

As the reconstructed SE is concave function with respect to e variable. The strong Duality holds for the Lagrangian dual problem which is:

$$\mathcal{L}(w,\lambda) = \log\left(w\right) - \lambda\left(1-\frac{1}{we}\right) \tag{8.18}$$

Maximizing the dual problem with respect to w provides the optimal value of λ which is:

$$\lambda* = we \tag{8.19}$$

Therefore Applying the optimal $\lambda*$ and the optimal value of w which is $\frac{1}{e}$ to the dual problem the SE can be reconstructed as:

$$\begin{aligned} \mathrm{SE} &= \underset{w,u}{\text{Maximize}} \ \log(w) - we \text{ where,} && (8.20) \\ e &= |1-u\left(\sqrt{\Delta}\right)|^2 + |u|^2\left(\Omega\right) \text{ and} && (8.21) \\ w &= \frac{1}{e}. && (8.22) \end{aligned}$$

In the formulation a constant 1 is eliminated as it is the optimization and it has no responsibility in it. Hence the proof.

8.2 Optimization of SUM SE

The SUM SE is a characterization of the Channel capacity of the homogeneous system. Furthermore, maximizing the SUM SE results in the improved data transmission between the BS and the User. However, it is impossible to ensure that each user will have improved data transfer. In the multi-cell multi-relay NOMA system, the sum SE is maximized by altering the power allocated to each user from the BS. In the sections that follow, we will attempt to maximise the sum SE through power allocation for UATF 8.2.1 and Ergodic 8.2.2 system models by iteratively optimising using closed-form solutions. *Previous works on NOMA mMIMO systems have not worked on optimizing the sum SE using iterative closed form solutions.* This work can be extended to optimizing GEE and WSEE which is a better characterization and also acts as the trade-off between Channel Capacity and the power consumption in the homogeneous and heterogeneous network respectively.

8.2.1 UATF: SUM SE problem formulation and Optimization

The optimization of sum SE consists of series of sub-problems which are solved iteratively with closed-from solutions. The objective function of the sum SE optimzation is:

$$\mathbf{P1}: \quad \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}}\ R^{c}_{\text{sum}} \tag{8.23}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l\ \forall l,\ q_{lk} \leq Q_{lk}\ \forall l,\ \forall k \text{ and } p_{lk,n}, q_{lk} \geq 0 \tag{8.24}$$

where P_l is the maximum transmit power of the base station of lth cell, Q_{lk} is the maximum power transmitted from relay R_{lk} and $\mathbf{P} = [\mathcal{P}_1, .., \cdots, \mathcal{P}_L] \in \mathbb{R}^{L\times\mathcal{U}_{lk}K}$ and $\mathbf{Q} = [\mathbf{q}_1, \cdots, \mathbf{q}_L] \in \mathbb{R}^{K\times L}$ where $\mathcal{P}_l = [p_{l1,1}, \cdots, p_{lK,\mathcal{U}_{lk}}]$ and $\mathbf{q}_l = [q_{l1}, \cdots, q_{lK}]$ and also

$$R^c_{\text{sum}} = \frac{1}{2}\left(\frac{\tau_c - \tau_p}{\tau_c}\right)\sum_{(l,k)}\sum_{n=1}^{\mathcal{U}_{lk}} \log_2\left(1 + \frac{\overline{\Delta}^c_{lk,n}}{\overline{\Omega}^c_{lk,n}}\right). \tag{8.25}$$

The constant term $\frac{1}{2}\left(\frac{\tau_c-\tau_p}{\tau_c}\right)$ is eliminated as it is irrelevant in the optimization. Also. the $\log_2$ function is converted into natural algorithm (ln) which removes any constant appearing while performing differentiation. The optimization of sum SE comes under the ambit of multi-ratio(sum of log-ratios) fractional programming framework and are therefore challenging-convex problem.

The problem can be solved by novel optimization framework which approximates the non-convex functions into pseudo concave function at a point using the relationship between SINR and *mean square error*(MSE) where the algorithm knowingly termed as WMMSE algorithm. The algorithm converts non-convex hard problem like in **P1** and translates into pseudo concave function, which can be maximized using simple iterative closed form solutions.

For this optimization problem, instead of using $_{lk,n}$ as an optimization variable, we use $\bar{\mu}_{lk}$ as the replacement optimization variable, one can understand that this is an important replacement to attain closed form solutions from the solution for the following sub-problems. As $\hat{\mu}_{lk}$ depends on the power variables too, the constraints for $\bar{\mu}_{lk}$ are iteratively updated to satisfy all the constraints of the optimization. Therefore the problem **P1** is restructured as:

$$\mathbf{P2}: \quad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}} \sum_{(l,k)}\sum_{n=1}^{\mathcal{U}_{lk}} \log\left(1 + \frac{\overline{\Delta}^c_{lk,n}}{\overline{\Omega}^c_{lk,n}}\right) \tag{8.26}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l\,\forall l,\ \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')} \kappa_{l'k',lk} p_{l'k'} + \sum\limits_{l'=1}^{L} \rho_{l'k,lk} p_{l'k} + 1\right)}}. \tag{8.27}$$

With the restructured problem formulation **P2**, the optimization can proceed with modified WMMSE algorithm. Using Proposition 2, one can equivalently reconstruct the write the SE of the nth user of the cluster $\mathcal{U}_{lk}$ as:

$$\log\left(1+\mathrm{SINR}^{c}_{lk,n}\right) \;=\; \log\left(1+\frac{\overline{\Delta}^{c}_{lk,n}}{\overline{\Omega}^{c}_{lk,n}}\right) \;=\; \max_{u^{c}_{lk,n},w^{c}_{lk,n}} \frac{1}{w^{c}_{lk,n}} - w^{c}_{lk,n}e^{c}_{lk,n} \text{ where,} \tag{8.28}$$

$$e^{c}_{lk,n} = |1-u^{c}_{lk,n}\sqrt{\overline{\Delta}^{c}_{lk,n}}|^{2} + |u^{c}_{lk,n}|^{2}\overline{\Omega}^{c}_{lk,n} \tag{8.29}$$

The important aspect of $\mathrm{e}_{lk,n}$ is that the equation is concave in nature with respect to transmit power $\mathbf{P}$ and relay amplitude factor $\boldsymbol{\mu}$. Using the equation (8.28) to the Problem **P2**, the problem is reconstructed as:

$$\mathbf{P3}: \quad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}} \sum_{(l,k)}\sum_{n=1}^{\mathcal{U}_{lk}} \underset{u^{c}_{lk,n},w^{c}_{lk,n}}{\text{Maximize}} \log\left(w^{c}_{lk,n}\right) - w^{c}_{lk,n}e^{c}_{lk,n} \tag{8.30}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \le P_l\,\forall l,\ \bar{\mu}_{lk} \le \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum_{(l',k')}\kappa_{l'k',lk}p_{l'k'} + \sum_{l'=1}^{L}\rho_{l'k,lk}p_{l'k} + 1\right)}}.. \tag{8.31}$$

The auxiliary variables $u^{c}_{lk,n}, w^{c}_{lk,n}$ can be pushed out of the summation as the each $u^{c}_{lk,n}$ and $w^{c}_{lk,n}$ does not have any inter-dependencies for all l,k,n. Therefore the problem **P3** is reformulated as:

$$\mathbf{P4}: \quad \underset{\mathbf{P},\mathbf{U},\mathbf{W},\boldsymbol{\mu}}{\text{Maximize}} \sum_{(l,k)}\sum_{n=1}^{\mathcal{U}_{lk}} \log\left(w^{c}_{lk,n}\right) - w^{c}_{lk,n}e^{c}_{lk,n} \tag{8.32}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \le P_l\,\forall l,\ \bar{\mu}_{lk} \le \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum_{(l',k')}\kappa_{l'k',lk}p_{l'k'} + \sum_{l'=1}^{L}\rho_{l'k,lk}p_{l'k} + 1\right)}}. \tag{8.33}$$

Here the matrix $\mathbf{U} = [\mathbf{u}_1, \cdots, \mathbf{u}_L] \in \mathbb{R}^{L\times\mathcal{U}_{lk}K}$ where $\mathbf{u}_l = [u^{c}_{l1,1}, \cdots, u^{c}_{lK,\mathcal{U}_{lk}}]$ and $\mathbf{W} = [\mathbf{w}_1, \cdots, \mathbf{w}_L] \in \mathbb{R}^{L\times\mathcal{U}_{lk}K}$ where $\mathbf{w}_l = [w^{c}_{l1,1}, \cdots, w^{c}_{lK,\mathcal{U}_{lk}}]$. Now, expanding the auxiliary variable $e^{t}_{lk,n}$ by expanding the Numerator and Denominator of the SINR becomes:

$$\begin{aligned}
e^c_{lk,n} &= 1 + |u^c_{lk,n}|^2 A_{lk,n}\overline{\mu}^2_{lk} p_{lk,n} - 2(u^c_{lk,n}\sqrt{A_{lk,n}\overline{\mu}^2_{lk} p_{lk,n}}) \\
&+|u^c_{lk,n}|^2 \Bigg(C^{(0)}_{lk,n}\overline{\mu}^2_{lk} p_{lk,n} + \sum_{n'=1}^{n-1} C^{(1)}_{lk,n,n'}\overline{\mu}^2_{lk} p_{lk,n'} + \sum_{n'=n+1}^{\mathcal{U}_{lk}} C^{(2)}_{lk,n,n'} p_{lk,n'}\overline{\mu}^2_{lk} + \sum_{(l',k')\neq(l,k)} \sum_{n'=1}^{\mathcal{U}_{l'k'}} C^{(3)}_{l'k',lk,n} p_{l'k',n'}\mu^2_{l'k'} \\
&+ \sum_{(l',k')} \sum_{(l'',k'')\neq(l',k')} \sum_{n'=1}^{\mathcal{U}_{l''k''}} C^{(4)}_{\substack{l''k'',l'k',\\ lk,n}} p_{l''k'',n'}\overline{\mu}^2_{l'k'} + \sum_{l'=1}^{L}\sum_{k'=1}^{K} C^{(5)}_{l'k',lk,n}\overline{\mu}^2_{l'k'} + 1 \Bigg) \qquad (8.34)
\end{aligned}$$

Grouping the above equation transmit power-wise and multiplying with $w^t_{lk,n}$, the equation becomes:

$$w^c_{lk,n} e^c_{lk,n} = const(w^c_{lk,n}) + \sum_{(l',k')} \sum_{n'=1}^{\mathbb{U}_{lk}} \alpha^c_{l'k',n',lk,n} p_{l'k',n'} + \beta^c_{lk,n}\sqrt{p_{lk,n}}. \qquad (8.35)$$

All the constants including the amplitude factor $\boldsymbol{\mu}$ are combined and termed as α and β and are termed as:

$$\alpha^c_{l'k',n',lk,n} = w^c_{lk,n}|u^c_{lk,n}|^2 * \begin{cases}
A_{lk,n}\overline{\mu}^2_{lk} + C^{(0)}_{lk,n}\overline{\mu}^2_{lk} + \sum_{l''k''\neq lk} C^{(4)}_{\substack{lk,l''k'',\\ lk,n}}\overline{\mu}^2_{l''k''} & (l'k',n') = (l,k,n) \\
C^{(1)}_{lk,n,n'}\overline{\mu}^2_{lk} + \sum_{l''k''\neq lk} C^{(4)}_{\substack{lk,l''k'',\\ lk,n}}\overline{\mu}^2_{l''k''} & \substack{(l'k')=(l,k)\\ n'\leq n-1} \\
C^{(2)}_{lk,n,n'}\overline{\mu}^2_{lk} + \sum_{l''k''\neq lk} C^{(4)}_{\substack{lk,l''k'',\\ lk,n}}\overline{\mu}^2_{l''k''} & (l'k') = (l,k) \\
n' \geq n+1 & \\
C^{(3)}_{l'k',lk,n}\overline{\mu}^2_{lk} + \sum_{l''k''\neq l'k'} C^{(4)}_{\substack{l'k',l''k'',\\ lk,n}}\overline{\mu}^2_{l''k''} & (l'k') \neq (l,k)
\end{cases} \qquad (8.36)$$

$$\beta^c_{lk,n} = -2w^c_{lk,n}|u^c_{lk,n}|\sqrt{A_{lk,n}\overline{\mu}^2_{lk}} \qquad (8.37)$$

Therefore, the objective function of problem **P4** becomes:

$$\sum_{(l,k)} \sum_{n=1}^{\mathcal{U}_{lk}} \log\left(w^c_{lk,n}\right) - w^c_{lk,n} e^c_{lk,n} =$$
$$\sum_{(l,k)} \sum_{n=1}^{\mathcal{U}_{lk}} \left(const(\mathbf{W}) - \sum_{(l',k')} \sum_{n'=1}^{\mathbb{U}_{lk}} \alpha^c_{l'k',n',lk,n} p_{l'k',n'} + \beta^c_{lk,n} \sqrt{p_{lk,n}} \right) \tag{8.38}$$

Therefore the problem formulation **P4** is reconstructed as:

$$\mathbf{P5}: \quad \underset{\mathbf{P},\mathbf{U},\mathbf{W},\boldsymbol{\mu}}{\text{Maximize}} \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \left(const(\mathbf{W}) - \sum_{(l',k')} \sum_{n'=1}^{\mathbb{U}_{lk}} \alpha^c_{l'k',n',lk,n} p_{l'k',n'} + \beta^c_{lk,n} \sqrt{p_{lk,n}} \right) \tag{8.39}$$

$$s.t \quad \sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l \, \forall l, \;\; \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$
$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum_{(l',k')} \kappa_{l'k',lk} p_{l'k'} + \sum_{l'=1}^{L} \rho_{l'k,lk} p_{l'k} + 1 \right)}}. \tag{8.40}$$

Now, after formulating the objective function **P5**, all the optimizing variables $\mathbf{P}, \mathbf{U}, \mathbf{W}, \boldsymbol{\mu}$ are iteratively optimized. Firstly, the auxiliary variable $u^c_{lk,n}$ is optimized by first order differentiation of $e^c_{lk,n}$ and equating to zero which is:

$$u^c_{lk,n}{}^* = \frac{\sqrt{\overline{\Delta}^c_{lk,n}}}{\overline{\Omega}^c_{lk,n}} \tag{8.41}$$

The optimization of the variable $w^c_{lk,n}$ can be easily known as it acts as the auxiliary variable for $e_{lk,n}$. Therefore, the optimal $w^c_{lk,n}$ is:

$$w^{c*}_{lk,n} = \frac{1}{e^{c*}_{lk,n}} \tag{8.42}$$

The only optimizing variables left are the transmit power and relay amplitude factor. As mentioned earlier that $e^c_{lk,n}$ is concave in nature with respect to the optimizing variables. *The variable $e^c_{lk,n}$ couldn't have been concave in nature, if continued with* *variables* $\mathbf{P}$ *and* $\mathbf{Q}$. Therefore, with the auxiliary variables $w^{c*}_{lk,n}$ and $u^{c*}_{lk,n}$ acting as fixed point equations, the problem **P5** acts as a pseudo concave optimization with respect to the $\mathbf{P}$ and $\boldsymbol{\mu}$ and provides a optimal value. Therefore, the optimal $p_{lk,n}$ are obtained through first-order differentiation of the objective function **P5** and equating to zero which provides:

$$p^*_{lk,n} = \left\{ \frac{\boldsymbol{\beta}^c_{lk,n}}{\left(\sum\limits_{(l',k')} \sum\limits_{n'=1}^{\mathbb{U}_{lk}} \boldsymbol{\alpha}^c_{lk,n,l'k',n'} \right) + \lambda^*_l} \right\}^2 \tag{8.43}$$

where λ^*_l is an internal auxiliary variable which can optimized through bisection algorithm which helps to satisfy $\sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{lk}} p^*_{lk,n} \leq P_l$. Similarly, if the expanding the auxiliary variable $e^t_{lk,n}$ by expanding the Numerator and Denominator of the SINR and group with respect to the μ_{lk} and multiplying by $w^c_{lk,n}$ becomes:

$$w^c_{lk,n} e^c_{lk,n} = const(w^c_{lk,n}) + \sum_{(l',k')} \sum_{n'=1}^{\mathcal{U}_{l'k'}} \gamma^c_{l'k'n',lk} \mu^2_{l'k'} + \omega^t_{lk} \mu_{lk}. \tag{8.44}$$

All the constants including the transmit power $\mathbf{P}$ are combined and termed as ω and γ and are termed as:

$$\gamma^c_{l'k'n',lk} = \sum_{n=1}^{\mathcal{U}_{lk}} w^c_{lk,n} |u^c_{lk,n}|^2 * D_{l'k'n',lkn} \text{ where,} \tag{8.45}$$

$$D_{l'k'n',lkn} = \begin{cases} A_{lk,n} p_{lk,n} + C^{(0)}_{lk,n} p_{lk,n} + C^{(5)}_{lk,lk,n} + \sum\limits_{l''k'' \neq lk} \sum\limits_{n'=1}^{\mathcal{U}_{l''k''}} C^{(4)}_{\substack{l''k'',lk,\\ lk,n}} p_{l''k'',n'} & (l'k',n') = (l,k,n) \\ C^{(1)}_{lk,n,n'} p_{lk,n'} + C^{(5)}_{lk,lk,n} + \sum\limits_{l''k'' \neq lk} \sum\limits_{n'=1}^{\mathcal{U}_{l''k''}} C^{(4)}_{\substack{l''k'',lk,\\ lk,n}} p_{l''k'',n'} & \substack{(l'k')=(l,k)\\ n' \leq n-1} \\ C^{(2)}_{lk,n,n'} p_{lk,n'} + C^{(5)}_{lk,lk,n} + \sum\limits_{l''k'' \neq lk} \sum\limits_{n'=1}^{\mathcal{U}_{l''k''}} C^{(4)}_{\substack{l''k'',lk,\\ lk,n}} p_{l''k'',n'} & \substack{(l'k')=(l,k)\\ n' \geq n+1} \\ \sum\limits_{n'=1}^{\mathcal{U}_{l'k'}} C^{(3)}_{l'k',lk,n} p_{lk,n'} + C^{(5)}_{l'k',lk,n} + \sum\limits_{l''k'' \neq l'k'} \sum\limits_{n'=1}^{\mathcal{U}_{l''k''}} C^{(4)}_{\substack{l''k'',l'k',\\ lk,n}} p_{l''k'',n'} & (l'k') \neq (l,k) \end{cases}$$

$$\omega^c_{lk} = \sum_{n=1}^{\mathcal{U}_{lk}} -2 w^c_{lk,n} |u^c_{lk,n}| \sqrt{A_{lk,n} p_{lk,n}} \tag{8.46}$$

Therefore the problem formulation **P4** is reconstructed as:

$$\textbf{P6}: \quad \underset{\mathbf{P},\mathbf{U},\mathbf{W},\boldsymbol{\mu}}{\text{Maximize}} \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \left(const(\mathbf{W}) - \sum_{(l',k')} \gamma^{c}_{l'k',lk} \mu^{2}_{lk} + \omega^{t}_{lk} \mu_{lk} \right) \tag{8.47}$$

$$s.t \quad \sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l \, \forall l, \; \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')} \kappa_{l'k',lk} p_{l'k'} + \sum\limits_{l'=1}^{L} \rho_{l'k,lk} p_{l'k} + 1 \right)}}. \tag{8.48}$$

Using the problem **P6**, the $\mu_{lk,n}$ is optimized through first-order differentiation and equated to zero as:

$$\bar{\mu}^{*}_{lk} = \min \left\{ \frac{\boldsymbol{\omega}^{c}_{lk}}{\sum\limits_{(l',k')} \sum\limits_{n'=1}^{\mathcal{U}_{lk}} \boldsymbol{\gamma}^{c}_{lkn,l'k'}}, \hat{\mu_{lk}} \right\} \tag{8.49}$$

where,

$$\hat{\mu_{lk}} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')} \kappa_{l'k',lk} p^{*}_{l'k'} + \sum\limits_{l'=1}^{L} \rho_{l'k,lk} p^{*}_{l'k} + 1 \right)}}. \tag{8.50}$$

We observe that (8.43) and (8.49) are fixed-point equations, where the RHS expression depend themselves on $p_{lk,n}$ and μ_{lk}. The constraints $\hat{\mu_{lk}}$ are iteratively updated with the optimal power values. We therefore develop an iterative algorithm to solve the problem **P1** by starting from a feasible transmit power and relay transmit power and iteratively updating the auxiliary variables and transmit powers and relay amplitude factors with the solutions provided below. The resulting formal procedure to solve **P1** in (8.23) is provided in Algorithm 3.

Algorithm 3: sum SE maximization using Deterministic WMMSE approach

Input: Given $\epsilon > 0$, the max iterations N and max power P_l for UE U_{lk} and max power Q_{lk} for the relay. Calculate the initial values $p_{lk,n}, \mu_{lk}$ with random power allocation for all relay and users i.e., $\mathbf{p}_{lk,n}^{(1)} \sim \mathbb{U}[0, P_l^{max}]$ and $\mu_{lk}^{(1)} \sim \mathbb{U}[0, Q_{max}^{lk}]$

Output: $p_{lk,n}^{*}$ and μ_{lk}^{*}.

for $n \leftarrow 1$ **to** N **do**
- Given a feasible $p_{lk,n}^{(i)}$ and $\mu_{lk}^{(i)}$, update auxiliary variable $u_{lk,n}$ using (8.41)
- Update the auxiliary variable $w_{lk,n}$ using (8.42)
- Compute $p_{jk}^{(n+1)}$ using (8.43)
- Update the $\hat{\mu}_{lk}$ constraint variable using (8.50)
- Compute $\mu_{lk}^{(n+1)}$ using (8.49)
- Do until convergence If$\sum_{l=1}^{L} \sum_{l=1}^{K} \sum_{n=1}^{\mathbb{U}_{lk}} \left(p_{lk,n}^{(n+1)} - p_{lk,n}^{(n)} \right) < \epsilon$ break.

return $\mathbf{p}^*$,μ^*;

8.2.2 Ergodic: SUM SE problem formulation and optimization

In the previous section of this chapter, we solved a deterministic SE issue by making use of a UATF SE expression. Although calculating these expectations in closed-form for MRC is a simple process, doing so for the ZF and MMSE combining schemes is a non-trivial endeavour. Inorder to solve it, with using deterministic algorithm, the optimum power allocation scheme for ZF and MMSE schemes makes use of statistical averages (see equation(15) of [4]), which necessitates the gathering of a large number of random channel realisations prior to the updating of the transmit powers. The deterministic sum SE maximisation has a larger computational complexity, and as a result, it calls for a greater amount of memory, as well as longer time to store the samples. We are going to now recast the sum SE problem that was in P1 as a stochastic optimization problem and then optimize it using a low-complexity stochastic modified WMMSE framework. The optimization process uses both the stochastic sequential upper-bound minimization technique (SSUM) algorithm [47] and the weighted minimum mean squared error (WMMSE) algorithm. This will help us lower the memory required as well as the computing complexity. The summary of this section is that, after every realization, the surrogate function is formed with instantaneous sum SE and the optimizing variables are updated. The problem formulation follows a pattern quite similar to that of the section before it. The following is the objective function for the optimization of the ergodic sum SE:

$$\mathbf{P1}_{sto}: \qquad \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}}\ \mathbb{E}\left[g(\mathbf{P},\mathbf{Q},\mathcal{F})\right] \triangleq R^{e}_{sum} \tag{8.51}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l\,\forall l,\ q_{lk} \leq Q_{lk}\ \forall l,\ \forall k \text{ and } p_{lk,n}, q_{lk} \geq 0. \tag{8.52}$$

Here $g(\mathbf{P},\mathbf{Q},\mathcal{F})$ denotes the instantaneous sum SE of the system, with $\mathcal{F} = \{\mathbf{h}^{l'}_{lk}, g^{l'k'}_{lk,n}\}, \forall l, k, l', k', n$. It is defined as

$$g(\mathbf{P},\mathbf{Q},\mathcal{F}) \;=\; \frac{1}{2}\left(\frac{\tau_c-\tau_p}{\tau_c}\right)\sum_{(l,k)}\sum_{n=1}^{\mathcal{U}_{lk}} \log_2\left(1+\frac{\overline{\Delta}^{i}_{lk,n}(\mathcal{F})}{\overline{\Omega}^{i}_{lk,n}(\mathcal{F})}\right) \tag{8.53}$$

where $\overline{\Delta}^{i}_{lk,n}(\mathcal{F}), \overline{\Omega}^{i}_{lk,n}(\mathcal{F})$ are instantaneous SINR, Numerator of SINR, Denominator of the SINR of the nth user in $\mathcal{U}_{lk}$. The expectation is due to the random channels $\mathcal{F}$ generated. While reconstructing **P2** of the deterministic optimization it is mentioned that the optimization variable q_{lk} is replaced by $\bar{\mu}_{lk}$ inorder to attain the concavity of the sub-objective functions. And the constraint $\hat{\mu}_{lk}$ is iteratively updated to satisfy the constraint of the relay transmit power. Considering the aforementioned reasons, and also eliminating the constants and converting $\log_2$ into log function the problem is reconstructed as:

$$\mathbf{P2}_{sto}: \qquad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}}\ \mathbb{E}\left[\sum_{(l,k)}\sum_{n=1}^{\mathcal{U}_{lk}} \log_2\left(1+\frac{\overline{\Delta}^{i}_{lk,n}((\mathcal{F})}{\overline{\Omega}^{i}_{lk,n}(\mathcal{F})}\right)\right] \tag{8.54}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l\,\forall l,\ \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')}\kappa_{l'k',lk}p_{l'k'} + \sum\limits_{l'=1}^{L}\rho_{l'k,lk}p_{l'k}+1\right)}}. \tag{8.55}$$

To solve this stochastic non-convex optimization problem we propose a modified SSUM-WMMSE algorithm to solve $\mathbf{P2}_{sto}$. Using Proposition 2, we can reconstruct the instantaneous SE and therefore the problem $\mathbf{P2}_{sto}$ as:

$$\mathbf{P3}_{sto}: \quad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}} \lim_{T \Rightarrow \infty} \frac{1}{T} \sum_{t=1}^{T} \left[\underset{\mathbf{U}^t,\mathbf{W}^t}{\text{Maximize}} \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} w_{lk,n}^t e_{lk,n}^t - \log\big(w_{lk,n}^t\big) \right] \tag{8.56}$$

$$s.t \quad \sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l \, \forall l, \;\; \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\Big(\sum\limits_{(l',k')} \kappa_{l'k',lk} p_{l'k'} + \sum\limits_{l'=1}^{L} \rho_{l'k,lk} p_{l'k} + 1 \Big)}}. \tag{8.57}$$

where,

$$e_{lk,n}^t = |1 - u_{lk,n}^t \sqrt{\overline{\Delta}_{lk,n}^i(\mathcal{F}^t)}|^2 + |u_{lk,n}^t|^2 \, \overline{\Omega}_{lk,n}^i(\mathcal{F}^t) \tag{8.58}$$

As there are no inter-dependence in the auxiliary variables $w_{lk,n}^t$ and $u_{lk,n}^t$ they are grouped together into matrices $\mathbf{W}, \mathbf{U}$ respectively. As we know that $w_{lk,n}^t$ is the auxiliary variable to epigraph the $e_{lk,n}^i$ variable its optimal value is expressed as:

$$w_{lk,n}^{t*} = \frac{1}{e_{lk,n}^{t*}} \tag{8.59}$$

Minimizing $e_{lk,n}^t$ auxiliary variable under $u_{lk,n}^t$ through first order differentiation and equating to zero provides the optimal value of:

$$u_{lk,n}^{t*} = \frac{\sqrt{\overline{\Delta}_{lk,n}^i(\mathcal{F}^t)}}{\overline{\Omega}_{lk,n}^i(\mathcal{F}^t)} \tag{8.60}$$

After maximizing the inner optimization $(\mathbf{U}^t, \mathbf{W}^t)$ for each instantaneous SE, the problem $\mathbf{P3}_{sto}$ is deconstructed as:

$$\mathbf{P4}_{sto}: \quad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}} \lim_{T \Rightarrow \infty} \frac{1}{T} \sum_{t=1}^{T} \left[\sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \log\big(w^{t**}_{lk,n}\big) - w^{t**}_{lk,n} e^{t**}_{lk,n} \right] \tag{8.61}$$

$$s.t \quad \sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l \, \forall l, \; \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\Big(\sum_{(l',k')} \kappa_{l'k',lk} p_{l'k'} + \sum_{l'=1}^{L} \rho_{l'k,lk} p_{l'k} + 1 \Big)}}. \tag{8.62}$$

where,

$$e^{t**}_{lk,n} = |1 - u^{t**}_{lk,n} \sqrt{\overline{\Delta}^{i}_{lk,n}(\mathcal{F}^t)}|^2 + |u^{t**}_{lk,n}|^2 \, \overline{\Omega}^{i}_{lk,n}(\mathcal{F}^t) \tag{8.63}$$

and the auxiliary variables $u^{t**}_{lk,n}, w^{t**}_{lk,n}$ are the optimized values after inner optimization. The problem $\mathbf{P4}_{sto}$ is the outer optimization $(\mathbf{P}, \boldsymbol{\mu})$ which is a case of stochastic optimization problem and is solved by the SSUM algorithm. The algorithm is summarized below.

Summary of the SSUM Algorithm: In the paper [47], the proposed SSUM algorithm is that, at each iteration t, a new realization of channel $\mathcal{F}^{\sqcup} = \mathbf{H}^t$,$\mathbf{g}^t$ are obtained and the surrogate function is update to tth realization $(\mathbf{R}^{1:t}(\mathbf{P}, \boldsymbol{\mu}, \mathbf{P}^t, \boldsymbol{\mu}^t))$. Using this surrogate function, the optimization variables are updated which is expressed as:

$$\mathbf{P5}_{sto}: [\mathbf{P}^{t+1}, \boldsymbol{\mu}^{t+1}] \Leftarrow \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Argmax}} \frac{1}{t} \left[\mathbf{R}^{1:t}(\mathbf{P}, \boldsymbol{\mu}, \mathbf{P}^t, \boldsymbol{\mu}^t) \right] \tag{8.64}$$

$$s.t \quad \sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l \, \forall l, \; \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\Big(\sum_{(l',k')} \kappa_{l'k',lk} p_{l'k'} + \sum_{l'=1}^{L} \rho_{l'k,lk} p_{l'k} + 1 \Big)}}. \tag{8.65}$$

Formation of Surrogate function: At each iteration t, a new realization of channel $\mathcal{F}^{\sqcup} = \mathbf{H}^t$,$\mathbf{g}^t$ are obtained and the surrogate function upto $t-1$th realization ($\mathbf{R}^{1:t-1}(\mathbf{P},\boldsymbol{\mu})$) are updated with the convex approximate function ($\tilde{R}^i_{sum}(\mathbf{P},\boldsymbol{\mu},\mathcal{F}^t)$) of the instantaneous sum SE to surrogate function upto tth realization ($\mathbf{R}^{1:t}(\mathbf{P},\boldsymbol{\mu})$) through recursive surrogate function from [48], which also guarantee the convergence of the algorithm and expressed as:

$$\mathbf{R}^{1:t}(\mathbf{P},\boldsymbol{\mu},\mathbf{P}^t,\boldsymbol{\mu}^t) = \frac{1}{t}\left[\tilde{R}^i_{sum}(\mathbf{P},\boldsymbol{\mu},\mathcal{F}^t)\right] + \frac{t-1}{t}\left[\mathbf{R}^{1:t-1}(\mathbf{P},\boldsymbol{\mu},\mathbf{P}^{t-1},\boldsymbol{\mu}^{t-1})\right] \quad (8.66)$$

Coming back to the stochastic optimization of Ergodic sum SE, we need to find the surrogate function and the convex approximate at the t realization and is found using the following proposition.

Proposition 3: The objective function in $\mathbf{P4}_{sto}$ can act as a convex approximate for the recursive surrogate function. In mathematical way,

$$\tilde{R}^i_{sum}(\mathbf{P},\boldsymbol{\mu},\mathcal{F}^t) = \sum_{(l,k)}\sum_{n=1}^{\mathbb{U}_{lk}} \log\big(w^{t**}_{lk,n}\big) - w^{t**}_{lk,n}e^{t**}_{lk,n} \quad (8.67)$$

Proof: Expanding the auxiliary variable $e^{t**}_{lk,n}$ by expanding the Numerator and Denominator of the SINR becomes:

$$\begin{aligned} e^t_{lk,n} &= 1 + |u^{t**}_{lk,n}|^2 \hat{A}^t_{lk,n}\tilde{\mu}^2_{lk}p_{lk,n} - 2Re(u^{t**}_{lk,n}\sqrt{\hat{A}^t_{lk,n}\tilde{\mu}^2_{lk}p_{lk,n}}) \\ &+|u^{t**}_{lk,n}|^2\Bigg(\sum_{n'=1}^{n-1}\hat{C}^{(1)^t}_{lk,n,n'}\tilde{\mu}^2_{lk}p_{lk,n'} + \sum_{n'=n+1}^{\mathcal{U}_{lk}}\hat{C}^{(2)^t}_{lk,n,n'}p_{lk,n'}\tilde{\mu}^2_{lk} + \sum_{(l',k')\neq(l,k)}\sum_{n'=1}^{\mathcal{U}_{l'k'}}\hat{C}^{(3)^t}_{l'k',lk,n}p_{l'k',n'}\mu^2_{l'k'} \\ &\quad + \sum_{(l',k')}\sum_{(l'',k'')\neq(l',k')}\sum_{n'=1}^{\mathcal{U}_{l''k''}} C^{(4)^t}_{\substack{l''k'',l'k',\\ lk,n}} p_{l''k'',n'}\tilde{\mu}^2_{l'k'} + \sum_{l'=1}^{L}\sum_{k'=1}^{K}\hat{C}^{(5)^t}_{l'k',lk,n}\tilde{\mu}^2_{l'k'} + 1\Bigg) \end{aligned} \quad (8.68)$$

Grouping the above equation transmit power-wise and multiplying with $w^{t**}_{lk,n}$, the equation becomes:

$$w^{t**}_{lk,n}e^{t**}_{lk,n} = const(w^{t**}_{lk,n}) + \sum_{(l',k')}\sum_{n'=1}^{\mathbb{U}_{lk}}\alpha^t_{l'k',n',lk,n}p_{l'k',n'} + \beta^t_{lk,n}\sqrt{p_{lk,n}}. \quad (8.69)$$

All the constants including the amplitude factor $\boldsymbol{\mu}$ are combined and termed as α and β and are similar to (8.45). Therefore, the objective function of problem $\mathbf{P4}_{sto}$ becomes:

$$\sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \log\left(w_{lk,n}^{t**}\right) - w_{lk,n}^{t**} e_{lk,n}^{t**} =$$

$$\sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \left(const(\mathbf{W}) - \sum_{(l',k')} \sum_{n'=1}^{\mathbb{U}_{lk}} \alpha^t_{l'k',n',lk,n} p_{l'k',n'} + \beta^t_{lk,n} \sqrt{p_{lk,n}} \right) \tag{8.70}$$

If the equation is carefully noticed, the function is concave in transmit power variables $\mathbf{P}$ and also the $t+1$th channel realization will also have the exact objective sum SE function. *The aforementioned reasons has proved that the instantaneous this objective function can be considered as the instantaneous convex approximate for the surrogate function.*

Hence the proof.

By induction method one can understand that the surrogate function upto tth realization $\mathbf{R}^{1:t}(\mathbf{P}, \boldsymbol{\mu}, \mathbf{P}^t, \boldsymbol{\mu}^t)$ will have similar kind of constants with them as in α and β. Therefore the surrogate function upto tth realization is written as:

$$\mathbf{R}^{1:t}(\mathbf{P}, \boldsymbol{\mu}, \mathbf{P}^t, \boldsymbol{\mu}^t) = \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \left(const(\mathbf{W}) - \sum_{(l',k')} \sum_{n'=1}^{\mathbb{U}_{lk}} \boldsymbol{\alpha}^{1:t}_{l'k',n',lk,n} p_{l'k',n'} + \boldsymbol{\beta^{1:t}}_{lk,n} \sqrt{p_{lk,n}} \right) \tag{8.71}$$

In order to update the constants of surrogate function $\boldsymbol{\alpha}^{1:t}_{l'k',n',lk,n}$ and $\boldsymbol{\beta}^{lk,n}_{1:t}$for the tth iteration, they are updated as:

$$\begin{aligned} \boldsymbol{\alpha}^{1:t}_{l'k',n',lk,n} &= \frac{1}{t}\alpha^t_{l'k',n',lk,n} + \frac{t-1}{t}\boldsymbol{\alpha}^{1:t-1}_{l'k',n',lk,n} \\ \boldsymbol{\beta}^{1:t}_{lk,n} &= \frac{1}{t}\beta^t_{lk,n} + \frac{t-1}{t}\boldsymbol{\beta}^{1:t-1}_{lk,n} \end{aligned} \tag{8.72}$$

The surrogate functions can be viewed as a pseudo concave function, in hindsight, it can be said that WMMSE algorithm provides pseudo concave functions. At each iteration t, the optimal solution $p_{lk,n}^{t+1}$ and μ_{lk}^{t+1} is obtained.

For this optimization, it follows similar pattern of maximizing of deterministic sum SE but with the history of the previous terms. With the problem formulation $\mathbf{P5}_{sto}$ when update with the constants of surrogate function becomes

$$\mathbf{P6}_{sto} : [\mathbf{P}^{t+1}, \boldsymbol{\mu}^{t+1}] \Leftarrow \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Argmax}} \; -\frac{1}{t}\left[\sum_{(l,k)}\sum_{n=1}^{\mathbb{U}_{lk}}\left(\sum_{(l',k')}\sum_{n'=1}^{\mathbb{U}_{lk}}\boldsymbol{\alpha}_{l'k',n',lk,n}^{1:t}p_{l'k',n'} + \boldsymbol{\beta^{1:t}}_{lk,n}\sqrt{p_{lk,n}}\right)\right] \tag{8.73}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l \, \forall l, \;\; \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')}\kappa_{l'k',lk}p_{l'k'} + \sum\limits_{l'=1}^{L}\rho_{l'k,lk}p_{l'k} + 1\right)}}. \tag{8.74}$$

With the above objective function, the optimal $p_{lk,n}$ is provided through first order maximization of the objective function and equating to zero which is expressed as:

$$p_{lk,n}^{t+1} = \left\{\frac{\boldsymbol{\beta}_{lk,n}^{1:t}}{\left(\sum\limits_{(l',k')}\sum\limits_{n'=1}^{\mathbb{U}_{lk}}\boldsymbol{\alpha}_{lk,n,l'k',n'}^{1:t}\right) + \lambda_l^*}\right\}^2 \tag{8.75}$$

where λ_l^* found through bisection algorithm which satisfies $\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n}^{t+1} \leq P_l$.

Similarly, instead of grouping the variable $\mathbf{P}$ wise, if the equation (8.102) is grouped with respect to $\boldsymbol{\mu}$ wise and multiplied by $w_{lk,n}^{t**}$ the term becomes:

$$w_{lk,n}^{t**}e_{lk,n}^{t**} = const(w_{lk,n}^{t**}) + \sum_{(l',k')}\sum_{n'=1}^{\mathcal{U}_{l'k'}}\gamma_{l'k'n',lk}^{t}\mu_{lk}^2 + \omega_{lk}^{t}\mu_{lk}. \tag{8.76}$$

Using the similar steps of updating the surrogate function with the surrogate constants $\boldsymbol{\gamma}_{lk,l'k'}^{1:t}$ and $\boldsymbol{\omega}_{lk,l'k'}^{1:t}$ forming the objective function as in $\mathbf{P6}_{sto}$, as is constructed as:

$$\mathbf{P7}_{sto} : [\mathbf{P}^{t+1}, \boldsymbol{\mu}^{t+1}] \Leftarrow \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Argmax}} \; -\frac{1}{t}\left[\sum_{(l,k)}\sum_{n=1}^{\mathbb{U}_{lk}}\left(\sum_{(l',k')}\sum_{n'=1}^{\mathbb{U}_{lk}}\gamma^{1:t}_{l'k'n',lk}\overline{\mu}^{2}_{l'k'} + \boldsymbol{\gamma}^{\mathbf{1:t}}{}_{lk}\overline{\mu}_{lk}\right)\right] \tag{8.77}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l \,\forall l, \;\; \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')}\kappa_{l'k',lk}p_{l'k'} + \sum\limits_{l'=1}^{L}\rho_{l'k,lk}p_{l'k} + 1\right)}}. \tag{8.78}$$

With the above objective function, the μ_{lk} is optimized through first-order differentiation and equated to zero as:

$$\mu_{lk}^{t+1} = \min\left\{\frac{\boldsymbol{\omega}^{1:t}_{lk}}{\sum\limits_{(l',k')}\sum_{n'=1}^{\mathcal{U}_{lk}}\gamma^{1:t}_{lkn,l'k'}}, \hat{\mu_{lk}}\right\} \quad \text{where,}$$

$$\hat{\mu_{lk}} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')}\kappa_{l'k',lk}p^{t+1}_{l'k'} + \sum\limits_{l'=1}^{L}\rho_{l'k,lk}p^{t+1}_{l'k} + 1\right)}}. \tag{8.79}$$

Overall, the modified stochastic WMMSE algorithm is summarized in Algorithm 4.

8.3 sum SE maximization with optimal allocation of precoder

In the previous sections of the multi-cell multi-relay downlink NOMA systems, we solved the optimization of sum SE with optimal power allocation deterministically and stochastically with different precoding combiners i.e MRC, Ia-ZF, Ia-MMSE. The precoding combiners with this approach depends on the instantaneous channel realizations of $\mathbf{H}$ (7.10). And especially precoding combiners such as Ia-ZF or Ia-MMSE requires a considerably large computation complexity and also requires a

Algorithm 4: sum SE maximization using Stochastic WMMSE approach

Input: Given $\epsilon > 0$, the max iterations N and max power P_l for UE U_{lk} and max power Q_{lk} for the relay. Calculate the initial values $p_{lk,n}, \mu_{lk}$ with random power allocation for all relay and users i.e., $\mathbf{p}_{lk,n}^{(1)} \sim \mathbb{U}[0, P_l^{max}]$ and $\mu_{lk}^{(1)} \sim \mathbb{U}[0, Q_{lk}]$

Output: $p_{lk,n}^{*}$ and μ_{lk}^{*}.

for $i \leftarrow 1$ **to** N **do**

  Obtain the ith channel realization.

  for $j \leftarrow 1$ **to** N **do**

    Given a feasible $p_{lk,n}^{i}$ and μ_{lk}^{i}, update auxiliary variable $u_{lk,n}^{ij}$ using (8.60)

    Update the auxiliary variable $w_{lk,n}^{ij}$ using (8.59).

    Do until convergence

    If $\left(\hat{R}_{\text{sum}}^{i}(p,\mu) - \left\{ \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} w_{lk,n}^{ij+1} e_{lk,n}^{ij+1} - \log\left(w_{lk,n}^{ij+1} \right) \right\} \right) < \epsilon$

    break.

    Output: $w_{lk,n}^{i**}$ and w_{lk}^{i**}.

  Compute p_{jk}^{i+1} using (8.43)

  Update the $\hat{\mu}_{lk}$ constraint variable using (8.50)

  Compute μ_{lk}^{i+1} using (8.49)

  Update $\bar{\alpha}^{i}, \bar{\beta}^{i}, \bar{\gamma}^{i}, \bar{\omega}^{i}$ using equations

  Do until convergence If $\sum_{l=1}^{L} \sum_{l=1}^{K} \sum_{n=1}^{\mathbb{U}_{lk}} \left(p_{lk,n}^{(i)} - p_{lk,n}^{(i-1)} \right) < \epsilon$

  break.

return $\mathbf{p}^{*}, \mu^{*}$;

greater amount of memory. Therefore, the question arising that instead of using instantaneous percoder schemes, what if we observe deterministic precoder based on second-order statistics with samples of instantaneous channel realization. This section tries to understand the transmit signal representation, SE representation using the above transmitted signal precoders, the sum SE problem formulation and solving it all while optimally allocating precoders deterministacally with statistics from $\mathbf{H}$ and $\mathbf{g}$. *This is section is a initial study of precoder application and requires lot of further study into it. Previous works have never worked on stochastic sum SE precoder allocation.*

Data Transmission BS-Relay

In the data transmission mentioned in the above system model, it is considered that the precoder is unit norm and is used for decoding at relay w_{lk}. Also the transmitted signal is multiplied with the square root of power transmitter $p_{lk,n}$ of the users in the clusters (7.5).

In this section, we use the precoders which are for the the users $w_{lk,n}$ and we also eliminate the power variable $p_{lk,n}$ in the data transmission model by constraining the precoders with transmit powers. The precoded NOMA transmit signal with $s_{lk,n}$ being the signal of the nth user in cluster $\mathcal{U}_{lk}$ where broadcasted by the lth BS is

$$\mathbf{x}^l = \sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{jk}} \mathbf{w}_{lk,n} s_{lk,n} \tag{8.80}$$

Here $\mathbf{w}_{lk,n} \in \mathbb{C}^{N \times 1}$ is the precoder for the NOMA signal for the users in cluster $\mathcal{U}_{jk}$.

Ergodic SE

Using a simple traditional comparison method, it is possible to get the SINR and the SE of the user and of the system. Looking at the equations (7.5) and (8.80), the former can be converted to later by eliminating $\mathbf{w}_{lk}$ and replacing $\sqrt{p_{lk,n}}$ by $\mathbf{w}_{lk,n}$ which is given in the following equation:

$$\mathbf{x}^l = \sum_{k=1}^{K} \cancel{\mathbf{w}_{lk}} \sum_{n=1}^{\mathcal{U}_{jk}} \frac{\mathbf{w}_{lk,n}}{\cancel{\sqrt{p_{lk,n}}}} s_{lk,n} \tag{8.81}$$

Following the similar trend of eliminating and replacing the precoders in the (??), the Genie-Bounded Ergodic Sum SE of the user is obtained as:

$$\overline{R}^p_{lk,n\text{sum}} = \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \mathbb{E}\left[\frac{1}{2}\left(1 - \frac{\tau}{\tau_c}\right) \log_2\left(1 + \frac{\overline{\Delta}^p_{lk,n}}{\overline{\Omega}^p_{lk,n}}\right)\right], \text{ where } \overline{\Omega}^p_{lk,n} = \sum_{m=0}^{7} \overline{I}^{(m)^t}_{lk,n} + 1, \tag{8.82}$$

$$\overline{\Delta}^p_{lk,n} = |\overline{A}_{lk,n} \tilde{\mu}_{lk} \mathbf{w}_{lk,n}|^2, \ \overline{I}^{(1)}_{lk,n} = \sum_{n'=1}^{n-1} |\overline{C}^{(1)}_{lk,n,n'} \tilde{\mu}_{lk} \mathbf{w}_{lk,n'}|^2, \ \overline{I}^{(2)}_{lk,n} = \sum_{n'=n+1}^{\mathcal{U}_{lk}} |\overline{C}^{(2)}_{lk,n,n'} \mathbf{w}_{lk,n'} \tilde{\mu}_{lk}|^2,$$
$$\sum_{m=3}^{6} \overline{I}^{(m)}_{lk,n} = \sum_{(l',k') \neq (l,k)} \sum_{n'=1}^{\mathcal{U}_{l'k'}} |\overline{C}^{(3)}_{l'k',lk,n} \mathbf{w}_{l'k',n'} \mu_{l'k'}|^2 + \sum_{(l',k')} \sum_{(l'',k'') \neq (l',k')} \sum_{n'=1}^{\mathcal{U}_{l''k''}} |\overline{C}^{(4)}_{\substack{l''k'',l'k',\\ lk,n}} \mathbf{w}_{l''k'',n'} \tilde{\mu}_{l'k'}|^2,$$
$$\overline{I}^{(7)}_{k,n} = \sum_{l'=1}^{L} \sum_{k'=1}^{K} \overline{C}^{(5)}_{l'k',lk,n} \tilde{\mu}^2_{l'k'}, \text{ and}$$
$$\overline{\mu}_{lk} = \sqrt{\frac{q_{lk}}{\left(\sum\limits_{(l',k')} \kappa_{l'k',lk} \sum_{n'=1}^{\mathcal{U}_{l'k'}} |\mathbf{w}_{l'k',n'}|^2 + \sum\limits_{l'=1}^{L} \rho_{l'k,lk} \sum_{n'=1}^{\mathcal{U}_{l'k}} |\mathbf{w}_{l'k,n'}|^2 + 1\right)}}. \tag{8.83}$$

where the terms $\overline{A}_{lk,n}$, $\overline{C}^{(1)}_{lk,n}$, $\overline{C}^{(2)}_{lk,n}$, $\overline{C}^{(3)}_{l'k',lk,n}$, $\overline{C}^{(4)}_{\substack{l''k'',l'k',\\ lk,n}}$, and $\overline{C}^{(5)}_{l'k',lk,n}$ are functions of instantaneous channel realizations which are given as

$$\overline{A}_{lk,n} = f_{lk,n} g^{lk}_{lk,n} \mathbf{h}^{l^H}_{lk}, \ \overline{C}^{(1)}_{lk,n} = f_{lk,n} g^{lk}_{lk,n} \mathbf{h}^{l^H}_{lk}, \ \overline{C}^{(2)}_{lk,n} = f_{lk,n} g^{lk}_{lk,n} \mathbf{h}^{l^H}_{lk} - f_{lk,n} \hat{g}^{lk}_{lk,n} \hat{\mathbf{h}}^{l^H}_{lk},$$
$$\overline{C}^{(3)}_{l'k',lk,n} = f_{lk,n} g^{l'k'}_{lk,n} \mathbf{h}^{l'^H}_{l'k'}, \ \overline{C}^{(4)}_{\substack{l''k'',l'k',\\ lk,n}} = f_{lk,n} g^{l'k'}_{lk,n} \mathbf{h}^{l''^H}_{l'k'}, \ \overline{C}^{(5)}_{l'k',lk,n} = |f_{lk,n} g^{l'k'}_{lk,n}|^2 \tag{8.84}$$

Sum SE formulation and Optimization

In this section, we solve the stochastic SE optimization by optimizing the precoder allocation, following the similar trend of the previous section. We use SSUM and WMMSE algorithm for optimization where at every realization the instantaneous sum SE is learnt through recursive surrogate function (SSUM) and the optimizing variables along with auxiliary variables are updated regularly. Before formulating the objective function, we also know that, inorder to attain the closed form solution, we replace the transmit relay power q_{lk} to amplitude factor μ_{lk} as the optimization variable to attain iterative closed form solutions. We also know that the stochastic optimization is based on the samples of the channel realization and the practical mean. Using all the known factors, the objective function for the optimization of the Ergodic sum SE is constructed as:

$$\mathbf{P1}_{pre}: \quad \underset{\overline{\mathbf{W}},\boldsymbol{\mu}}{\text{Maximize}} \lim_{T\Rightarrow\infty} \frac{1}{T}\sum_{t=1}^{T}\left[\sum_{(l,k)}\sum_{n=1}^{\mathcal{U}_{lk}} \log_2\left(1+\frac{\overline{\Delta}^{p}_{lk,n}(\mathcal{F}^t)}{\overline{\Omega}^{p}_{lk,n}(\mathcal{F}^t)}\right)\right] \tag{8.85}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} |\mathbf{w}_{lk,n}|^2 \leq P_l \,\forall l,\ \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')} \kappa_{l'k',lk}\sum_{n'=1}^{\mathcal{U}_{l'k'}} |\mathbf{w}_{l'k',n'}|^2 + \sum\limits_{l'=1}^{L} \rho_{l'k,lk}\sum_{n'=1}^{\mathcal{U}_{l'k}} |\mathbf{w}_{l'k,n'}|^2 + 1\right)}}. \tag{8.86}$$

where $\mathcal{F} = \{\mathbf{h}^{l'}_{lk}, g^{l'k'}_{lk,n}\}, \forall l, k, l', k', n$. Using the methods followed to construct double optimization using auxiliary variables in $\mathbf{P4}_{sto}$ from $\mathbf{P3}_{sto}$, the above problem formulation $\mathbf{P1}_{pre}$ can reconstructed with auxiliary variables as:

$$\mathbf{P2}_{pre}: \quad \underset{\overline{\mathbf{W}},\boldsymbol{\mu}}{\text{Maximize}} \lim_{T \Rightarrow \infty} \frac{1}{T} \sum_{t=1}^{T} \left[\underset{\mathbf{U}^t,\mathbf{W}^t}{\text{Maximize}} \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} w_{lk,n}^t e_{lk,n}^t - \log\big(w_{lk,n}^t\big) \right] \tag{8.87}$$

$$s.t \quad \sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{lk}} |\mathbf{w}_{lk,n}|^2 \leq P_l \,\forall l,\ \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\Big(\sum\limits_{(l',k')} \kappa_{l'k',lk} \sum_{n'=1}^{\mathcal{U}_{l'k'}} |\mathbf{w}_{l'k',n'}|^2 + \sum\limits_{l'=1}^{L} \rho_{l'k,lk} \sum_{n'=1}^{\mathcal{U}_{l'k}} |\mathbf{w}_{l'k,n'}|^2 + 1 \Big)}}. \tag{8.88}$$

where the auxiliary variables $w_{lk,n}^t$ and $u_{lk,n}^t$ are grouped into matrix $\mathbf{W}$,$\mathbf{U}^t$ respectively. The auxiliary variable $e_{lk,n}$ is defined as:

$$e_{lk,n}^t = |1 - u_{lk,n}^t \sqrt{\overline{\Delta}_{lk,n}^p(\mathcal{F}^t)}|^2 + |u_{lk,n}^t|^2\, \overline{\Omega}_{lk,n}^p(\mathcal{F}^t), \forall\, l, k, n.$$

And the optimal values of $w_{lk,n}^t$ and $u_{lk,n}^t$ are given as:

$$u_{lk,n}^{t*} = \frac{\sqrt{\overline{\Delta}_{lk,n}^i(\mathcal{F}^t)}}{\overline{\Omega}_{lk,n}^i(\mathcal{F}^t)} \tag{8.89}$$

$$w_{lk,n}^{t*} = \frac{1}{e_{lk,n}^{t*}} \text{ where,} \tag{8.90}$$

$$e_{lk,n}^{t*} = |1 - u_{lk,n}^{t*} \sqrt{\overline{\Delta}_{lk,n}^p(\mathcal{F}^t)}|^2 + |u_{lk,n}^{t*}|^2\, \overline{\Omega}_{lk,n}^p(\mathcal{F}^t)$$

Once the inner optimization (i.e optimizing the matrices $\mathbf{W}^t, \mathbf{U}^t$) is completed. With the optimal auxiliary variables $(\mathbf{W}^{t**}, \mathbf{U}^{t**})$ the problem formulation $\mathbf{P2}_{pre}$ is deconstructed as:

$$\mathbf{P3}_{pre}: \quad \underset{\overline{\mathbf{W}},\boldsymbol{\mu}}{\text{Maximize}} \lim_{T\Rightarrow\infty}\frac{1}{T}\sum_{t=1}^{T}\left[\sum_{(l,k)}\sum_{n=1}^{\mathbb{U}_{lk}} w_{lk,n}^{t**}e_{lk,n}^{t**}-\log\left(w_{lk,n}^{t**}\right)\right] \tag{8.91}$$

$$s.t \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}}|\mathbf{w}_{lk,n}|^2 \leq P_l\,\forall l,\ \bar{\mu}_{lk}\leq\hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk}=\sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')}\kappa_{l'k',lk}\sum_{n'=1}^{\mathcal{U}_{l'k'}}|\mathbf{w}_{l'k',n'}|^2+\sum\limits_{l'=1}^{L}\rho_{l'k,lk}\sum_{n'=1}^{\mathcal{U}_{l'k}}|\mathbf{w}_{l'k,n'}|^2+1\right)}}. \tag{8.92}$$

where the variables $e_{lk,n}^{t**}$, $w_{lk,n}^{t**}$ and $u_{lk,n}^{t**}$ are the optimal variables after several iterations. As the above objective function (8.91) is pseudo concave function from the proposition 2. Therefore the problem $\mathbf{P3}_{pre}$ becomes stochastic convex optimization problem. Inorder to apply the SSUM algorithm, we require a convex approximation of the objective function of the tth realization which helps in forming the surrogate function. From the proposition 2, that:

$$\tilde{R}_{\text{sum}}(\overline{\mathbf{W}},\boldsymbol{\mu},\mathcal{F}^t)=\sum_{(l,k)}\sum_{n=1}^{\mathbb{U}_{lk}}\log\left(w_{lk,n}^{t**}\right)-w_{lk,n}^{t**}e_{lk,n}^{t**} \tag{8.93}$$

is a valid convex approximate of the objective function of $\mathbf{P2}_{pre}$ only at the t realization. Using the recursive surrogate function [], the surrogate function $\tilde{\mathbb{R}}_{\text{sum}}^{1:t}$ is expressed as:

$$\mathbf{R}^{1:t}(\bar{\mathbf{W}},\boldsymbol{\mu})=\frac{1}{t}\left[\mathbf{R}^{1:t-1}(\bar{\mathbf{W}},\boldsymbol{\mu})\right]+\frac{t-1}{t}\left[\tilde{R}_{\text{sum}}(\overline{\mathbf{W}},\boldsymbol{\mu},\mathcal{F}^t)\right] \tag{8.94}$$

To make the above equation simpler, we try to group the auxiliary variable $e_{lk,n}^{t**}$ present in the equation (8.96), based on percoder $\mathbf{w}_{lk,n}$ which is presented as:

$$w_{lk,n}^{t**}e_{lk,n}^{t**}=cons(w_{lk,n}^{t**})+\sum_{(l',k')}\sum_{n'=1}^{\mathbb{U}_{lk}}\mathbf{w}_{l'k',n'}^{H}\boldsymbol{\alpha}_{l'k',n',lk,n}^{t}\mathbf{w}_{l'k',n'}+\boldsymbol{\beta}_{lk,n}^{t}\mathbf{w}_{lk,n}. \tag{8.95}$$

All the constants including the amplitude factor $\boldsymbol{\mu}$ are combined and termed as $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$. With the above grouping the convex approximate function $\tilde{R}_{\text{sum}}(\overline{\mathbf{W}},\boldsymbol{\mu},\mathcal{F}^t)$

is written as:

$$\tilde{R}_{\text{sum}}(\overline{\mathbf{W}}, \boldsymbol{\mu}, \mathcal{F}^t) = \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \left(const(w) - \sum_{(l',k')} \sum_{n'=1}^{\mathbb{U}_{lk}} \mathbf{w}_{l'k',n'}^H \boldsymbol{\alpha}_{l'k',n',lk,n}^t \mathbf{w}_{l'k',n'} + \boldsymbol{\beta}_{lk,n}^t \mathbf{w}_{lk,n} \right) \tag{8.96}$$

As one can understand that for any t realization, the convex approximate tend to be the same. Therefore using the mathematical induction and keeping the amplitude factor $\boldsymbol{\mu}$ as fixed and using the grouped auxiliary variable, one can simplify the surrogate function to be:

$$\mathbf{R}^{1:t}(\bar{\mathbf{W}}, \boldsymbol{\mu}) = \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \left(const(w) - \sum_{(l',k')} \sum_{n'=1}^{\mathbb{U}_{lk}} \mathbf{w}_{l'k',n'}^H \check{\boldsymbol{\alpha}}_{l'k',n',lk,n}^{1:t} \mathbf{w}_{l'k',n'} + \check{\boldsymbol{\beta}}_{1:t}^{lk,n} \mathbf{w}_{lk,n} \right) \text{ where,} \tag{8.97}$$

$$\check{\boldsymbol{\alpha}}_{l'k',n',lk,n}^{1:t} = \frac{1}{t} \boldsymbol{\alpha}_{l'k',n',lk,n}^t + \frac{t-1}{t} \check{\boldsymbol{\alpha}}_{l'k',n',lk,n}^{1:t-1} \tag{8.98}$$

$$\check{\boldsymbol{\beta}}_{lk,n}^{1:t} = \frac{1}{t} \boldsymbol{\beta}_{lk,n}^t + \frac{t-1}{t} \check{\boldsymbol{\beta}}_{lk,n}^{1:t-1} \tag{8.99}$$

Using the construction of the objective function mentioned in $\mathbf{P6}_{sto}$, the problem formulation for the stochastic optimization with the surrogate function is constructed as:

$$\mathbf{P4}_{pre} : [\overline{\mathbf{W}}^{t+1}, \boldsymbol{\mu}^{t+1}] \Leftarrow \underset{\bar{\mathbf{W}}, \boldsymbol{\mu}}{\text{Argmax}} \ \frac{1}{t} \left[\mathbf{R}^{1:t-1}(\bar{\mathbf{W}}, \boldsymbol{\mu}) \right] + \frac{t-1}{t} \left[\tilde{R}_{\text{sum}}(\overline{\mathbf{W}}, \boldsymbol{\mu}, \mathcal{F}^t) \right] \triangleq \mathbf{R}^{1:t}(\bar{\mathbf{W}}, \boldsymbol{\mu}) \tag{8.100}$$

$$s.t \quad \sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{lk}} |\mathbf{w}_{lk,n}|^2 \leq P_l \, \forall l, \ \bar{\mu}_{lk} \leq \hat{\mu}_{lk} \quad \text{where,}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')} \kappa_{l'k',lk} \sum_{n'=1}^{\mathcal{U}_{l'k'}} |\mathbf{w}_{l'k',n'}|^2 + \sum\limits_{l'=1}^{L} \rho_{l'k,lk} \sum_{n'=1}^{\mathcal{U}_{l'k}} |\mathbf{w}_{l'k,n'}|^2 + 1 \right)}}.$$

the optimized $\mathbf{w}_{lk,n}$ is provided through the first order differentiation of the objective function of $\mathbf{P4}_{pre}$, which surprisingly follows the similar pattern to that of optimizing the Ergodic sum SE and is given below:

$$\mathbf{w}_{lk,n}^{t+1} = \left\{ \left(\sum_{(l',k')} \sum_{n'=1}^{\mathbb{U}_{lk}} \check{\boldsymbol{\alpha}}_{lk,n,l'k',n'}^{1:t} \right) + \lambda_l^* I_N \right\}^{-1} \left\{ \left(\boldsymbol{\beta}_{lk,n}^{1:t} \right) \right\} \tag{8.101}$$

where λ_l^* found through bisection algorithm which satisfies $\sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{lk}} |\mathbf{w}_{lk,n}^{t+1}|^2 \leq P_l$.

Similarly, instead of grouping the variable $\overline{\mathbf{W}}$ wise, if the equation (8.95) is grouped with respect to $\boldsymbol{\mu}$ wise, the term becomes:

$$w_{lk,n}^{t**} e_{lk,n}^{t**} = const(w) + \sum_{(l',k')} \gamma_{l'k',lk}^{t} \mu_{lk}^2 + \omega_{lk}^{t} \mu_{lk}. \tag{8.102}$$

Keeping the updated precoder fixed and following the similar steps from problem formulation $\mathbf{P3}_{pre}$ to $\mathbf{P4}_{pre}$ and by forming the simpler surrogate function (8.97) and updation of the surrogate variables (8.98), the $\mu_{lk,n}$ is optimized through first-order differentiation of new surrogate variable keeping precoder constant and equating to zeros is expressed as:

$$\overline{\mu}_{lk}^{t+1} = \min \left\{ \frac{\omega_{lk}^{1:t}}{\sum\limits_{(l',k')} \gamma_{lk,l'k'}^{1:t}}, \hat{\mu_{lk}} \right\} \text{ where,} \tag{8.103}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{(l',k')} \kappa_{l'k',lk} \sum_{n'=1}^{\mathcal{U}_{l'k'}} |\mathbf{w}_{l'k',n'}|^2 + \sum\limits_{l'=1}^{L} \rho_{l'k,lk} \sum_{n'=1}^{\mathcal{U}_{l'k}} |\mathbf{w}_{l'k,n'}|^2 + 1 \right)}}. \tag{8.104}$$

The overall Algorithm, the modified stochastic successive convex approximation is summarized in Algorithm 5.

Algorithm 5: sum SE maximization with Precoder allocation

Input: Given $\epsilon > 0$, the max iterations N and max power P_l for UE U_{lk} and max power Q_{lk} for the relay. Calculate the initial values $\mathbf{w}_{lk,n}, \mu_{lk}$ with random precoder allocation for all relay and users i.e., $|\mathbf{w}_{lk,n}^{(1)}|^2 \sim \mathbb{U}[0, P_l^{max}]$ and $\mu_{lk}^{(1)} \sim \mathbb{U}[0, Q_{lk}]$

Output: $\mathbf{w}_{lk,n}^*$ and μ_{lk}^*.

for $i \leftarrow 1$ **to** N **do**

  Obtain the ith channel realization.

  for $j \leftarrow 1$ **to** N **do**

    Given a feasible $\mathbf{w}_{lk,n}^i$ and μ_{lk}^i, update auxiliary variable $u_{lk,n}^{i^j}$ using (**??**)

    Update the auxiliary variable $w_{lk,n}^{i^j}$ using (8.90)

    Do until convergence

    If$\left\{ \sum_{(l,k)} \sum_{n=1}^{\mathbb{U}_{lk}} \left(w_{lk,n}^{i^j} e_{lk,n}^{i^j} - \log\left(w_{lk,n}^{i^j} \right) \right) - \left(w_{lk,n}^{i^{j+1}} e_{lk,n}^{i^{j+1}} - \log\left(w_{lk,n}^{i^{j+1}} \right) \right) \right\} < \epsilon$

    break.

    Output: $w_{lk,n}^{i**}$ and u_{lk}^{i**}.

  Compute p_{jk}^{i+1} using (8.101)

  Update the $\hat{\mu}_{lk}$ constraint variable using (8.104)

  Compute μ_{lk}^{i+1} using (8.103)

  Update $\bar{\alpha}^i, \bar{\beta}^i, \bar{\gamma}^i, \bar{\omega}^i$ using equations (8.98)

  Do until convergenceIf$\sum_{l=1}^{L} \sum_{l=1}^{K} \sum_{n=1}^{\mathbb{U}_{lk}} \left(|\mathbf{w}_{lk,n}^{(i)}|^2 - |\mathbf{w}_{lk,n}^{(i-1)}|^2 \right) < \epsilon$

  break.

return $\mathbf{w}^*$,μ^*

Chapter 9

Multi-Relay Multi-Cell mMIMO NOMA system : Simulation Results and Conclusion

9.1 Simulation results

In this part, we assess the multi-cell multi-relay mMIMO NOMA Downlink system's optimizations' evaluation, where the BS and users estimate the local CSI and the users deploy imperfect SIC. Consider a 20-MHz system with 4 cell, 60 BS antennas, 3 relays, 12 users with each relay of 4 users and remains throughout the simulation unless its mentioned. Also each coherence period has 200 symbols and a K-symbol pilot transmission interval with $p_p = 20dBm$. The system model model is simulated such that the BS is at the center and the relays are uniformly distributed around 200m circle and the users are randomly distributed around 100m circle. We model the large scale fading coefficient from the l'th BS to k th relay in lth cell and k' th relay in l'th cell to nth user in cluster $\mathcal{U}_{lk}$ as $\beta_{lk}^{l\prime}[\text{dB}] = \Upsilon - 10\alpha \log_{10}\left(d_{lk}^{l''}\right) + F_{lk}^{l''}$ and $\beta_{lk,n}^{l'k'}[\text{dB}] = \Upsilon - 10\alpha \log_{10}\left(d_{lk,n}^{l'k'}\right) + F_{lk,n}^{l'k'}$. Here Υ denotes the median channel gain at a reference distance of 30.18, α is the path loss exponent of 2.6, $d_{\underset{lk}{l'}}$ and $d_{lk,n}^{l'k'}$ is the separation distance between the BS-and the relay and relay and user in metres, and the scalars $F_{lk}^{l''}$ and $F_{lk,n}^{l'k'}$ are the shadow fading terms which models the log-normal random variations. The Rician factors of the random channels are modelled as $K_{lk}^{l'}$ and $K_{lk,n}^{l'k'}$ can be calculated as $K_{lk}^{l'}[\mathrm{dB}] = 13 - 0.03 d_{lk}^{l'}$ and $K_{lk,n}^{l'k'}[\mathrm{dB}] = 13 - 0.03 d_{lk,n}^{l'k'}$ respectively.The other parameters is the assumption of ULA at the base station and the ASD of 10^o.

9.1.1 Validation of optimization techniques

As the work of the paper is of optimization techniques, it is necessary to understand that the optimal value has been achieved. The ϵ is set to be as 0.001 and the framework converges smoothly. In the Fig.9.1, we plot the sum SE (UATF and Ergodic) of users vs the maximum transmit power allocation to the BS P_T and to the relay Q_T. The plots are compared between Optimized Power Allocation (OPA) and Full Power Allocation (FPA) where the users are allocated power equally to maximum. $P_T/(K * \mathcal{U}_{lk})$. The Fig.9.1a, provides the OPA plot of the converged value of deterministic WMMSE optimization of the UATF sum SE. Similarly, the Figs.9.1b-d provides the FPA plot of the converged value of stochastic WMMSE optimization of the Ergodic sum SE with the precoders of MRC, ZF,MMSE respectively. At higher powers, you can clearly see that MRC precoders can maximize the sum SE much better that the MMSE precoder, which should be true because the MMSE precoder is built based on maximizing the SINR at FPA. Thus strengthening the work. The Fig.9.1e, provides the OPA plot of the converged value of stochastic WMMSE optimization of the Ergodic sum SE with optimal precoder allocation. The FPA of this plot is the random precoder which satisfies the power constraint and is optimized. Inorder to plot the OPA, the paper has taken different channel realizations than the one used for optimizing, thus proving that the optimized precoder is solely based on second-order statistics of the channel. The paper has taken 10 User Setups and have averaged it for the uniformity between optimization techniques.

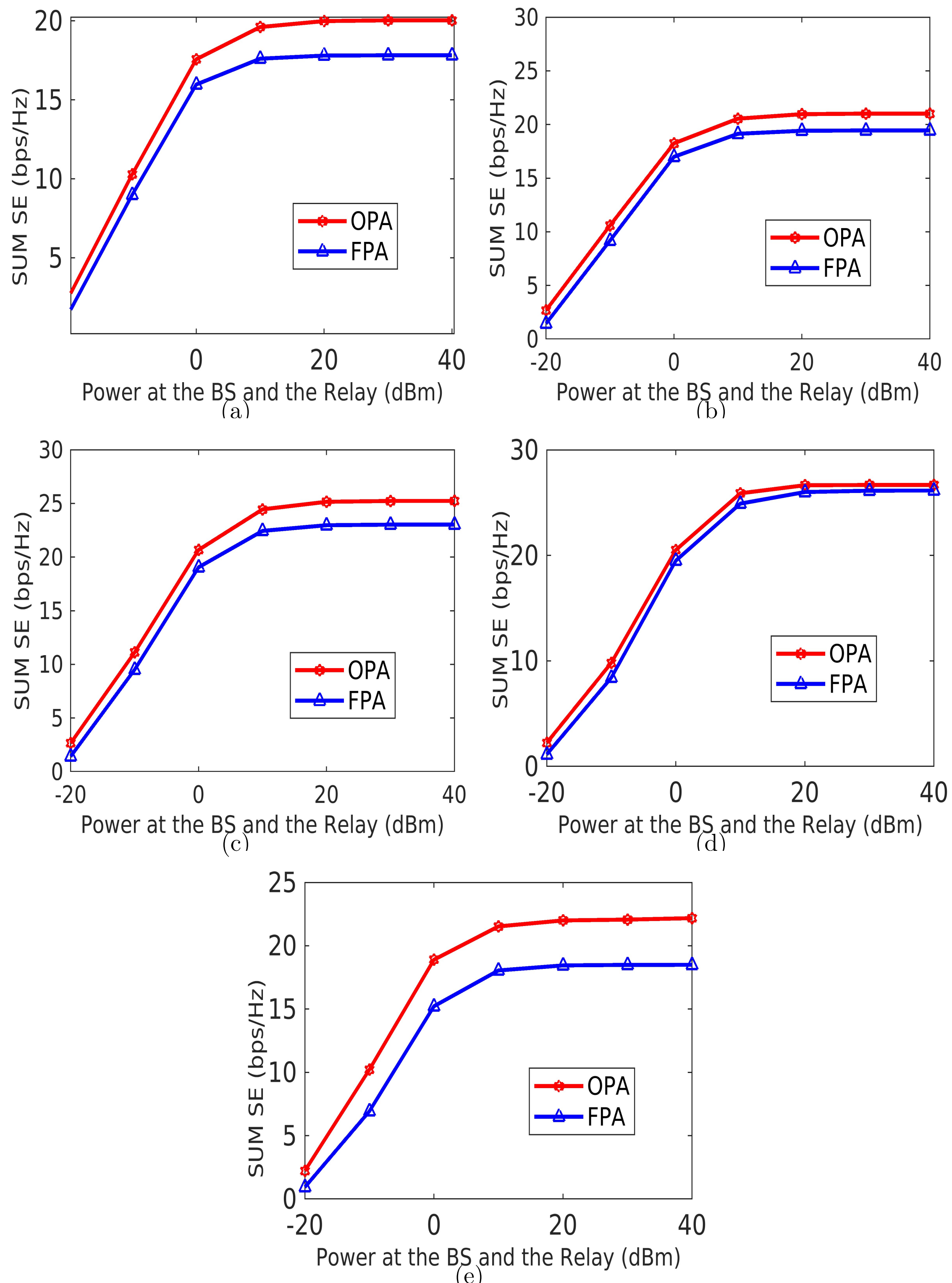

Figure 9.1: a) UATF SUM SE versus BS transmit power P_T Ergodic Sum SE with precoder b) MRC; c) ZF and; d) MMSE e)Ergodic Sum SE with precoder allocation

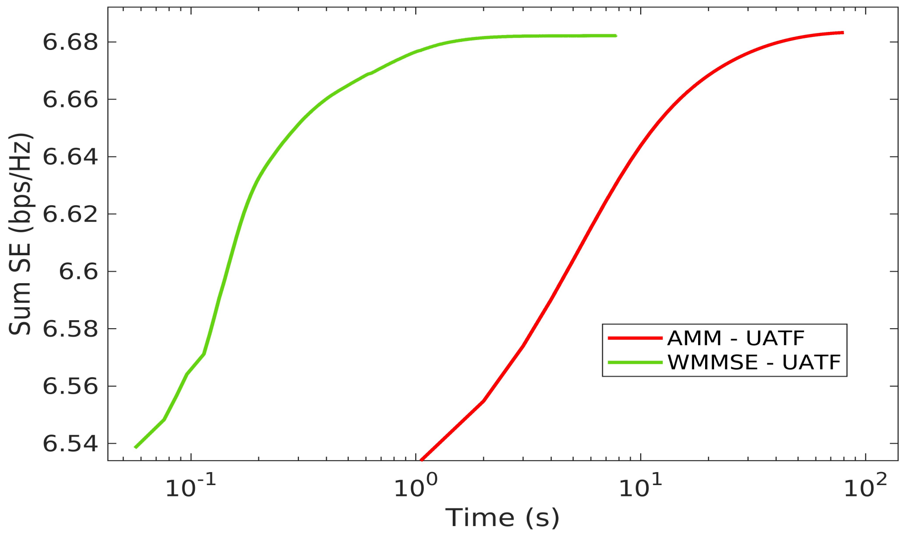

Figure 9.2: UATF Sum SE vs Time.

9.1.2 Algorithm Time Complexity vs Convergence

The base paper for this work [4], provides an exciting optimizing work which uses the modified AMM approach with QT and LDT. In the Fig.9.2 we plot the CPU time while optimizing the UATF sum SE at each iterations with maximum power allocation at the BS and at the relays to be $20dBm$. The setup is compared in a Intel® Core™ i5-8250U CPU @ 1.60GHz ×8 with a 8GB Ram. Inorder to provide a perspective, the no. of cell taken to be 1 to reduce the convergence time of the existing work. In this plot, we compare the time taken to converge, with an best, low-complexity, existing framework (AMM) [4] and the current framework (Deterministic WMMSE approach) by optimizing the UATF sum SE and the iterations for former is 1050 and the later is 60. The time scale is a log scale inorder to provide better information from the plot. As you can see, the current optimization technique is way faster than the existing best technique by a large difference as the existing framework uses cvx for optimization. Also the current framework does take large iterations to converge at a inconsiderable time. With the faster convergence there is always a trade-off, here its the sub-optimality of the current framework as it does harsher convex approximation to bring iterative closed-form solution. As the famous saying goes *"You live by the sword, you die by the sword."*

9.1.3 Inference from the Optimization Techniques

Inorder to comparing the optimization techniques, we need to make sure the initial values are equal for all the optimizations. In the Fig.9.3, we plot the Ergodic sum SE vs the maximum power allocation of the BS and the relay. We compare how the optimal precoder allocation has fared with the MR and ZF precoders. From the Fig.9.3a, which is optimized power/precoder allocation, one can infer that at low SNR, all the three are closely converged and at high SNR you can clearly see that the OPA of ZF $\geq$ Precoder Allocation $\geq$ MRC. But Fig.9.3b, at high SNR, shows that the FPA of ZF $\geq$ MRC $\geq$ Precoder Allocation. The above observation provides the details that the optimal precoder allocation has better optimized which is due the random initialization of the precoder with the power constraints. It also gives a interesting fact that, being a deterministic precoder, it fares better than MRC but not to ZF where these have precoders that are instantaneous. It is an exciting future prospect to move towards optimal precoder allocation for two reasons: 1) Precoders will be deterministic, reduction in the time complexity in the formation of the precoders. 2) Reduction in cost of both optimizing power and formation of precoder. These aforementioned reasons will also improves the throughput of the system which is need of the hour.

9.2 Conclusion

The paper has worked on a existing system model, to obtain three optimization frameworks with the WMMSE approach at a very low cost, which was observed with the time complexity plot. The paper has provided Deterministic WMMSE approach, Stochastic WMMSE approach. The paper also extended its work to precoder allocation, a new way of inexpensive stochastic optimizations with a very insightful future prospect to it.

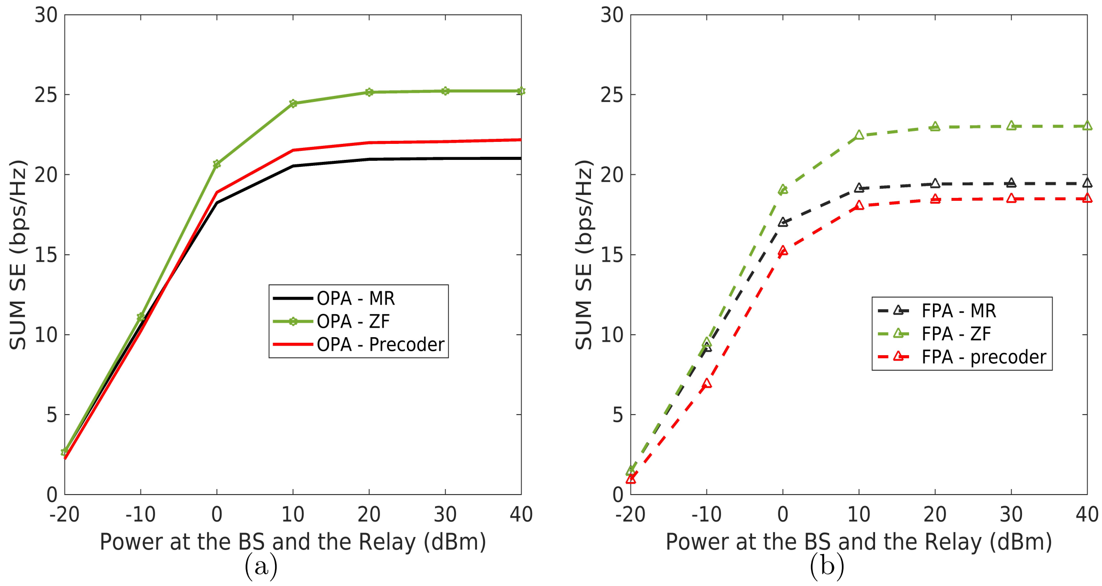

Figure 9.3: Ergodic Sum SE versus BS transmit power P_T with a) OPA b) FPA.

Appendices

Appendix A

Appendix 1

A.1 Proof of Theorem 3

We begin by differentiating the objective of problem **P11b** in (**??**) with respect to p_{jk} as follows

$$\frac{\partial f}{\partial p_{j'k'}} = -\sum_{d=1}^{D} y_d^2 \frac{\partial \Delta_d^{UL}(\mathbf{p})}{\partial p_{j'k'}} + y_{j',k'} \frac{\sqrt{1+\nu_{jk} - \frac{\mathrm{C}\sqrt{\mu_{jk}}}{2}}}{\sqrt{\Lambda_{jk}^{UL}(\mathbf{p})}} \frac{\partial \Lambda_{j'k'}}{\partial p_{j'k'}} - \sum_{j,k=1}^{L,K} y_{jk}^2 (\frac{\partial \Lambda_{j'k'} + \Delta_{jk}^{UL}(\mathbf{p})}{\partial p_{j'k'}}) - z\sqrt{\Delta} p_{j',k'} \quad \text{(A.1)}$$

Differentiation of each term is given below:

$$\frac{\partial \Delta_d^{UL}(\mathbf{p})}{\partial p_{j'k'}} = \left(\left| \overline{g}_{j',k'}^{d} \right|^2 + \beta_{j',k'}^{d} \right) \quad \text{(A.2)}$$

$$\frac{\partial \Lambda_{j'k'}}{\partial p_{j'k'}} = \left(p_{j',k'}^{p} \tau_p \mathrm{tr} \left(\mathbf{R}_{j',k'}^{j} \mathbf{\Psi}_{j',k'}^{j} \mathbf{R}_{j',k'}^{j'} \right) + \left\| \overline{\mathbf{h}}_{j',k'}^{j'} \right\|^2 \right)^2 \quad \text{(A.3)}$$

$$\frac{\partial \Delta_{jk}^{UL}(\mathbf{p})}{\partial p_{j'k'}} = \begin{cases} \hat{W}_{j',k'}^{jk} - \hat{V}_{jk} & \text{if } \{j,k\} = \{j',k'\} \\ \hat{W}_{j',k'}^{jk} & \text{if } (j,k) \in \mathcal{P}_{j',k'} \backslash (j',k') \\ \hat{W}_{j',k'}^{jk} + \hat{T}_{j',k'}^{jk} & \text{Otherwise} \end{cases} \quad \text{(A.4)}$$

Using the equations A.2-A.4, we can form the closed form of the power of CU users which is bounded by [0,P_{max}] becomes,

$$p*_{jk} = \min\left(\frac{y_{jk}^2(1+\nu_{jk}-\frac{\mathrm{C}\sqrt{\mu_{jk}}}{2})\left(p_{jk}^p\tau_p\mathrm{tr}\left(\mathbf{R}_{jk}^j\mathbf{\Psi}_{jk}^j\mathbf{R}_{jk}^j\right)+p_{jk}^p\left\|\overline{\mathbf{h}}_{jk}^j\right\|^2\right)}{\left(\sum_{l=1}^{L}\sum_{i=1}^{K}y_{l,i}^2W_{jk}^{l,i}+\sum_{(l,i)\in\mathcal{P}_{jk}}y_{l,i}^2T_{jk}^{l,i}+\sum_{d=1}^{D}y_d^2(|g_{jk}^d|^2+\beta_{jk}^d)+z\sqrt{\Delta}\right)^2}, P_{max}\right) \tag{A.5}$$

Similarly, We proceed by differentiating the objective of problem P10b in (??) with respect to $p_{d'}$ as follows:

$$\frac{\partial f}{\partial p_{d'}} = y_d'\frac{\sqrt{(1+\nu_d-\frac{C\mu_d^{\frac{1}{2}}}{2})}}{\sqrt{\Lambda_d^{UL}(\mathbf{p})'}}\frac{\partial\Lambda_d^{UL}(\mathbf{p})'}{\partial p_{d'}}-\sum_{d=1}^{D}y_d^2\frac{\partial(\Lambda_d^{UL}(\mathbf{p})+\Delta_d^{UL}(\mathbf{p}))}{\partial p_{d'}}$$
$$-\sum_{j,k=1}^{L,K}y_{jk}^2\frac{\partial\Delta_{jk}^{UL}(\mathbf{p})}{\partial p_{d'}}-z\sqrt{\Delta}p_{d'}$$
$$\frac{\partial\Lambda_d^{UL}(\mathbf{p})}{\partial p_{d'}} = \gamma_d^d+\left|\overline{g}_d^d\right|^2, \frac{\partial(\Lambda_d^{UL}(\mathbf{p})+\Delta_d^{UL}(\mathbf{p}))}{\partial p_{d'}} = \left|\overline{g}_{d'}^d)\right|^2+\beta_{d'}^d,, \frac{\partial\Delta_{jk}^{UL}(\mathbf{p})}{\partial p_{d'}} = \hat{U}_{jk}^{d'}, \tag{A.6}$$

Using (A.6), we can form the closed form of the power for D2D users which is bounded by [0,P_{max}] becomes,

$$p*_d = \min\left(\frac{y_d^2(1+\nu_d-\frac{C\sqrt{\mu_d}}{2})\left(\gamma_d^d+\left|\overline{g}_d^d\right|^2\right)}{\left(\sum_{l=1}^{L}\sum_{k=1}^{K}y_{l,k}^2U_{jk}^d+\sum_{d'=1}y_{d'}^2\left(\left|\overline{g}_d^{d'})\right|^2+\beta_d^{d'}\right)+z\sqrt{\Delta}\right)^2}, P_{max}\right) \tag{A.7}$$

Bibliography

[1] Yongming Huang, Cheng Zhang, Jiaheng Wang, Yindi Jing, Luxi Yang, and Xiaohu You. Signal processing for MIMO-NOMA: present and future challenges. IEEE Wireless Commun., 25(2):32–38, 2018.

[2] Arthur Sousa de Sena, Francisco Rafael Marques Lima, Daniel Benevides da Costa, Zhiguo Ding, Pedro H. J. Nardelli, Ugo Silva Dias, and Constantinos B. Papadias. Massive MIMO-NOMA networks with imperfect SIC: design and fairness enhancement. IEEE Trans. Wireless Commun., 19(9):6100–6115, 2020.

[3] Rohit Budhiraja Rakshit Verma. Spectral efficiency characterization of massive mimo systems with correlated rician fading channels. `https://etd.iitk.ac.in:8443/jspui/handle/123456789/19658`. Accessed: 2021-08-13.

[4] Dheeraj Naidu Amudala, Bibhor Kumar, and Rohit Budhiraja. Spatially-correlated rician-faded multi-relay multi-cell massive mimo noma systems. IEEE Transactions on Communications, pages 1–1, 2022.

[5] Luca Sanguinetti, Emil Björnson, and Jakob Hoydis. Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination. IEEE Trans. Commun., 68(1):232–257, 2020.

[6] Özgecan Özdogan, Emil Björnson, and Erik G Larsson. Massive MIMO with spatially correlated Rician fading channels. IEEE Trans. on Commun., 67(5):3234–3250, 2019.

[7] Xingqin Lin, Robert W. Heath, and Jeffrey G. Andrews. The interplay between massive MIMO and underlaid D2D networking. IEEE Trans. Wireless Commun., 14(6):3337–3351, 2015.

[8] Huy T. Nguyen, Hoang Duong Tuan, Dusit Niyato, Dong In Kim, and H. Vincent Poor. Improper gaussian signaling for D2D communication coexisting MISO cellular networks. IEEE Trans. Wireless Commun., 20(8):5186–5198, 2021.

[9] Hao Xu, Nuo Huang, Zhaohui Yang, Jianfeng Shi, Bingyang Wu, and Ming Chen. Pilot allocation and power control in D2D underlay massive MIMO systems. IEEE Commun. Lett., 21(1):112–115, 2017.

[10] Hao Xu, Wei Xu, Zhaohui Yang, Jianfeng Shi, and Ming Chen. Pilot reuse among D2D users in D2D underlaid massive MIMO systems. IEEE Trans. Veh. Technol., 67(1):467–482, 2018.

[11] Anqi He, Lifeng Wang, Yue Chen, Kai-Kit Wong, and Maged Elkashlan. Spectral and energy efficiency of uplink D2D underlaid massive MIMO cellular networks. IEEE Trans. Commun., 65(9):3780–3793, 2017.

[12] Zezhong Zhang, Yang Li, Rui Wang, and Kaibin Huang. Rate adaptation for downlink massive MIMO networks and underlaid D2D links: A learning approach. IEEE Trans. Wireless Commun., 18(3):1819–1833, 2019.

[13] Amin Ghazanfari, Emil Björnson, and Erik G. Larsson. Optimized power control for massive MIMO with underlaid D2D communications. IEEE Trans. Commun., 67(4):2763–2778, 2019.

[14] Hyoungju Ji, Sunho Park, Jeongho Yeo, Younsun Kim, Juho Lee, and Byonghyo Shim. Ultra-reliable and low-latency communications in 5G downlink: Physical layer aspects. IEEE Wireless Commun., 25(3):124–130, 2018.

[15] Chengjian Sun, Changyang She, Chenyang Yang, Tony Q. S. Quek, Yonghui Li, and Branka Vucetic. Optimizing resource allocation in the short blocklength

regime for ultra-reliable and low-latency communications. IEEE Trans. Wireless Commun., 18(1):402–415, 2019.

[16] Giuseppe Durisi, Tobias Koch, and Petar Popovski. Toward massive, ultrareliable, and low-latency wireless communication with short packets. Proceedings of the IEEE, 104(9):1711–1726, 2016.

[17] Yury Polyanskiy, H. Vincent Poor, and Sergio Verdu. Channel coding rate in the finite blocklength regime. IEEE Trans. Inf. Theory, 56(5):2307–2359, 2010.

[18] Johan Östman, Alejandro Lancho, Giuseppe Durisi, and Luca Sanguinetti. URLLC with massive MIMO: Analysis and design at finite blocklength. IEEE Trans. Wireless Commun., 20(10):6387–6401, 2021.

[19] Hong Ren, Cunhua Pan, Yansha Deng, Maged Elkashlan, and Arumugam Nallanathan. Joint pilot and payload power allocation for massive-MIMO-enabled URLLC IIoT networks. IEEE J. Sel. Areas Commun., 38(5):816–830, 2020.

[20] Ali Arshad Nasir, Hoang Duong Tuan, Hien Quoc Ngo, Trung Q. Duong, and H. Vincent Poor. Cell-free massive MIMO in the short blocklength regime for URLLC. IEEE Trans. Wireless Commun., 20(9):5861–5871, 2021.

[21] Mehdi Monemi and Hina Tabassum. Performance of UAV-assisted D2D networks in the finite block-length regime. IEEE Trans. Commun., 68(11):7270–7285, 2020.

[22] Haojun Yang, Kuan Zhang, Kan Zheng, and Yi Qian. Joint frame design and resource allocation for ultra-reliable and low-latency vehicular networks. IEEE Trans. Wireless Commun., 19(5):3607–3622, 2020.

[23] Özgecan Özdogan, Emil Björnson, and Erik G. Larsson. Massive MIMO with spatially correlated Rician fading channels. IEEE Trans. Commun., 67(5):3234–3250, 2019.

[24] Keshav Singh, Meng-Lin Ku, and Mark F. Flanagan. Energy-efficient precoder design for downlink multi-user MISO networks with finite blocklength codes. IEEE Trans. on Green Commun. and Netw., 5(1):160–173, 2021.

[25] Emil Björnson, Jakob Hoydis, and Luca Sanguinetti. Massive MIMO networks: Spectral, energy, and hardware efficiency. Foundations and Trends in Signal Processing, 11(3-4):154–655, 2017.

[26] Alessio Zappone and Eduard Jorswieck. Energy Efficiency in Wireless Networks via Fractional Programming Theory, volume 11. Now Publishers Inc., Hanover, MA, USA, June 2015.

[27] An Liu, Vincent K. N. Lau, and Borna Kananian. Stochastic successive convex approximation for non-convex constrained stochastic optimization. IEEE Trans. Signal Process., 67(16):4189–4203, 2019.

[28] Quang-Doanh Vu, Le-Nam Tran, and Markku Juntti. On spectral efficiency for multiuser MISO systems under imperfect channel information. IEEE Trans. Veh. Technol., 70(2):1946–1951, 2021.

[29] Jiayi Zhang, Linglong Dai, Ziyan He, Bo Ai, and Octavia A. Dobre. Mixed-ADC/DAC multipair massive MIMO relaying systems: Performance analysis and power optimization. IEEE Trans. Commun., 67(1):140–153, 2019.

[30] Technical specification group radio access network; spatial channel model for multiple input multiple output (MIMO) simulations. document 3GPP TR 25.996 V14.0.0, 3rd Generation Partnership Project, Mar 2017.

[31] Trinh Van Chien, Christopher Mollen, and Emil Björnson. Large-scale-fading decoding in cellular massive MIMO systems with spatially correlated channels. IEEE Trans. Commun., 67(4):2746–2762, 2019.

[32] Lei Liu, Yuhao Chi, Chau Yuen, Yong Liang Guan, and Ying Li. Capacity-achieving MIMO-NOMA: Iterative LMMSE detection. IEEE Trans. Signal Process., 67(7):1758–1773, 2019.

[33] V. Mandawaria, E. Sharma, and R. Budhiraja. WSEE maximization of mmwave NOMA systems. IEEE Commun. Lett., 23(8):1413 – 1417, 2019.

[34] Dhanushka Kudathanthirige and Gayan Amarasuriya Aruma Baduge. NOMA-aided multicell downlink massive MIMO. IEEE J. Sel. Topics Signal Process., 13(3):612–627, 2019.

[35] Yikai Li and Gayan Amarasuriya Aruma Baduge. NOMA-aided cell-free massive MIMO systems. IEEE Wireless Commun. Lett., 7(6):950–953, 2018.

[36] Manijeh Bashar, Kanapathippillai Cumanan, Alister G. Burr, Hien Quoc Ngo, Lajos Hanzo, and Pei Xiao. On the performance of cell-free massive MIMO relying on adaptive NOMA/OMA mode-switching. IEEE Trans. Commun., 68(2):792–810, 2020.

[37] Mojtaba Vaezi, Gayan Amarasuriya Aruma Baduge, Yuanwei Liu, Ahmed Arafa, Fang Fang, and Zhiguo Ding. Interplay between NOMA and other emerging technologies: A survey. IEEE Trans. on Cogn. Commun. Netw., 5(4):900–919, 2019.

[38] Di Zhang, Yuanwei Liu, Zhiguo Ding, Zhenyu Zhou, Arumugam Nallanathan, and Takuro Sato. Performance analysis of non-regenerative massive-MIMO-NOMA relay systems for 5G. IEEE Trans. Commun., 65(11):4777–4790, 2017.

[39] Xiaoming Chen, Rundong Jia, and Derrick Wing Kwan Ng. The application of relay to massive non-orthogonal multiple access. IEEE Trans. Commun., 66(11):5168–5180, 2018.

[40] Mandawaria, Ekant Sharma, and Rohit Budhiraja. Energy-efficient massive MIMO multi-relay NOMA systems with CSI errors. IEEE Trans. Commun., 68(12):7410–7428, 2020.

[41] Yikai Li and Gayan Amarasuriya. Multiple relay-aided massive MIMO NOMA. In 2019 IEEE Global Commun. Conf. (GLOBECOM), pages 1–6, 2019.

[42] Yikai Li and Gayan Amarasuriya Aruma Baduge. Relay-aided downlink massive MIMO NOMA with estimated CSI. IEEE Trans. Veh. Technol., 70(3):2258–2271, 2021.

[43] Di Zhang, Yuanwei Liu, Zhiguo Ding, Zhenyu Zhou, Arumugam Nallanathan, and Takuro Sato. Performance analysis of non-regenerative massive-MIMO-NOMA relay systems for 5G. IEEE Trans. Commun., 65(11):4777–4790, 2017.

[44] Søren Skovgaard Christensen, Rajiv Agarwal, Elisabeth De Carvalho, and John M. Cioffi. Weighted sum-rate maximization using weighted mmse for mimo-bc beamforming design. IEEE Transactions on Wireless Communications, 7(12):4792–4799, 2008.

[45] H. Zhang, F. Fang, J. Cheng, K. Long, et al. Energy-efficient resource allocation in NOMA heterogeneous networks. IEEE Wireless Commun., 25(2):48–53, April 2018.

[46] Milton Abramowitz and Irene A. Stegun. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover, New York, ninth Dover printing edition, 1964.

[47] Meisam Razaviyayn, Maziar Sanjabi, and Zhi-Quan Luo. A stochastic successive minimization method for nonsmooth nonconvex optimization with applications to transceiver design in wireless communication networks, 2013.

[48] An Liu, Vincent K. N. Lau, and Borna Kananian. Stochastic successive convex approximation for non-convex constrained stochastic optimization. IEEE Transactions on Signal Processing, 67(16):4189–4203, 2019.