Optimization of SUM SE and GEE for Ergodic and UATF of Spatially-Correlated Rician-Faded Multi-Relay Multi-Cell Massive MIMO NOMA Systems with WMMSE and other Algorithms

Index Terms

NOMA, multiple relays, WMMSE, Dinkelbach, Inexact SAA Algorithm

I. System Model

We consider, as shown in Fig. 1, the downlink of an L-cell mMIMO system. In each cell, an N-antenna mMIMO BS serve clusters of cell-edge users which, due to high path loss, have an extremely weak direct link with it. The BS serves such coverage-limited users via single-antenna half-duplex amplify-and-forward (AF) relays. These relays are installed at a high altitude such that the BS-relay and users-relay channels have both LoS and NLoS components. These channels, therefore, have Rician probability density function (pdf). Further, the lth BS communicates with a cluster of $\mathcal{U}_{lk}$ single-antenna users via the relay R_{lk} by employing NOMA for $l = 1$ to L and $k = 1$ to K. The users associate with a particular relay based on their spatial locations i.e., users close to a relay form a cluster.

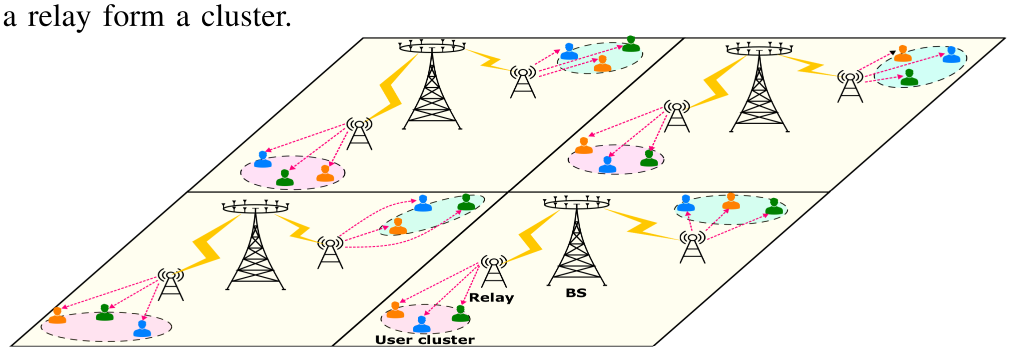

Fig. 1: Multi-cell relay-aided downlink mMIMO system model.

We next explain the communication protocol with all nodes operating in time division duplex mode. Here a τ_c symbol long channel coherence interval is divided into channel estimation (CE) and data transmission phases of $\tau \leq \tau_c$ and $(\tau_c$ - $\tau)$ symbols, respectively. In the CE phase, the relays will transmit pilots using which the BS and users estimate the BS-to-relay, and the relay-to-user channels, respectively. In the data transmission phase, the BS employs NOMA to serve users via relays. Before explaining the CE and data transmission phases, we model different channels in the system.

BS - relay channel: We denote the kth relay in the lth cell as R_{lk}. The channel from the jth BS to R_{lk} is denoted as $\mathbf{h}_{lk}^{j} \in \mathbb{C}^{N\times 1}$. Due to the lack of rich scattering around the BS, and the limited antenna spacing, the channel $\mathbf{h}_{lk}^{j}$ is spatially-correlated [1]. The presence of LoS link between the BS and relays leads to $\mathbf{h}_{lk}^{j}$ having a Rician pdf. It is, accordingly, mathematically expressed as follows

$$\mathbf{h}_{lk}^{j} = \overline{\mathbf{h}}_{lk}^{j} + (\mathbf{R}_{lk}^{j})^{\frac{1}{2}}\mathbf{h}_{lk}^{j,\mathrm{NLoS}}, \;\text{ where } \overline{\mathbf{h}}_{k} = \sqrt{\frac{K_{lk}^{j}\beta_{lk}^{j}}{1+K_{lk}^{j}}}\mathbf{h}_{lk}^{j,\mathrm{LoS}} \text{ and } \mathbf{R}_{lk}^{j} = \frac{\beta_{lk}^{j}}{1+K_{lk}^{j}}\overline{\mathbf{R}}_{lk}^{j}. \tag{1}$$

Here K_{lk}^{j} is the Rician factor and β_{lk}^{j} is the large-scale fading coefficient. The vector $\mathbf{h}_{lk}^{j,\mathrm{NLoS}}$ with pdf $\mathcal{CN}\left(0,\mathbf{I}_N\right)$, and $\mathbf{h}_{lk}^{j,\mathrm{LoS}} = \left[1, e^{j2\pi d_\lambda \sin(\varphi_{lk}^{j})}, \ldots, e^{j2\pi d_\lambda (N-1)\sin(\varphi_{lk}^{j})}\right]^T$ model the NLoS and LoS components, respectively. Here φ_{lk}^{j} is the nominal angle between the jth BS and the relay R_{lk}, and d_λ is the inter-antenna distance (in fractions of wavelengths). The matrix $\overline{\mathbf{R}}_{lk}^{j}$ characterizes the spatial correlation of the NLoS component. This is unlike [2] which considered only uncorrelated Rayleigh-faded channels. The spatial correlation can dramatically alter the system insights, and is therefore crucial to consider [1], [3].

Relay - user channel: The channel vector from the relay $R_{jk'}$ to the nth user in the cluster $\mathcal{U}_{lk}$ is denoted as $g_{lk,n}^{jk'} \in \mathbb{C}^{1\times 1}$. The channels $g_{lk,n}^{jk'}$ are also Rician-faded such that

$$g_{lk,n}^{jk'} = \overline{g}_{lk,n}^{jk'} + (\gamma_{lk,n}^{jk'})^{\frac{1}{2}} g_{lk,n}^{jk',\mathrm{NLoS}}, \;\text{ where } \overline{g}_{lk,n}^{jk'} = \sqrt{\frac{K_{lk,n}^{jk'}\beta_{lk,n}^{jk'}}{1+K_{lk,n}^{jk'}}} \text{ and } \gamma_{lk,n}^{jk'} = \frac{\beta_{lk,n}^{jk'}}{1+K_{lk,n}^{jk'}}. \tag{2}$$

Here $K_{lk,n}^{jk'}$ is the Rician factor and $\beta_{lk,n}^{jk'}$ is the large-scale fading coefficient. The scalar $g_{lk,n}^{jk',\mathrm{NLoS}} \sim \mathcal{CN}\left(0,1\right)$ model the small-scale fading in the NLoS component.

Channel Estimation: Recall that the users in the current system, due to large path loss and shadowing, do not have a direct link with the BSs [2], [4]. This prohibits the users and the BSs to estimate the end-to-end downlink and uplink CSI, respectively. It is, therefore, crucial to design various precoders, combiners and optimization algorithms based on the local CSI available with the BSs and users. The K relays in each cell, which are connected to the BSs and users, can help them in estimating their first-hop BS-to-relay and the second-hop relay-to-users CSI, respectively. The proposed design uses only this CSI. To estimate the CSI, the kth relay in lth cell R_{lk}, transmits $\tau \geq K$-length pilot $_k$ to the BSs and users. The pilots are mutually orthogonal in each cell i.e., ${}^H_k{}_j = 0$ for $k \neq j$ and $\|{}_k\|^2 = \tau$. The kth relay in each cell shares the same pilot sequence, which causes pilot contamination. The lth BS uses this pilot to estimate $\mathbf{h}^l_{lk}$ i.e., its first-hop uplink CSI from the relay R_{lk}. The nth user associated with the kth relay in the lth cell, which belongs to the cluster $\mathcal{U}_{lk}$, uses the same pilot to estimate $g^{lk}_{lk,n}$, which is its second-hop downlink CSI from the relay R_{lk}. The pilot signals received by the lth BS and the nth user in cluster $\mathcal{U}_{lk}$ are given respectively as follows:

$$\mathbf{Y}^l = \sqrt{p_p}\sum_{l'=1}^{L}\sum_{k'=1}^{K} \mathbf{h}^l_{l'k'}{}^T_{k'} + \mathbf{N}^l \text{ and } \mathbf{y}^p_{lk,n} = \sqrt{p_p}\sum_{l'=1}^{L}\sum_{k'=1}^{K} g^{l'k'}_{lk,nk'}{}^T + \mathbf{n}^p_{lk,n}. \tag{3}$$

Here p_p is the pilot transmit power and $\mathbf{N}^l$ (resp. $\mathbf{n}^p_{lk,n}$) is the additive white Gaussian noise (AWGN) at the lth BS (resp. nth user in cluster $\mathcal{U}_{lk}$) with independent and identically distributed (i.i.d) $\mathcal{CN}(0,1)$ elements. The BS estimates $\mathbf{h}^l_{lk}$ by projecting $\mathbf{Y}^l$ on to ${}_k$ as follows

$$\tilde{\mathbf{y}}^l_k = \mathbf{Y}^l{}^*_k = \sum_{l'=1}^{L} \sqrt{p_p}\tau \mathbf{h}^l_{l'k} + \mathbf{N}^l{}^*_k. \tag{4}$$

The MMSE estimate of the channel $\mathbf{h}^l_{lk}$ is therefore obtained using (4) as

$$\widehat{\mathbf{h}}^l_{lk} = \overline{\mathbf{h}}^l_{lk} + \sqrt{p_p}\mathbf{R}^l_{lk}\mathbf{\Psi}_{lk}\left[\tilde{\mathbf{y}}^l_k - \overline{\mathbf{y}}^l_k\right], \tag{5}$$

where $\mathbf{\Psi}_{lk} = \left(\mathbf{I}_N + \tau p_p \sum\limits_{l'=1}^{L} \mathbf{R}^l_{l'k}\right)^{-1}$ and $\overline{\mathbf{y}}^l_k = \sum\limits_{l'=1}^{L} \sqrt{p_p}\tau\overline{\mathbf{h}}^l_{l'k}$. The channel estimation error ${}^l_{lk} = \mathbf{h}^l_{lk} - \widehat{\mathbf{h}}^l_{lk}$ has pdf $\mathcal{CN}(0, \mathbf{R}^l_{lk} - \widehat{\mathbf{R}}^l_{lk})$, where $\widehat{\mathbf{R}}^l_{lk} = \tau p_p \mathbf{R}^l_{lk}\mathbf{\Psi}_{lk}\mathbf{R}^l_{lk}$ [5].

To estimate $g^{lk}_{lk,n}$, the nth user in cluster $\mathcal{U}_{lk}$ projects its received signal $\mathbf{y}^p_{lk,n}$ onto ${}_k$ as $\tilde{y}^p_{lk,n} = \mathbf{y}^p_{lk,n}\boldsymbol{\psi}^*_k = \sum\limits_{l'=1}^{L} \sqrt{p_p}\tau g^{l'k}_{lk,n} + \mathbf{n}^p_{lk,n}\boldsymbol{\psi}^*_k$. The MMSE estimate of $g^{lk}_{lk,n}$ is [5]

$$\widehat{g}^{lk}_{lk,n} = \overline{g}^{lk}_{lk,n} + \frac{\sqrt{p_p}\gamma^{l'k}_{lk,n}}{1 + \sum\limits_{l'=1}^{L} \tau p_p \gamma^{l'k}_{lk,n}}\left[\tilde{y}^p_{lk,n} - \overline{y}^p_{lk,n}\right], \quad \text{where } \overline{y}^p_{lk,n} = \sum_{l'=1}^{L} \sqrt{p_p}\tau\overline{g}^{l'k}_{lk,n}. \tag{6}$$

The channel estimation error $e^{lk}_{lk,n} = g^{lk}_{lk,n} - \widehat{g}^{lk}_{lk,n}$ is statistically independent from the MMSE estimate $\widehat{g}^{lk}_{lk,n}$ and is distributed as $e^{lk}_{lk,n} \sim \mathcal{CN}(0, \gamma^{lk}_{lk,n} - \upsilon^{lk}_{lk,n})$ with $\upsilon^{lk}_{lk,n} = \frac{\tau p_p(\gamma^{lk}_{lk,n})^2}{\sum\limits_{l'=1}^{L}\tau p_p \gamma^{l'k}_{lk,n} + 1}$ [5].

Downlink data transmission phase: It is divided into two following time slots:

1) First time slot $-$ BS to relay transmission: A BS first uses NOMA to superpose transmit signals of users in its cell, and then precodes and transmits it to the relays. Let $s_{jk,n}$ be the signal of the nth user in cluster $\mathcal{U}_{jk}$. The precoded NOMA signal transmitted by the jth BS is

$$\mathbf{x}^j = \sum_{k=1}^{K} \mathbf{w}_{jk} \sum_{n=1}^{\mathcal{U}_{jk}} \sqrt{p_{jk,n}} s_{jk,n} \triangleq \sum_{k=1}^{K} \mathbf{w}_{jk} x_{jk}. \tag{7}$$

Here $x_{jk} = \sum\limits_{n=1}^{\mathcal{U}_{jk}} \sqrt{p_{jk,n}} s_{jk,n}$ is the NOMA signal for the users in cluster $\mathcal{U}_{jk}$, $\mathbf{w}_{jk} \in \mathbb{C}^{N\times 1}$ is the transmit precoder and $p_{jk,n}$ is the transmit power corresponding to the data of nth user in the cluster $\mathcal{U}_{jk}$, respectively. The precoder $\mathbf{w}_{jk}$ is designed based on MR transmission as

$\mathbf{w}_{jk} = \frac{(\widehat{\mathbf{h}}_{jk}^{j})^{*}}{\sqrt{\delta_{jk}}}$, with $\delta_{jk} = \mathbb{E}\big(\big\|\widehat{\mathbf{h}}_{jk}^{j}\big\|^2\big)$. The signal received by the kth relay in lth cell is given as

$$y_{R_{lk}} = \underbrace{\sum_{l''=1}^{L}\sum_{k''=1}^{K}(\mathbf{h}_{lk}^{l''})^{T}\mathbf{w}_{l''k''}x_{l''k''}}_{\tilde{y}_{R_{lk}}} + z_{R_{lk}}. \tag{8}$$

The scalar $z_{R_{lk}}$ with pdf $\mathcal{CN}(0,1)$ is the AWGN at the kth relay in lth cell.

2) Second time slot – relay to user transmission: The AF relay R_{lk} amplifies its received precoded NOMA signal as $x_{R_{lk}} = \mu_{lk}y_{R_{lk}}$, and then broadcasts it to the users in its cluster. Here μ_{lk} is the amplification factor designed to constrain the maximum relay transmit power to q_{lk} i.e., $\mathbf{E}[|x_{R_{lk}}|^2] = q_{lk}$. The expression for amplification factor is therefore given as

$$\mathbf{E}(|x_{R_{lk}}|^2) = q_{lk} \implies \mu_{lk}^2\mathbf{E}(|y_{R_{lk}}|^2) = q_{lk} \implies \mu_{lk} = \sqrt{\frac{q_{lk}}{\mathbb{E}\left[|\tilde{y}_{R_{lk}} + z_{R_{lk}}|^2\right]}}. \tag{9}$$

The transmit signals of all the LK relays interfere with each other. The nth user associated with the relay R_{lk}, i.e., the user in the cluster $\mathcal{U}_{lk}$, receives a sum-signal $\hat{y}_{lk,n} = \sum\limits_{l'=1}^{L}\sum\limits_{k'=1}^{K} g_{lk,n}^{l'k'}x_{R_{l'k'}} + z_{lk,n}$, with $z_{lk,n}$ being the AWGN. The nth user in the cluster $\mathcal{U}_{lk}$ uses the estimated CSI $\widehat{g}_{lk,n}^{lk}$ to design an equalizer as $f_{lk,n} = (\widehat{g}_{lk,n}^{lk})^{*}/|\widehat{g}_{lk,n}^{lk}|$. It then equalizes its received signal as follows:

$$y_{lk,n} = f_{lk,n}\hat{y}_{lk,n} = \sum_{l'=1}^{L}\sum_{k'=1}^{K} f_{lk,n}g_{lk,n}^{l'k'}x_{R_{l'k'}} + f_{lk,n}z_{lk,n} \tag{10}$$

The equalized user signal $y_{lk,n}$ is re-expressed by substituting $x_{R_{l'k'}} = \mu_{l'k'}y_{R_{l'k'}}$ and using (8) as

$$\begin{aligned} y_{lk,n} = {} & \underbrace{f_{lk,n}g_{lk,n}^{lk}\mu_{lk}\mathbf{h}_{lk}^{lT}\mathbf{w}_{lk}\sqrt{p_{lk,n}}s_{lk,n}}_{\text{desired signal}} + \underbrace{f_{lk,n}g_{lk,n}^{lk}\mu_{lk}\mathbf{h}_{lk}^{lT}\mathbf{w}_{lk}\sum_{n'\neq n}^{\mathcal{U}_{lk}}\sqrt{p_{lk,n'}}s_{lk,n'}}_{\text{intra-relay interference}} \\ & + \underbrace{\sum_{l''\neq l}^{L}\sum_{n'=1}^{\mathcal{U}_{lk}} f_{lk,n}g_{lk,n}^{lk}\mu_{lk}\mathbf{h}_{lk}^{l''T}\mathbf{w}_{l''k}\sqrt{p_{l''k,n'}}s_{l''k,n'}}_{\text{1st hop PS inter-relay interference}} + \underbrace{\sum_{l''=1}^{L}\sum_{k''\neq k}^{K}\sum_{n'=1}^{\mathcal{U}_{lk}} f_{lk,n}g_{lk,n}^{lk}\mu_{lk}\mathbf{h}_{lk}^{l''T}\mathbf{w}_{l''k''}\sqrt{p_{l''k'',n'}}s_{l''k'',n'}}_{\text{1st hop nPS inter-relay interference}} \\ & + \underbrace{\sum_{l'\neq l}^{L} f_{k,n}g_{lk,n}^{l'k}\mu_{l'k}\tilde{y}_{R_{l'k}}}_{\substack{\text{2nd hop PS} \\ \text{inter-relay interference}}} + \underbrace{\sum_{l'=1}^{L}\sum_{k'\neq k}^{K} f_{lk,n}g_{lk,n}^{l'k'}\mu_{l'k'}\tilde{y}_{R_{l'k'}}}_{\substack{\text{2nd hop nPS} \\ \text{inter-relay interference}}} + \underbrace{\sum_{l'=1}^{L}\sum_{k'=1}^{K} f_{lk,n}g_{lk,n}^{l'k'}\mu_{l'k'}z_{R_{l'k'}}}_{\text{forwarding noise}} + \underbrace{f_{k,n}z_{k,n}}_{\text{receiver noise}}. \end{aligned} \tag{11}$$

In (11), i) intra-relay interference is caused by the data signals of other users served by the relay R_{lk}; ii) 1st hop pilot-shared (PS)/non pilot-sharing (nPS) inter-relay interference is because the BS uses MR precoding for NOMA signals of multiple relays, and the relay R_{lk} amplifies the NOMA signal of the PS/nPS relays; and ii) 2nd hop PS/nPS inter-relay interference is caused by the amplified transmit signal of the PS/nPS relays that serve their respective user clusters.

In a NOMA system, users served by the kth relay in lth cell mitigate the intra-relay interference by performing SIC. To enable successful SIC, we assume that the users in cluster $\mathcal{U}_{lk}$ are ordered in the descending order of their channel statistics. The nth user associated in the cluster first cancels the intra-relay interference from $\forall\ n' > n$ users by employing SIC [2], [6], and then decodes its own signal while treating the signal from the first $n-1$ users as inherent intra-relay interference [2], [6]. We observe from (11) that for a user to perform SIC and cancel intra-relay interference, it should also have instantaneous CSI $(\mathbf{h}_{lk}^{l})^T\mathbf{w}_{lk}$ along with $\hat{g}_{lk,n}^{lk}$. It is difficult for a user to have $(\mathbf{h}_{lk}^{l})^T\mathbf{w}_{lk}$. The user, therefore, uses $\mathbf{E}[(\mathbf{h}_{lk}^{l})^T\mathbf{w}_{lk}]$ to perform SIC [6]. We, similar to [2], assume that the users employ statistical value $\mathbf{E}[(\mathbf{h}_{lk}^{l})^T\mathbf{w}_{lk}]$ and its channel estimate $\widehat{g}_{lk,n}^{lk}$ to perform SIC. The intra-relay interference in (11) after performing SIC is given as

$$\underbrace{f_{lk,n}g_{lk,n}^{lk}\mu_{lk}\mathbf{h}_{lk}^{l^T}\mathbf{w}_{lk}\sum_{n'=1}^{n-1}\sqrt{p_{k,n'}}s_{k,n'}}_{\text{inherent intra-relay interference}}+\underbrace{\sum_{n'=n+1}^{\mathcal{U}_{lk}}\mu_{lk}\left[f_{lk,n}g_{lk,n}^{lk}\mathbf{h}_{lk}^{l^T}\mathbf{w}_{lk}-f_{lk,n}\widehat{g}_{lk,n}^{lk}\mathbb{E}\left[\mathbf{h}_{lk}^{l^T}\mathbf{w}_{lk}\right]\right]\sqrt{p_{k,n'}}s_{k,n'}}_{\text{residual intra-relay interference due to imperfect SIC}}.$$

The first-term is the inherent intra-relay interference, and the second-term is the residual intra-relay interference due to imperfect SIC. The post-SIC receive signal at the nth user associated with the relay R_{lk}, denoted as $\bar{y}_{lk,n}$, can be derived as

$$\begin{aligned}\bar{y}_{lk,n} &= f_{lk,n}g_{lk,n}^{lk}\mu_{lk}\mathbf{h}_{lk}^{l^T}\mathbf{w}_{lk}\sqrt{p_{lk,n}}s_{lk,n} + f_{lk,n}g_{lk,n}^{lk}\mu_{lk}\mathbf{h}_{lk}^{l^T}\mathbf{w}_{lk}\sum_{n'=1}^{n-1}\sqrt{p_{lk,n'}}s_{lk,n'} \qquad (12)\\ &+\sum_{n'=n+1}^{\mathcal{U}_{lk}}\mu_{lk}\left[f_{lk,n}g_{lk,n}^{lk}\mathbf{h}_{lk}^{l^T}\mathbf{w}_{lk}-f_{lk,n}\widehat{g}_{lk,n}^{lk}\mathbf{E}\left[\mathbf{h}_{lk}^{l^T}\mathbf{w}_{lk}\right]\right]\sqrt{p_{lk,n'}}s_{lk,n'}\\ &+\sum_{l''\neq l}^{L}\sum_{n'=1}^{\mathcal{U}_{lk}}f_{lk,n}g_{lk,n}^{lk}\mu_{lk}\mathbf{h}_{lk}^{l''^T}\mathbf{w}_{l''k}\sqrt{p_{l''k,n'}}s_{l''k,n'}+\sum_{l''=1}^{L}\sum_{k''\neq k}^{K}\sum_{n'=1}^{\mathcal{U}_{lk}}f_{lk,n}g_{lk,n}^{lk}\mu_{lk}\mathbf{h}_{lk}^{l''^T}\mathbf{w}_{l''k''}\sqrt{p_{l''k'',n'}}s_{l''k'',n'}\\ &+\sum_{l'\neq l}^{L}f_{k,n}g_{lk,n}^{l'k}\mu_{l'k}\tilde{y}_{R_{l'k}}+\sum_{l'=1}^{L}\sum_{k'\neq k}^{K}f_{lk,n}g_{lk,n}^{l'k'}\mu_{l'k'}\tilde{y}_{R_{l'k'}}+\sum_{l'=1}^{L}\sum_{k'=1}^{K}f_{lk,n}g_{lk,n}^{l'k'}\mu_{l'k'}z_{R_{l'k'}}+f_{k,n}z_{k,n}.\end{aligned}$$

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

$$\overline{\mathbf{W}}_l = \begin{cases} \mathbf{H}_l, & \text{for MRC} \\ \mathbf{H}_l \left[\mathbf{H}_l^H \mathbf{H}_j\right]^{-1}, & \text{for ZF} \\ \left[\sum\limits_{l'=1}^{L} \mathbf{H}_{l'} \overline{\mathbf{P}}_{l'}^{cd} \mathbf{H}_{l'}^H + \mathbf{I}_M\right]^{-1} \mathbf{H}_j \overline{\mathbf{P}}_j, & \text{for MMSE.} \end{cases} \tag{13}$$

Here $\overline{\mathbf{W}^l} = [\mathbf{w}_{l1}, \cdots, \mathbf{w}_{lK}] \in \mathbf{C}^{N\times K}$ denotes the set of combiners used by lth BS for the relays in the lth cell. The matrices $\mathbf{H}_l = [\mathbf{h}_{l1}^l, \cdots, \mathbf{h}_{lK}^l] \in \mathbf{C}^{N\times K}$ 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 $\mathcal{U}_{l'k}$.

II. SE ANALYSIS

1) Ergodic SE Analysis: From the equation (12) of the base paper [7], The Ergodic sum SE of the system using Genie-bound for a finite number of BS antennas with MMSE channel Estimation and with imperfect user SIC, is given as

$$\widetilde{R}_{\text{sum}}^{\text{e}} = \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} \mathbb{E}\left[\frac{1}{2}\left(1 - \frac{\tau}{\tau_c}\right)\log_2\left(1 + \frac{\overline{\Delta}_{lk,n}^i}{\overline{\Omega}_{lk,n}^i}\right)\right], \text{ where } \overline{\Omega}_{lk,n}^i = \sum_{m=1}^{5} \hat{I}_{lk,n}^{(m)} + 1, \tag{14}$$

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

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

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

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

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

$$\hat{C}_{l'k',lk,n}^{(3)} = |f_{lk,n} g_{lk,n}^{l'k'} \mathbf{h}_{l'k'}^{l'H} \mathbf{w}_{l'k'}|^2,\ \hat{C}_{\substack{l''k'',l'k',\\ lk,n}}^{(4)} = |f_{lk,n} g_{lk,n}^{l'k'} \mathbf{h}_{l'k'}^{l''H} \mathbf{w}_{l''k''}|^2,\ \hat{C}_{l'k',lk,n}^{(5)} = |f_{lk,n} g_{lk,n}^{l'k'}|^2 \tag{15}$$

2) UATF closed-form SE: The hardening bound technique to derive a UATF closed-form SE expression with MRC precoder for the multi-cell relay-aided mMIMO NOMA system for a finite number of BS antennas relying on MMSE channel estimation, and with imperfect user

SIC which is derived in the base paper [7] is given as

$$\overline{R}^{c}_{\text{sum}}=\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}}\frac{1}{2}\left(1-\frac{\tau}{\tau_c}\right)\log_2\left(1+\frac{\overline{\Delta}^{c}_{lk,n}}{\overline{\Omega}^{c}_{lk,n}}\right),\ \text{where}\ \ \overline{\Omega}_{lk,n}=\sum_{m=0}^{7}\overline{I}^{(m)}_{lk,n}+1, \tag{16}$$

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

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

$$\overline{I}^{(7)}_{k,n}=\sum_{l'=1}^{L}\sum_{k'=1}^{K}C^{(5)}_{l'k',lk,n}\overline{\mu}^{2}_{l'k'},\ \text{and}\ \overline{\mu}_{lk}=\sqrt{\frac{q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}p_{l''k''}+\sum\limits_{l''=1}^{K}\xi_{l''k,lk}p_{l''k}+1\right)}}. \tag{17}$$

Here $p_{lk}=\sum_{n=1}^{\mathcal{U}_{lk}}p_{lk,n}$, $A_{lk,n}=\frac{\pi\upsilon^{lk}_{lk,n}\delta_{lk}}{4}\left[L_{1/2}\left(-\left|\overline{g}^{lk}_{lk,n}\right|^{2}/\upsilon^{lk}_{lk,n}\right)\right]^{2}$, with $L_{\frac{1}{2}}(\cdot)$ being Laguerre polynomial [8]. 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 of long term channel statistics, which are given in Appendix A of paper [].

III. Preface to WMMSE Algorithm

The WMMSE algorithm was first proposed in [9] 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.

$$\text{SINR}=\max_{u}\gamma=\max_{u}\frac{1}{e}-1 \tag{18}$$

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:

$$y_k=h_k\sqrt{p_k}s_k+\sum_{\substack{i=1\\ i\neq k}}^{N}h_i\sqrt{p_i}s_i+n \tag{19}$$

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:

$$\text{SE}_k^{CU} = C\log(1+\text{SINR}_k^{CU}) = \frac{\Delta_k}{\Omega_k}, \text{ where}$$

$$\text{SINR}_k^{CU} = \frac{|h_k|^2 p_k}{\sum\limits_{\substack{i=1\\ i\neq k}}^{N} |h_i|^2 p_i + 1}. \tag{20}$$

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;

$$\hat{s_k} = u_k y_k = u_k h_k \sqrt{p_k} s_k + \sum_{\substack{i=1\\ i\neq k}}^{N} u_k h_i \sqrt{p_i} s_i + u_k n, \tag{21}$$

the SINR of CU k of the decoded signal is:

$$\gamma_k = \frac{|u_k h_k|^2 p_k}{\sum_{\substack{i=1\\ i\neq k}}^{N} |u_k h_i|^2 p_i + |u_k|^2}, \tag{22}$$

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

$$e_k = E(|\hat{s_k} - s_k|^2) = |1 - u_k h_k \sqrt{p_k}|^2 + \sum_{\substack{i=1\\ i\neq k}}^{N} |u_k h_i|^2 p_i + |u_k|^2 \tag{23}$$

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

$$\text{SINR}_k^{CU} \;=\; \frac{1}{\min\limits_{u_k} e_k} - 1 \;=\; \max_{u_k} \frac{1}{e_k} - 1 \tag{24}$$

which follows the equation mentioned earlier (18). 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,

$$\text{SINR} \;=\; \max_{u} \frac{1}{e} - 1 \text{ where,} \tag{25}$$

$$e \;=\; |1 - u\left(\sqrt{\Delta}\right)|^2 + |u|^2 \,(\Omega) \tag{26}$$

Reason: If we look carefully the e_k variable in (23), it can be interpreted as,

$$e_k = |1 - u_k \underbrace{h_k\sqrt{p_k}}_{\sqrt{\Delta_k}}|^2 + |u_k|^2 \underbrace{\left(\sum_{\substack{i=1\\ i\neq k}}^{N} |h_i|^2 p_i + 1\right)}_{\Omega_k} \tag{27}$$

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:

$$\text{SE} = \log\left(1+\frac{\Delta}{\Omega}\right) = \underset{w,u}{\text{Maximize}} \ \log(w) - we \text{ where,} \tag{28}$$

$$e = |1-u\left(\sqrt{\Delta}\right)|^2 + |u|^2\left(\Omega\right) \text{ and} \tag{29}$$

$$w = \frac{1}{e}. \tag{30}$$

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

$$\text{SE} = \log\left(1+\frac{\Delta}{\Omega}\right) = \log\left(\max_u \frac{1}{e}\right) \text{ where,} \tag{31}$$

$$e = |1-u\left(\sqrt{\Delta}\right)|^2 + |u|^2\left(\Omega\right) \tag{32}$$

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:

$$\text{SE} = \log\left(1+\frac{\Delta}{\Omega}\right) = \max_u \log\left(w\right) \text{ where,} \tag{33}$$

$$w \leq \frac{1}{e}. \tag{34}$$

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

$$\mathcal{L}(w,\lambda) = \log\left(w\right) - \lambda\left(1-\frac{1}{we}\right) \tag{35}$$

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

$$\lambda* = we \tag{36}$$

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:

$$\text{SE} = \underset{w,u}{\text{Maximize}} \ \log(w) - we \text{ where,} \tag{37}$$

$$e = |1-u\left(\sqrt{\Delta}\right)|^2 + |u|^2\left(\Omega\right) \text{ and} \tag{38}$$

$$w = \frac{1}{e}. \tag{39}$$

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

IV. 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 IV-A and Ergodic IV-B 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.

A. 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:

$$\mathbf{P1}: \qquad \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}}\ R^c_{\text{sum}} \tag{40}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l\ \forall l,\ \ q_{lk} \leq Q_{lk}\ \forall l,\ \forall k \text{ and } p_{lk,n}, q_{lk} \geq 0 \tag{41}$$

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

$$R^c_{\text{sum}} = \frac{1}{2}\left(\frac{\tau_c - \tau_p}{\tau_c}\right)\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} \log_2\left(1 + \frac{\overline{\Delta}^c_{lk,n}}{\overline{\Omega}^c_{lk,n}}\right). \tag{42}$$

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 SINR expression in (99), contains fractional functions of transmit user power and transmit relay power sum, the optimization variables in both numerator and denominator. 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:

$$\textbf{P2}: \quad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}} \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} \log\left(1+\frac{\overline{\Delta}^{c}_{lk,n}}{\overline{\Omega}^{c}_{lk,n}}\right) \tag{43}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\Big(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}p_{l''k''} + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}p_{l''k}+1\Big)}}. \tag{44}$$

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:

$$\log(1+\text{SINR}^{c}_{lk,n}) \;=\; \log(1+\frac{\overline{\Delta}^{c}_{lk,n}}{\overline{\Omega}^{c}_{lk,n}}) \;=\; \max_{u^{c}_{lk,n},w^{c}_{lk,n}} \frac{1}{w^{c}_{lk,n}} - w^{c}_{lk,n}e^{c}_{lk,n} \text{ where,} \tag{45}$$

$$e^{c}_{lk,n} = |1-u^{c}_{lk,n}\sqrt{\overline{\Delta}^{c}_{lk,n}}|^2 + |u^{c}_{lk,n}|^2\overline{\Omega}^{c}_{lk,n} \tag{46}$$

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 (45) to the Problem **P2**, the problem is reconstructed as:

$$\textbf{P3}: \quad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}} \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} \underset{u^{c}_{lk,n},w^{c}_{lk,n}}{\text{Maximize}}\ \log\left(w^{c}_{lk,n}\right) - w^{c}_{lk,n}e^{c}_{lk,n} \tag{47}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\Big(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}p_{l''k''} + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}p_{l''k}+1\Big)}}.. \tag{48}$$

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:

$$\textbf{P4}: \quad \underset{\mathbf{P},\mathbf{U},\mathbf{W},\boldsymbol{\mu}}{\text{Maximize}} \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} \log\left(w^{c}_{lk,n}\right) - w^{c}_{lk,n}e^{c}_{lk,n} \tag{49}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K} \rho_{l''k'',lk}p_{l''k''} + \sum\limits_{l''=1}^{L} \xi_{l''k,lk}p_{l''k} + 1\right)}}. \tag{50}$$

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:

$$\begin{aligned}
e^c_{lk,n} &= 1 + |u^c_{lk,n}|^2 A_{lk,n}\overline{\mu}^2_{lk}p_{lk,n} - 2Re(u^c_{lk,n}\sqrt{A_{lk,n}\overline{\mu}^2_{lk}p_{lk,n}}) \\
&+|u^c_{lk,n}|^2\Bigg(A_{lk,n}\overline{\mu}^2_{lk}p_{lk,n} + \sum_{n'=1}^{n-1} C^{(1)}_{lk,n,n'}\overline{\mu}^2_{lk}p_{lk,n'} + \sum_{n'=n+1}^{\mathcal{U}_{lk}} C^{(2)}_{lk,n,n'}p_{lk,n'}\overline{\mu}^2_{lk} + \sum_{(l',k')\neq(l,k)}\sum_{n'=1}^{\mathcal{U}_{l'k'}} C^{(3)}_{l'k',lk,n}p_{l'k',n'}\mu^2_{l'k'} \\
&+ \sum_{(l',k')}\sum_{(l'',k'')\neq(l',k')}\sum_{n'=1}^{\mathcal{U}_{l''k''}} C^{(4)}_{\substack{l''k'',l'k',\\ lk,n}}p_{l''k'',n'}\overline{\mu}^2_{l'k'} + \sum_{l'=1}^{L}\sum_{k'=1}^{K} C^{(5)}_{l'k',lk,n}\overline{\mu}^2_{l'k'} + 1\Bigg)
\end{aligned} \tag{51}$$

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

$$w^c_{lk,n}e^c_{lk,n} = w^c_{lk,n} + \sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}} \alpha^c_{l'k',n',lk,n}p_{l'k',n'} + \beta^c_{lk,n}\sqrt{p_{lk,n}}. \tag{52}$$

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

$$\alpha^c_{l'k',n',lk,n} = \tag{53}$$

$$\beta^c_{lk,n} = \tag{54}$$

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

$$\begin{aligned}
&\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}} \log(w^c_{lk,n}) - w^c_{lk,n}e^c_{lk,n} = \\
&\qquad \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\left(const(\mathbf{W}) - \sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}} \alpha^c_{l'k',n',lk,n}p_{l'k',n'} + \beta^c_{lk,n}\sqrt{p_{lk,n}}\right)
\end{aligned} \tag{55}$$

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

$$\textbf{P5}: \quad \underset{\mathbf{P},\mathbf{U},\mathbf{W},\boldsymbol{\mu}}{\text{Maximize}} \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}}\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}} \left(const(\mathbf{W}) - \sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}} \alpha^{c}_{l'k',n',lk,n} p_{l'k',n'} + \beta^{c}_{lk,n}\sqrt{p_{lk,n}} \right) \tag{56}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K} \rho_{l''k'',lk} p_{l''k''} + \sum\limits_{l''=1}^{L} \xi_{l''k,lk} p_{l''k} + 1\right)}}. \tag{57}$$

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:

$$u^{c}_{lk,n}{}^{*} = \frac{\sqrt{\overline{\Delta}^{c}_{lk,n}}}{\overline{\Omega}^{c}_{lk,n}} \tag{58}$$

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:

$$w^{c*}_{lk,n} = \frac{1}{e^{c*}_{lk,n}} \tag{59}$$

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:

$$p^{*}_{lk,n} = \left\{ \frac{\boldsymbol{\beta}^{c}_{lk,n}}{\left(\sum\limits_{l'=1}^{L}\sum\limits_{k'=1}^{K}\sum\limits_{n'=1}^{\mathbb{U}_{lk}} \boldsymbol{\alpha}^{c}_{lk,n,l'k',n'}\right) + \lambda^{*}_{l}} \right\}^{2} \tag{60}$$

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:

$$w^{c}_{lk,n} e^{c}_{lk,n} = w^{c}_{lk,n} + \sum_{l'=1}^{L}\sum_{k'=1}^{K} \gamma^{c}_{l'k',lk} \mu^{2}_{lk} + \omega^{t}_{lk} \mu_{lk}. \tag{61}$$

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

$$\gamma^{c}_{l'k',lk} = \tag{62}$$

$$\omega^{c}_{lk} = \tag{63}$$

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

$$\textbf{P6}: \quad \underset{\mathbf{P},\mathbf{U},\mathbf{W},\boldsymbol{\mu}}{\text{Maximize}} \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}}\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\left(const(\mathbf{W}) - \sum_{l'=1}^{L}\sum_{k'=1}^{K}\gamma^{c}_{l'k',lk}\mu^{2}_{lk} + \omega^{t}_{lk}\mu_{lk}\right) \tag{64}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}p_{l''k''} + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}p_{l''k} + 1\right)}}. \tag{65}$$

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

$$\bar{\mu}^{*}_{lk} = \min\left\{\frac{\boldsymbol{\omega}^{c}_{lk}}{\sum\limits_{l'=1}^{L}\sum\limits_{k'=1}^{K}\boldsymbol{\gamma}^{c}_{lk,l'k'}}, \hat{\mu_{lk}}\right\} \tag{66}$$

where,

$$\hat{\mu_{lk}} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}p^{*}_{l''k''} + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}p^{*}_{l''k} + 1\right)}}. \tag{67}$$

We observe that (60) and (66) 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 (40) is provided in Algorithm 1.

B. 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 [7]), which necessitates the gathering of a large number of random channel realisations prior to the updating

Algorithm 1: sum SE maximization using Deterministic WMMSE approach

Input: Given a tolerance $\epsilon > 0$, the maximum number of iterations N and maximum power constraint P_l^{max} for UE U_{lk} and maximum power constraint Q_{lk}^{max} 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 (58)
- Update the auxiliary variable $w_{lk,n}$ using (59)
- Compute $p_{jk}^{(n+1)}$ using (60)
- Update the $\hat{\mu}_{lk}$ constraint variable using (67)
- Compute $\mu_{lk}^{(i+1)}$ using (66)
- Do until convergence **if** $\sum_{l=1}^{L}\sum_{l=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\left(p_{lk,n}^{(i)} - p_{lk,n}^{(i-1)}\right) < \epsilon$ **then** break.

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

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 [10] 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:

$$\mathbf{P1}_{sto}: \qquad \underset{\mathbf{P},\mathbf{Q}}{\text{Maximize}}\ \mathbf{E}\left[g(\mathbf{P},\mathbf{Q},\mathcal{F})\right] \triangleq R_{sum}^{e} \tag{68}$$

$$s.t \quad \sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} p_{lk,n} \leq P_l\ \forall l,\ \ q_{lk} \leq Q_{lk}\ \forall l,\ \forall k \text{ and } p_{lk,n}, q_{lk} \geq 0. \tag{69}$$

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

$$g(\mathbf{P},\mathbf{Q},\mathcal{F}) = \frac{1}{2}\left(\frac{\tau_c - \tau_p}{\tau_c}\right)\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} \log_2\left(1 + \frac{\overline{\Delta}_{lk,n}^{i}(\mathcal{F})}{\overline{\Omega}_{lk,n}^{i}(\mathcal{F})}\right) \tag{70}$$

where $\overline{\Delta}_{lk,n}^{i}(\mathcal{F}), \overline{\Omega}_{lk,n}^{i}(\mathcal{F})$ are instantaneous SINR, Numerator of SINR, Denominator of the SINR of the nth user of the cluster $\mathcal{U}_{lk}$. 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:

$$\mathbf{P2}_{sto}: \quad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}} \ \mathbf{E}\left[\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}} \log_2\left(1+\frac{\overline{\Delta}^{i}_{lk,n}((\mathcal{F})}{\overline{\Omega}^{i}_{lk,n}(\mathcal{F})}\right)\right] \tag{71}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}p_{l''k''} + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}p_{l''k}+1\right)}}. \tag{72}$$

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:

$$\mathbf{P3}_{sto}: \quad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}} \ \lim_{T\Rightarrow\infty}\frac{1}{T}\sum_{t=1}^{T}\left[\underset{\mathbf{U}^t,\mathbf{W}^t}{\text{Maximize}} \ \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}} w^{t}_{lk,n}e^{t}_{lk,n} - \log(w^{t}_{lk,n})\right] \tag{73}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}p_{l''k''} + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}p_{l''k}+1\right)}}. \tag{74}$$

where,

$$e^{t}_{lk,n} = |1-u^{t}_{lk,n}\sqrt{\overline{\Delta}^{i}_{lk,n}(\mathcal{F}^t)}|^2 + |u^{t}_{lk,n}|^2\,\overline{\Omega}^{i}_{lk,n}(\mathcal{F}^t) \tag{75}$$

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

$$w^{t*}_{lk,n} = \frac{1}{e^{t*}_{lk,n}} \tag{76}$$

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

$$u^{t*}_{lk,n} = \frac{\sqrt{\overline{\Delta}^{i}_{lk,n}(\mathcal{F}^t)}}{\overline{\Omega}^{i}_{lk,n}(\mathcal{F}^t)} \tag{77}$$

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

$$\mathbf{P4}_{sto}: \quad \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Maximize}} \lim_{T \Rightarrow \infty} \frac{1}{T} \sum_{t=1}^{T} \left[\sum_{l=1}^{L} \sum_{k=1}^{K} \sum_{n=1}^{\mathbb{U}_{lk}} \log(w_{lk,n}^{t**}) - w_{lk,n}^{t**} e_{lk,n}^{t**} \right] \tag{78}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\Big(\sum\limits_{l''=1}^{L} \sum\limits_{k''=1}^{K} \rho_{l''k'',lk} p_{l''k''} + \sum\limits_{l''=1}^{L} \xi_{l''k,lk} p_{l''k} + 1 \Big)}}. \tag{79}$$

where,

$$e_{lk,n}^{t**} = |1 - u_{lk,n}^{t**} \sqrt{\overline{\Delta}_{lk,n}^{i}(\mathcal{F}^t)}|^2 + |u_{lk,n}^{t**}|^2 \, \overline{\Omega}_{lk,n}^{i}(\mathcal{F}^t) \tag{80}$$

and the auxiliary variables $u_{lk,n}^{t**}, w_{lk,n}^{t**}$ 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 [10], 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:

$$\mathbf{P5}_{sto}: [\mathbf{P}^{t+1}, \boldsymbol{\mu}^{t+1}] \Leftarrow \underset{\mathbf{P},\boldsymbol{\mu}}{\text{Argmax}} \; \frac{1}{t} \left[\mathbf{R}^{1:t}(\mathbf{P}, \boldsymbol{\mu}, \mathbf{P}^t, \boldsymbol{\mu}^t) \right] \tag{81}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\Big(\sum\limits_{l''=1}^{L} \sum\limits_{k''=1}^{K} \rho_{l''k'',lk} p_{l''k''} + \sum\limits_{l''=1}^{L} \xi_{l''k,lk} p_{l''k} + 1 \Big)}}. \tag{82}$$

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 $(\tilde{R}_{sum}^{i}(\mathbf{P}, \boldsymbol{\mu}, \mathcal{F}^t))$ of the instantaneous sum SE to surrogate function upto tth realization $(\mathbf{R}^{1:t}(\mathbf{P}, \boldsymbol{\mu}))$ through recursive surrogate function from [11], which also guarantee the convergence of the algorithm and expressed as:

$$\mathbf{R}^{1:t}(\mathbf{P}, \boldsymbol{\mu}, \mathbf{P}^t, \boldsymbol{\mu}^t) = \frac{1}{t} \left[\tilde{R}_{sum}^{i}(\mathbf{P}, \boldsymbol{\mu}, \mathcal{F}^t) \right] + \frac{t-1}{t} \left[\mathbf{R}^{1:t-1}(\mathbf{P}, \boldsymbol{\mu}, \mathbf{P}^{t-1}, \boldsymbol{\mu}^{t-1}) \right] \tag{83}$$

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,

$$\tilde{R}^{i}_{sum}(\mathbf{P},\boldsymbol{\mu},\mathcal{F}^t)=\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\log(w^{t**}_{lk,n})-w^{t**}_{lk,n}e^{t**}_{lk,n} \tag{84}$$

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

$$\begin{aligned}e^{t}_{lk,n}&=1+|u^{t**}_{lk,n}|^2\hat{A}^{t}_{lk,n}\tilde{\mu}^2_{lk}p_{lk,n}-2Re(u^{t**}_{lk,n}\sqrt{\hat{A}^{t}_{lk,n}\tilde{\mu}^2_{lk}p_{lk,n}})\\&+|u^{t**}_{lk,n}|^2\Bigg(\sum_{n'=1}^{n-1}\hat{C}^{(1)^t}_{lk,n,n'}\tilde{\mu}^2_{lk}p_{lk,n'}+\sum_{n'=n+1}^{\mathcal{U}_{lk}}\hat{C}^{(2)^t}_{lk,n,n'}p_{lk,n'}\tilde{\mu}^2_{lk}+\sum_{(l',k')\neq(l,k)}\sum_{n'=1}^{\mathcal{U}_{l'k'}}\hat{C}^{(3)^t}_{l'k',lk,n}p_{l'k',n'}\mu^2_{l'k'}\\&+\sum_{(l',k')}\sum_{(l'',k'')\neq(l',k')}\sum_{n'=1}^{\mathcal{U}_{l''k''}}C^{(4)^t}_{\substack{l''k'',l'k',\\lk,n}}p_{l''k'',n'}\tilde{\mu}^2_{l'k'}+\sum_{l'=1}^{L}\sum_{k'=1}^{K}\hat{C}^{(5)^t}_{l'k',lk,n}\tilde{\mu}^2_{l'k'}+1\Bigg)\end{aligned} \tag{85}$$

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

$$w^{t**}_{lk,n}e^{t**}_{lk,n}=w^{t**}_{lk,n}+\sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}}\alpha^{t}_{l'k',n',lk,n}p_{l'k',n'}+\beta^{t}_{lk,n}\sqrt{p_{lk,n}}. \tag{86}$$

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

$$\alpha_{l'k',n',lk,n}= \tag{87}$$

$$\beta^{t}_{lk,n}= \tag{88}$$

Therefore, the objective function of problem $\mathbf{P4}_{sto}$ becomes:

$$\begin{aligned}&\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\log(w^{t**}_{lk,n})-w^{t**}_{lk,n}e^{t**}_{lk,n}=\\&\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\left(const(\mathbf{W})-\sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}}\alpha^{t}_{l'k',n',lk,n}p_{l'k',n'}+\beta^{t}_{lk,n}\sqrt{p_{lk,n}}\right)\end{aligned} \tag{89}$$

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 (87). Therefore the surrogate function upto

tth realization is written as:

$$\mathbf{R}^{1:t}(\mathbf{P},\boldsymbol{\mu},\mathbf{P}^t,\boldsymbol{\mu}^t)=\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\left(const(\mathbf{W})-\sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}}\boldsymbol{\alpha}^{1:t}_{l'k',n',lk,n}p_{l'k',n'}+\boldsymbol{\beta^{1:t}}_{lk,n}\sqrt{p_{lk,n}}\right) \tag{90}$$

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:

$$\begin{aligned}\boldsymbol{\alpha}^{1:t}_{l'k',n',lk,n}&=\frac{1}{t}\alpha^{t}_{l'k',n',lk,n}+\frac{t-1}{t}\boldsymbol{\alpha}^{1:t-1}_{l'k',n',lk,n}\\ \boldsymbol{\beta}^{1:t}_{lk,n}&=\frac{1}{t}\beta^{t}_{lk,n}+\frac{t-1}{t}\boldsymbol{\beta}^{1:t-1}_{lk,n}\end{aligned} \tag{91}$$

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 $p^{t+1}_{lk,n}$ and μ^{t+1}_{lk} 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

$$\mathbf{P6}_{sto}:[\mathbf{P}^{t+1},\boldsymbol{\mu}^{t+1}]\Leftarrow\underset{\mathbf{P},\boldsymbol{\mu}}{\text{Argmax}}\ -\frac{1}{t}\left[\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\left(\sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}}\boldsymbol{\alpha}^{1:t}_{l'k',n',lk,n}p_{l'k',n'}+\boldsymbol{\beta^{1:t}}_{lk,n}\sqrt{p_{lk,n}}\right)\right] \tag{92}$$

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

$$\hat{\mu}_{lk}=\sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}p_{l''k''}+\sum\limits_{l''=1}^{L}\xi_{l''k,lk}p_{l''k}+1\right)}}. \tag{93}$$

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:

$$p^{t+1}_{lk,n}=\left\{\frac{\boldsymbol{\beta}^{1:t}_{lk,n}}{\left(\sum\limits_{l'=1}^{L}\sum\limits_{k'=1}^{K}\sum\limits_{n'=1}^{\mathbb{U}_{lk}}\boldsymbol{\alpha}^{1:t}_{lk,n,l'k',n'}\right)+\lambda^*_l}\right\}^2 \tag{94}$$

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

Similarly, instead of grouping the variable **P** wise, if the equation (119) is grouped with respect to $\boldsymbol{\mu}$ wise and multiplied by $w^{t**}_{lk,n}$ the term becomes:

$$w^{t**}_{lk,n}e^{t**}_{lk,n}=w^{t**}_{lk,n}+\sum_{l'=1}^{L}\sum_{k'=1}^{K}\gamma^{t}_{l'k',lk}\mu^2_{lk}+\omega^{t}_{lk}\mu_{lk}. \tag{95}$$

Using the similar steps of updating the surrogate function with the constants forming the objective function as in $\mathbf{P6}_{sto}$, the $\mu_{lk,n}$ is optimized through first-order differentiation and equated to zero as:

$$\mu_{lk}^{t+1} = \min\left\{\frac{\boldsymbol{\omega}_{lk}^{1:t}}{\sum\limits_{l'=1}^{L}\sum\limits_{k'=1}^{K}\boldsymbol{\gamma}_{lk,l'k'}^{1:t}}, \hat{\mu_{lk}}\right\} \quad \text{where,}$$

$$\hat{\mu_{lk}} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}p_{l''k''}^{t+1} + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}p_{l''k}^{t+1} + 1\right)}}. \tag{96}$$

Overall, the modified stochastic WMMSE algorithm is summarized in Algorithm 2.

Algorithm 2: sum SE maximization using Stochastic WMMSE approach

---

Input: Given a tolerance $\epsilon > 0$, the maximum number of iterations N and maximum power constraint P_l^{max} for UE U_{lk} and maximum power constraint Q_{lk}^{max} 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 $\mu_{lk}^{(1)} \sim \mathbb{U}[0, Q_{lk}^{max}]$

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 (77)

    Update the auxiliary variable $w_{lk,n}^{ij}$ using (76) Do until convergence **if** $\left(\hat{R}_{sum}^i(p,\mu) - \left\{\sum\limits_{l=1}^{L}\sum\limits_{k=1}^{K}\sum\limits_{n=1}^{\mathbb{U}_{lk}} w_{lk,n}^{ij+1}e_{lk,n}^{ij+1} - \log(w_{lk,n}^{ij+1})\right\}\right) < \epsilon$ **then**

     break.

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

  Compute p_{jk}^{i+1} using (60)

  Update the $\hat{\mu}_{lk}$ constraint variable using (67)

  Compute μ_{lk}^{i+1} using (66)

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

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

  break.

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

V. 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 **H** (13). And especially precoding combiners such as Ia-ZF or Ia-MMSE requires a considerably large computation complexity and also requires a 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.*

1) 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).

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

$$\mathbf{x}^l = \sum_{k=1}^{K} \sum_{n=1}^{\mathcal{U}_{jk}} \mathbf{w}_{lk,n} s_{lk,n} \tag{97}$$

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

2) Ergodic SE analysis: 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) and (97), 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:

$$\mathbf{x}^l = \sum_{k=1}^{K} \cancel{\mathbf{w}_{lk}} \sum_{n=1}^{\mathcal{U}_{jk}} \frac{\mathbf{w}_{lk,n}}{\cancel{\sqrt{p_{lk,n}}}} s_{lk,n} \tag{98}$$

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

$$\overline{R}^p_{lk,n\text{sum}} = \sum_{l=1}^{L} \sum_{k=1}^{K} \sum_{n=1}^{\mathbb{U}_{lk}} \mathbb{E}\left[\frac{1}{2}\left(1 - \frac{\tau}{\tau_c}\right) \log_2\left(1 + \frac{\overline{\Delta}^p_{lk,n}}{\overline{\Omega}^p_{lk,n}}\right)\right], \text{ where } \overline{\Omega}^p_{lk,n} = \sum_{m=0}^{7} \overline{I}^{(m)^t}_{lk,n} + 1, \tag{99}$$

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

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

$$\overline{I}^{(7)}_{k,n} = \sum_{l'=1}^{L}\sum_{k'=1}^{K}\overline{C}^{(5)}_{l'k',lk,n}\tilde{\mu}^2_{l'k'},\ \text{and}$$

$$\overline{\mu}_{lk} = \sqrt{\frac{q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}\sum_{n'=1}^{\mathcal{U}_{l''k''}}|\mathbf{w}_{l''k'',n'}|^2 + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}\sum_{n'=1}^{\mathcal{U}_{l''k}}|\mathbf{w}_{l''k,n'}|^2 + 1\right)}}. \tag{100}$$

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

$$\overline{A}_{lk,n} = f_{lk,n}g^{lk}_{lk,n}\mathbf{h}^{l^H}_{lk},\ \overline{C}^{(1)}_{lk,n} = f_{lk,n}g^{lk}_{lk,n}\mathbf{h}^{l^H}_{lk},\ \overline{C}^{(2)}_{lk,n} = f_{lk,n}g^{lk}_{lk,n}\mathbf{h}^{l^H}_{lk} - f_{lk,n}\hat{g}^{lk}_{lk,n}\hat{\mathbf{h}}^{l^H}_{lk},$$

$$\overline{C}^{(3)}_{l'k',lk,n} = f_{lk,n}g^{l'k'}_{lk,n}\mathbf{h}^{l'^H}_{l'k'},\ \overline{C}^{(4)}_{\substack{l''k'',l'k',\\ lk,n}} = f_{lk,n}g^{l'k'}_{lk,n}\mathbf{h}^{l''^H}_{l'k'},\ \overline{C}^{(5)}_{l'k',lk,n} = |f_{lk,n}g^{l'k'}_{lk,n}|^2 \tag{101}$$

3) 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:

$$\mathbf{P1}_{pre}:\quad \underset{\overline{\mathbf{W}},\boldsymbol{\mu}}{\text{Maximize}}\ \lim_{T\Rightarrow\infty}\frac{1}{T}\sum_{t=1}^{T}\left[\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathcal{U}_{lk}}\log_2\left(1+\frac{\overline{\Delta}^p_{lk,n}(\mathcal{F}^t)}{\overline{\Omega}^p_{lk,n}(\mathcal{F}^t)}\right)\right] \tag{102}$$

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

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}\sum_{n'=1}^{\mathcal{U}_{l''k''}}|\mathbf{w}_{l''k'',n'}|^2 + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}\sum_{n'=1}^{\mathcal{U}_{l''k}}|\mathbf{w}_{l''k,n'}|^2 + 1\right)}}. \tag{103}$$

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:

$$\mathbf{P2}_{pre}: \quad \underset{\overline{\mathbf{W}},\boldsymbol{\mu}}{\text{Maximize}} \lim_{T\Rightarrow\infty}\frac{1}{T}\sum_{t=1}^{T}\left[\underset{\mathbf{U}^t,\mathbf{W}^t}{\text{Maximize}}\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}} w_{lk,n}^{t}e_{lk,n}^{t}-\log(w_{lk,n}^{t})\right] \tag{104}$$

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

$$\hat{\mu}_{lk}=\sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}\sum_{n'=1}^{\mathcal{U}_{l''k''}}|\mathbf{w}_{l''k'',n'}|^2+\sum\limits_{l''=1}^{L}\xi_{l''k,lk}\sum_{n'=1}^{\mathcal{U}_{l''k}}|\mathbf{w}_{l''k,n'}|^2+1\right)}}. \tag{105}$$

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:

$$u_{lk,n}^{t*}=\frac{\sqrt{\overline{\Delta}_{lk,n}^{i}(\mathcal{F}^t)}}{\overline{\Omega}_{lk,n}^{i}(\mathcal{F}^t)} \tag{106}$$

$$w_{lk,n}^{t*}=\frac{1}{e_{lk,n}^{t*}} \text{ where,} \tag{107}$$

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

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:

$$\mathbf{P3}_{pre}: \quad \underset{\overline{\mathbf{W}},\boldsymbol{\mu}}{\text{Maximize}} \lim_{T\Rightarrow\infty}\frac{1}{T}\sum_{t=1}^{T}\left[\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}} w_{lk,n}^{t**}e_{lk,n}^{t**}-\log(w_{lk,n}^{t**})\right] \tag{108}$$

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

$$\hat{\mu}_{lk}=\sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}\sum_{n'=1}^{\mathcal{U}_{l''k''}}|\mathbf{w}_{l''k'',n'}|^2+\sum\limits_{l''=1}^{L}\xi_{l''k,lk}\sum_{n'=1}^{\mathcal{U}_{l''k}}|\mathbf{w}_{l''k,n'}|^2+1\right)}}. \tag{109}$$

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 (108) 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:

$$\tilde{R}_{\text{sum}}(\overline{\mathbf{W}},\boldsymbol{\mu},\mathcal{F}^t)=\sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\log(w_{lk,n}^{t**})-w_{lk,n}^{t**}e_{lk,n}^{t**} \tag{110}$$

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 $\tilde{\mathbb{R}}^{1:t}_{\text{sum}}$ is expressed as:

$$\mathbf{R}^{1:t}(\bar{\mathbf{W}},\boldsymbol{\mu}) = \frac{1}{t}\left[\mathbf{R}^{1:t-1}(\bar{\mathbf{W}},\boldsymbol{\mu})\right] + \frac{t-1}{t}\left[\tilde{R}_{\text{sum}}(\overline{\mathbf{W}},\boldsymbol{\mu},\mathcal{F}^t)\right] \tag{111}$$

To make the above equation simpler, we try to group the auxiliary variable $e^{t**}_{lk,n}$ present in the equation (113), based on percoder $\mathbf{w}_{lk,n}$ which is presented as:

$$w^{t**}_{lk,n}e^{t**}_{lk,n} = cons(w^{t**}_{lk,n}) + \sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}} \mathbf{w}^H_{l'k',n'}\boldsymbol{\alpha}^t_{l'k',n',lk,n}\mathbf{w}_{l'k',n'} + \boldsymbol{\beta}^t_{lk,n}\mathbf{w}_{lk,n}. \tag{112}$$

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:

$$\tilde{R}_{\text{sum}}(\overline{\mathbf{W}},\boldsymbol{\mu},\mathcal{F}^t) = \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\left(const(w) - \sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}} \mathbf{w}^H_{l'k',n'}\boldsymbol{\alpha}^t_{l'k',n',lk,n}\mathbf{w}_{l'k',n'} + \beta^t_{lk,n}\mathbf{w}_{lk,n}\right) \tag{113}$$

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:

$$\mathbf{R}^{1:t}(\bar{\mathbf{W}},\boldsymbol{\mu}) = \sum_{l=1}^{L}\sum_{k=1}^{K}\sum_{n=1}^{\mathbb{U}_{lk}}\left(const(w) - \sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}} \mathbf{w}^H_{l'k',n'}\check{\boldsymbol{\alpha}}^{1:t}_{l'k',n',lk,n}\mathbf{w}_{l'k',n'} + \check{\boldsymbol{\beta}}^{lk,n}_{1:t}\mathbf{w}_{lk,n}\right) \text{ where,} \tag{114}$$

$$\check{\boldsymbol{\alpha}}^{1:t}_{l'k',n',lk,n} = \frac{1}{t}\boldsymbol{\alpha}^t_{l'k',n',lk,n} + \frac{t-1}{t}\check{\boldsymbol{\alpha}}^{1:t-1}_{l'k',n',lk,n} \tag{115}$$

$$\check{\boldsymbol{\beta}}^{1:t}_{lk,n} = \frac{1}{t}\boldsymbol{\beta}^t_{lk,n} + \frac{t-1}{t}\check{\boldsymbol{\beta}}^{1:t-1}_{lk,n} \tag{116}$$

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:

$$\mathbf{P4}_{pre}:[\overline{\mathbf{W}}^{t+1},\boldsymbol{\mu}^{t+1}] \Leftarrow \underset{\overline{\mathbf{W}},\boldsymbol{\mu}}{\text{Argmax}}\ \frac{1}{t}\left[\mathbf{R}^{1:t-1}(\bar{\mathbf{W}},\boldsymbol{\mu})\right] + \frac{t-1}{t}\left[\tilde{R}_{\text{sum}}(\overline{\mathbf{W}},\boldsymbol{\mu},\mathcal{F}^t)\right] \triangleq \mathbf{R}^{1:t}(\bar{\mathbf{W}},\boldsymbol{\mu}) \tag{117}$$

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

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

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:

$$\mathbf{w}_{lk,n}^{t+1} = \left\{\left(\sum_{l'=1}^{L}\sum_{k'=1}^{K}\sum_{n'=1}^{\mathbb{U}_{lk}} \check{\boldsymbol{\alpha}}_{lk,n,l'k',n'}^{1:t}\right) + \lambda_l^* I_N\right\}^{-1} \left\{\left(\boldsymbol{\beta}_{lk,n}^{1:t}\right)\right\} \tag{118}$$

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 (112) is grouped with respect to $\boldsymbol{\mu}$ wise, the term becomes:

$$w_{lk,n}^{t**} e_{lk,n}^{t**} = const(w) + \sum_{l'=1}^{L}\sum_{k'=1}^{K} \gamma_{l'k',lk}^{t} \mu_{lk}^2 + \omega_{lk}^t \mu_{lk}. \tag{119}$$

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 (114) and updation of the surrogate variables (115), the $\mu_{lk,n}$ is optimized through first-order differentiation of new surrogate variable keeping precoder constant and equating to zeros is expressed as:

$$\overline{\mu}_{lk}^{t+1} = \min\left\{\frac{\omega_{lk}^{1:t}}{\sum\limits_{l'=1}^{L}\sum\limits_{k'=1}^{K}\gamma_{lk,l'k'}^{1:t}}, \hat{\mu_{lk}}\right\} \text{ where,} \tag{120}$$

$$\hat{\mu}_{lk} = \sqrt{\frac{Q_{lk}}{\left(\sum\limits_{l''=1}^{L}\sum\limits_{k''=1}^{K}\rho_{l''k'',lk}\sum_{n'=1}^{\mathcal{U}_{l''k''}}|\mathbf{w}_{l''k'',n'}|^2 + \sum\limits_{l''=1}^{L}\xi_{l''k,lk}\sum_{n'=1}^{\mathcal{U}_{l''k}}|\mathbf{w}_{l''k,n'}|^2 + 1\right)}}. \tag{121}$$

The overall Algorithm, the modified stochastic successive convex approximation is summarized in Algorithm 3.

VI. Simulation results

In this section, We now evaluate the optimization techniques of the multiple-relay-aided mMIMO NOMA system, where the BS and users estimate CSI, and the users perform imperfect SIC. We consider a 20 MHz system with $N_T = 100$ BS antennas, $K = 5$ relays and a total of 20 users. We assume that i) relays and users are randomly allocated on a circle of the radius 500m, with the BS as center; and ii) each relay is allocated four users, i.e. $U_k = 4$. Each coherence interval is of $\tau_c = 200$ symbols, with a pilot transmission interval of $\tau = K$ symbols. The pilot power is $p_p = 20$ dBm. 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, $d_{\substack{l\prime\\lk}}$ 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 $K_{lk,n}^{l'k'}$

Algorithm 3: sum SE maximization with Precoder allocation

Input: Given a tolerance $\epsilon > 0$, the maximum number of iterations N and maximum power constraint P_l^{max} for UE U_{lk} and maximum power constraint Q_{lk}^{max} 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 $\mu_{lk}^{(1)} \sim \mathbb{U}[0, Q_{lk}^{max}]$

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 $u_{lk,n}^{ij}$ using (**??**)

    Update the auxiliary variable $w_{lk,n}^{ij}$ using (107) Do until convergence **if** $\left\{ \sum_{l=1}^{L} \sum_{k=1}^{K} \sum_{n=1}^{\mathbb{U}_{lk}} \left(w_{lk,n}^{ij} e_{lk,n}^{ij} - \log(w_{lk,n}^{ij}) \right) \right) - \left(w_{lk,n}^{ij+1} e_{lk,n}^{ij+1} - \log(w_{lk,n}^{ij+1}) \right) \right\} < \epsilon$ **then**

     break.

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

  Compute p_{jk}^{i+1} using (118)

  Update the $\hat{\mu}_{lk}$ constraint variable using (121)

  Compute μ_{lk}^{i+1} using (120)

  Update $\bar{\alpha}^i, \bar{\beta}^i, \bar{\gamma}^i, \bar{\omega}^i$ using equations (115)

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

   break.

return $\mathbf{w}^{*}$,μ^{*}

can be calculated as $K_{lk}^{l'}[\text{dB}] = 13 - 0.03 d_{lk}^{l'}$ and $K_{lk,n}^{l'k'}[\text{dB}] = 13 - 0.03 d_{lk,n}^{l'k'}$ respectively.The other parameters is the assumption of ULA at the base station and the ASD of 10^o.

1) Deterministic sum SE optimization:

2) Stochastic sum SE optimization:

3) sum SE optimization with Precoder allocation:

VII. Conclusion

References

[1] L. Sanguinetti, E. Björnson, and J. Hoydis, "Toward massive MIMO 2.0: Understanding spatial correlation, interference suppression, and pilot contamination," *IEEE Trans. Commun.*, vol. 68, no. 1, pp. 232–257, Jan. 2020.

[2] Mandawaria, E. Sharma, and R. Budhiraja, "Energy-efficient massive MIMO multi-relay NOMA systems with CSI errors," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7410–7428, 2020.

[3] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO Networks: Spectral, energy, and hardware efficiency," *Found. and Trends® in Signal Process.*, vol. 11, no. 3-4, pp. 154–655, 2017.

[4] X. Chen, R. Jia, and D. W. K. Ng, "On the design of massive non-orthogonal multiple access with imperfect successive interference cancellation," *IEEE Trans. Commun.*, vol. 67, no. 3, pp. 2539–2551, 2019.

[5] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1997.

[6] Y. Li and G. A. A. Baduge, "Relay-aided downlink massive MIMO NOMA with estimated CSI," *IEEE Trans. Veh. Technol.*, vol. 70, no. 3, pp. 2258–2271, 2021.

[7] D. N. Amudala, B. Kumar, and R. Budhiraja, “Spatially-correlated rician-faded multi-relay multi-cell massive mimo noma systems,” *IEEE Transactions on Communications*, pp. 1–1, 2022.

[8] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, ninth Dover printing ed. New York: Dover, 1964.

[9] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, “Weighted sum-rate maximization using weighted mmse for mimo-bc beamforming design,” *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792–4799, 2008.

[10] M. Razaviyayn, M. Sanjabi, and Z.-Q. Luo, “A stochastic successive minimization method for nonsmooth nonconvex optimization with applications to transceiver design in wireless communication networks,” 2013. [Online]. Available: https://arxiv.org/abs/1307.4457

[11] A. Liu, V. K. N. Lau, and B. Kananian, “Stochastic successive convex approximation for non-convex constrained stochastic optimization,” *IEEE Transactions on Signal Processing*, vol. 67, no. 16, pp. 4189–4203, 2019.