IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received July 19, 2019, accepted August 8, 2019, date of publication August 19, 2019, date of current version September 24, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2935990

Interference Management in LTE-Advanced Cooperative Relay Networks: Decentralized Transceiver Design With Channel Estimation

ARMELINE DEMBO MAFUTA[1], TOM WALINGO[1], AND FAMBIRAI TAKAWIRA[2], (Member, IEEE)

[1]School of Electrical, Electronic, and Computer Engineering, University of KwaZulu-Natal, Durban 4000, South Africa
[2]School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg 2050, South Africa

Corresponding author: Armeline Dembo Mafuta (armelinedembo@gmail.com)

ABSTRACT Wireless networks improve indoor communications by deploying femtocell networks into the macrocell coverage. This results in spectrum sharing with the consequences of cross-tier interference from the macrocell user equipment (MUEs) to the femtocell access points (FAPs). This work considers the uplink cross-tier interference management for the cell-edge MUEs (CUEs) in cooperative multi-user multiple input multiple output (MU-MIMO) systems. For better interference management, the CUEs are grouped into clusters and communicate to the macrocell base station (MBS) through a relay node (RN). The linear pre-coders and decoders algorithms for the FAPs, MUEs and CUEs are proposed for effective interference management to minimize the sum mean square error (MSE), subject to the total transmit power constraints. The designed pre-coders and decoders use the pilot-assisted channel estimation to improve the accuracy of the acquired channel state information (CSI). The least-square (LS) and minimum MSE (MMSE) channel estimators are considered. The performance of the system is investigated in terms of the bit error rate (BER) for the linear pre-coders and decoders algorithms with the pilot-assisted channel estimators.

INDEX TERMS Channel estimation, decentralized transceiver, femtocells, long term evolution-advanced, relay node, uplink communication.

I. INTRODUCTION

Femtocell networks are deployed into macrocell networks to improve the indoor coverage and provide high data rates to end users while reducing their load. Femtocells do not require specific infrastructure, they are easily installed by the end users. They operate in the licensed band of the macrocell and in some cases, they are imposed on the same frequency spectrum [1]. This results in the challenge of cross-tier interference between macrocell and femtocells when both transmit on the same frequency band simultaneously. Furthermore, the macrocell users located at the cell-edge, referred in this paper, as cell-edge macrocell user equipment (CUEs), experience performance degradation due to the long distance between the CUEs and the macrocell base station (MBS). The management of cross-tier interference from CUEs to the femtocell access point (FAPs) is of paramount importance and is part of the focus of this work.

The associate editor coordinating the review of this article and approving it for publication was Ahmed Mohamed Ahmed Almradi.

Several techniques have been employed to mitigate this interference. They include: Interference alignment, where the signals are constrained into the same subspaces at the unintended receivers and the desired signals are retrieved at each receiver by eliminating the aligned interferences, using a decoding matrix [2]; interference avoidance, where the allocation of various system resources to users is controlled to ensure that the interference remains within acceptable limits [3]; and interference cancellation, where the suppression of the interference can be done at the transmitter or receiver side [4]. Another effective interference management technique is the pre-coding approach, which provides reliable high data rate communication in multi-user multiple input multiple output (MU-MIMO) systems. A linear pre-coder and decoder design, also known as transceiver design, is an effective way to reduce or mitigate multi-user interference in femtocell networks while improving the performance [5]. It can be employed to enhance the bit error rate (BER) performance and increase the information rate of spatial multiplexed MU-MIMO systems. Moreover, minimum mean

square error (MMSE) or zero-forcing (ZF) can be used to design the pre-coder at the source and decoder at the destination to estimate the transmitted signal. In this paper, MMSE is employed in the transceiver design, due to its simplicity and effectiveness [6]. It is also known that MMSE mitigates both interference and noise compared to other linear detectors such as ZF, which cancels interference but enhances noise power [7].

The performance of the linear pre-coder and decoder is heavily dependent on the availability of timely channel state information (CSI) at both transmitter and receiver [8]. Although, a non-linear pre-coder and decoder can provide good performance, it is complex in its design [6]. At the receiver, the channel is estimated using known pilot symbols. Channel estimation techniques can either be least square (LS) based, MMSE based or maximum likelihood (ML) based [9].

To improve the coverage area, network capacity and reliability of the link in a MU-MIMO system [6], [10], cooperative relays are incorporated into the system as one of the interference management techniques. With this technique, the relay node (RN) acts as a bridge that facilitates the cooperative communication and retransmits the signal received from a CUE to the MBS, interference-free. The CUE's high signal power causes interference to the neighbouring FAPs. Hence, an effective interference management technique is required to optimize the network lifetime. The RNs improve the transmit signal of the CUEs and maintain good communication to the MBS within the cluster by relaying their signals to the MBS. They should mitigate the cross-tier interference to the FAPs to enhance the performance of the CUEs and FAPs in the MU-MIMO relay systems. Cooperative technique enables a FAP to gather information about its neighbouring femtocells and performs its allocation by considering its effects on the neighbours. This increases the average femtocells' throughput and quality of services (QoS) as well as its global performance, which are locally optimized. The MU-MIMO relay networks are heterogeneous featuring FAP, RN and MBS networks with their respective users, the half-duplex communication network with the CUEs, RNs, FUEs, MUEs in multi-slot transmission. For such a distributed network, a centralized or joint pre-coder and decoder design is not appropriate or feasible. Therefore, decentralized algorithms for the FAPs, MUEs, RNs and CUEs are considered in the design of optimal pre-coders and decoders based on the CSI. Each FAP manages its own sub-channels for suitable performance. The pre-coder for the RNs is also designed based on the CSI for the MU-MIMO relay networks considering amplify-and-forward (A-F) technique at the RN. In this work, the pilot-assisted LS and MMSE channel estimators are employed in the MU-MIMO networks, due to their inherent advantages of low complexity and good mean square error (MSE) performance.

A. RELATED RESEARCH

Different techniques for interference mitigation in femtocells have been proposed. The authors in [11] studied the interference mitigation techniques in femtocells/macrocell networks where a frequency reuse mechanism that increases overall system performance was proposed. Clustering algorithm and cooperative relay schemes have been used in multiple interference management schemes for a better radio resource allocation, interference management or power control [12]. The authors in [13], [14] investigated a management of cross-tier interference, where a novel femtocell clustering based on interference cancellation (IC) was introduced. A distributed antenna system was also used to mitigate cross-tier interference between macrocell and femocells. The authors in [15] proposed a scheme called IC-relay time division multiple access (TDMA), which allows multi-user concurrent transmission in the source relay link. They aimed to cancel interference at the multi-antenna relay by linear IC techniques. A MIMO relaying system with fixed relay networks was introduced in [16] for IC. Authors in [17] analyzed and designed an A-F, decode-forward and demodulation-forward relay protocol and discussed IC in the MU-MIMO environment.

Several pilot-aided schemes have been investigated to enhance the channel estimation accuracy in the MU-MIMO systems. This guarantees the performance of linear pre-coder and decoder designs. However, pilot contamination is one of the limitations of this technique. Authors in [18] provided an explicit expression of the massive MIMO user capacity in the pilot contaminated regime where the number of users is larger than the pilot sequence length. Authors in [19] proposed a channel estimation scheme for a massive MIMO which does not require the knowledge of the inter-cell large fading coefficient, thus no overload. An iterative soft decision IC has been investigated in multi-cell multi-user massive MIMO with pilot contamination [20] where only the MMSE was considered. Authors in [21] designed a pilot contamination pre-coding which maximizes the minimum SINR subject to the network sum power constraint for interference reduction. Most of these works assumed the perfect CSI at the transceiver side, whereas in practice CSI is prone to errors due to different factors. In [22] and [23], the authors did not only consider perfect CSI, but also the channel uncertainty and/or Imperfect CSI. This work considers imperfect CSI for a more realistic scenario and, the pilot-aided scheme is also considered for synchronisation and channel estimation purposes to design the transceivers.

The linear pre-coder and decoder designs have also been considered as interference management techniques. In [24], with the assumption of perfect CSI, the authors employed an optimization technique for the design of optimal source, relay and receiver in uplink MU-MIMO relay communication systems in order to minimize the MSE of the estimated signal at the destination. The pilot-aided channel estimation is not done in [24]. Channel uncertainties were considered in [25], where a robust transceiver design for a general MU-MIMO relay was studied in the presence of statistical CSI errors. Imperfect CSI was also considered in [26], where proposed joint linear pre-coder and decoder designs

for downlink and uplink were compared to a conventional joint linear pre-coder and decoder design in MU-MIMO systems. The authors in [27] and [28] designed algorithms that converge to the optimum pre-coders and decoders for users in a MU-MIMO system. Moreover, the authors in [28] introduced interference alignment to help the femtocell user equipment (FUEs) to eliminate the cross-tier interference by aligning the MUE interference signal, subject to individual SINR constraints at their MBS. Interference alignment transceiver is designed in [29] to minimize the maximum MSE for multicell MU-MIMO wireless communication systems where a robust Min-Max MSE algorithm is proposed to counter the channel uncertainty. Transceiver designs with imperfect CSI were addressed in [30] for a MIMO relaying system where a near-optimal closed-form solution is provided for the source-to-relay-destination transceiver designs with imperfect CSI at all nodes. Beamforming has been considered as one of the approach for interference management in [31]. The authors in [32] proposed semi-decentralized beamforming design to minimize the total transmit power under SINR and interference constraints on multiple-input single-output (MISO) channels in HetNets where the beamformer is designed in a distributed fashion using the signal-to-leakage-and-noise ratio (SLNR) criteria. However, in the aforementioned research works, clustering, relaying and decentralized algorithms for the pre-coders and decoders design are not considered together. In this paper, the advantages of each of the aspects is employed to achieve a better interference management transceiver for the proposed distributed system and well instigated.

B. MAIN CONTRIBUTIONS

In this paper, we present a cooperative relay interference management technique where decentralized algorithms for linear pre-coders and decoders design based on pilot-assisted channel estimation are employed. The main contributions of this paper are summarized as follows:

- This paper extends the system model presented in [28] and [33] by introducing cooperative RN systems to manage the cross-tier interference caused by the CUEs to FAPs [10], while providing further performance enhancements for the CUEs as well as increasing the coverage of macrocell networks. Numerical evaluations are provided to prove the benefit of this cooperative RN extension over a simple MU-MIMO system.
- In pratical networks, a perfect CSI is impossible to obain due to limitation of channel estimation [34]. The authors in [28] considered a perfect CSI, which is not a practical scenario. This motivates us to investige the each sub-optimal transceiver design (FUEs, MUEs, CUEs and RNs) in the presence of channel estimation errors.
- The MMSE methodology considered is similar to [28]. However, instead of considering only a joint design for all the pre-coders, we consider decentralized transceiver designs for the MU-MIMO relay systems. We divide the optimization problem into sub-optimal problems where we consider four different transceiver designs, the FUEs, MUEs, CUEs and RNs. Therefore, decentralized approach is considered in the design of pre-coders and decoders at the FAPs and MBS during the first and second time slots. Furthermore, the transceivers are designed with the estimated channels and are iteratively updated until their optimal values are found. Finding the optimal values for these pre-coders and decoders depends on the Lagrange multipliers.

C. ORGANISATION, NOTATIONS AND LIST OF SYMBOLS

Organisation: The remaining sections of the paper are organised as follows: Section II describes the system model, problem formulation, proposed network architecture, uplink training and channel estimation as well as the uplink transmission designs. The decentralized algorithms for the linear transceiver designs with MMSE approach for the FAPs, MUEs and CUEs are presented in Section III. Section IV describes the performance evaluation. Section V presents the conclusion of the paper.

Notations: We use upper-case bold letters for matrices and lower-case bold letters for vectors. $(\cdot)^H$, $(\cdot)^T$, $(\cdot)^*$ and $(\cdot)^{-1}$ denote the Hermitian, transpose, optimal and inverse of matrices respectively. $\mathbf{I}_N$ is a $N \times N$ identity matrix and $\mathbb{E}[\cdot]$ denote the expectation. $\|\cdot\|$ is the norm of a vector or complex scalar and tr(.) represents the trace of a matrix.

List of Symbols: Table 1 gives the list of some symbols used in this paper. The omitted symbols are all defined in the paper.

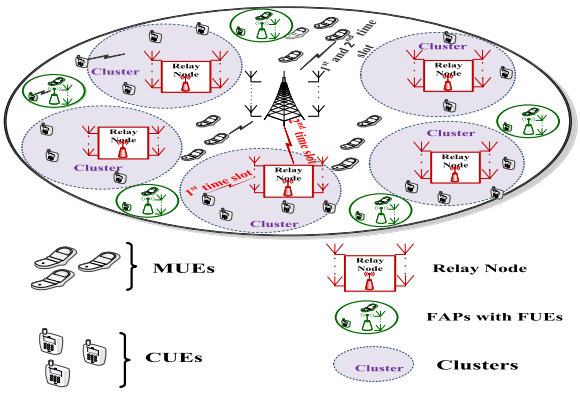

FIGURE 1. The network architecture with a single MBS, MUEs, RNs clusters with CUEs, FAPs with FUEs.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. THE NETWORK ARCHITECTURE

The network architecture consists of half-duplex multi-user relay Long Term Evolution-Advanced (LTE-Advanced) femtocell networks deployed into a macrocell network. The CUEs considered are grouped into clusters and communicate to the MBS through a RN. The FUEs and the CUEs transmit during the first time slot while the RNs transmit to the MBS during the second time slot. The MUEs, on the other hand, transmit continuously to the MBS during both first and second time slot. The network architecture is illustrated in Fig. 1.

TABLE 1. Table of symbols.

Symbols	Description	Symbols	Description
N_s, N_f, N_R, N_B	Number of antennas at the users (FUEs, MUEs and CUEs), at the FAPs, at RNs and at the MBS, respectively	$\hat{H}_{ji}^{\text{FAP-MMSE}}, \hat{H}_{rk}^{\text{CUE-MMSE}}$	MMSE Channel estimators for the i^{th} FUE of the j^{th} FAP and the k^{th} CUE of the r^{th} RN, respectively
F, U	Number of FAPs and number of FUEs in the f^{th} femtocells, respectively	R, K	Set of RNs and number of CUEs in the r^{th} RN, respectively
$\hat{H}_{ji}^{\text{FAP-LS}}, \hat{H}_{rk}^{\text{CUE-LS}}, \hat{H}_{om}^{\text{MUE-LS}}$	LS estimators for the i^{th} FUE of the j^{th} FAP, the k^{th} CUE of the r^{th} RN and the m^{th} MUE of the MBS, respectively	$n_j^{\psi}, n_r^{\psi}, n_o^{\psi}$	Vectors representing additive white Gaussian noise (AWGN) at the j^{th} FAP, at the r^{th} RN and at the MBS, respectively
M	Number of MUEs outside clusters	$\psi_{om}^{\text{MUE}}, \psi_{rk}^{\text{CUE}}$	Pilot symbols of the m^{th} MUE of the MBS and k^{th} CUE of the r^{th} RN, respectively
$y_j^{\Psi}, y_r^{\psi}, y_o^{\psi}$	Received pilot signals at the j^{th} FAP, at the r^{th} RN and at the MBS, respectively	$\hat{H}_{om}^{\text{MUE-MMSE}}$	MMSE Channel estimator for the m^{th} MUE of the MBS
$P_{ji}^{\text{FAP}}, P_{fu}^{\text{FAP}}$	Transmit powers of the i^{th} FUE of the j^{th} FAP and the u^{th} FUE of the f^{th} FAP, respectively	$y_j^{\text{FAP}'}, y_o^{\text{MBS}'}, y_r^{\text{RN}'}$	Received signal at the j^{th} FAP, received signal of the MBS and the received signal at the r^{th} RN during the first time slot with the ZF pre-coder assumption and an estimation error
$P_{om}^{\text{MUE}}, P_{rk}^{\text{CUE}}$	Transmit powers of the m^{th} MUE of the MBS and the k^{th} CUE of the r^{th} RN, respectively	$s_{ji}^{\text{FAP}}, s_{om}^{\text{MUE}}, s_{fu}^{\text{FAP}}, s_{rk}^{\text{CUE}}$	Messages of the i^{th} user of the j^{th} FAP, m^{th} MUE of the MBS, of the u^{th} FUE of the f^{th} FAP and of the k^{th} CUE of the r^{th} RN, respectively
$H_{ji}^{\text{FAP}}, H_{jfu}^{\text{FAP}}$	Channel matrices of the i^{th} user of the j^{th} FAP and from the u^{th} user of the f^{th} femtocell to the j^{th} FAP, respectively	$w_{ji}^{\text{FAP}}, w_{om}^{\text{MUE}}, w_{fu}^{\text{FAP}}, w_{rk}^{\text{CUE}}$	Pre-coding vectors of the i^{th} FUE of the j^{th} femtocell, of the m^{th} MUE of the MBS, of the u^{th} FUE of the f^{th} FAP and of the k^{th} CUE of the r^{th} RN, respectively
$\mathbf{H}_{jm}^{\text{MUE}}, H_{jrk}^{\text{CUE}}$	Channel matrices from m^{th} MUE to the j^{th} FAP and from k^{th} CUE of the r^{th} RN to the j^{th} FAP, respectively	F_{or}, x_{or}, H_{or}	RN pre-coders during the second time slot, amplified r^{th} transmit signal to the MBS during the second time slot and channel matrix from the r^{th} RN to the MBS, respectively
$\mathbf{H}_{rk}^{\text{CUE}}, \mathbf{H}_{rfu}^{\text{FAP}}, \mathbf{H}_{rm}^{\text{MUE}}$	Channel matrices of the k^{th} CUE of the r^{th} RN, from the u^{th} FUE of the f^{th} femtocell to the r^{th} RN and from the m^{th} MUE to r^{th} RN, respectively	$L_{ji}^{\text{FAP}}, L_{fu}^{\text{FAP}}, L_{om}^{\text{MUE}}, L_{rk}^{\text{CUE}}$	Model the distance in slow fading of the i^{th} FUE of the j^{th} FAP, the u^{th} FUE of the f^{th} FAP, the m^{th} MUE of the MBS and k^{th} CUE of the r^{th} RN, respectively
P_r, L_r	The transmit power at the RN and the propagation loss at the r^{th} RN, respectively	$d_{jx}^{\text{FAP}}, d_{ol}^{\text{MUE}}$	The decoding vectors for x^{th} FUE of j^{th} FAP and for l^{th} MUE of MBS, respectively
$\mathbf{H}_{om}^{\text{MUE}}, H_{ork}^{\text{CUE}}, H_{ofu}^{\text{FAP}}$	Channel matrix from the m^{th} MUE of the MBS and channel gain from k^{th} CUE of the r^{th} RN to the MBS and channel matrix from the u^{th} user of the f^{th} FAP to the MBS, respectively	D_o, F_o, W^{UE}	The decoder matrices, relay pre-coder matrices for the UEs (CUEs and MUEs) and UE pre-coder matrices during the second time slot.
$y_o^{\text{MBS-nd}'}$	The received signal at the MBS during the second time slot	$\psi_{ji}^{\text{FAP}}, \psi_{fu}^{\text{FAP}}$	Pilot symbols of the i^{th} FUE of the j^{th} FAP and u^{th} FUE of the f^{th} FAP, respectively

The RN creates a cooperative communication between the CUEs and the MBS. Furthermore, RNs enable a cross-tier interference management to the neighbouring FAPs.

B. UPLINK TRAINING AND CHANNEL ESTIMATION

We consider an uplink transmission where all users share the same time-frequency resource. To detect the transmitted signal from the users, the base stations use the CSI knowledge acquired through uplink training. We assume that the channel remains constant during the training phase in order to analyze the system performance. For the channel estimation purposes, we use pilot symbols (a set of symbols whose location and values are known to the receiver) multiplexed with the information-bearing data [35]. During the uplink training

phase, the users (FUEs, MUEs and CUEs), transmitting to the base stations (FAPs, MBS and RNs), are assigned pilot sequences.

Let N_s, N_f, N_R and N_B represent the antennas at the users (FUEs, MUEs and CUEs), FAPs, RNs and MBS, respectively. We denote F as the number of FAP and U the number of FUEs in the f^{th} femtocells. Let R be the set of RN in each cluster, K the number of CUEs in the r^{th} RN and M the number of MUEs outside the cluster. Ψ denotes the pilot sequence matrix transmitted from the users to their base stations or access point. The pilot sequence matrix Ψ satisfies $\Psi\Psi^H = I$. Let U_i be the number of FUEs in the j^{th} FAP. Thus, the received pilot signal y_j^{ψ} at the j^{th} FAP is written as

$$y_j^{\psi} = \underbrace{\sum_{i=1}^{U_i} \sqrt{P_{ji}^{\text{FAP}} L_{ji}^{\text{FAP}}} H_{ji}^{\text{FAP}} \psi_{ji}^{\text{FAP}}}_{\text{pilot signal from } j^{th} \text{ FAP users}} + \underbrace{\sum_{\substack{f=1 \\ f\neq j}}^{F} \sum_{u=1}^{U} \sqrt{P_{fu}^{\text{FAP}} L_{fu}^{\text{FAP}}} H_{jfu}^{\text{FAP}} \psi_{fu}^{\text{FAP}}}_{\text{pilot from other femtocell}} + \underbrace{\sum_{m=1}^{M} \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} \mathbf{H}_{jm}^{\text{MUE}} \psi_{om}^{\text{MUE}}}_{\text{pilot from MUE outside the cluster}} + \underbrace{\sum_{r=1}^{R} \sum_{k=1}^{K} \sqrt{P_{rk}^{\text{CUE}} L_{rk}^{\text{CUE}}} H_{jrk}^{\text{CUE}} \psi_{rk}^{\text{CUE}}}_{\text{pilot from the CUE in the RN}} + n_j^{\psi}, \quad (1)$$

where P_{ji}^{FAP} and P_{fu}^{FAP} are the transmit powers of the i^{th} FUE of the j^{th} FAP and the u^{th} FUE of the f^{th} FAP while P_{om}^{MUE} and P_{rk}^{CUE} are the transmit powers of the m^{th} MUE of the MBS and k^{th} CUE of the r^{th} RN. H_{ji}^{FAP} denotes the channel matrix of the i^{th} user of the j^{th} FAP. H_{jfu}^{FAP} is considered as the channel matrix from the u^{th} user of the f^{th} femtocell to the j^{th} FAP, $\mathbf{H}_{jm}^{\text{MUE}}$ is the channel matrix from m^{th} MUE to the j^{th} FAP and H_{jrk}^{CUE} is the channel matrix from k^{th} CUE of the r^{th} RN to the j^{th} FAP. It is worth mentioning that $L_{ji}^{\text{FAP}} H_{ji}^{\text{FAP}}$ is the propagation loss of the i^{th} FUE of the j^{th} FAPs while $L_{fu}^{\text{FAP}} H_{jfu}^{\text{FAP}}$ is the propagation loss of the u^{th} FUE of the f^{th} FAP. Similarly, $L_{om}^{\text{MUE}} \mathbf{H}_{jm}^{\text{MUE}}$ is the propagation loss of the m^{th} MUE of the MBS and $L_{rk}^{\text{CUE}} H_{jrk}^{\text{CUE}}$ is the propagation loss of the k^{th} CUE of the r^{th} RN. However, L_{ji}^{FAP}, L_{fu}^{FAP}, L_{om}^{MUE} and L_{rk}^{CUE} model the distance in slow fading while H_{ji}^{FAP}, H_{jfu}^{FAP}, $\mathbf{H}_{jm}^{\text{MUE}}$ and H_{jrk}^{CUE} model the fast Rayleigh fading. ψ_{ji}^{FAP} and ψ_{fu}^{FAP} are the pilot symbol of the i^{th} FUE of the j^{th} FAP and u^{th} FUE of the f^{th} FAP, respectively. ψ_{om}^{MUE} and ψ_{rk}^{CUE} are the pilot symbol of the m^{th} MUE of the MBS and k^{th} CUE of the r^{th} RN, respectively. n_j^{ψ} is the vector representing additive white Gaussian noise (AWGN) at the j^{th} FAP, where the AWGN vector satisfies $\mathbb{E}\{n_j^{\psi} n_j^{\psi H}\} = (\sigma_j^{\text{FAP}})^2 I_{N_f}$ in which $n_j^{\psi H}$ is the conjugate transpose of n_j and I_{N_f} denotes the identity matrix. The received pilot signal y_r^{ψ} at the r^{th} RN is written as

$$y_r^{\psi} = \underbrace{\sum_{k=1}^{K} \sqrt{P_{rk}^{\text{CUE}} L_{rk}^{\text{CUE}}} \mathbf{H}_{rk}^{\text{CUE}} \psi_{r,k}^{\text{CUE}}}_{\text{pilot from the } r^{th} \text{ RN users}} + \underbrace{\sum_{f=1}^{F} \sum_{u=1}^{U} \sqrt{P_{fu}^{\text{FAP}} L_{fu}^{\text{FAP}}} \mathbf{H}_{rfu}^{\text{FAP}} \psi_{fu}^{\text{FAP}}}_{\text{pilot from the femtocells}} + \underbrace{\sum_{m=1}^{M} \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} \mathbf{H}_{rm}^{\text{MUE}} \psi_{om}^{\text{MUE}}}_{\text{pilot from the other MUEs}} + n_r^{\psi}, \quad (2)$$

where $\mathbf{H}_{rk}^{\text{CUE}}$ is the channel matrix from the k^{th} CUE of the r^{th} RN. $\mathbf{H}_{rfu}^{\text{FAP}}$ is the channel matrix from the u^{th} FUE of the f^{th} femtocell to the r^{th} RN and $\mathbf{H}_{rm}^{\text{MUE}}$ is the channel matrix from m^{th} MUE to the r^{th} RN. n_r^{ψ} is the AWGN vector at the r^{th} RN that satisfies $\mathbb{E}\{n_r^{\psi} n_r^{\psi H}\} = (\sigma_r^{\text{CUE}})^2 I_{N_R}$. The received pilot signal y_o^{ψ} at the MBS is written as:

$$y_o^{\psi} = \underbrace{\sum_{m=1}^{M} \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} \mathbf{H}_{om}^{\text{MUE}} \psi_{om}^{\text{MUE}}}_{\text{pilot from all MUEs}} + \underbrace{\sum_{f=1}^{F} \sum_{u=1}^{U} \sqrt{P_{fu}^{\text{FAP}} L_{fu}^{\text{FAP}}} H_{ofu}^{\text{FAP}} \psi_{fu}^{\text{FAP}}}_{\text{pilot from all femtocells}} + \underbrace{\sum_{r=1}^{R} \sum_{k=1}^{K} \sqrt{P_{rk}^{\text{CUE}} L_{rk}^{\text{CUE}}} \mathbf{H}_{ork}^{\text{CUE}} \psi_{rk}^{\text{CUE}}}_{\text{pilot from all CUEs in the RNs}} + n_o^{\psi}, \quad (3)$$

where $\mathbf{H}_{om}^{\text{MUE}}$ is the channel matrix from the m^{th} MUE of the MBS. H_{ork}^{CUE} is the channel gain from k^{th} CUE of the r^{th} RN to the MBS and H_{ofu}^{FAP} is the channel matrix from the u^{th} user of the f^{th} FAP to the MBS. n_o^{ψ} is the AWGN vector at the MBS that satisfies $\mathbb{E}\{n_o^{\psi} n_o^{\psi H}\} = (\sigma_o^{\text{MUE}})^2 I_{N_B}$.

1) LS CHANNEL ESTIMATOR

The LS channel estimation $\hat{H}_{ji}^{\text{FAP-LS}}$ for the j^{th} FAP is given as

$$\hat{H}_{ji}^{\text{FAP-LS}} = \frac{y_j^{\psi}}{\psi_{ji}^{\text{FAP}} . \sqrt{P_{ji}^{\text{FAP}} L_{ji}^{\text{FAP}}}}. \quad (4)$$

The LS channel estimate $\hat{H}_{rk}^{\text{CUE-LS}}$ for r^{th} RN is obtained as follows

$$\hat{H}_{rk}^{\text{CUE-LS}} = \frac{y_r^{\psi}}{\psi_{rk}^{\text{CUE}} . \sqrt{P_{rk}^{\text{CUE}} L_{rk}^{\text{CUE}}}}. \quad (5)$$

Similarly, the LS channel estimate $\hat{H}_{om}^{\text{MUE-LS}}$ for MBS is obtained as

$$\hat{H}_{om}^{\text{MUE-LS}} = \frac{y_o^{\psi}}{\psi_{om}^{\text{MUE}}.\sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}}}, \tag{6}$$

where we assume that in (4, 5, 6), P_{ji}^{FAP}, P_{rk}^{CUE}, P_{om}^{MUE} and L_{ji}^{FAP}, L_{rk}^{CUE}, L_{om}^{MUE} are known.

2) MMSE CHANNEL ESTIMATOR

Considering (4), the MMSE channel estimator $\hat{H}_{ji}^{\text{FAP-MMSE}}$ for j^{th} FAP is given as

$$\hat{H}_{ji}^{\text{FAP-MMSE}} = \hat{H}_{ji}^{\text{FAP-LS}}.Q_{ji}^{\text{FAP}}.R_{H_{ji}^{\text{FAP}}\hat{H}_{ji}^{\text{FAP-LS}}}, \tag{7}$$

where $R_{H_{ji}^{\text{FAP}}\hat{H}_{ji}^{\text{FAP-LS}}}$ represents the covariance matrix of $N \times N$ matrices H_{ji}^{FAP} and $\hat{H}_{ji}^{\text{FAP-LS}}$, i.e. $R_{H_{ji}^{\text{FAP}}\hat{H}_{ji}^{\text{FAP-LS}}} = \mathbb{E}\left\{\left(H_{ji}^{\text{FAP}}\right)\left(\hat{H}_{ji}^{\text{FAP-LS}}\right)^{H}\right\}$ and $Q_{ji}^{\text{FAP}} = \left(\sum_{i=1}^{U_i} R_{\mathbf{H}_{ji}^{\text{FAP}}\mathbf{H}_{ji}^{\text{FAP}}} + (\sigma_j^{\text{FAP}})^2 I_{N_f}\right)^{-1}$.

The MMSE channel estimator $\hat{H}_{rk}^{\text{CUE-MMSE}}$ for r^{th} RN with equation (5) is obtained as

$$\hat{H}_{rk}^{\text{CUE-MMSE}} = \hat{H}_{rk}^{\text{CUE-LS}}.Q_{rk}^{\text{CUE}}.R_{H_{rk}^{\text{CUE}}\hat{H}_{rk}^{\text{CUE-LS}}}, \tag{8}$$

where $R_{H_{rk}^{\text{CUE}}\hat{H}_{rk}^{\text{CUE-LS}}}$ denotes the covariance matrix of $N \times N$ matrices H_{rk}^{CUE} and $\hat{H}_{rk}^{\text{CUE-LS}}$, i.e. $R_{H_{rk}^{\text{CUE}}\hat{H}_{rk}^{\text{CUE-LS}}} = \mathbb{E}\left\{\left(H_{rk}^{\text{CUE}}\right)\left(\hat{H}_{rk}^{\text{CUE-LS}}\right)^{H}\right\}$ and $Q_{rk}^{\text{CUE}} = \left(\sum_{k=1}^{K} R_{\mathbf{H}_{rk}^{\text{CUE}}\mathbf{H}_{rk}^{\text{CUE}}} + (\sigma_r^{\text{CUE}})^2 I_{N_R}\right)^{-1}$.

Similarly with (6), the MMSE channel estimation $\hat{H}_{om}^{\text{MUE-MMSE}}$ for MBS is given as

$$\hat{H}_{om}^{\text{MUE-MMSE}} = \hat{H}_{om}^{\text{MUE-LS}}.Q_{om}^{\text{MUE}}.R_{H_{om}^{\text{MUE}}\hat{H}_{om}^{\text{MUE-LS}}}, \tag{9}$$

where $R_{H_{om}^{\text{MUE}}\hat{H}_{om}^{\text{MUE-LS}}}$ represents the covariance matrix of $N \times N$ matrices H_{om}^{MUE} and $\hat{H}_{om}^{\text{MUE-LS}}$, i.e. $R_{H_{om}^{\text{MUE}}\hat{H}_{om}^{\text{MUE-LS}}} = \mathbb{E}\left\{\left(H_{om}^{\text{MUE}}\right)\left(\hat{H}_{om}^{\text{MUE-LS}}\right)^{H}\right\}$ and $Q_{om}^{\text{MUE}} = \left(\sum_{m=1}^{M} R_{\mathbf{H}_{om}^{\text{MUE}}\mathbf{H}_{om}^{\text{MUE}}} + (\sigma_o^{\text{MUE}})^2 I_{N_B}\right)^{-1}$.

C. UPLINK TRANSMISSION DESIGN

The signals are transmitted from the CUEs to the MBS through the RN during the first time slot. Similarly, the channels from the RN to the MBS are estimated at the MBS and fed back to the RN, which then forwards the estimates back to the MUE. The FUEs also transmit to their respective FAPs during the first time slot while the MUEs transmit to the MBS during both time slots. The relaying operates in a half-duplex mode, in the first time slot, and the CUEs use transmit precoding to broadcast to the RN and in the second time slot, the RNs cooperatively form a distributed relay beam-former to forward the signals to MBS. Direct links between CUEs and MBS are not assumed due to severe attenuation [25]. The Rayleigh flat-fading channel and noise have independent and identically distributed (i.i.d.) complex Gaussian entries with zero mean and unit variance $\mathcal{CN}(0, 1)$.

The complex received signal vector at the j^{th} FAP during the first time slot y_j is defined as

$$\begin{aligned} y_j^{\text{FAP}} = & \underbrace{\sum_{i=1}^{U_i} \sqrt{P_{ji}^{\text{FAP}} L_{ji}^{\text{FAP}}} H_{ji}^{\text{FAP}} w_{ji}^{\text{FAP}} s_{ji}^{\text{FAP}}}_{\text{FUEs signal of the } j^{th} \text{ FAP}} \\ & + \underbrace{\sum_{m=1}^{M} \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} H_{jm}^{\text{MUE}} w_{om}^{\text{MUE}} s_{om}^{\text{MUE}}}_{\text{MUEs interference}} \\ & + \underbrace{\sum_{\substack{f=1 \\ f \neq j}}^{F} \sum_{u=1}^{U} \sqrt{P_{fu}^{\text{FAP}} L_{fu}^{\text{FAP}}} H_{jfu}^{\text{FAP}} w_{fu}^{\text{FAP}} s_{fu}^{\text{FAP}}}_{\text{other femtocells interference}} \\ & + \underbrace{\sum_{r=1}^{R} \sum_{k=1}^{K} \sqrt{P_{rk}^{\text{CUE}} L_{rk}^{\text{CUE}}} H_{jrk}^{\text{CUE}} w_{rk}^{\text{CUE}} s_{rk}^{\text{CUE}}}_{\text{CUEs interference}} + \underbrace{n_j}_{\text{noise}}, \end{aligned} \tag{10}$$

where s_{ji}^{FAP} is the message of the i^{th} user of the j^{th} FAP and w_{ji}^{FAP} is the pre-coding vector of the i^{th} FUE of the j^{th} femtocell, while s_{om}^{MUE} and w_{om}^{MUE} are the message and pre-coding vector of the m^{th} MUE of the MBS, respectively. s_{fu}^{FAP} is the message of the u^{th} FUE of the f^{th} FAP and w_{fu}^{FAP} is the pre-coding vector of the u^{th} FUE of the f^{th} FAP. s_{rk}^{CUE} and w_{rk}^{CUE} are the message and the pre-coding vector of the k^{th} CUE of the r^{th} RN, respectively. n_j is the AWGN vector at the j^{th} FAP and that satisfies $\mathbb{E}\{n_j n_j^H\} = (\sigma_j^{\text{FAP}})^2 I_{N_f}$.

In order to design the pre-coder w_{ji}^{FAP} in (10), the knowledge of w_{om}^{MUE}, w_{fu}^{FAP} and w_{rk}^{CUE} is required. This can be done by joint design which is computationally complex. To simplify the pre-coder design problem, the design of the pre-coders in (10) is divided into four different pre-coder designs where in each design, we assume that the pre-coder variable that is not currently being designed is represented by ZF and is independent of each other. Hence, to find the optimal w_{ji}^{FAP} in (10) of the i^{th} FUE at the j^{th} FAP, w_{om}^{MUE}, w_{fu}^{FAP} and w_{rk}^{CUE} are found using ZF pre-coder assumption and an estimation error as follows

$$w_{om}^{\text{MUE}'} = \left[(\hat{H}_{om}^{\text{MUE-Est}})^H \left((\hat{H}_{om}^{\text{MUE-Est}})^H \hat{H}_{om}^{\text{MUE-Est}}\right)^{-1} + \varepsilon\right], \tag{11}$$

$$w_{fu}^{\text{FAP}'} = \left[(\hat{H}_{fu}^{\text{FAP-Est}})^H \left((\hat{H}_{fu}^{\text{FAP-Est}})^H \hat{H}_{fu}^{\text{FAP-Est}}\right)^{-1} + \varepsilon\right], \tag{12}$$

$$w_{rk}^{\text{CUE}'} = \left[(\hat{H}_{rk}^{\text{CUE-Est}})^H \left((\hat{H}_{rk}^{\text{CUE-Est}})^H \hat{H}_{rk}^{\text{CUE-Est}}\right)^{-1} + \varepsilon\right], \tag{13}$$

where ε represents an estimation error, a Gaussian random number of zero mean and σ_ε^2. Equation (10) is rewritten with the ZF assumption design as

$$y_j^{\text{FAP}'} = \sum_{i=1}^{U_i} \sqrt{P_{ji}^{\text{FAP}} L_{ji}^{\text{FAP}}} H_{ji}^{\text{FAP}} w_{ji}^{\text{FAP}} s_{ji}^{\text{FAP}} + \sum_{\substack{f=1 \\ f\neq j}}^{F} \sum_{u=1}^{U} \sqrt{P_{fu}^{\text{FAP}} L_{fu}^{\text{FAP}}} H_{jfu}^{\text{FAP}} w_{fu}^{\text{FAP}'} s_{fu}^{\text{FAP}} + \sum_{m=1}^{M} \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} H_{jm}^{\text{MUE}} w_{om}^{\text{MUE}'} s_{om}^{\text{MUE}} + \sum_{r=1}^{R} \sum_{k=1}^{K} \sqrt{P_{rk}^{\text{CUE}} L_{rk}^{\text{CUE}}} H_{jrk}^{\text{CUE}} w_{rk}^{\text{CUE}'} s_{rk}^{\text{CUE}} + n_j. \quad (14)$$

Similarly, the received signal $y_o^{\text{MBS}'}$ of the MBS during the first time slot is assumed to be

$$y_o^{\text{MBS}'} = \sum_{m=1}^{M} \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} H_{o,m}^{\text{MUE}} w_{om}^{\text{MUE}} s_{o,m}^{\text{MUE}} + \sum_{r=1}^{R} \sum_{k=1}^{K} \sqrt{P_{rk}^{\text{CUE}} L_{rk}^{\text{CUE}}} H_{ork}^{\text{CUE}} w_{rk}^{\text{CUE}'} s_{rk}^{\text{CUE}} + \sum_{f=1}^{F} \sum_{u=1}^{U} \sqrt{P_{fu}^{\text{FAP}} L_{fu}^{\text{FAP}}} H_{ofu}^{\text{FAP}} w_{fu}^{\text{FAP}'} s_{fu}^{\text{FAP}} + n_o, \quad (15)$$

where n_o is the AWGN vector with variance $(\sigma_o^{\text{MBS}})^2$ distributed according to $\mathcal{CN}(0, (\sigma_o^{\text{MBS}})^2)$. Similarly, the received signal $y_r^{\text{RN}'}$ at the RN during the first time slot is

$$y_r^{\text{RN}'} = \sum_{k=1}^{K} \sqrt{P_{rk}^{\text{CUE}} L_{rk}^{\text{CUE}}} H_{rk}^{\text{CUE}} w_{rk}^{\text{CUE}} s_{rk}^{\text{CUE}} + \sum_{m=1}^{M} \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} H_{rm}^{\text{MUE}} w_{om}^{\text{MUE}'} s_{om}^{\text{MUE}} + \sum_{f=1}^{F} \sum_{u=1}^{U} \sqrt{P_{fu}^{\text{FAP}} L_{fu}^{\text{FAP}}} H_{rfu}^{\text{FAP}} w_{fu}^{\text{FAP}'} s_{fu}^{\text{FAP}} + n_r, \quad \forall r = 1, \ldots, R \quad (16)$$

where n_r is the AWGN vector with variance $(\sigma_r^{\text{CUE}})^2$ which is distributed according to $\mathcal{CN}(0, (\sigma_r^{\text{CUE}})^2)$. It is assumed that the channels are i.i.d. complex Gaussian random variables.

The RN receives the signal from the K-CUEs and interference from the FUEs and MUEs. It amplifies and forwards the signal vector multiplied by the RN pre-coder F_{or} during the second time slot. The amplified r^{th} transmit signal x_{or} to the MBS during the second time slot is expressed as

$$x_{or} = F_{or} \times y_r^{\text{RN}'}, \quad \forall\, r = 1, \cdots, R. \quad (17)$$

The MUEs continuously transmit signals to the MBS in both time slots. The FUEs transmit only during the first time slot. Therefore, the received signal $y_o^{\text{MBS-nd}'}$ at the MBS is written as:

$$y_o^{\text{MBS-nd}'} = \sum_{r=1}^{R} \sqrt{P_r L_r} H_{or} x_{or} + \underbrace{\sum_{m=1}^{M} \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} H_{om}^{\text{MUE}} w_{om}^{\text{MUE}} s_{om}^{\text{MUE}}}_{\text{MUEs signal at the MBS}} + n_o, \quad (18)$$

where H_{or} is the channel matrix from the r^{th} RN to the MBS and P_r is the transmit power at the RN. L_r is the propagation loss at the r^{th} RN. n_o is the AWGN vector at the MBS with variance $(\sigma_o^{\text{MBS}})^2$ which is distributed according to $\mathcal{CN}(0, (\sigma_o^{\text{MBS}})^2)$. After substitution and calculation, the received signal at the MBS during the second time signal is written as

$$y_o^{\text{MBS-nd}'} = \underbrace{\sum_{r=1}^{R} \sum_{k=1}^{K} \sqrt{P_r L_r} H_{or} F_{or} H_{rk}^{\text{CUE}} w_{rk}^{\text{CUE}} s_{rk}^{\text{CUE}}}_{1^{st}\text{term}} + \underbrace{\sum_{m=1}^{M} \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} H_{o,m}^{\text{MUE}} w_{om}^{\text{MUE}} s_{om}^{\text{MUE}}}_{2^{nd}\text{term}} + z_o, \quad (19)$$

where $z_o = \sum_{r=1}^{R} H_{o,r} F_{or} \tilde{n}_r + n_o$ and $\tilde{n}_r = \left(\sum_{f=1}^{F} \sum_{u=1}^{U} H_{rfu}^{\text{FAP}} w_{fu}^{\text{FAP}'} s_{fu}^{\text{FAP}} + n_r \right)$. The femtocell interferences during the first time slot are considered as noise at the MBS. The 1^{st} and 2^{nd} terms of equation (19) are the signals to be decoded at the MBS during the second time slot, and need to be combined as one term. *Note*: N_s is the number of antennas for users. N_R is the number of equipped antenna for RNs, N_B is the number of MBS antennas and d_s is the data stream. The following is assumed:

$\hat{d} = \sum_{k=1}^{K} d_s$, $\hat{N}_s = \sum_{k=1}^{K} N_s$, $N_R \geq \hat{d}$ and $\hat{N}_s > \hat{d}$. We assume that $N_B \geq N_R$. With the above assumptions, the signals, channels matrices and pre-coders for CUEs and MUEs (during the second time slot) are combined as in (20)–(22), as shown at the top of the next page. respectively. The received signal $y_o^{\text{MBS-nd}}$ of the MBS in (19) during the second time slot can be rewritten as

$$y_o^{\text{MBS-nd}} = H_o F_o H^{\text{UE}} W^{\text{UE}} s^{\text{UE}} + z_o, \quad (23)$$

where $H_o = [\sqrt{P_1 L_1} H_{o,1}, \cdots, \sqrt{P_R L_R} H_{oR}] \in \mathbb{C}^{N_B \times \hat{N}_R(R)}$ and $F_o = [F_{o,1}, \cdots, F_{o,R}]^T \in \mathbb{C}^{N_R \times \hat{N}_R(R)}$. The combination of CUEs and MUEs will be referred to as UEs throughout this article. z_o is a $\mathbb{C}^{N_B \times 1}$ is the AWGN vector with variance $\sigma_{z_o}^2$ which is distributed according to $\mathcal{CN}(0, \sigma_{z_o}^2)$.

$$s^{\text{UE}} = \Big[\underbrace{s_{1,1}^{\text{CUE}}, \cdots, s_{1,K}^{\text{CUE}} \mid \cdots \mid s_{R,1}^{\text{CUE}}, \cdots, s_{R,K}^{\text{CUE}}}_{\text{CUE transmit signals}} \mid \underbrace{s_{o,1}^{\text{MUE}}, \cdots, s_{o,M}^{\text{MUE}}}_{\substack{\text{MUE transmit signals} \\ \text{at } 2^{nd} \text{ time slot}}}\Big]^T \in \mathbb{C}^{1\times(RK+M)\hat{d}} \tag{20}$$

$$H^{\text{UE}} = \Big[\underbrace{\sqrt{P_{11}^{\text{CUE}}L_{11}^{\text{CUE}}}H_{11}^{\text{CUE}}, \cdots, \sqrt{P_{1K}^{\text{CUE}}L_{1K}^{\text{CUE}}}H_{1K}^{\text{CUE}} \mid \cdots \mid \sqrt{P_{R1}^{\text{CUE}}L_{R1}^{\text{CUE}}}H_{R1}^{\text{CUE}}, \cdots, \sqrt{P_{RK}^{\text{CUE}}L_{RK}^{\text{CUE}}}H_{RK}^{\text{CUE}}}_{\text{CUE channels matrices}}$$
$$\mid \underbrace{\sqrt{P_{o1}^{\text{MUE}}L_{o1}^{\text{MUE}}}H_{o1}^{\text{MUE}}, \cdots, \sqrt{P_{oM}^{\text{MUE}}L_{oM}^{\text{MUE}}}H_{oM}^{\text{MUE}}}_{\text{MUE channel matrix}}\Big] \in \mathbb{C}^{N_R\times\hat{N}_s(RK+M)} \tag{21}$$

$$W^{\text{UE}} = \begin{pmatrix} w_{11}^{\text{CUE}} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & w_{RK}^{\text{CUE}} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & w_{o1}^{\text{MUE}} & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & w_{oM}^{\text{MUE}} \end{pmatrix} \in \mathbb{C}^{\hat{N}_s(RK+M)\times\hat{d}(RK+M)} \tag{22}$$

III. DECENTRALIZED ALGORITHMS FOR LINEAR TRANSCEIVER DESIGNS

In this section, the decentralized transceiver optimization algorithms for the FAPs, MUEs and CUEs are designed with the coordinated MMSE approach during the first and second time slots.

A. COORDINATED MMSE APPROACH FOR FEMTOCELL AND MUES TRANSCEIVER DURING THE FIRST TIME SLOT

1) OPTIMIZATION OF THE FAPS PRE-CODING AND DECODING VECTORS

The algorithm starts with initialized random pre-coders and decoders w_{jx}^{FAP}. In each iteration, the FUE pre-coders and decoders are updated alternatively. Considering the MMSE receiver, we apply the MMSE decoding for j^{th} FAP such that the interference is received from the MUEs and neighbouring femtocells during the first time slot. The decoded information $\hat{s}_{jx}^{\text{FAP}}$ for the x^{th} FUE of the j^{th} FAP can be expressed as

$$\hat{s}_{jx}^{\text{FAP}} = (d_{jx}^{\text{FAP}})^H . y_j^{\text{FAP}'} \tag{24}$$

where $d_{j,x}^{\text{FAP}}$ is the decoding vector for x^{th} FUE of j^{th} FAP and $y_j^{\text{FAP}'}$ is as in (14). In order to minimize the power noise component, we employ the coordinated MMSE algorithm that minimizes the sum MSE at the j^{th} FAP estimated as

$$\min_{\substack{w_{j1}^{\text{FAP}},\ldots,w_{jU_i}^{\text{FAP}} \\ d_{j1}^{\text{FAP}},\cdots,d_{jU_i}^{\text{FAP}}}} \sum_{x=1}^{U_i} \mathbb{E}\left[\|\hat{s}_{j,x}^{\text{FAP}} - s_{j,x}^{\text{FAP}}\|^2\right]$$

$$\text{s. t.: } (w_{jx}^{\text{FAP}})^H (w_{jx}^{\text{FAP}}) \leq \frac{P_{jx}^{\text{FAP}}}{L_{jx}^{\text{FAP}}}, \quad x = 1, \cdots, U_i, \tag{25}$$

where P_{jx}^{FAP} is the maximum transmit power of the x^{th} FUE of the j^{th} FAP. We consider the estimated channels to rewrite the minimum sum MSE at the x^{th} FUE of the j^{th} FAP. This is rewritten as

$$\begin{aligned} \min_{\substack{w_{j1}^{\text{FAP}},\ldots,w_{jU_i}^{\text{FAP}} \\ d_{j1}^{\text{FAP}},\cdots,d_{jU_i}^{\text{FAP}}}} \sum_{x=1}^{U_i} \Bigg[& \|(d_{j,x}^{\text{FAP}})^H \sqrt{P_{jx}^{\text{FAP}}L_{jx}^{\text{FAP}}}\hat{H}_{jx}^{\text{FAP-Est}} w_{jx}^{\text{FAP}} - 1\|^2 \\ & + \sum_{\substack{i=1 \\ i\neq x}}^{U_i} \|(d_{jx}^{\text{FAP}})^H \sqrt{P_{ji}^{\text{FAP}}L_{ji}^{\text{FAP}}}\hat{H}_{ji}^{\text{FAP-Est}} w_{ji}^{\text{FAP}'}\|^2 \\ & + \sum_{m=1}^{M} \|(d_{jx}^{\text{FAP}})^H \sqrt{P_{om}^{\text{MUE}}L_{om}^{\text{MUE}}}\hat{H}_{jm}^{\text{MUE-Est}} w_{om}^{\text{MUE}'}\|^2 \\ & + \sum_{\substack{f=1 \\ f\neq j}}^{F} \sum_{u=1}^{U} \|(d_{jx}^{\text{FAP}})^H \sqrt{P_{fu}^{\text{FAP}}L_{fu}^{\text{FAP}}}\hat{H}_{jfu}^{\text{FAP-Est}} w_{fu}^{\text{FAP}'}\|^2 \\ & + \sum_{r=1}^{R} \sum_{k=1}^{K} \|(d_{j,x}^{\text{FAP}})^H \sqrt{P_{rk}^{\text{CUE}}L_{rk}^{\text{CUE}}}\hat{H}_{jrk}^{\text{CUE-Est}} w_{rk}^{\text{CUE}'}\|^2 \\ & + \|(d_{j,x}^{\text{FAP}})\|^2\sigma^2 \Bigg] \end{aligned}$$

$$\text{s. t.: } (w_{jx}^{\text{FAP}})^H (w_{jx}^{\text{FAP}}) \leq \frac{P_{jx}^{\text{FAP}}}{L_{jx}^{\text{FAP}}}, \quad x = 1, \ldots, U_i. \tag{26}$$

The minimum sum MSE problem in (26) is convex in $w_{j,x}^{\text{FAP}}, x = 1, \ldots, U_i$, if all $d_{j,x}^{\text{FAP}}$ are fixed and convex in $d_{j,x}^{\text{FAP}}, x = 1, \ldots, U_i$, if all $w_{j,x}^{\text{FAP}}$ are also fixed. This enables obtaining the FUE pre-coding vectors of the j^{th} FAP when the FUE decoding vectors of the j^{th} FAP are fixed and vice

versa [28]. When the $d_{j,x}^{\text{FAP}}$ are fixed, the sum MSE optimization problem with respect to the $w_{j,x}^{\text{FAP}}$ pre-coder can be reformulated as

$$\begin{aligned}\min_{w_{j1}^{\text{FAP}},\ldots,w_{jU_i}^{\text{FAP}}} \sum_{x=1}^{U_i} \Bigg[& \|(d_{jx}^{\text{FAP}})^H \sqrt{P_{jx}^{\text{FAP}} L_{jx}^{\text{FAP}}} \hat{H}_{jx}^{\text{FAP-Est}} w_{jx}^{\text{FAP}} - 1\|^2 \\ & + \sum_{\substack{i=1 \\ i\neq x}}^{U_i} \|(d_{jx}^{\text{FAP}})^H \sqrt{P_{ji}^{\text{FAP}} L_{ji}^{\text{FAP}}} \hat{H}_{ji}^{\text{FAP-Est}} w_{ji}^{\text{FAP}'}\|^2 \\ & + \sum_{m=1}^{M} \|(d_{jx}^{\text{FAP}})^H \sqrt{P_{om}^{\text{MUE}} L_{om}^{\text{MUE}}} \hat{H}_{jm}^{\text{MUE-Est}} w_{om}^{\text{MUE}'}\|^2 \\ & + \sum_{\substack{f=1 \\ f\neq j}}^{F} \sum_{u=1}^{U} \|(d_{jx}^{\text{FAP}})^H \sqrt{P_{fu}^{\text{FAP}} L_{fu}^{\text{FAP}}} \hat{H}_{jfu}^{\text{FAP-Est}} w_{fu}^{\text{FAP}'}\|^2 \\ & + \sum_{r=1}^{R} \sum_{k=1}^{K} \|(d_{jx}^{\text{FAP}})^H \sqrt{P_{rk}^{\text{CUE}} L_{rk}^{\text{CUE}}} \hat{H}_{jrk}^{\text{CUE-Est}} w_{r,k}^{\text{CUE}'}\|^2 \\ & + \|(d_{j,x}^{\text{FAP}})\|^2 \sigma^2 \Bigg] \\ \text{s. t.: } & (w_{jx}^{\text{FAP}})^H (w_{jx}^{\text{FAP}}) \leq \frac{P_{jx}^{\text{FAP}}}{L_{jx}^{\text{FAP}}}, \quad x = 1, \ldots, U_i. \end{aligned} \tag{27}$$

The Lagrange duality and Karush-Kuhn-Tucker (KKT) conditions are employed to efficiently solve the FAP optimization problem. The KKT conditions are given by

- *Stationarity:*

$$\sum_{i=1}^{U_i} \hat{H}_{jx}^{\text{FAP-Est}} (d_{ji}^{\text{FAP}})^H d_{ji}^{\text{FAP}} (\hat{H}_{jx}^{\text{FAP-Est}})^H w_{ji}^{\text{FAP}*} + \mu_{jx} w_{jx}^{\text{FAP}*} - (\hat{H}_{jx}^{\text{FAP-Est}})^H d_{jx}^{\text{FAP}} = 0,$$

- *Primary feasibility:*

$$(w_{jx}^{\text{FAP}})^H (w_{jx}^{\text{FAP}}) \leq \frac{P_{jx}^{\text{FAP}}}{L_{jx}^{\text{FAP}}},$$

- *Complementary Slackness:*

$$\mu_{jx} \left((w_{jx}^{\text{FAP}})^H (w_{jx}^{\text{FAP}}) - \frac{P_{jx}^{\text{FAP}}}{L_{jx}^{\text{FAP}}} \right) = 0,$$

- *Dual feasibility:*

$$\mu_{jx} \geq 0,$$

where $w_{jx}^{\text{FAP}*}$ is the optimal value for the FAP pre-coder. Using the KKT conditions the optimal MMSE pre-coding vector $w_{jx}^{\text{FAP}*}$ is obtained as

$$w_{jx}^{\text{FAP}*} = \left(\sum_{i=1}^{U_i} (\hat{H}_{jx}^{\text{FAP-Est}})^H (d_{ji}^{\text{FAP}}) (d_{ji}^{\text{FAP}})^H \hat{H}_{jx}^{\text{FAP-Est}} + \mu_{jx} \mathbf{I}_{U_i} \right)^{-1} \times (\hat{H}_{jx}^{\text{FAP-Est}})^H d_{jx}^{\text{FAP}}, \tag{28}$$

where μ_{jx} represents the satisfaction of the transmit power constraint $(w_{jx}^{\text{FAP}})^H (w_{jx}^{\text{FAP}}) \leq \frac{P_{jx}^{\text{FAP}}}{L_{jx}^{\text{FAP}}}$. Similarly, we fix the pre-coding vectors w_{jx}^{FAP} and obtain the KKT conditions for the optimization problem with respect to the decoder d_{jx}^{FAP}, from which the optimal decoding matrix $d_{ji}^{\text{FAP}*}$ can be obtained. The decoding vector $d_{jx}^{\text{FAP}*}$ can be expressed as

$$\begin{aligned} & d_{jx}^{\text{FAP}*} \\ & = \Bigg(\sum_{m=1}^{M} (\hat{H}_{jm}^{\text{MUE-Est}} w_{om}^{\text{MUE}'}) (\hat{H}_{jm}^{\text{MUE-Est}} w_{om}^{\text{MUE}'})^H \\ & \quad + \sum_{\substack{f=1 \\ f\neq j}}^{F} \sum_{u=1}^{U} (\hat{H}_{jfu}^{\text{FAP-Est}} w_{fu}^{\text{FAP}'}) (\hat{H}_{jfu}^{\text{FAP-Est}} w_{fu}^{\text{FAP}'})^H \\ & \quad + \sum_{r=1}^{R} \sum_{k=1}^{K} (\hat{H}_{jrk}^{\text{CUE-Est}} w_{rk}^{\text{CUE}'}) (\hat{H}_{jrk}^{\text{CUE-Est}} w_{rk}^{\text{CUE}'})^H + \sigma_j^2 \mathbf{I}_{U_i} \Bigg)^{-1} \\ & \quad \times \hat{H}_{jx}^{\text{FAP-Est}} w_{jx}^{\text{FAP}}. \end{aligned} \tag{29}$$

It is assumed that each pre-coder is updated instantaneously when the decoder is updated. This is accomplished by inserting the resulting pre-coder (28) in decoder (29). The detailed optimization process is presented in Algorithm 1.

Algorithm 1 Coordinated MMSE for FAPs During the First Time Slot

1: Initialize and construct the estimated channels using the LS and MMSE estimators as (4) - (9), respectively.
2: Initialize the FUEs pre-coders w_{jx}^{FAP} with each element drawn i.i.d. from the standard Gaussian distribution $\mathcal{CN}(0, 1)$.
3: **for** $j = 1, \cdots, F$ **do**
4: Initialize the FUEs decoder $d_{j1}^{\text{FAP}}, \cdots, d_{jx}^{\text{FAP}}$ by $\mathcal{CN}(0, 1)$.
5: Compute the $w_{om}^{\text{MUE}'}$, $w_{fu}^{\text{FAP}'}$, $w_{rk}^{\text{CUE}'}$ as (11) - (13)
6: Calculate the sum MSE $\epsilon = \sum_{x=1}^{U_i} \mathbb{E}\left[\|\hat{s}_{jx}^{\text{FAP}} - s_{jx}^{\text{FAP}}\|^2 \right]$ in (25), (26)
7: Set $n = 0$ and $\delta_o = \epsilon$
8: **repeat**
9: Update the decoder $d_{j1}^{\text{FAP}}, \cdots, d_{jx}^{\text{FAP}}$ as (29).
10: Calculate the FUE pre-coder $w_{j1}^{\text{FAP}}, \cdots, w_{jx}^{\text{FAP}}$ (28) with the updated decoder.
11: Calculate ϵ with the new pre-coder and decoder
12: set $n = n + 1$
13: $\delta_n = \delta_{n-i} - \epsilon$
14: **until** $\epsilon \approx 0$
15: **end for**

2) OPTIMIZATION OF THE MUES PRE-CODING AND DECODING MATRICES

Similar to the FAPs algorithm, the MUE algorithm starts with initialized random pre-coders and decoders w_{ol}^{MUE}. In each

iteration, the MUE pre-coders and decoders are updated alternatively. The decoded information $\hat{s}_{ol}^{\text{MUE}}$ is expressed as

$$\hat{s}_{ol}^{\text{MUE}} = (d_{ol}^{\text{MUE}})^H . \text{y}_o^{\text{MBS}'} \tag{30}$$

where $(d_{ol}^{\text{MUE}})^H$ is the decoding vector for l^{th} MUE of MBS and $\text{y}_o^{\text{MBS}'}$ is as in (18). We describe the minimum sum MSE problem of the l^{th} user at the MBS during the second time slot as follows

$$\min_{\substack{w_{o1}^{\text{MUE}},\ldots,w_{oM}^{\text{MUE}} \\ d_{o1}^{\text{MUE}},\cdots,d_{oM}^{\text{MUE}}}} \sum_{l=1}^{M} \mathbb{E}\{\|\hat{s}_{ol}^{\text{MUE}} - s_{ol}^{\text{MUE}}\|^2\}$$

$$\text{s. t.: } w_{ol}^{\text{MUE}}(w_{ol}^{\text{MUE}})^H \leq \frac{P_{ol}^{\text{MUE}}}{L_{ol}^{\text{MUE}}}, \quad l = 1, \cdots, M, \tag{31}$$

where $\text{P}_{ol}^{\text{MUE}}$ is the maximum transmit power of the l^{th} MUE of the MBS. The sum MSE optimization problem, Lagrange duality and KKT conditions are simularly described as the optimization FAP transceiver. Thus, the optimal MMSE pre-coding vector $w_{ol}^{\text{MUE}*}$ for the l^{th} MUE of the MBS is given by

$$w_{ol}^{\text{MUE}*} = \left(\sum_{m=1}^{M} (\hat{H}_{ol}^{\text{MUE-Est}})^H (d_{om}^{\text{MUE}})(d_{om}^{\text{MUE}})^H (\hat{H}_{ol}^{\text{MUE-Est}}) + \mu_{ol}\mathbf{I}_M \right)^{-1} (\hat{H}_{ol}^{\text{MUE-Est}})^H d_{ol}^{\text{MUE}}, \tag{32}$$

where μ_{ol} represents the satisfaction of transmit power constraint $(w_{ol}^{\text{MUE}})^H (w_{ol}^{\text{MUE}}) \leq \text{P}_{ol}^{\text{MUE}}$. Considering the same process of fixing the w_{ol}^{MUE} MUE pre-coder and obtaining the KKT conditions of the resulting sum-MSE problem, the optimal decoding vector $d_{ol}^{\text{MUE}*}$ of the l^{th} MUE of the MBS can be formulated as

$$d_{ol}^{\text{MUE}*} = \left(\sum_{r=1}^{R}\sum_{k=1}^{K} (\hat{H}_{ork}^{\text{CUE-Est}} w_{rk}^{\text{CUE}'})(\hat{H}_{ork}^{\text{CUE-Est}} w_{rk}^{\text{CUE}'})^H + \sum_{f=1}^{F}\sum_{u=1}^{U} (\hat{H}_{ofu}^{\text{FAP-Est}} w_{fu}^{\text{FAP}'})(\hat{H}_{ofu}^{\text{FAP-Est}} w_{fu}^{\text{FAP}'})^H + \sigma_{ol}^2 I_M \right)^{-1} \times \hat{H}_{ol}^{\text{MUE-Est}} w_{ol}^{\text{MUE}}. \tag{33}$$

The details coordinated MMSE algorithm for the MUEs is presented in Algorithm 2.

3) CONVERGENCE ANALYSIS

In this subsection, we analyze some details of the algorithms proposed for the FAPs and MUEs during the first time slot, Algorithm 1 and 2, respectively. Let the minimum sum MSE be denoted $\mathcal{H}(\mathcal{W}, \mathcal{D})$ where $\mathcal{W}$ represents the variables recomputed in the precoder update and $\mathcal{D}$ is recomputed in the decoder update. It is noted that when the value of $\mathcal{H}(\mathcal{W}, \mathcal{D})$ approaches zero, the minimum sum MSE of the system is achieved. We assumed to obtained the optimal solutions in the t^{th}-iteration of the proposed iterative algorithms (Algorithm 1 and 2) as $\left\{\mathcal{W}^{(t)}, \mathcal{D}^{(t)}\right\}$.

Algorithm 2 Coordinated MMSE for MUEs During the First Time Slot

1: Initialize and construct the estimated channels using the LS and MMSE estimators as (4) - (9), respectively.
2: Initialize the MUEs pre-coders w_{ol}^{MUE} with each element drawn i.i.d. from the standard Gaussian distribution $\mathcal{CN}(0, 1)$.
3: Initialize the MUEs decoder $d_{o1}^{\text{MUE}}, \cdots, d_{ol}^{\text{MUE}}$ by $\mathcal{CN}(0, 1)$.
4: Compute the $w_{fu}^{\text{FAP}'}$, $w_{rk}^{\text{CUE}'}$ as (12), (13)
5: Calculate the sum MSE $\epsilon = \sum_{l=1}^{M} \mathbb{E}\left[\|\hat{s}_{ol}^{\text{MUE}} - s_{ol}^{\text{MUE}}\|^2\right]$ as in (31)
6: Set $i = 0$ and $\delta_o = \epsilon$
7: **repeat**
8: Update the MUE decoder $d_{o1}^{\text{MUE}}, \cdots, d_{ol}^{\text{MUE}}$ as (33).
9: Calculate the MUE pre-coder $w_{o1}^{\text{MUE}}, \cdots, w_{ol}^{\text{MUE}}$ (32) with the updated decoder.
10: Calculate ϵ with the new pre-coder and decoder
11: set $i = i + 1$
12: $\delta_i = \delta_{i-i} - \epsilon$
13: **until** $\epsilon \approx 0$

In the pre-coder update at the t^{th}-iteration, $\mathcal{W}^{(t)}$ are chosen to minimize the MSE for a given $\mathcal{D}^{(t-1)}$. Thus, $\mathcal{H}(\mathcal{W}^{(t)}, \mathcal{D}^{(t-1)}) \leq \mathcal{H}(\mathcal{W}, \mathcal{D}^{(t-1)})$ for any $\mathcal{W}$ and in particular,

$$\mathcal{H}\left(\mathcal{W}^{(t)}, \mathcal{D}^{(t-1)}\right) \leq \mathcal{H}\left(\mathcal{W}^{(t-1)}, \mathcal{D}^{(t-1)}\right).$$

Similarly, the variables in $\mathcal{D}^{(t)}$ are chosen to minimize the sum MSE for fixed $\mathcal{W}^{(t)}$. We definitely have $\mathcal{H}(\mathcal{W}^{(t)}, \mathcal{D}^{(t)}) \leq \mathcal{H}(\mathcal{W}^{(t)}, \mathcal{D})$ for any $\mathcal{D}$ and therefore,

$$\mathcal{H}\left(\mathcal{W}^{(t)}, \mathcal{D}^{(t)}\right) \leq \mathcal{H}\left(\mathcal{W}^{(t)}, \mathcal{D}^{(t-1)}\right).$$

After combining these two results, we can see that the minimum sum MSE is monotonically decreasing during the iteration,

$$\mathcal{H}\left(\mathcal{W}^{(t)}, \mathcal{D}^{(t)}\right) \leq \mathcal{H}\left(\mathcal{W}^{(t)}, \mathcal{D}^{(t-1)}\right) \leq \mathcal{H}\left(\mathcal{W}^{(t-1)}, \mathcal{D}^{(t-1)}\right)$$

Then, we obtain

$$\begin{aligned} \mathcal{H}\left(\mathcal{W}^{(t)}, \mathcal{D}^{(t)}\right) &\leq \mathcal{H}\left(\mathcal{W}^{(t-1)}, \mathcal{D}^{(t-1)}\right) \\ &\leq \quad \vdots \\ &\leq \mathcal{H}\left(\mathcal{W}^{(1)}, \mathcal{D}^{(1)}\right) \\ &\leq \mathcal{H}\left(\mathcal{W}^{(0)}, \mathcal{D}^{(0)}\right) \end{aligned}$$

The transceivers of FUEs and MUEs are obtained iteratively by solving a problem of minimising the sum MSE. We know that the objective function, sum MSE is bounded below zero and is decreasing at each iteration. Since the minimum sum

MSE is lower bounded, this indicates that it is non-negative, therefore the proposed algorithms converge.

B. COORDINATED MMSE APPROACH FOR UES (CUES AND MUES) DURING THE SECOND TIME SLOT

Generally, the transceiver design for cooperative RN system with multiple users is a difficult task since the RN is shared by multiple users and multi-users interference exists at both RN and MBS. In the following, an iterative design algorithm is proposed based on convex quadratic optimization theory. Specifically, the algorithm iteratively computes the decoder matrices D_o, relay pre-coder matrices F_o and UE pre-coder matrices W^{UE}, starting with initial values for W^{UE} and F_o. The decoded information $\hat{s}^{\mathrm{UE}}$ of the UE through the RN is expressed as

$$\hat{s} = (D_o)^H . y_o^{\text{MBS-nd}} \tag{34}$$

where D_o is the decoding matrix for UEs. The same process of minimizing the sum MSE for the UEs to the MBS during the second time slot is applied and estimated as

$$\begin{aligned} &\min_{W^{\mathrm{UE}},D_o} \mathbb{E}\{\|\hat{s}^{\mathrm{UE}} - s^{\mathrm{UE}}\|^2\} \\ &\text{s. t. } (W^{\mathrm{UE}})(W^{\mathrm{UE}})^H \le \frac{\mathrm{P}^{\mathrm{UE}}}{L^{\mathrm{UE}}}, \\ &\left(F_o\left(\hat{H}^{\text{UE-Est}} W^{\mathrm{UE}} (W^{\mathrm{UE}})^H (\hat{H}^{\text{UE-Est}})^H \right.\right. \\ &\left.\left. + \sigma^2_{\mathrm{UE}} I_{N_B}\right) F_o^H\right) \le \frac{\mathrm{P}_R}{L_R}, \end{aligned} \tag{35}$$

where $\mathrm{P}^{\mathrm{UE}}_{max}$ is the maximum transmit power of the UE and P_R is the maximum transmit power at the RN. The sum MSE problem for UE in (35) can be rewritten as

$$\begin{aligned} &\min_{W^{\mathrm{UE}},D_o} \left[\|(D_o)^H \hat{H}_o^{\mathrm{Est}} F_o \hat{H}^{\text{UE-Est}} W^{\mathrm{UE}} - 1\|^2 + \|(D_o)^H\|^2 C_{z_o}\right] \\ &\text{s. t. } (W^{\mathrm{UE}})(W^{\mathrm{UE}})^H \le \frac{\mathrm{P}^{\mathrm{UE}}}{L^{\mathrm{UE}}}, \\ &\left(F_o\left(\hat{H}^{\text{UE-Est}} W^{\mathrm{UE}} (W^{\mathrm{UE}})^H (\hat{H}^{\text{UE-Est}})^H \right.\right. \\ &\left.\left. + \sigma^2_{\mathrm{UE}} I_{N_B}\right) F_o^H\right) \le \frac{\mathrm{P}_R}{L_R}, \end{aligned} \tag{36}$$

where C_{z_o} is the equivalent noise covariance matrix given by

$$\begin{aligned} C_{z_o} &= \mathbb{E}\left[z_o z_o^H\right] \\ &= \sum_{r=1}^{R} \mathbb{E}\left[\left(\hat{H}^{\mathrm{Est}}_{o,r} F_{o,r} \tilde{n}_r + n_o\right)\left(\hat{H}^{\mathrm{Est}}_{o,r} F_{o,r} \tilde{n}_r + n_o\right)^H\right] \\ &= \sum_{r=1}^{R} \hat{H}^{\mathrm{Est}}_{o,r} F_{o,r} (\hat{H}^{\mathrm{Est}}_{o,r})^H (F_{o,r})^H + I_{N_B}. \end{aligned} \tag{37}$$

1) DESIGN OF THE RN PRE-CODING MATRIX

In order to evaluate the RN pre-coder F_o, we consider the fixed MMSE decoder D_o and pre-coder W^{UE}, the sum-MSE optimization pre-coder with respect to the RN pre-coder can be formulated as

$$\begin{aligned} &\min_{F_o} \|(D_o)^H \hat{H}_o^{\mathrm{Est}} F_o \hat{H}^{\mathrm{UE}} W^{\mathrm{UE}} - 1\|^2 + \|(D_o)^H\|^2 C_{z_o} \\ &\text{s. t. } \left(F_o\left(\hat{H}^{\text{UE-Est}} W^{\mathrm{UE}} (W^{\mathrm{UE}})^H (\hat{H}^{\text{UE-Est}})^H \right.\right. \\ &\left.\left. + \sigma^2_{\mathrm{UE}} I_{N_B}\right) F_o^H\right) \le \frac{\mathrm{P}_R}{L_R}. \end{aligned} \tag{38}$$

The sum-MSE optimization problem is solve with the Lagrange function and the KKT conditions. The Lagrange function is written as

$$\begin{aligned} \mathcal{L}^{\mathrm{RN}}(\lambda, F_o) = &\left[\|(D_o)^H \hat{H}_o^{\mathrm{Est}} F_o \hat{H}^{\text{UE-Est}} W^{\mathrm{UE}} - 1\|^2 \right. \\ &\left. + \|(D_o)^H\|^2 C_{z_o}\right] \\ &+ \lambda\left[F_o\left(\hat{H}^{\text{UE-Est}} W^{\mathrm{UE}} (W^{\mathrm{UE}})^H (\hat{H}^{\text{UE-Est}})^H \right.\right. \\ &\left.\left. + \sigma^2_{\mathrm{UE}} I_{N_B}\right) F_o^H - \frac{\mathrm{P}_R}{L_R}\right], \end{aligned} \tag{39}$$

where λ is the non-negative Lagrange multiplier, $\lambda \ge 0$. The KKT conditions are given as

- *Stationarity:*

$$\frac{\partial \mathcal{L}^{\mathrm{RN}}(\lambda, F_o)}{\partial F_o} = 0,$$

- *Primary feasibility:*

$$\left[F_o\left(\hat{H}^{\text{UE-Est}} W^{\mathrm{UE}} (W^{\mathrm{UE}})^H (\hat{H}^{\text{UE-Est}})^H + \sigma^2_{\mathrm{UE}} I_{N_B}\right) F_o^H\right] \le \frac{\mathrm{P}_R}{L_R},$$

- *Complementary Slackness:*

$$\lambda\left[\left(F_o\left(\hat{H}^{\text{UE-Est}} W^{\mathrm{UE}} (W^{\mathrm{UE}})^H (\hat{H}^{\text{UE-Est}})^H + \sigma^2_{\mathrm{UE}} I\right) F_o^H\right) - \mathrm{P}_R\right] = 0,$$

- *Dual feasibility:*

$$\lambda \ge 0.$$

Through the mathematical manipulation, we can derive the optimal RN pre-coder as

$$\begin{aligned} F_o = &\left(D_o^H \hat{H}_o^{\mathrm{Est}} F_o \hat{H}^{\text{UE-Est}} W^{\mathrm{UE}} D_o (\hat{H}_o^{\mathrm{Est}})^H F_o^H (\hat{H}^{\text{UE-Est}})^H \right. \\ &\left. (W^{\mathrm{UE}})^H + \lambda\phi\right)^{-1} \times D_o (\hat{H}_o^{\mathrm{Est}})^H F_o^H (\hat{H}^{\text{UE-Est}})^H (W^{\mathrm{UE}})^H, \end{aligned} \tag{40}$$

where $\phi = \left((\hat{H}^{\text{UE-Est}})^H W^{\mathrm{UE}} (W^{\mathrm{UE}})^H \hat{H}^{\text{UE-Est}} + \sigma^2_{\mathrm{UE}} I_{N_B}\right)$.

2) DESIGN OF THE PRE-CODING AND DECODING MATRICES

To solve the optimization problem in (36), we consider the Lagrange duality and KKT conditions. The Lagrange function is formulated as

$$\begin{aligned}\mathcal{L}^{\mathrm{UE}}&(\mu_1,\mu_2,D_o,W^{\mathrm{UE}})\\&=\Bigg[\|(D_o)^H\hat{H}_o^{\mathrm{Est}}F_o\hat{H}^{\text{UE-Est}}W^{\mathrm{UE}}-1\|^2\\&\quad+\|(D_o)^H\|^2C_{z_o}\Bigg]\\&\quad+\mu_1\left[(W^{\mathrm{UE}})(W^{\mathrm{UE}})^H-\frac{P^{\mathrm{UE}}}{L^{\mathrm{UE}}}\right]\\&\quad+\mu_2\Bigg[\left(F_o\left(\hat{H}^{\text{UE-Est}}W^{\mathrm{UE}}(W^{\mathrm{UE}})^H(\hat{H}^{\text{UE-Est}})^H\right.\right.\\&\quad\left.\left.+\sigma^2_{\mathrm{UE}}I_{N_B}\right)F_o^H\right)-\frac{\mathrm{P}_R}{L_R}\Bigg],\end{aligned}\tag{41}$$

where μ_1,μ_2 are the non-negative Lagrange multipliers, $\mu_1,\mu_2\geq 0$. The KKT conditions for UEs pre-coders are given as

- *Stationarity:*

$$\frac{\partial\mathcal{L}^{\mathrm{UE}}(\mu_1,\mu_2,D_o,W^{\mathrm{UE}})}{\partial W^{\mathrm{UE}}}=0,$$

- *Primary feasibility:*

$$\begin{aligned}(W^{\mathrm{UE}})(W^{\mathrm{UE}})^H&\leq\frac{P^{\mathrm{UE}}}{L^{\mathrm{UE}}},\\\left[F_o\left(\hat{H}^{\text{UE-Est}}W^{\mathrm{UE}}(W^{\mathrm{UE}})^H(\hat{H}^{\text{UE-Est}})^H\right.\right.&\\\left.\left.+\sigma^2_{\mathrm{UE}}I_{N_B}\right)F_o^H\right]&\leq\frac{\mathrm{P}_R}{L_R},\end{aligned}$$

- *Complementary Slackness:*

$$\begin{aligned}\mu_1\left[(W^{\mathrm{UE}})(W^{\mathrm{UE}})^H-\frac{P^{\mathrm{UE}}}{L^{\mathrm{UE}}}\right]&=0,\\\mu_2\Bigg[F_o\left(\hat{H}^{\text{UE-Est}}W^{\mathrm{UE}}(W^{\mathrm{UE}})^H(\hat{H}^{\text{UE-Est}})^H\right.&\\\left.+\sigma^2_{\mathrm{UE}}I_{N_B}\right)F_o^H-\frac{\mathrm{P}_R}{L_R}\Bigg]&=0,\end{aligned}$$

- *Dual feasibility:*

$$\mu_1,\mu_2\geq 0.$$

To evaluate the derivation of the Lagrange function given in (41) the matrices D_o and D_o^H are treated independently. This is also applied to W^{UE} and $(W^{\mathrm{UE}})^H$. Furthermore, it can be seen that the optimization problem in (36) is convex with respect to W^{UE}. The Lagrange duality function can be defined as

$$f(\mu_1,\mu_2)=\min_{W^{\mathrm{UE}}}\mathcal{L}^{\mathrm{UE}}(\mu_1,\mu_2,D_o,W^{\mathrm{UE}}).\tag{42}$$

Moreover, the dual problem is defined as

$$\max_{\mu_1,\mu_2\geq 0}f(\mu_1,\mu_2).\tag{43}$$

Using the KKT conditions for the resulting problem, the optimal decoder D_o^* is obtained with the fixed pre-coders as

$$\begin{aligned}D_o^*=\Bigg(\hat{H}_o^{\mathrm{Est}}F_o\hat{H}^{\text{UE-Est}}W^{\mathrm{UE}}(\hat{H}_o^{\mathrm{Est}})^HF_o^H(\hat{H}^{\text{UE-Est}})^H(W^{\mathrm{UE}})^H&\\+C_{z_o}\Bigg)^{-1}\hat{H}_o^{\mathrm{Est}}F_o\hat{H}^{\text{UE-Est}}W^{\mathrm{UE}}.&\end{aligned}\tag{44}$$

Hence, the MMSE pre-coding vector W^{UE} for UE during the second time slot is obtained as

$$\begin{aligned}W^{\mathrm{UE}*}=\Bigg(\mu_1I_{\mathrm{UE}}+\mu_2\left((\hat{H}^{\text{UE-Est}})^HF_oF_o^H\hat{H}^{\text{UE-Est}}\right)\Bigg)^{-1}&\\\times D_o(\hat{H}_o^{\mathrm{Est}})^H(\hat{H}^{\text{UE-Est}})^HF_o^H,&\end{aligned}\tag{45}$$

where μ_1 represents the satisfaction of UE transmit power constraint $(W^{\mathrm{UE}})^H(W^{\mathrm{UE}})\leq\frac{P^{\mathrm{UE}}}{L^{\mathrm{UE}}}$ and μ_2 represents the satisfaction of the RN transmit power $\left[\left(F_o\left(\hat{H}^{\text{UE-Est}}W^{\mathrm{UE}}(W^{\mathrm{UE}})^H(\hat{H}^{\text{UE-Est}})^H+\sigma^2_{\mathrm{UE}}I_{N_B}\right)F_o^H\right)-\frac{P_R}{L_R}\right]$. The details of the proposed decentralized algorithm for the UEs is presented in Algorithm 3.

Algorithm 3 Coordinated MMSE for CUEs and MUEs (During the Second Time Slot)

1: Initialize and construct the estimated channels $\hat{H}_o^{\mathrm{Est}}$ and $\hat{H}^{\text{UE-Est}}$ using the LS and MMSE channel estimators.
2: Initialize the UEs and RN pre-coders W^{UE} and F_o with each element drawn i.i.d. from the $\mathcal{CN}(0,1)$.
3: Compute the $w_{om}^{\mathrm{MUE}'}$, $w_{fu}^{\mathrm{FAP}'}$ as (11), (12)
4: Initialize the UEs decoder D_o by $\mathcal{CN}(0,1)$.
5: Calculate the sum MSE $\epsilon=\mathbb{E}\left[\|\hat{s}^{\mathrm{UE}}-s^{\mathrm{UE}}\|^2\right]$ in (35), (36)
6: Set $n=0$ and $\delta_o=\epsilon$
7: **repeat**
8: Update the UE decoder D_o as (44).
9: Calculate the RN pre-coder F_o (40) with the updated decoder.
10: Obtain the UE pre-coder with the updated decoder and RN pre-coder as 45.
11: Calculate ϵ with the new UE, RN pre-coder and decoder
12: set $n=n+1$
13: **until** $\epsilon\approx 0$

IV. PERFORMANCE EVALUATION

In this section, we present the performance evaluation of the proposed schemes for FAPs, MUEs and UEs in the MU-MIMO relay system through numerical simulations. The simulated model is illustrated in Fig. 2. This figure shows a macrocell of dimension $2km\times 2km$ with a MBS placed at the center of the area at coordinates $(1km, 1km)$. There are 12 MUEs distributed near the cell edge, considered as the CUEs. The CUEs are grouped into clusters where each cluster has 1 RN and 3 CUEs. The RNs are strategically placed at the

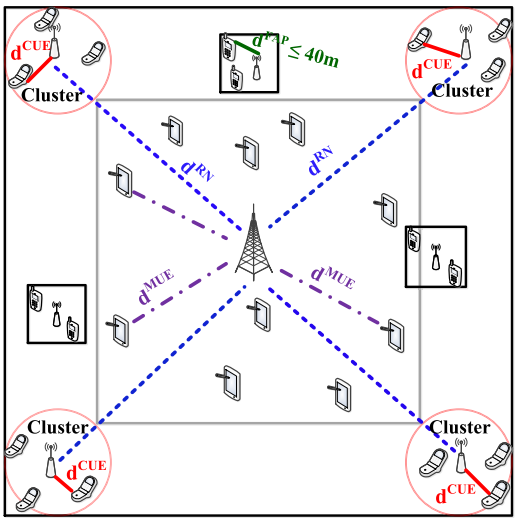

FIGURE 2. Simulation scenario with fixed MBS, FAPs with their FUEs, MUEs and CUEs randomly distributed with a certain distance.

TABLE 2. Coordinate parameters.

Parameters	Values
MBS	(1km, 1km)
MUEs	MUEs are uniformly distributed within the following interval range $x = [> 0.25\text{km} - < 1.75\text{km}]$, $y = [> 0.25\text{km} - < 1.75\text{km}]$
1. Cluster 1:	CUEs are randomly distributed with the following interval range in RN 1 for (x,y)
• RN 1 for R=1	(0.125 km × 0.125 km)
• CUE 1, CUE 2, CUE 3 in RN 1	$x = [0 - \leq 0.25\text{km}]$, $y = [0 - \leq 0.25\text{km}]$
2. Cluster 2:	CUEs are randomly distributed with the following interval range in RN 2 for (x,y)
• RN 2 for R=2	(0.125 km × 1.875 km)
• CUE 1, CUE 2, CUE 3 in RN 2	$x = [0 - \leq 0.25\text{km}]$, $y = [1.75\text{km} - \leq 2\text{km}]$
3. Cluster 3:	CUEs are randomly distributed with the following interval range in RN 3 for (x,y)
• RN 3 for R=3	(1.875 km × 0.125 km)
• CUE 1, CUE 2, CUE 3 in RN 3	$x = [\geq 1.75\text{km} - 2\text{km}]$, $y = [0 - \leq 0.25\text{km}]$
4. Cluster 4:	CUEs are randomly distributed with the following interval range in RN 4 for (x,y)
• RN 4 for R=4	(1.875 km × 1.875 km)
• CUE 1, CUE 2, CUE 3 in RN 4	$x = [\geq 1.75\text{km} - 2\text{km}]$, $y = [\geq 1.75\text{km} - 2\text{km}]$

coordinates detailed in the Table 2. The CUEs are located at distance d^{CUE} from their respective RN. The RNs are located at d^{RN} from the MBS. The users (CUEs, FUEs and MUEs) are randomly placed in the area at the coordinated detailed in Table 2. There are 10 MUEs deployed in the macrocell area located at a distance d^{MUE} from the MBS. We consider 3 FAPs with 2 FUEs each uniformly distributed with the distance d^{FAP}. All the channel coefficients are assumed to be Rayleigh fading channels complex Gaussian random variables with zero mean and variance one. The propagation loss is modelled for each FUEs, MUEs and UEs based on their respective distance. The simulation parameters are similar to [28], [36] and are given in Table 3.

TABLE 3. Simulation parameters.

Parameters	Values
- Number of MUEs and RNs, M and R	10 and 4 respectively
- Number of CUEs, K	12 with 3 per RN
- Number of FAPs, F	3
- Number of FUEs per FAP, U	2 per FAP
- Propagation Loss for $d \geq 0.75$km	$148.1 + 37.6 \log_{10}(d)$ dB
- Propagation Loss for $d \leq 0.75$km	$127 + 30 \log_{10}(d)$ dB
- Channel model	Rayleigh fading
- The maximum transmit power for MUEs and FUEs	$\geq$ 1W and 1mW, respectively
- Noise power	-110dB
- Transmitted signal	BPSK

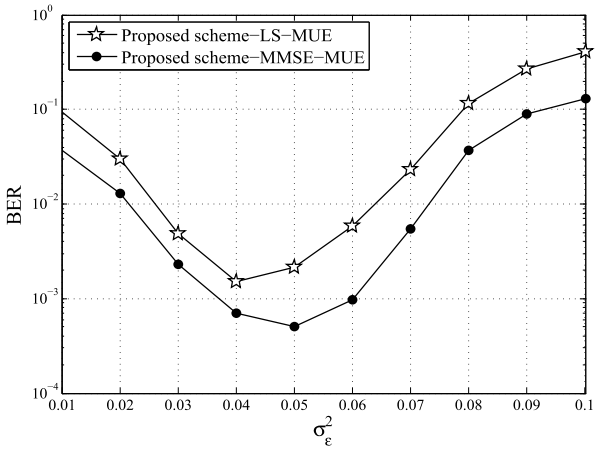

FIGURE 3. BER performance of the proposed schemes for MUEs with different values of σ_ε^2, SNR = 15 dB.

In several cases, the users and RN pre-coders are derived by assuming that all channel matrices are perfectly known at each node. In practical systems, such assumption may not always be realistic. In this regard, the BER evaluation of the channel estimation errors effect is further conducted for the proposed transceiver design for the FAPs, MUEs and UEs at both time slots. Just like [34], we consider the case of imperferct CSI scenario with LS and MMSE channel estimators for a realistic network scenario. The BER performance evaluation versus different value of σ_ε^2 for the MUEs during the first time slot is shown in Fig. 3. The effect of the LS and MMSE channel estimators is considered and compared. It can be observed that for a chosen value of SNR, as the estimation value of σ_ε^2 increases, the BER performance increases as well until a maximum estimation value is achieved. It then starts decreasing as the estimation value decreases. This indicates that the maximum estimation value σ_ε^2 improves

the performance of the MUEs in terms of BER evaluation versus the SNR and can achieve the optimal performance. For the LS and MMSE channel estimators, the maximum estimation value is achieved at $\sigma_\varepsilon^2 = 0.04$ and $\sigma_\varepsilon^2 = 0.05$, respectively. The BER performance versus the different values of σ_ε^2 is illustrated in Fig. 4 for FAPs with SNR = 15dB. Although, the effect of the channel estimation errors is considered, the proposed scheme for FAP still performed well. As observed, the BER performance increases when σ_ε^2 increases and decreases after a certain value of σ_ε^2. For the proposed scheme with LS estimator, the maximum value for σ_ε^2 to achieve an optimal BER performance is $\sigma_\varepsilon^2 = 0.06$ while for the proposed scheme with MMSE estimator is $\sigma_\varepsilon^2 = 0.05$.

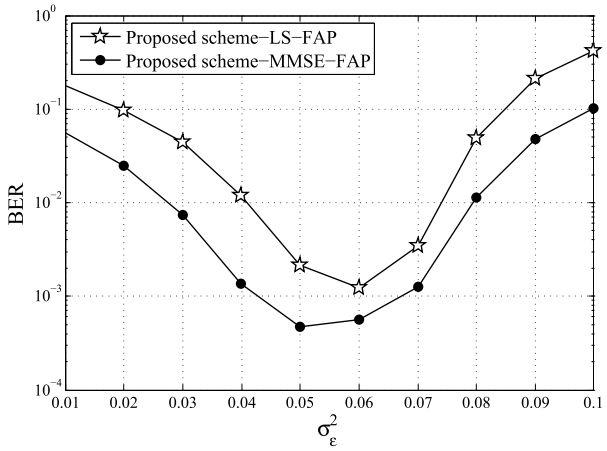

FIGURE 4. BER performance of the proposed schemes for FAP with different values of σ_ε^2, SNR = 15 dB.

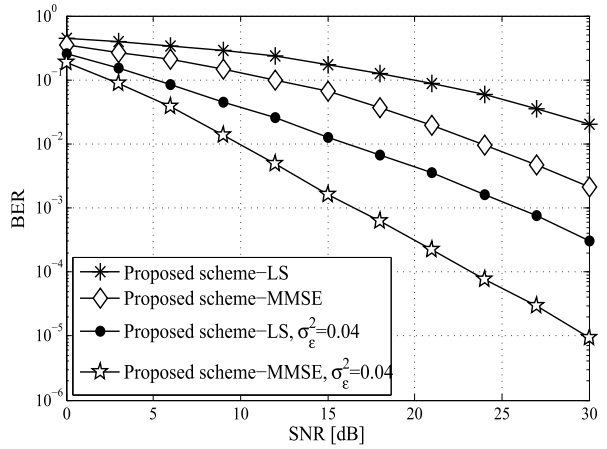

FIGURE 5. BER performance versus SNR for the FAP during the first time slot.

The BER performance as a function of SNR for FAPs is illustrated in Fig. 5, for the LS and MMSE channel estimators with $\sigma_\varepsilon^2 = 0$ and $\sigma_\varepsilon^2 = 0.04$. It can be observed that, the proposed scheme with MMSE channel estimator outperforms the BER performance with the LS estimator effect regardless the value of σ_ε^2. Another interesting observation is that the value of σ_ε^2 affects the BER of the proposed schemes such that it can increase the performance at the maximum value and decrease after the maximum value. With the value of $\sigma_\varepsilon^2 = 0.04$ for LS and MMSE estimators at SNR = 15dB, it can be seen that the achieved BER performances in Fig. 5 are similar to the ones in Fig. 4.

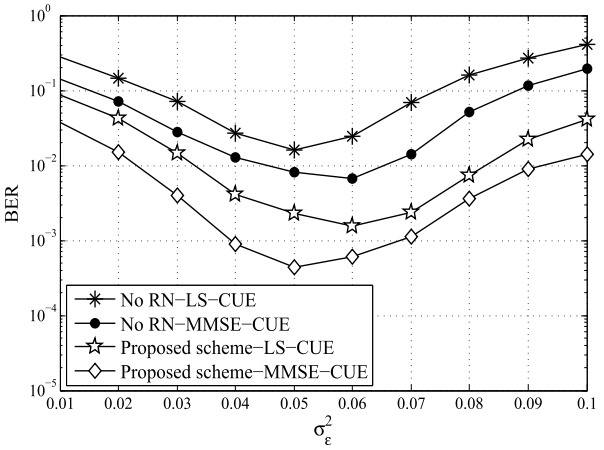

FIGURE 6. BER performance versus different values σ_ε^2 for the CUEs with and without the RN.

Figure 6 illustrates the BER performance as a function of σ_ε^2 for the CUE with and without cooperative RN during the first time slot with SNR = 20dB. It can be observed that, the BER performance of a system without cooperative RNs is not as good as a system with cooperative RNs. However, the performance of the proposed scheme with *"No RN"* is improved when adding estimation values to the ZF assumptions. For LS and MMSE estimators of the proposed scheme with *"No RN"*, the maximum σ_ε^2 is achieved between 0.05 and 0.06, respectively. The proposed scheme, on the other hand, achieves far better BER performance when $\sigma_\varepsilon^2 = 0.06$ for LS estimator and $\sigma_\varepsilon^2 = 0.05$ for the MMSE estimator. Therefore, with the parameters considered, we showed that very significant performances are obtained by adding estimation values to the ZF assumptions.

In Fig. 7, we consider the maximum σ_ε^2 in the case of CUEs with and without RN. For the LS $\sigma_\varepsilon^2 = 0.05$ and MMSE estimators $\sigma_\varepsilon^2 = 0.06$ for *"No RN"*, the same σ_ε^2 are considered for the proposed scheme. We can observe that the proposed schemes for the CUEs still provide significant improvement with the effect of MMSE and LS channel estimators than when a RN is not considered. Fig. 9 shows the BER evaluation versus different values of σ_ε^2 for the UEs during the second time slot with the effect of the LS and MMSE channel estimators. As observed, the proposed scheme with the MMSE channel estimator is better than the LS estimator performance. Interestingly, the maximum value of σ_ε^2 to achieve an optimal BER performance for the proposed scheme with LS estimator is $\sigma_\varepsilon^2 = 0.05$ while $\sigma_\varepsilon^2 = 0.06$ for the proposed scheme with MMSE estimator. Fig. 8 illustrates the BER performance as a function of the SNR for

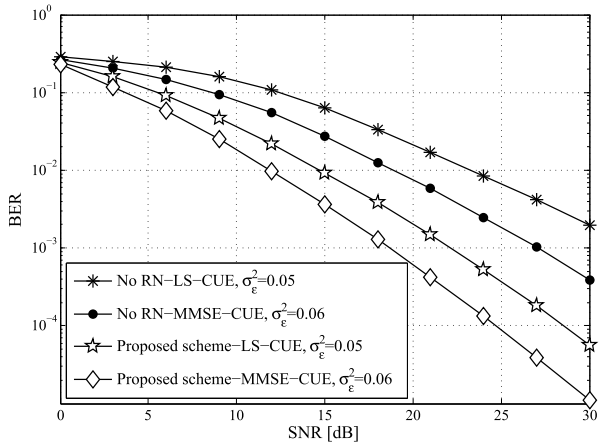

FIGURE 7. BER evaluation as function of the SNR for the CUEs with and without the RN during the first time slot, for $\sigma_{\varepsilon}^2 = 0.05$ and $\sigma_{\varepsilon}^2 = 0.06$.

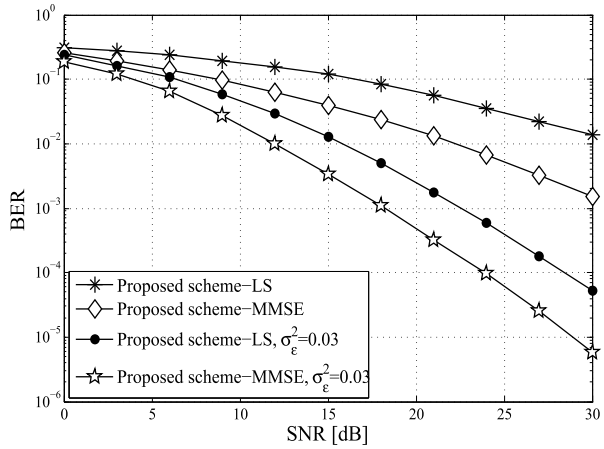

FIGURE 8. BER performance of the proposed schemes for the UEs during the second time slot, for $\sigma_{\varepsilon}^2 = 0.03$.

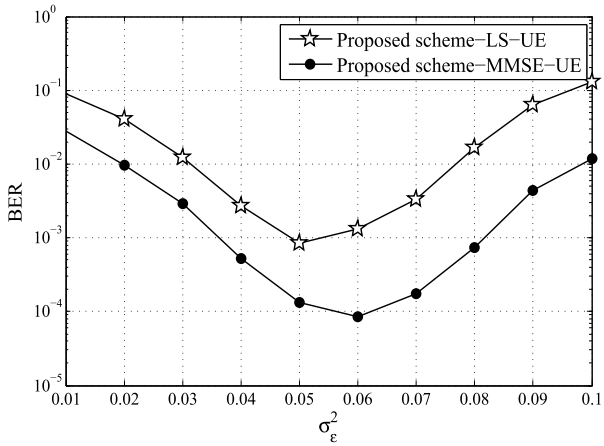

FIGURE 9. BER performance of the proposed schemes for UEs with different values of σ_{ε}^2, SNR = 15 dB.

the UEs during the second time slot. This considers the UEs, which the signal coming from the RN and the MUEs during the second time slot to the MBS. As observed in the figure, the proposed scheme with the effect channel estimator errors still perform well. The reason is that, the proposed schemes update the Lagrange multiplier at each iteration in addition to the UEs and RN matrices. Interestingly, the BER performance of the proposed scheme with the MMSE estimator is better than the LS channel estimator. Furthermore, the proposed scheme with $\sigma_{\varepsilon}^2 = 0.03$, obviously outperforms the proposed schemes with channel estimators when only the ZF $\sigma_{\varepsilon}^2 = 0$ is considered without adding the estimation error σ_{ε}^2.

V. CONCLUSION

In this paper, optimal transceivers for the FUEs, MUEs, CUEs and RN (amplifying matrix) with channel estimators in the MU-MIMO relay systems have been designed for interference management. We considered a decentralized transceiver design instead of a centralized design which essentially is computationally impossible to do in this MU-MIMO relay system due to the unknown interfering terms. The interfering terms have been assumed to be generated as ZF solutions. Due to the inaccuracy of the ZF solutions, estimation values were added to the ZF assumptions in order to achieve better performance. The simulation results demonstrate a much better performance of the proposed schemes in terms of BER when estimation values are added to the ZF assumptions. The proposed decentralized schemes further enhance performance with respect to non-cooperative MU-MIMO systems. This confirms the importance of including cooperative RNs into MU-MIMO systems. Future lines of research could consider the application of the proposed algorithms in a massive MIMO system.

REFERENCES

[1] N. Saquib, E. Hossain, L. B. Le, and D. I. Kim, "Interference management in OFDMA femtocell networks: Issues and approaches," *IEEE Wireless Commun.*, vol. 19, no. 3, pp. 86–95, Jun. 2012.
[2] N. Zhao, F. R. Yu, M. Jin, Q. Yan, and V. C. M. Leung, "Interference alignment and its applications: A survey, research issues, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 1779–1803, 3rd Quart., 2016.
[3] A. S. Hamza, S. S. Khalifa, H. S. Hamza, and K. Elsayed, "A survey on inter-cell interference coordination techniques in OFDMA-based cellular networks," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1642–1670, 4th Quart., 2013.
[4] N. Miridakis and D. D. Vergados, "A survey on the successive interference cancellation performance for single-antenna and multiple-antenna OFDM systems," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 1, pp. 312–335, 1st Quart., 2013.
[5] S. Serbetli and A. Yener, "Iterative transceiver optimization for multiuser MIMO systems," in *Proc. Annu. Allerton Conf. Commun., Control Comput.*, 2002, vol. 40, no. 2, pp. 926–935.
[6] C.-T. Lin, F.-S. Tseng, and W.-R. Wu, "MMSE transceiver design for full-duplex MIMO relay systems," *IEEE Trans. Veh. Technol.*, vol. 66, no. 8, pp. 6849–6861, Aug. 2017.
[7] N. I. Miridakis and T. A. Tsiftsis, "On the joint impact of hardware impairments and imperfect CSI on successive decoding," *IEEE Trans. Veh. Technol.*, vol. 66, no. 6, pp. 4810–4822, Jun. 2017.
[8] N. Khaled, G. Leus, C. Desset, and H. De Man, "A robust joint linear precoder and decoder MMSE design for slowly time-varying MIMO channels," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process. (ICASSP)*, vol. 4, May 2004, p. 4.
[9] P. Sure and C. M. Bhuma, "A survey on OFDM channel estimation techniques based on denoising strategies," *Eng. Sci. Technol., Int. J.*, vol. 20, no. 2, pp. 629–636, 2017.
[10] C. Liu, W. Xia, S. Xing, and L. Shen, "Optimal power allocation assisted with relay in open access femtocell network with the registered users protection," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2015, pp. 1147–1152.

[11] C. Bouras, G. Kavourgias, V. Kokkinos, and A. Papazois, "Interference management in LTE femtocell systems using an adaptive frequency reuse scheme," in *Proc. IEEE Wireless Telecommun. Symp. (WTS)*, Apr. 2012, pp. 1–7.
[12] A. Hatoum, R. Langar, N. Aitsaadi, R. Boutaba, and G. Pujolle, "Cluster-based resource management in OFDMA femtocell networks with Qos guarantees," *IEEE Trans. Veh. Technol.*, vol. 63, no. 5, pp. 2378–2391, Jun. 2014.
[13] M. Ndong and T. Fujii, "Cross-tier interference management with a distributed antenna system for multi-tier cellular networks," *EURASIP J. Wireless Commun. Netw.*, vol. 2014, no. 1, 2014, Art. no. 73.
[14] M. Ndong and T. Fujii, "Joint femtocell clustering and cross-tier interference mitigation with distributed antenna system in small cell networks," in *Proc. 9th IEEE Int. Symp. Commun. Syst., Netw. Digit. Signal Process. (CSNDSP)*, Jul. 2014, pp. 652–657.
[15] L. Li, Y. Jing, and H. Jafarkhani, "Interference cancellation at the relay for multi-user wireless cooperative networks," *IEEE Trans. Wireless Commun.*, vol. 10, no. 3, pp. 930–939, Mar. 2011.
[16] C.-B. Chae, T. Tang, R. W. Heath, Jr., and S. Cho, "MIMO relaying with linear processing for multiuser transmission in fixed relay networks," *IEEE Trans. Signal Process.*, vol. 56, no. 2, pp. 727–738, Feb. 2008.
[17] S. Malik, S. Moon, B. Kim, C. You, H. Liu, and I. Hwang, "Design and analysis of an interference cancellation algorithm for AF, DF and DMF relay protocol in multiuser MIMO scenario based on the LTE-advanced system," *Wireless Pers. Commun.*, vol. 75, no. 1, pp. 775–797, 2014.
[18] J. C. Shen, J. Zhang, and K. B. Letaief, "Downlink user capacity of massive MIMO under pilot contamination," *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3183–3193, Jun. 2015.
[19] A. Khansefid and H. Minn, "On channel estimation for massive MIMO with pilot contamination," *IEEE Commun. Lett.*, vol. 19, no. 9, pp. 1660–1663, Sep. 2015.
[20] Y. Huang, W. Tang, H. Wei, J. Li, D. Wang, and X. Su, "On the performance of iterative receivers in massive MIMO systems with pilot contamination," in *Proc. 9th IEEE Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2014, pp. 52–57.
[21] L. Li, A. Ashikhmin, and T. Marzetta, "Pilot contamination precoding for interference reduction in large scale antenna systems," in *Proc. 51st IEEE Annu. Allerton Conf. Commun., Control, Comput. (Allerton)*, Oct. 2013, pp. 226–232.
[22] J. Liu, F. Gao, and Z. Qiu, "Robust transceiver design for downlink multiuser MIMO AF relay systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2218–2231, Apr. 2015.
[23] A. C. Cirik, S. Biswas, S. Vuppala, and T. Ratnarajah, "Robust transceiver design for full duplex multiuser MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 5, no. 3, pp. 260–263, Jun. 2016.
[24] M. R. A. Khandaker and Y. Rong, "Joint transceiver optimization for multiuser MIMO relay communication systems," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5977–5986, Nov. 2012.
[25] J. Yang and B. Champagne, "Joint transceiver optimization for MIMO multiuser relaying networks with channel uncertainties," in *Proc. 80th IEEE Veh. Technol. Conf. (VTC Fall)*, Sep. 2014, pp. 1–6.
[26] M. Raja and P. Muthuchidambaranathan, "Multiuser MIMO transceiver design for uplink and downlink with imperfect CSI," *Wireless Pers. Commun.*, vol. 75, no. 2, pp. 1215–1234, 2014.
[27] S. Serbetli and A. Yener, "Transceiver optimization for multiuser MIMO systems," *IEEE Trans. Signal Process.*, vol. 52, no. 1, pp. 214–226, Jan. 2004.
[28] B. Guler and A. Yener, "Uplink interference management for coexisting MIMO femtocell and macrocell networks: An interference alignment approach," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 2246–2257, Apr. 2014.
[29] A. Dong, H. Zhang, D. Yuan, and X. Zhou, "Interference alignment transceiver design by minimizing the maximum mean square error for MIMO interfering broadcast channel," *IEEE Trans. Veh. Technol.*, vol. 65, no. 8, pp. 6024–6037, Aug. 2016.
[30] B. D. Antwi-Boasiako, C. Song, and K.-J. Lee, "Transceiver designs for MIMO relaying systems with imperfect channel state information," *Mobile Inf. Syst.*, vol. 2018, Jul. 2018, Art. no. 1429427.
[31] Y. Xu, X. Zhao, and Y.-C. Liang, "Robust power control and beamforming in cognitive radio networks: A survey," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1834–1857, 4th Quart., 2015.
[32] Y. Jeong, T. Q. S. Quek, and H. Shin, "Beamforming optimization for multiuser two-tier networks," *J. Commun. Netw.*, vol. 13, no. 4, pp. 327–338, Aug. 2011.
[33] S. Gong, C. Xing, N. Yang, Y.-C. Wu, and Z. Fei, "Energy efficient transmission in multi-user MIMO relay channels with perfect and imperfect channel state information," *IEEE Trans. Wireless Commun.*, vol. 16, no. 6, pp. 3885–3898, Jun. 2017.
[34] X.-M. Chen, J.-X. Su, Q.-M. Zhu, X.-J. Hu, and Z. Fang, "Linear precoding scheme design for MIMO two-way relay systems with imperfect channel state information," *Math. Problems Eng.*, vol. 2017, Apr. 2017, Art. no. 5164681.
[35] C. Qi, G. Yue, L. Wu, Y. Huang, and A. Nallanathan, "Pilot design schemes for sparse channel estimation in OFDM systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1493–1505, Apr. 2015.
[36] E. Björnson, M. Kountouris, and M. Debbah, "Massive MIMO and small cells: Improving energy efficiency by optimal soft-cell coordination," in *Proc. IEEE 20th Int. Conf. Telecommun. (ICT)*, May 2013, pp. 1–5.

ARMELINE DEMBO MAFUTA received the B.Sc. degree in computer science from the University of Kinshasa, Democratic Republic of Congo, the Postgraduate Diploma degree in mathematical sciences from the African Institute for Mathematical Sciences, University of Cape Town, in 2012, the M.Sc. degree in computer science from the University of KwaZulu-Natal (UKZN), South Africa, in 2014, and the Ph.D. degree in electronic engineering from UKZN. She has published several journals/conferences in her field. Her current research interests include wireless communications, more specifically cooperative communications, MIMO systems, deep learning, and applied cognitive Internet of Things. She is a Reviewer for several journals and conferences.

TOM WALINGO received the B.Sc. degree in electrical and communications engineering from Moi University, Kenya, the M.Sc. degree in electronic engineering from the University of Natal, and the Ph.D. degree in electronic engineering from the University of KwaZulu-Natal, Durban, South Africa. He is currently a Researcher with the Center of Radio Access and Rural Technology and a Lecturer in electrical, electronic and computer engineering with the University of KwaZulu-Natal. He has authored many peer-reviewed journals and conference papers. His current research interests include digital and wireless communications, satellite communications, and wireless sensor networks. He was a recipient of several research awards, including the SAIEE ARJ 2013 Award. He is an Editor/Reviewer for several journals in his field.

FAMBIRAI TAKAWIRA (M'96) received the B.Sc. degree (Hons.) in electrical and electronic engineering from The University of Manchester, Manchester, U.K., in 1981, and the Ph.D. degree from the University of Cambridge, Cambridge, U.K. in 1984. At the University of KwaZulu-Natal (UKZN), he held various academic positions, including the Head of the School of Electrical, Electronic, and Computer Engineering, and just before his departure, he was the Dean of the Faculty of Engineering. He has also held appointments at the University of Zimbabwe, the University of California at San Diego, British Telecom Research Laboratories, and the National University of Singapore. He joined the University of the Witwatersrand, Johannesburg, South Africa, in 2012, after 19 years at UKZN. His research interest includes wireless communication systems and networks. He has served on several conference organizing committees. He served as the Communications Society Director of the Europe, Middle East, and Africa region for the 2012–2013 term. He is a past Editor of the IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS.