IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 8 April 2024, accepted 22 April 2024, date of publication 1 May 2024, date of current version 8 May 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3395585*

RESEARCH ARTICLE

# Centralized and Decentralized IDD Schemes for Cell-Free Massive MIMO Systems: AP Selection and LLR Refinement

**TONNY SSETTUMBA[1], ZHICHAO SHAO[2], (Member, IEEE), LUKAS T. N. LANDAU[1], (Senior Member, IEEE), MICHELLE SOARES PEREIRA FACINA[3], PAULO RICARDO BRANCO DA SILVA[3], AND RODRIGO C. DE LAMARE[1], (Senior Member, IEEE)**

[1]Department of Electrical Engineering (DEE), Pontifical Catholic University of Rio de Janeiro (PUC-Rio), 22451-900 Rio de Janeiro, Brazil
[2]Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China, Quzhou 324003, China
[3]Fundação Centro de Pesquisa e Desenvolvimento em Telecomunicações (CPQD), 13086-902 Campinas, Brazil

Corresponding author: Tonny Ssettumba (tssettumba@gmail.com)

**ABSTRACT** In this paper, we propose iterative interference cancellation schemes with access points selection (APs-Sel) for cell-free massive multiple-input multiple-output (CF-mMIMO) systems. Closed-form expressions for centralized and decentralized linear minimum mean square error (LMMSE) receive filters with APs-Sel are derived assuming imperfect channel state information (CSI). Based on the derived expressions, insights are drawn and general expressions are devised for several cases, namely: firstly, MMSE-SIC filter for the non-scalable CF-mMIMO that uses all APs. Secondly, an MMSE-SIC filter assuming perfect channel state information. Thirdly, in this case we assume no interference cancellation and the linear MMSE filter is obtained. Additionally, we formulate a new Gaussian approximation of the likelihood function by deriving new closed-form expressions for the second order statistics (mean and variance) of the detected signal parameters in presence of channel estimation errors, APs-Sel matrix and multi-user interference (MUI). Since the MMSE-SIC filter experiences error propagation due to erroneous decisions from the previous stages, we develop a list-based detector based on LMMSE receive filters for CF-mMIMO systems that exploits interference cancellation and the constellation points to mitigate the error propagation that occurs in conventional MMSE-SIC receivers. An iterative detection and decoding (IDD) scheme that employs low-density parity-check codes is then developed. Moreover, log-likelihood ratio (LLR) refinement strategies based on censoring and a linear combination of local LLRs are proposed to improve the network performance. We assess the proposed centralized and decentralized IDD schemes against existing approaches in terms of bit error rate performance, complexity, and signaling under perfect CSI and imperfect CSI and verify the superiority of the distributed IDD architecture with LLR refinements.

**INDEX TERMS** Cell-free massive MIMO systems, centralized processing, decentralized processing, iterative detection and decoding, list-based detectors.

## I. INTRODUCTION

Cell-free massive multiple-input multiple-output (CF-mMIMO) networks exploit the distributed nature of large-scale multiple-antenna systems to improve the quality of service, yielding very high throughput [1]. In such networks, a user equipment (UE) is served by a large number of access points (APs), which are equipped with either a single antenna or multiple antennas [2], [3]. Also, the distributed nature of the network with extra spatial degrees of freedom makes the channel between UEs and APs almost orthogonal, which reduces the level of interference compared to standard cellular systems [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Stefan Schwarz.

### A. PRIOR AND RELATED WORKS

In the traditional CF-mMIMO systems, all APs serve users in the entire service area [1], [2], [4], [5]. However, such a setting is highly impractical and non-scalable since it requires increased front-haul link connections between the APs and the central processing unit (CPU). Moreover, the hardware complexity of the network grows exponentially with the increase in radio frequency units and signaling [6], [7], [8], [9]. Recently, APs selection (APs-Sel) techniques have been proposed to reduce signaling and the number of front-haul connections, whose performance is close to that of the traditional CF-mMIMO system [1], [5], [6]. For example, the work in [5] proposed a sort and connect algorithm between UEs and APs based on the effective channel gain (ECG) and channel quality. A joint power allocation algorithm is proposed to minimize the total energy consumption and reduce overhead signaling for APs-Sel [6]. This approach is compared with other APs-Sel algorithms based on the largest large-scale fading (LLSF) coefficients [2].

Moreover, another factor that influences the performance of CF-mMIMO systems is multi-user interference (MUI) due to their broadcast nature, pilot contamination, and overlapping of the uplink transmitted signals [2]. To address this issue, an efficient receiver design is necessary [1], [9], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21], [22]. Prior works on receiver design for CF-mMIMO [1], [2], [9], [10] have studied linear receivers such as receive matched filter (RMF) and minimum mean square error (MMSE) receivers for uncoded systems. However, the MUI remains constant even in high SNR regimes. To further reduce the MUI, channel codes with iterative detection and decoding (IDD) schemes have been reported in [11], [12], [13], [15], [21], [22], [27]. In which channel codes with IDD can correct unreliable detection decisions as the number of iterations increases. Moreover, using long code words yields improved performance at the expense of increased complexity in terms of time and computing power [29], [30]. IDD techniques use the message-passing principle by exchanging the soft beliefs in the form of log-likelihood ratios (LLRs) between the detector and the decoder [11], [12], [22], [23], [24], [26].

In IDD schemes, soft interference cancellation (soft-IC) detectors are generally used due to their low complexity. For example, the list-based detection techniques that are capable of eliminating error propagation are presented in [22], [23]. In [24], a local partial marginalization detector based on turbo iterations was proposed for uplink CF-mMIMO systems. The proposed detector was compared with other baseline schemes, such as MMSE and MMSE with SIC (MMSE-SIC) detectors. The authors in [25] proposed a distributed expectation detector for CF-mMIMO systems, where the CPU is equipped with a non-linear module and the APs are equipped with a linear module. Before sending the posterior mean and variance to the CPU, the APs first detect the symbols using local CSI. The extrinsic data for each AP is then generated and integrated at the CPU using maximum-ratio combining (MRC). Shaik et al. [28] performed a theoretical analysis on the distributed computation of LLRs based on the optimal maximum a posterior (MAP) detector for the sequential architecture of CF-mMIMO. However, the complexity of such an optimal filter grows exponentially with the increase in the number of UEs and antennas, which is prevalent in CF-mMIMO networks.

### B. MOTIVATION AND CONTRIBUTIONS

In the literature, few works are devoted to the bit error rate (BER) performance analysis for coded CF-mMIMO networks using message passing for various implementation architectures and cooperation levels [32], [33], [34]. There is still a need for low-complexity LLR processing schemes for decentralized processing in CF-mMIMO networks to facilitate BER performance close to that of the centralized schemes. Furthermore, list-based detection methods like MF-SIC have not been investigated for CF-mMIMO systems, despite the fact that they have the potential to improve performance by eliminating the error propagation that occurs in traditional MMSE-SIC schemes. Additionally, the use of MMSE receivers with soft-IC (MMSE-soft-IC) detectors and LPDC codes can produce a very simple, efficient, and practical implementation compared to optimal detectors. Therefore, this work develops iterative centralized and decentralized CF-mMIMO architectures, which is different from the work in [24] that considers non-iterative processing. We also propose three LLR processing strategies for low-complexity detection schemes. The main contributions of this paper can be summarized as follows:

- IDD schemes with LDPC channel codes for scalable and non-scalable centralized and decentralized CF-mMIMO schemes are developed.
- New closed-form expressions for the MMSE-Soft-IC detectors are derived for the case with APs-Sel while taking the channel estimation errors into account. Based on the derived equations, general expressions are devised for a system that uses all the APs (All-APs). Furthermore, we draw insights and derive the uplink MMSE receive filters based on the a priori information of transmitted bits.
- A new Gaussian approximation of the likelihood function is formulated by deriving novel closed-form expressions for the mean and variance of the detected signal parameters in the presence of channel estimation errors, APs selection matrix, and MUI.
- A List-MMSE-Soft-IC detector to reduce the error propagation in the IC step is proposed. The List-MMSE-Soft-IC performance is compared with other baseline detectors, such as LMMSE-Soft-IC and soft LMMSE.
- Three novel LLR processing schemes are proposed for decentralized CF-mMIMO: a) one based on decoding the LLRs at each AP (Standard LLR processing); b) another based on censoring the LLRs by decoding each user equipment (UE) information at the AP, where it achieves the largest mean absolute value of LLRs (LLR Censoring). This censoring of LLRs helps to

reduce redundant processing at the CPU,and c) the last that provides refinements in the LLRs by performing a linear combination of LLRs from the different APs by assuming statistical independence of the APs.

- The proposed local LLR processing strategies are compared with the traditional network, which is based on centralized processing schemes. The impact of different IDD iterations is also examined. Furthermore, the performance of the decentralized and centralized schemes is compared in terms of computational cost and the amount of required signaling between the APs and CPU.

### C. PAPER ORGANIZATION AND NOTATION

The rest of this paper is organized as follows: Section II presents the proposed signal model, the channel estimation procedure, the received signal statistics, and the APs selection scheme. It also presents the proposed centralized IDD scheme and receiver design. The decentralized IDD processing scheme and receiver designs are presented in Section IV. The proposed List-based detector is presented in section V. The iterative processing, LLR statistics, refinement, computational complexity, signaling loads, and the decoding algorithm are presented in Section VI. Numerical results, network setup, assumptions, and remarks are presented in Section VII. Finally, Section VIII gives the conclusions.

*Notation:* Lower and upper bold case symbols represent vectors and matrices, respectively. The operator $(\cdot)^T$, $(\cdot)^H$ denote the transpose and the Hermitian transpose, respectively.

## II. PROPOSED SYSTEM AND SIGNAL MODEL

In this section, we consider the uplink of a CF-mMIMO system with $L$ APs and $K$ single-antenna user equipment (UE), where each AP is equipped with $N$ receive antennas. The system is assumed to have imperfect channel estimates. The received signal statistics and channel estimation procedures are given below.

### A. UPLINK PILOT TRANSMISSION AND CHANNEL ESTIMATION

We assume that $\tau_p$ mutually orthogonal pilot sequences $\boldsymbol{\psi}_1,\ldots,\boldsymbol{\psi}_{\tau_p}$ with $||\boldsymbol{\psi}_t||^2 = \tau_p$ are used to estimate the channel. Furthermore, $K > \tau_p$ is such that more than one UE can be assigned per pilot. The index of UE $k$ that uses the same pilot is denoted as $t_k \in \{1,\ldots,\tau_p\}$ with $\vartheta_k \subset \{1,\ldots,K\}$ as the subset of UEs that use the same pilot as UE $k$ inclusive. The complex received signal at the $l$-th AP after the UE transmission, [1], [2], [33] $\mathbf{Y}_l$, with dimensions $N \times \tau_p$, is given by

$$\mathbf{Y}_l = \sum_{j=1}^{K} \sqrt{\eta_j}\mathbf{g}_{jl}\boldsymbol{\psi}_{t_j}^T + \mathbf{N}_l, \tag{1}$$

where $\eta_j$ is the transmit power from UE $j$, $\mathbf{N}_l$ is the received noise signal with independent $\mathcal{N}_{\mathbb{C}} \sim (0, \sigma^2)$ entries and noise power $\sigma^2$, $\mathbf{g}_{jl} \sim \mathcal{N}_{\mathbb{C}}(0, \boldsymbol{\Omega}_{jl})$, and $\boldsymbol{\Omega}_{jl} \in \mathbb{C}^{N\times N}$ is the spatial correlation matrix that describes the channel's spatial properties between the $k$-th UE and the $l$-th AP, $\beta_{k,l} \triangleq \frac{\mathrm{tr}(\boldsymbol{\Omega}_{jl})}{N}$ is the large-scale (LS) fading coefficient. The first AP correlates the received signal with the associated normalized pilot signal $\boldsymbol{\psi}_{t_k}/\sqrt{\tau_p}$ to $\mathbf{y}_{t_{kl}} \triangleq \frac{1}{\sqrt{\tau_p}}\mathbf{Y}_l\boldsymbol{\psi}_{t_k}^* \in \mathbb{C}^N$ to estimate the channel $\mathbf{g}_{jl}$ given by

$$\mathbf{y}_{t_{kl}} = \sum_{j\in\vartheta_k} \sqrt{\eta_j\tau_p}\mathbf{g}_{jl} + \mathbf{n}_{t_{kl}}, \tag{2}$$

where $\mathbf{n}_{t_{kl}} \triangleq \frac{1}{\sqrt{\tau_p}}\mathbf{N}_l\boldsymbol{\psi}_{t_k}^* \sim \mathcal{N}_c(0, \sigma^2\mathbf{I}_N)$ is the obtained noise sample after estimation. From [1], the MMSE estimate of $\mathbf{g}_{kl}$ is given by

$$\hat{\mathbf{g}}_{kl} = \sqrt{\eta_k\tau_p}\boldsymbol{\Omega}_{kl}\Psi_{t_{kl}}^{-1}\mathbf{y}_{t_{kl}}, \tag{3}$$

where $\Psi_{t_{kl}} = \mathbb{E}\{\mathbf{y}_{t_{kl}}\mathbf{y}_{t_{kl}}^H\} = \sum_{j\in\vartheta_k}\eta_j\tau_p\Omega_{jl} + \mathbf{I}_N$ is the received signal vector correlation matrix. The channel estimate $\hat{\mathbf{g}}_{kl}$ and the estimation error $\tilde{\mathbf{g}}_{kl} = \mathbf{g}_{kl} - \hat{\mathbf{g}}_{kl}$ are independent with distributions $\hat{\mathbf{g}}_{kl} \sim \mathcal{N}_c(0, \eta_k\tau_p\boldsymbol{\Omega}_{kl}\Psi^{-1}\boldsymbol{\Omega}_{kl})$ and $\tilde{\mathbf{g}}_{kl} \sim \mathcal{N}_c(0, \mathbf{C}_{kl})$, where the matrix $\mathbf{C}_{kl}$ is given by

$$\mathbf{C}_{kl} = \mathbb{E}\left\{\tilde{\mathbf{g}}_{kl}\tilde{\mathbf{g}}_{kl}^H\right\} = \boldsymbol{\Omega}_{kl} - \eta_k\tau_p\boldsymbol{\Omega}_{kl}\Psi^{-1}\boldsymbol{\Omega}_{kl}. \tag{4}$$

It should be noted that the pilot contamination is caused by the mutual interference made by UEs using the same pilot signals in (2), which lowers the system's performance [1].

From (1), the received signal vector after stacking the channel vectors from all the APs is given by

$$\mathbf{y} = \mathbf{G}\mathbf{s} + \mathbf{n}, \tag{5}$$

where the channel matrix $\mathbf{G} \in \mathbb{C}^{NL\times K}$ has both small scale and LS fading coefficients. Vector $\mathbf{s} = [s_1, .., s_K]^T$ is the transmitted symbols with $\mathbb{E}\{s_k s_k^*\} = \rho_k$, the average transmit power is given by $\boldsymbol{\rho} = [\rho_1, .., \rho_K]^T$, $\mathbf{n}$ is the additive white Gaussian noise (AWGN).

In CF-mMIMO networks, there are limitations on the complexity and amount of signaling the APs and CPU must exchange. Both of these issues make system modeling and design almost impracticable. To solve this problem, we adopt a scalable CF-mMIMO setup that takes the selection of APs into account. This is accomplished using the APs-selection technique described as follows:

### B. ACCESS POINT SELECTION PROCEDURE

The dynamic cooperative clustering (DCC) approach described in [2], [33] is considered when selecting the APs. Unlike existing approaches for APs-Sel, the proposed approach incorporates the APs-Sel in the receive filter expression, which facilitates its computation. By letting $\mathcal{M}_k \subset \{1,\ldots,L\}$ be the subset of APs in service of UE $k$, the matrix $\mathbf{D}_{kl}$ is defined as

$$\mathbf{D}_{kl} = \begin{cases} \mathbf{I}_N & \text{if } l \in \mathcal{M}_k \\ \mathbf{0}_N & \text{if } l \notin \mathcal{M}_k. \end{cases} \tag{6}$$

The APs that provide service to a specific UE are determined by the block diagonal matrix $\mathbf{D}_k = \text{diag}(\mathbf{D}_{k1}, .., \mathbf{D}_{kL}) \in \mathbb{C}^{NL\times NL}$. Specifically, when $\mathbf{D}_k = \mathbf{I}_{NL}$, all APs serve all the UEs. However, using all APs is not scalable and practical, and thus clustering approaches such as user-centric techniques can be adopted. Then, the set of UEs that are served by AP $l$ is

$$\mathcal{D}_l = \left\{k : \text{tr}(\mathbf{D}_{kl}) \geq 1, k \in \{1, .., K\}\right\}. \tag{7}$$

It is important to note that the DCC does not alter the received signal statistics since all APs receive the broadcast signal. An essential feature of such a selection process is limiting the number of APs that take part in signal detection. The joint APs selection criterion described in [2] determines which APs will provide service to a specific UE. In this scenario, the UE designates a master AP to coordinate uplink (UL) detection and decoding based on the LLSF. The CPU then establishes a threshold value $\beta_{th}$ for non-master APs to provide service to a certain UE. A detailed explanation of the operation of the DCC approach can be found in [2]. There is a need to demodulate the transmitted symbols at the receiver. Thus, the proposed centralized IDD receiver structure is detailed in the following section.

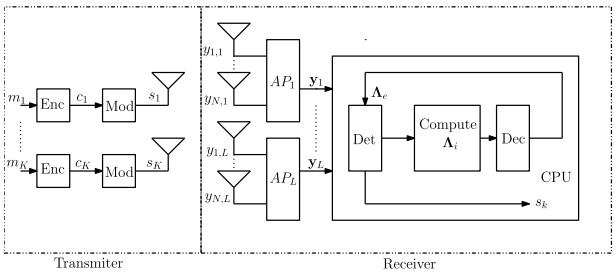



**FIGURE 1.** Block diagram for IDD scheme with centralized processing.

## III. PROPOSED CENTRALIZED IDD SCHEME

In this section, we present the proposed centralized processing architecture for coded CF-mMIMO systems. The block diagram of the proposed centralized IDD scheme is depicted in Figure 1. The codeword sequence $\boldsymbol{c}_k$ is created by first encoding the message sequence $\boldsymbol{m}_k$ for UE $k$ by an LDPC encoder (Enc) with a code rate of $R$. This encoded sequence is then modulated (Mod) to form complex symbols with a complex constellation of $2^{M_c}$ possible signal points. The $K$ UEs then send the modulated symbols to the APs. The APs serve as relays during data reception and transfer the information to the CPU, which has a joint detector (Det), an LLR computing module, and an LDPC decoder (Dec). Then, the detector forwards the data to a module that computes the LLRs $\Lambda_i$. These computed LLRs are then sent to the decoder. By providing extrinsic data $\Lambda_e$ to the joint detector, the decoder uses an iterative technique presented in Section VI that enhances the detection performance of the receiver.

### A. PROPOSED CENTRALIZED RECEIVER DESIGN

The proposed receiver configuration aims to cancel the MUI caused by the other $K-1$ UEs in the network. The receiver consists of an MMSE filter followed by a soft interference cancellation scheme, which may use either a successive or a list-based successive interference cancellation technique. The receiver first creates soft estimates of the transmitted symbols by computing the $j$th UE symbol mean $\bar{s}_j$ based on the soft beliefs from the LDPC decoder [13], [32], [33] given $\mathbb{E}\{s_j\} = \bar{s}_j$ as described by

$$\bar{s}_j = \sum_{s\in\mathcal{A}} sP(s_j = s), \tag{8}$$

where $\mathcal{A}$ is the set of complex constellations. The a priori probability of the extrinsic LLRs is given by [32], [33]

$$P(s_j = s) = \prod_{l=1}^{M_c}[1 + \exp(-s^{b_l}\Lambda_i(b_{(j-1)M_c+l}))]^{-1}, \tag{9}$$

where $\Lambda_i(b_i)$ is the extrinsic LLR of the $i$-th bit calculated by the LDPC decoder from the previous iteration, and $s^{b_l} \in (+1, -1)$ denotes the $l$-th bit of symbol $s$. The variance of the $j$-th UE symbol is calculated as [32], [33]

$$\sigma_j^2 = \sum_{s\in\mathcal{A}} |s - \bar{s}_j|^2 P(s_j = s). \tag{10}$$

After decomposition of (5) and using $\mathbf{G} = \hat{\mathbf{G}} + \tilde{\mathbf{G}}$, the received signal at the CPU is given by

$$\mathbf{y} = \hat{\mathbf{g}}_k s_k + \hat{\mathbf{G}}_{\mathrm{i}}\bar{\mathbf{s}}_{\mathrm{i}} + \sum_{m=1}^{K} \tilde{\mathbf{g}}_m s_m + \mathbf{n}, \tag{11}$$

where the first term on the right-hand side (RHS) is the desired signal, $\hat{\mathbf{g}}_k$ and $s_k$ are the approximate channel estimation vector and transmitted symbol for the desired signal, respectively. The second term is the interference from the other $K-1$ users; $\bar{\mathbf{s}}_{\mathrm{i}}$ is the vector with the interfering symbols except the $k$-th symbol, and $\hat{\mathbf{G}}_{\mathrm{i}}$ is the matrix with channel estimates of the other $K-1$ users. The third term denotes the interference due to channel estimation errors, and the fourth is the phase-rotated noise. After the estimated MUI has been removed and APs have been selected, the received symbol estimate of the $k$-th UE data stream at the CPU is given by

$$\tilde{s}_k = \mathbf{w}_k^H \mathbf{D}_k \left(\mathbf{y} - \hat{\mathbf{G}}_{\mathrm{i}}\bar{\mathbf{s}}_{\mathrm{i}}\right), \tag{12}$$

The optimization of the receive combining filter $\mathbf{w}_k$ is achieved by minimizing the mean square error (MSE) between the symbol estimate and the transmitted symbol conditioned on $\hat{\mathbf{G}}$. The formulation of the optimization problem is given by

$$\mathbf{w}_k = \arg\min_{(\mathbf{w}_k)} \mathbb{E}\left\{||\tilde{s}_k - s_k||^2 \mid \hat{\mathbf{G}}\right\}. \tag{13}$$

Differentiating the objective function on the RHS of (13) with respect to (w.r.t.) $\mathbf{w}_k^H$, the optimal MMSE receive filter $\mathbf{w}_k$ should satisfy the following relation:

$$\mathbf{D}_k \mathbb{E}\{\mathbf{y}_R \mathbf{y}_R^H\} \mathbf{D}_k^H \mathbf{w}_k - \mathbf{D}_k \mathbb{E}\{\mathbf{y}_R s_k^*\} = 0, \tag{14}$$

where $\mathbf{y}_R = \mathbf{y} - \hat{\mathbf{G}}_\mathrm{i} \bar{\mathbf{s}}_\mathrm{i}$. The solution to the filter is obtained by making $\mathbf{w}_k$ the subject of (14) as

$$\mathbf{w}_k = \left( \mathbf{D}_k \mathbb{E}\{\mathbf{y}_R \mathbf{y}_R^H\} \mathbf{D}_k^H \right)^{-1} \mathbf{D}_k \mathbb{E}\{\mathbf{y}_R s_k^*\}. \tag{15}$$

By using the orthogonality principle in [35] and assuming statistical independence of each term in the received signal $\mathbf{y}$, the terms $\mathbb{E}\{\mathbf{y}_R s_k^*\}$ and $\mathbb{E}\{\mathbf{y}_R \mathbf{y}_R^H\}$ are given by

$$\mathbb{E}\{\mathbf{y}_R s_k^*\} = \rho_k \hat{\mathbf{g}}_k, \tag{16}$$

and

$$\mathbb{E}\{\mathbf{y}_R \mathbf{y}_R^H\} = \rho_k \hat{\mathbf{g}}_k \hat{\mathbf{g}}_k^H + \hat{\mathbf{G}}_\mathrm{i} \Delta_\mathrm{i} \hat{\mathbf{G}}_\mathrm{i}^H + \sum_{m=1}^{K} \left( |\, s_m \,|^2 + \sigma_m^2 \right) \mathbf{C}_m + \sigma^2 \mathbf{I}_{NL}. \tag{17}$$

The matrix $\mathbf{\Delta}_\mathrm{i} = \mathrm{diag}\left[\sigma_1^2, \ldots, \sigma_{k-1}^2, \sigma_{k+1}^2, \ldots, \sigma_K^2\right]$ denotes the covariance matrix that consists of the entries computed in (10). By substituting (16) and (17) into (15), the centralized MMSE filter is given by

$$\mathbf{w}_k = \rho_k \left( \mathbf{D}_k \left( \rho_k \hat{\mathbf{g}}_k \hat{\mathbf{g}}_k^H + \hat{\mathbf{G}}_\mathrm{i} \Delta_\mathrm{i} \hat{\mathbf{G}}_\mathrm{i}^H \right) \mathbf{D}_k^H + \mathbf{D}_k \left( \sigma^2 \mathbf{I}_{NL} + \sum_{m=1}^{K} \left( |\, s_m \,|^2 + \sigma_m^2 \right) \mathbf{C}_m \right) \mathbf{D}_k^H \right)^{-1} \mathbf{D}_k \hat{\mathbf{g}}_k. \tag{18}$$

A detailed derivation of (18) can be found in Appendix A.

### B. INSIGHTS INTO THE CENTRALIZED MMSE FILTER

The major parameters that affect the performance of the derived receiver in (18) can be explained as follows.

- The channel estimation error, a high channel estimation error which results into imperfect channel side information (ICSI). This yields poor channel coefficients which results into wrong detection decisions, thus degrading the performance. This can be reduced by using longer pilots to estimate the channel but at the expense of increasing complexity in terms of training time. On the other hand if there is no channel estimation error the performance of the network improves and this is the case with perfect channel estimates (PCSI). For the case with PCSI, the channel estimation error $\tilde{\mathbf{g}}_m$ is 0, which makes $\mathbf{C}_m = 0$. This yields $\hat{\mathbf{g}}_k = \mathbf{g}_k$. Thus, the third term in (18) vanishes to zero, and we obtain

$$\mathbf{w}_k = \rho_k \left( \mathbf{D}_k \left( \rho_k \mathbf{g}_k \mathbf{g}_k^H + \mathbf{G}_\mathrm{i} \Delta_\mathrm{i} \mathbf{G}_\mathrm{i}^H + \sigma^2 \mathbf{I}_{NL} \right) \mathbf{D}_k^H \right)^{-1} \times \mathbf{D}_k \mathbf{g}_k. \tag{19}$$

- The number of APs and APs-Sel matrix $\mathbf{D}_k$. Increasing the number of APs in the network improves the detector performance which yields lower BERs. For the case with all APs, the selection matrix $\mathbf{D}_k = \mathbf{I}_{NL}$. Thus, the filter in (18) is given by

$$\mathbf{w}_k = \left( \rho_k \hat{\mathbf{g}}_k \hat{\mathbf{g}}_k^H + \hat{\mathbf{G}}_\mathrm{i} \Delta_\mathrm{i} \hat{\mathbf{G}}_\mathrm{i}^H + \sum_{m=1}^{K} \left( |\, s_m \,|^2 + \sigma_m^2 \right) \mathbf{C}_m + \sigma^2 \mathbf{I}_{NL} \right)^{-1} \rho_k \hat{\mathbf{g}}_k. \tag{20}$$

  Note that $\mathbf{D}_k = \mathbf{0}$, implies that $\mathbf{w}_k^H \mathbf{D}_k = \mathbf{0}$, where $\mathbf{w}_k^H$ is the receive vector. This means that only APs with $\mathbf{D}_k \neq \mathbf{0}$ will apply receive combining in the uplink detection.
- The number of IDD iterations and interference cancellation matrix $\Delta_i$. Increasing the number of iterations improves the performance. This is because more information is exchanged between the decoder and detector with improves on the interference cancellation capability of the receiver. Mathematically, this can be explained as: In the first iteration it is considered that $\bar{\mathbf{s}}_\mathrm{i} = \mathbf{0}$ in (8). In this case, we have a linear MMSE filter and the estimated signal in (12) is

$$\tilde{s}_k = \rho_k \hat{\mathbf{g}}_k^H \mathbf{D}_k^H \left( \rho_k \mathbf{D}_k \hat{\mathbf{g}}_k \hat{\mathbf{g}}_k^H \mathbf{D}_k^H + \mathbf{D}_k \hat{\mathbf{G}}_\mathrm{i}\, \mathrm{diag}\left(\boldsymbol{\rho}_\mathrm{i}\right) \hat{\mathbf{G}}_\mathrm{i}^H \mathbf{D}_k^H + \mathbf{D}_k \left( \sigma^2 \mathbf{I}_{NL} + \sum_{m=1}^{K} \rho_m \mathbf{C}_m \right) \mathbf{D}_k^H \right)^{-1} \mathbf{y}. \tag{21}$$

  The vector $\boldsymbol{\rho}_\mathrm{i}$ denotes the average transmit power vector for the other $K-1$ UEs. As the number of iterations increases, there is more a posterior information about the transmitted bit. This implies that mean of the interference symbol $\bar{\mathbf{s}}_\mathrm{i} \approx \mathbf{s}_\mathrm{i}$ in (8). Thus, the filter becomes a perfect interference canceler, and (12) yields

$$\tilde{s}_k = \rho_k \hat{\mathbf{g}}_k^H \mathbf{D}_k^H \left( \rho_k \mathbf{D}_k \hat{\mathbf{g}}_k \hat{\mathbf{g}}_k^H \mathbf{D}_k^H + \mathbf{D}_k \left( \sigma^2 \mathbf{I}_{NL} + \sum_{m=1}^{K} |\, s_m \,|^2 \mathbf{C}_m \right) \mathbf{D}_k^H \right)^{-1} \times \left( \mathbf{y} - \mathbf{D}_k \hat{\mathbf{G}}_\mathrm{i} \mathbf{s}_\mathrm{i} \right). \tag{22}$$

The centralized detection schemes experience high levels of complexity as the number of UEs, APs, and antennas at the APs increases. This makes the design of receivers more complicated, and the amount of required signaling between APs and the CPU increases. To alleviate the above problem, the decentralized IDD scheme is proposed as follows:

## IV. PROPOSED DECENTRALIZED IDD SCHEME

The proposed decentralized IDD scheme is shown in Figure 2. The transmitter operates in the same way as that of the centralized processor in Section III. At the receiver,

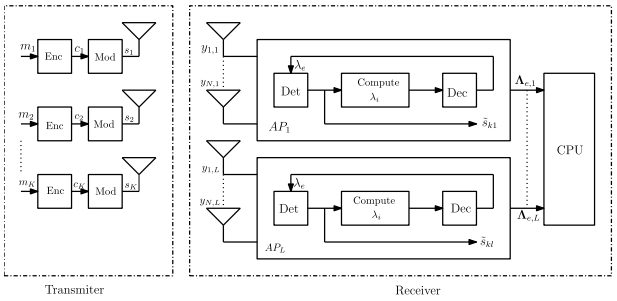



**FIGURE 2.** Block diagram for IDD scheme with decentralized processing.

each AP is equipped with a local detector, an LLR computing module, and an LDPC decoder. The APs use their local channel estimates to perform IDD on the received signals. The detector sends its symbol estimates to a module that computes the local soft information $\lambda_i$ in the form of LLRs. The computed LLRs are then sent to the decoder, which performs iterative processing by exchanging extrinsic information $\lambda_e$ with the local detector. In the decentralized operation, the APs act as compute-and-forward relays by sending their soft beliefs to the CPU for further processing. The challenge at the CPU is to design an intelligent way of processing these LLRs. We devised three techniques for processing these LLRs. The first scheme (standard LLR processing) is based on individual decisions from each AP, and an average BER is then computed based on these decisions from each AP. The second scheme considers censoring the LLRs (LLR censoring) and decoding each UE data at the AP, where it achieves the largest mean absolute value of LLRs. The third scheme is based on the linear combination of the LLRs (LLR Ref). A detailed explanation, operation, and analysis of the proposed LLR processing schemes are given in Section VI.

### A. PROPOSED DECENTRALIZED RECEIVER DESIGN

The channel statistics, estimation, and received signals follow the model introduced in Section II-A. The received signal at the $l$-th AP is given by

$$\mathbf{y}_l = \sum_{i=1}^{K} \mathbf{g}_{il} s_i + \mathbf{n}_l \in \mathbb{C}^{N\times 1}, \tag{23}$$

which can further be decomposed as

$$\mathbf{y}_l = \hat{\mathbf{g}}_{kl} s_k + \hat{\mathbf{G}}_{\mathrm{i}l}\mathbf{s}_{\mathrm{i}} + \sum_{m=1}^{K} \tilde{\mathbf{g}}_{ml} s_m + \mathbf{n}_l, \tag{24}$$

where the first term on the RHS is the desired signal, the second term is the interference from the other $K-1$ users, the third term denotes the interference due to channel estimation errors, and the fourth term is the phase-rotated noise. The received local symbol estimate of the $k$-th UE data stream at the $l$-th AP after removing the MUI is given by

$$\tilde{s}_{kl} = \mathbf{w}_{kl}^H \mathbf{D}_{kl} \left( \mathbf{y}_l - \hat{\mathbf{G}}_{\mathrm{i}l}\bar{\mathbf{s}}_{\mathrm{i}} \right), \tag{25}$$

where the notation $\mathbf{D}_{kl}$ implies that the $l$-th AP is among the selected APs. Here, the optimization of the received combining filter $\mathbf{w}_{kl}$ is obtained by minimizing the error between the estimated detected symbol and the transmitted symbol. The optimization problem is formulated as

$$\mathbf{w}_{kl} = \arg\min_{(\mathbf{w}_{kl})} \mathbb{E}\left\{ ||\tilde{s}_{kl} - s_k||^2 \mid \hat{\mathbf{G}}_l \right\}. \tag{26}$$

The derivation is similar to the centralized approach, and, for completeness, it is described in detail in Appendix B. The optimal receive filter $\mathbf{w}_{kl}$ should satisfy the following relation:

$$\mathbf{D}_{kl}\mathbb{E}\{\mathbf{y}_{Rl}\mathbf{y}_{Rl}^H\}\mathbf{D}_{kl}^H\mathbf{w}_{kl} - \mathbf{D}_{kl}\mathbb{E}\{\mathbf{y}_{Rl}s_k^*\} = 0. \tag{27}$$

Thus, the solution to the receive filter is given by

$$\mathbf{w}_{kl} = \left(\mathbf{D}_{kl}\mathbb{E}\{\mathbf{y}_{Rl}\mathbf{y}_{Rl}^H\}\mathbf{D}_{kl}^H\right)^{-1}\mathbf{D}_{kl}\mathbb{E}\{\mathbf{y}_{Rl}s_k^*\}. \tag{28}$$

The terms $\mathbb{E}\{\mathbf{y}_{Rl}s_k^*\}$ and $\mathbb{E}\{\mathbf{y}_{Rl}\mathbf{y}_{Rl}^H\}$ are respectively, given by

$$\mathbb{E}\{\mathbf{y}_{Rl}s_k^*\} = \rho_k\hat{\mathbf{g}}_{kl}, \tag{29}$$

$$\mathbb{E}\{\mathbf{y}_{Rl}\mathbf{y}_{Rl}^H\} = \rho_k\hat{\mathbf{g}}_{kl}\hat{\mathbf{g}}_{kl}^H + \hat{\mathbf{G}}_{\mathrm{i}l}\Delta_{\mathrm{i}l}\hat{\mathbf{G}}_{\mathrm{i}l}^H + \sum_{m=1}^{K}\left(|\,s_m\,|^2 + \sigma_m^2\right)\mathbf{C}_{ml} + \sigma^2\mathbf{I}_N, \tag{30}$$

where $\Delta_{\mathrm{i}l}$ denotes the covariance matrix that consists of the entries computed in (10), locally computed at the $l$-th AP. By substituting (29) and (30) into (28), the solution of the local receive filter $\mathbf{w}_{kl}^*$ is given by

$$\mathbf{w}_{kl} = \rho_k\left(\mathbf{D}_{kl}\left(\rho_k\hat{\mathbf{g}}_{kl}\hat{\mathbf{g}}_{kl}^H + \hat{\mathbf{G}}_{\mathrm{i}l}\Delta_{\mathrm{i}l}\hat{\mathbf{G}}_{\mathrm{i}l}^H + \sigma^2\mathbf{I}_N + \sum_{m=1}^{K}\left(|\,s_m\,|^2 + \sigma_m^2\right)\mathbf{C}_{ml}\right)\mathbf{D}_{kl}^H\right)^{-1}\mathbf{D}_{kl}\hat{\mathbf{g}}_{kl}. \tag{31}$$

### B. INSIGHTS INTO THE DECENTRALIZED MMSE FILTER

In what follows, we draw insights into the derived local MMSE filter expression to study the impact of the major parameters and scenarios that affect the receiver's performance for the cases with a selected AP, PCSI, and different number of iterations. When the local AP is selected with all its antennas, the selection matrix $\mathbf{D}_{kl} = \mathbf{I}_N$. The local filter in (31) is

$$\mathbf{w}_{kl} = \rho_k\left(\rho_k\hat{\mathbf{g}}_{kl}\hat{\mathbf{g}}_{kl}^H + \hat{\mathbf{G}}_{\mathrm{i}l}\Delta_{\mathrm{i}l}\hat{\mathbf{G}}_{\mathrm{i}l}^H + \left(\sum_{m=1}^{K}\left(|\,s_m\,|^2 + \sigma_m^2\right)\mathbf{C}_{ml} + \sigma^2\mathbf{I}_N\right)\right)^{-1}\hat{\mathbf{g}}_{kl}. \tag{32}$$

Assuming PCSI, the fourth term in (31) vanishes to zero, and we obtain the filter given by

$$\mathbf{w}_{kl} = \rho_k\left(\rho_k\mathbf{D}_{kl}\mathbf{g}_{kl}\mathbf{g}_{kl}^H\mathbf{D}_{kl}^H + \mathbf{D}_{kl}\mathbf{G}_{\mathrm{i}l}\Delta_{\mathrm{i}l}\mathbf{G}_{\mathrm{i}l}^H\mathbf{D}_{kl}^H + \sigma^2\mathbf{D}_{kl}\mathbf{I}_N\tilde{\mathbf{D}}_{kl}^H\right)^{-1}\mathbf{g}_{kl}\mathbf{D}_{kl}. \tag{33}$$

In the first iteration, $\bar{\mathbf{s}}_{\mathrm{i}} = \mathbf{0}$ in (8), which yields the linear MMSE filter and the estimated signal in (25) given by

$$\tilde{s}_{kl} = \rho_k \hat{\mathbf{g}}_{kl}^H \mathbf{D}_{kl}^H \Bigg( \mathbf{D}_{kl} \Bigg( \rho_k \hat{\mathbf{g}}_{kl} \hat{\mathbf{g}}_{kl}^H + \hat{\mathbf{G}}_{\mathrm{i}l} \mathrm{diag}\,(\rho)\, \hat{\mathbf{G}}_{\mathrm{i}l}^H + \sum_{m=1}^{K} \left( |\, \bar{s}_m \,|^2 + \sigma_m^2 \right) \mathbf{C}_{ml} + \sigma^2 \mathbf{I}_N \Bigg) \mathbf{D}_{kl}^H \Bigg)^{-1} \mathbf{y}_l, \tag{34}$$

where the parameter $\boldsymbol{\rho}$ denotes the average transmit power vector for the other $K-1$ UEs. The mean symbol $\bar{\mathbf{s}}_{\mathrm{i}} \approx \mathbf{s}_{\mathrm{i}}$ in (8) increases as the number of iterations increases. Thus, the filter becomes a perfect interference canceler. (25) yields

$$\tilde{s}_{kl} = \rho_k \hat{\mathbf{g}}_{kl}^H \mathbf{D}_{kl}^H \Bigg( \mathbf{D}_{kl} \Bigg( \rho_k \hat{\mathbf{g}}_{kl} \hat{\mathbf{g}}_{kl}^H + \sum_{m=1}^{K} |\, s_m \,|^2 \, \mathbf{C}_{ml} + \sigma^2 \mathbf{I}_N \Bigg) \mathbf{D}_{kl}^H \Bigg)^{-1} \left( \mathbf{y}_l - \mathbf{D}_{kl} \hat{\mathbf{G}}_{il} \mathbf{s}_{\mathrm{i}} \right). \tag{35}$$

The derived MMSE-SIC filters in (18) and (31) experience error propagation at each sequential step. In what follows, we describe the proposed list-based detector that is capable of suppressing the error propagation that occurs at each cancellation step.

## V. PROPOSED LIST-BASED DETECTOR

In this section, we detail the proposed list-based detection scheme shown in Figure 3, which is similar to the one adopted in our previous papers [32], [33].

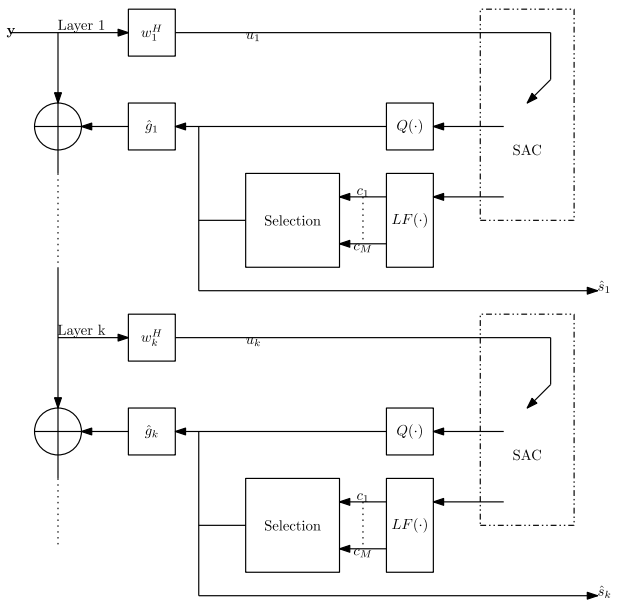



**FIGURE 3.** Block diagram of the proposed list-based detector.

The design takes advantage of list feedback (LF) diversity by selecting a list of constellation candidates if there is the unreliability of the previously detected symbols [32], [33]. A shadow area constraint (SAC) is employed to obtain an optimal feedback candidate. The SAC is capable of reducing the search space from exponentially growing as well as reducing computational complexity. The key idea of such a selection criterion is to avoid redundant processing when there is a reliable decision. The procedure for obtaining the detected symbol $\hat{s}_k$ of the $k$-th user is analogous to the steps presented in [22], [32], [33]. The $k$-th user soft estimate is obtained by $u_k = \mathbf{w}_k^H \mathbf{D}_k \check{\mathbf{y}}_k$. The filter $\mathbf{w}_k$ is the receive MMSE filter described in (15) and later in (28). The residual signal $\check{\mathbf{y}}_k = \mathbf{y} - \sum_{t=1}^{k-1} \hat{\mathbf{g}}_t \hat{s}_t$ is the received vector following the soft cancellation of the $k-1$ symbols that were previously detected. The SAC assesses the reliability of this decision using the soft estimate $u_k$ for each layer according to

$$d_k = |u_k - v_f|, \tag{36}$$

where $v_f = \arg\min_{v_{\mathrm{f}} \in \mathcal{A}} \{|u_k - v_f|\}$ denotes the closest constellation point to the $k$-th user soft estimate $u_k$. If $d_k > d_{\mathrm{th}}$, the chosen constellation point gets dumped into the shadow area of the constellation map since the choice is deemed to be unreliable. The parameter $d_{\mathrm{th}}$ is the predefined threshold on the Euclidean distance to ensure the reliability of the selected symbol [32], [33]. The list-based algorithm performs hard slicing for UE $k$ as in the soft-IC if the soft estimate $u_k$ is reliable. In this case, $\hat{s}_k = Q(u_k)$ is the estimated symbol, where $Q(\cdot)$ is the quantization notation which maps to the constellation symbol closest to $u_k$.

Otherwise, the decision is deemed unreliable, and a candidate list $\mathcal{L} = \{c_1, c_2, \ldots, c_m, \ldots, c_M\} \subseteq \mathcal{A}$ is generated, which is made up of the $M$ constellation points that are closest to $u_k$, where $M \leq 2^{M_c}$. The algorithm selects an optimal candidate $c_{m,\mathrm{opt}}$ from a list of $\mathcal{L}$ candidates. Thus, the unreliable choice $Q(u_k)$ is replaced by a hard decision, and $\hat{s}_k = c_{m,\mathrm{opt}}$ is obtained. The list-based detector first defines the selection vectors $\boldsymbol{\phi}^1, \boldsymbol{\phi}^2, \ldots, \boldsymbol{\phi}^m, \ldots \boldsymbol{\phi}^M$ whose size is equal to the number of the constellation candidates that are used every time a decision is considered unreliable. For example, for the $k$-th layer, a $K \times 1$ vector $\boldsymbol{\phi}^m = \left[\hat{s}_1, \ldots, \hat{s}_{k-1}, c_m, \phi_{k+1}^m, \ldots, \phi_q^m, \ldots, \phi_K^m\right]^T$ which is a potential choice corresponding to $c_m$ in the $k$-th user, comprises the following items: (a) the previously estimated symbols $\hat{s}_1, \hat{s}_2, \ldots, \hat{s}_{k-1}$; (b) the candidate symbol $c_m$ obtained from the constellation for subtracting a decision that was considered unreliable $Q(u_k)$ of the $k$-th user; (c) using (a) and (b) as the previous decisions, detection of the next user data $k+1, \ldots, q, \ldots, K$-th is performed by the soft-IC approach. Mathematically, the choice $\phi^m$ is given by [33]

$$\phi_q^m = Q(\mathbf{w}_q^H \hat{\mathbf{y}}_q^m), \tag{37}$$

where the index $q$ denotes a UE between the $k+1$-th and the $K$-th UE, $\hat{\mathbf{y}}_q^m = \check{\mathbf{y}}_k - \mathbf{D}_k \hat{\mathbf{g}}_k c_m - \mathbf{D}_k \sum_{p=k+1}^{q-1} \hat{\mathbf{g}}_p \phi_p^m$. A key attribute of the list-based detector is that the same MMSE filter $\mathbf{w}_k$ is used for all the constellation candidates. Therefore, it has a computational cost that is close to that of the conventional soft-IC. The optimal candidate $m_{\mathrm{opt}}$ is selected according to the local maximum likelihood (ML)

rule given by

$$m_{\text{opt}} = \arg \min_{1 \le m \le M} \left\| \mathbf{D}_k \boldsymbol{y} - \mathbf{D}_k \hat{\mathbf{G}} \boldsymbol{\phi}^m \right\|_2^2. \tag{38}$$

The derived centralized and decentralized filters suffer from interference due to the other $K-1$ users, channel estimation errors, and AWGN noise. This makes the derived filters non-Gaussian because the output of a Gaussian filter should be Gaussian for a Gaussian input. In the next section, we approximate the filters to be Gaussian by computing the mean and variances, present the LLR processing schemes, perform signaling and computational complexity analysis, and explain the considered decoding algorithm.

## VI. ITERATIVE PROCESSING AND REFINEMENT

This section presents the iterative processing of the IDD schemes for the studied MMSE-based detectors, which employ a detector and an LDPC decoder, and the proposed LLR refinement techniques. The received signal at the output of the receive filter contains the desired symbol, MUI, and noise. The parameter $u_k$ in Figure 3 is assumed to be an output of an AWGN channel [24], [32] given by

$$u_k = \omega_k s_k + z_k, \tag{39}$$

where $\mathbb{E}\{s_k^* u_k\} = \rho_k \mathbf{w}_k^H \tilde{\mathbf{D}}_k \hat{\mathbf{g}}_k$ and $\mathbb{E}\{s_k^* u_{kl}\} = \rho_k \mathbf{w}_{kl}^H \tilde{\mathbf{D}}_{kl} \hat{\mathbf{g}}_{kl}$ are for the centralized and decentralized schemes, respectively. The parameter $z_k$ is a zero-mean AWGN variable. Using similar procedures as in [24], [32], the variance $\kappa^2$ of $z_k$ is computed by $\kappa^2 = \mathbb{E}\left\{ \mid u_k - \omega_k s_k \mid^2 \right\}$:

$$\kappa^2 = \mathbf{w}_k^H \mathbf{D}_k \left( \hat{\mathbf{G}}_{\text{i}} \Delta_{\text{i}} \hat{\mathbf{G}}_{\text{i}}^H + \sum_{m=1}^{K} \rho_m \mathbf{C}_m + \sigma^2 \mathbf{I}_{NL} \right) \mathbf{D}_k^H \mathbf{w}_k, \tag{40}$$

and

$$\kappa_l^2 = \mathbf{w}_{kl}^H \mathbf{D}_{kl} \left( \hat{\mathbf{G}}_{\text{i}l} \Delta_{\text{i}} \hat{\mathbf{G}}_{\text{i}l}^H + \sum_{m=1}^{K} \rho_m \mathbf{C}_{ml} + \sigma^2 \mathbf{I}_N \right) \mathbf{D}_{kl}^H \mathbf{w}_{kl}, \tag{41}$$

are for the centralized and decentralized schemes, respectively.

Detailed derivations of (40) and (41) are presented in Appendices C and D for centralized and decentralized processing schemes, respectively. The extrinsic LLR computed by the detector for the $l$-th bit $l \in \{1, 2, \ldots, M_c\}$ of the symbol $s_k$ [32], [33] is given by

$$\begin{aligned} &\Lambda_e \left( b_{(k-1)M_c+l} \right) \\ &= \frac{\log P\left( b_{(k-1)M_c+l} = +1 | u_k \right)}{\log P\left( b_{(k-1)M_c+l} = -1 | u_k \right)} - \frac{\log P\left( b_{(k-1)M_{c+1}} = +1 \right)}{\log P\left( b_{(k-1)M_{c+1}} = -1 \right)} \\ &= \log \frac{\sum_{s \in A_l^{+1}} f\left( u_k | s \right) P\left( s \right)}{\sum_{s \in A_l^{-1}} f\left( u_k | s \right) P\left( s \right)} - \Lambda_i \left( b_{(k-1)M_c+l} \right), \end{aligned} \tag{42}$$

where the last equality of (42) follows from Bayesian rule. The parameter $A_l^{+1}$ is the set of $2^{Mc-1}$ hypotheses for which the $l$-th bit is $+1$. The a-priori probability $P(s)$ is given by (9). The approximation of the likelihood function [32], [33] $f(u_k|s)$ is given by

$$f\left( u_k | s \right) \simeq \frac{1}{\pi \kappa^2} \exp\left( -\frac{1}{\kappa^2} | u_k - \omega_k s |^2 \right). \tag{43}$$

After local processing, the CPU has to perform final decisions by using the LLRs from the different APs. This is accomplished by proposing three LLR processing strategies presented as follows:

### A. STANDARD LLR PROCESSING

In this strategy, each AP computes the BER based on decisions from its LLRs. After obtaining the BER from each AP, an average BER is calculated for the entire network. However, such an approach yields poor results, as some APs have very unreliable LLRs for particular UEs. We then discuss two proposed strategies to improve the performance of local detectors.

### B. LLR CENSORING

In this subsection, we present an LLR censoring technique that helps to reduce the redundant processing of LLRs at the CPU. First, the independent streams of LLRs are sent from the APs to the CPU with dimensions $KC_{\text{leng}}$. At each AP, we compute the mean absolute value of the LLRs, which is given by

$$\mu_{\Lambda_{kl,e}} = \frac{1}{C_{\text{leng}}} \sum_{c=1}^{C_{\text{leng}}} | \Lambda_{l,e} |. \tag{44}$$

Based on $\mu_{\Lambda_{kl,e}}$, the UE is decoded at the AP when this parameter is highest and the other LLRs are discarded. This is done for all APs, and a new matrix $\boldsymbol{\Lambda}_{\boldsymbol{k,e}}^{\text{new}}$ with the censored LLRs is formed and used in performing final decoding. The LLR censoring strategy is summarized in Algorithm 1.

**Algorithm 1** Algorithm for Censoring Local LLRs

$\boldsymbol{\Lambda}_e \in \mathbb{C}^{KC_{\text{leng}}L}$, $\boldsymbol{\Lambda}_{\boldsymbol{k,e}}^{\text{new}} = \mathbf{0}_{KC_{\text{leng}}}$
**for** l=1 to L **do**
  **for** k=1 to K **do**
    **if** $\mu_{\Lambda_{kl,e}} \ge \max\left(\mu_{\boldsymbol{\Lambda}_{k,e}}\right)$ **then**
      $\boldsymbol{\Lambda}_{\boldsymbol{k,e}}^{\text{new}} = \boldsymbol{\Lambda}_{kl,e}$
    **else**
      Continue
    **end if**
    $k \leftarrow k+1$
  **end for**
  $l \leftarrow l+1$
**end for**
Output $\boldsymbol{\Lambda}_{\boldsymbol{k,e}}^{\text{new}}$

### C. LLR REFINEMENT

We propose an LLR refinement strategy that computes the linear summation of the multiple streams of LLRs obtained from the locally computed joint IDD detectors.

Mathematically, the refined combination of LLRs at the CPU is given by

$$\Lambda_{avg,e}\left(b_{(k-1)M_c+l}\right) = \sum_{l=1}^{L} \Lambda_{l,e}\left(b_{(k-1)M_c+l}\right). \quad (45)$$

The idea of combining multiple streams of LLRs creates some diversity benefits from the LLRs and yields some performance improvement in the network. Another key advantage of decentralized processing is that each AP has accurate channel estimates; thus, it is better to perform the detection locally than at the CPU. The mean of the refined LLRs is given by

$$\mathbb{E}[\Lambda_{avg,e}\left(b_{(k-1)M_c+l}\right)] = \sum_{l=1}^{L} \mathbb{E}[\Lambda_{l,e}\left(b_{(k-1)M_c+l}\right)], \quad (46)$$

$$= \mu_{\Lambda_{avg,e}}, \quad (47)$$

where $\mathbb{E}[\Lambda_{l,e}\left(b_{(k-1)M_c+l}\right)] \to 0$ since

$$\mathbb{E}[\Lambda_{l,e}\left(b_{(k-1)M_c+l}\right)] = \log \int \Lambda_{l,e}\left(b_{(k-1)M_c+l}\right) \times p_{\Lambda_{l,e}|H_0} d\Lambda_{l,e},$$

$$= \log \int \frac{p_{\Lambda_{l,e}|H_1}}{p_{\Lambda_{l,e}|H_0}} p_{\Lambda_{l,e}|H_0} d\Lambda_{l,e} = 0,$$

where $p_{\Lambda_{l,e}|H_1}$ is the conditional probability density function (pdf) of the LLR of stream $l$ given bit 1 and $p_{\Lambda_{l,e}|H_0}$ is the conditional pdf of the LLR given bit 0. The variance of the refined LLRs is given by

$$\sigma^2_{\Lambda_{avg,e}} = \mathbb{E}[\left(\Lambda_{avg,e}\left(b_{(k-1)M_c+l}\right) - \mu_{\Lambda_{avg,e}}\right)^2]$$

$$= \frac{1}{L}\sum_{l=1}^{L}\Big(\left(\Lambda_{avg,e}\left(b_{(k-1)M_c+l}\right)\right)^2 - \sum_{n=1}^{L}\Lambda_{n,e}\left(b_{(k-1)M_c+l}\right)\Big).$$

This suggests that the refinement benefits come from enhancing the quality of the LLRs through their variance reduction, which shifts the LLRs with small values away from the origin.

### D. COMPUTATIONAL COMPLEXITY

We consider the worst-case with all APs to assess the computational complexity of obtaining the studied detection schemes. A key observation from the derived expressions is that decentralized detection reduces the complexity at the CPU in terms of computations since each AP locally detects its signal based on the available channel estimates, i.e., local detection only requires $N \times N$ matrix inversions. On the other hand, $NL \times NL$ matrix inversions are required for the centralized processing scenario since all the combined signal is detected as a whole at the CPU, which increases the complexity of the detectors. However, the CPU is designed to have high processing power to handle such complexity [1], [2]. A detailed complexity analysis for the considered detectors can be found in Table 1. It can be observed that the computational complexity of the decentralized and centralized detectors is of the order $\mathcal{O}(N^2LK)$ and $\mathcal{O}(N^2L^2K)$, respectively. Where $\mathcal{O}(\cdot)$ is the big O notation.

**TABLE 1. Computational complexity per detector.**

| Detector | Multiplications |
|---|---|
| Decentralized-MMSE | $2N^2LK + 2K^2NL + 8KNL + 4KL2^{M_c} + 2M_cKL2^{M_c} + KL$ |
| Centralized-MMSE | $2N^2L^2K + 8KNL + 2K^2NL + 4K2^{M_c} + 2M_cK2^{M_c} + K$ |
| Decentralized-SIC | $4N^2LK + 2K^2NL + 8KNL + 9KL2^{M_c} + 4M_cKL2^{M_c} + KL$ |
| Centralized-SIC | $2(NLK)^2 + 2(NL)^2K + K^2NL + 5NLK + K + 9K2^{M_c} + 4M_cK2^{M_c}$ |
| Decentralized-List | $4N^2LK + 5K^2NL + 12KNL + 9KL2^{M_c} + 4M_cKL2^{M_c} + 2KL$ |
| Centralized-List | $2(NLK)^2 + 2(NL)^2K + 3K^2NL + 9NLK + K + 9K2^{M_c} + 4M_cK2^{M_c}$ |

### E. SIGNALING ANALYSIS

In CF-mMIMO, the APs detect the signals locally or delegate the task fully or partially to the CPU. However, there should be a tradeoff between the required front-haul signaling amount and detection performance [1]. Both the centralized and the decentralized processing require $\left(\tau_c - \tau_p\right)N$ scalars for the uplink received data and $\tau_p N$ complex scalars for the pilot sequences. Additionally, the centralized processing requires the $\frac{KLN^2}{2}$-dimensional spatial correlation matrix $\mathbf{\Omega}_{kl}$. For decentralized processing, the CPU does not require any statistical parameters for the spatial correlation matrix since the local channel estimates exist at the APs. However, the CPU should know the $KC_{\text{leng}}L$-dimensional matrix of the LLRs in (45) to compute the average, where $C_{\text{leng}}$ is the code word length. Thus, the signaling is summarized in Table 2 and is analogous to the one in [1], with additional knowledge of the dimension of the LLR matrix for the final decoding of the LLRs received from local processors.

**TABLE 2. Number of complex sequences to share via fronthaul connections, from APs to CPU.**

| Processing Scenario | Each Coherence block | Statistical Parameters |
|---|---|---|
| Centralized | $\tau_c NL$ [1] | $KLN^2/2$ [1] |
| Decentralized | $(\tau_c - \tau_p)\,KL$ [1] | – |

### F. DECODING ALGORITHM

The proposed detectors and the decoder iteratively exchange soft beliefs. The tangent function degrades the performance of the conventional sum-product algorithm (SPA), especially in the error-rate region [33]. Since the box-plus SPA produces less complex approximations, we use it in this paper [32], [33]. The single parity check (SPC) stage and the repetition stage are two steps that make up the decoder. The LLR sent

from check node $(CN)_J$ to variable node $(VN)_i$ is computed as

$$\Lambda_{j\longrightarrow i} = \boxplus_{i'\in N(j)\setminus i}\Lambda_{i'\longrightarrow j}, \tag{48}$$

where $\boxplus$ denotes the pairwise "box-plus" operator given by

$$\Lambda_1 \boxplus \Lambda_2 = \log\left(\frac{1+e^{\Lambda_1+\Lambda_2}}{e^{\Lambda_1}+e^{\Lambda_2}}\right), \tag{49}$$

$$= \text{sign}(\Lambda_1)\text{sign}(\Lambda_2)\min(|\Lambda_1|,|\Lambda_2|) + \log\left(1+e^{-|\Lambda_1+\Lambda_2|}\right) - \log\left(1+e^{-|\Lambda_1-\Lambda_2|}\right). \tag{50}$$

The LLR from $VN_i$ to $CN_j$ is given by

$$\Lambda_{i\longrightarrow j} = \Lambda_i + \sum_{j'\in N(i)\setminus j}\Lambda_{j'\longrightarrow i}, \tag{51}$$

where the parameter $\Lambda_i$ denotes the LLR at $VN_i$, $j' \in N(i)\setminus j$ means that all CNs connected to $VN_i$ except $CN_j$

## VII. SIMULATION RESULTS

In this section, the bit error rate (BER) performance of the proposed soft detectors is presented for the CF-mMIMO settings. The CF-mMIMO channel exhibits high PL values due to LS fading coefficients. Thus, the SNR is expressed by

$$SNR = \frac{\sum_{l=1}^{L}(\mathbf{G}_l\ \text{diag}\ (\boldsymbol{\rho})\ \mathbf{G}_l^H)}{\sigma_w^2 NLK}. \tag{52}$$

The simulation parameters are varied according to Table 3, unless stated otherwise.

**TABLE 3. Simulation parameters.**

| Parameter | Value |
|---|---|
| Codeword length ($C_{\text{leng}}$) | 256 |
| Parity Check bits ($M$) | 128 |
| Message bits | $C_{\text{leng}} - M$ |
| Code rate $R$ | $\frac{1}{2}$ |
| Threshold Euclidean distance ($d_{\text{th}}$) | 0.38 |
| $\tau_u, \tau_p, \tau_c$ | 190, 10, 200 |
| $\eta_k$ | 100 mW |
| Threshold for non-master AP ($\beta_{\text{th}}$) | −20 dB |
| Maximum decoder iterations | 10 |
| Signal power $\rho$ | 1 W |
| Maximum decoder iterations | 10 |
| Square length of CF-mMIMO Network $D$ | 1 km |
| Angular standard deviation | 15° |
| Noise power | −96 dBm |
| Bandwidth | 20 $MHz$ |
| Number of channel realizations | 10000 |

*Network Setup, Assumptions and Remarks:* We consider a cell-free environment with a square of dimensions $D \times D$. The spatial correlation matrices $\boldsymbol{\Omega}_{jl}$ are assumed to be locally available at the APs, and their entries are generated using the Gaussian local scattering model [2], [36] with an angular standard deviation defined in Table 3. The modulation scheme used is QPSK. The LS fading coefficients are obtained according to the 3rd Generation Partnership Project (3GPP) Urban Microcell model in [2] given by

$$\beta_{k,l}\,[\text{dB}] = -30.5 - 36.7\log_{10}\left(\frac{d_{kl}}{1m}\right) + \Upsilon_{kl}, \tag{53}$$

where $d_{kl}$ is the distance between the $k$-th UE and $l$-th AP, $\Upsilon_{kl} \sim \mathcal{N}(0, 4^2)$ is the shadow fading. We believe that the considered propagation channel model is sufficiently general to allow simple changes and assessments of line of sight (LoS), pathloss, and shadowing distributions for evaluating CF-mMIMO networks, as recommended in the literature for microcell scenarios [2]. The simulation results are based on single antenna UEs for simplicity of analysis, the test of ideas, and to allow a shorter simulation time. However, this can be extended to multiple antenna UEs to fully address the most practical systems. Note should also be taken that the considered scenarios and network settings in terms of codeword length, number of APs, and antennas provide reasonably acceptable performances in terms of BER. However, improvements in the BER can be obtained by using longer channel codes and increasing the number of APs and antennas in the network at the expense of increased complexity.

In Table 4, we provide numerical values for the number of multiplications that should be done per iteration for each detector for different numbers of UEs $K$, APs $L$ and APs antennas $N$. It can be seen that the linear MMSE-based receivers have low complexity values as compared to the SIC and list-based receivers. Also, it is noteworthy that decentralized receivers have lower complexity values as compared to centralized receivers. Nevertheless, the proposed list-based receivers achieve costs that are slightly higher than those of the SIC-based receivers. However, using list-based detection improves performance, as shown in the numerical results.

**TABLE 4. Cost in number of multiplications for the detectors.**

| Detector | Cost for $K = 8$, $L = 16, N = 4$ | Cost for $K = 8$, $L = 32, N = 4$ |
|---|---|---|
| Decentralized-MMSE | $1.70 \times 10^3$ | $3.40 \times 10^4$ |
| Centralized-MMSE | $7.81 \times 10^4$ | $2.87 \times 10^5$ |
| Decentralized-SIC | $2.57 \times 10^4$ | $5.15 \times 10^4$ |
| Centralized-SIC | $5.97 \times 10^5$ | $2.37 \times 10^6$ |
| Decentralized-List | $3.84 \times 10^4$ | $7.68 \times 10^4$ |
| Centralized-List | $6.08 \times 10^5$ | $2.39 \times 10^6$ |

Figure 4 presents the BER versus the SNR for the cases (a) before LLR refinement (w/o-LLR-Ref) and (b) after LLR refinement (w-LLR-Ref) for the studied detectors. It can be

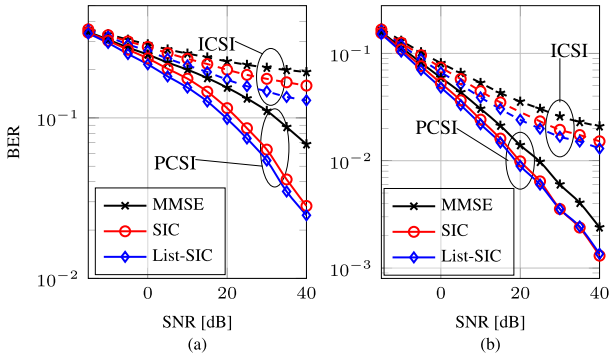



**FIGURE 4.** BER versus SNR while comparing detectors for decentralized processing for $L = 4$, $N = 4$, $K = 4$: (a) Before LLR Refinement and (b) After LLR Refinement.

noticed that there is a significant reduction in the BER for the case with LLR refinement as compared to the scenario without LLR refinement. This performance improvement is attributed to the linear combination of the multiple streams of LLRs from the different APs, which improves their reliability by shifting the LLRs with small values away from the origin. Secondly, some benefits arise due to the diversity of LLRs, which improves the system performance, whereas in the w/o-LLR-Ref case, the hard decisions are made based on the individual APs LLRs, and later an average BER is obtained. This naive approach leads to performance degradation as some APs have very unreliable estimates and hence poor LLRs. Thus, hard decisions made on LLRs from such APs compromise the entire network performance. Additionally, the figure compares the case with perfect CSI (PCSI) and imperfect CSI (ICSI). It can be observed that the detection based on the PCSI achieves lower BERs as compared to the case with the ICSI. This is because the channel estimation error and pilot contamination degrade the network performance, resulting in high BERs. Another key observation is that the proposed list-based detector achieves lower BER values than the SIC detector. This is because the convention SIC experiences error propagation that occurs due to erroneous decisions in the previous cancellation stages. To overcome this issue and improve performance, list-based detection provides multi-feedback (MF) diversity that helps to correct this error propagation as the number of iterations increases. Note also that the linear MMSE receiver has the worst performance among the studied detectors since it does not have the $\Delta_i$ matrix used for interference cancellation.

A comparison of the centralized and decentralized processing schemes in terms of BER versus SNR is presented in Figure 5. The results show that the case with centralized processing achieves lower BER values than the case with decentralized processing. This is because centralized processing takes the joint detection of all the received signals into account. Also, the case w/o-LLR-Ref achieves the worst performance since each AP performs its hard decisions locally based on the available LLRs and an average BER is obtained for the entire network, which yields a huge performance gap and degradation. The case w-LLR-Ref

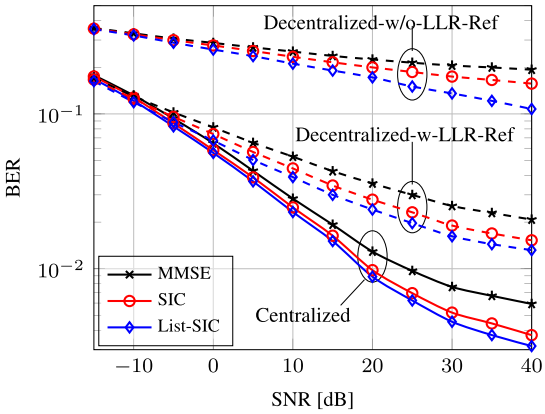



**FIGURE 5.** BER versus SNR for All APs comparing decentralized and centralized processing for the case with imperfect CSI with $L = 4$, $K = 4$, $N = 4$, IDD = 2.

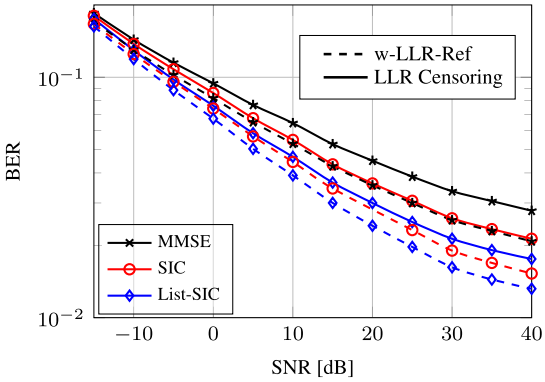



**FIGURE 6.** BER versus SNR for All APs comparing LLR Censoring and LLR Refinement for decentralized processing for the case with imperfect CSI with $L = 4$, $K = 4$, $N = 4$, IDD = 2.

outperforms the standard processing scheme since it takes advantage of LLR combining. This yields more reliable LLRs around the mean, which improves performance. This performance improvement is significant for CF-mMIMO architectures as it can yield less complex solutions in uplink detection schemes, i.e., for the decentralized processing, there is a substantial reduction in the computation complexity and the fronthaul signaling load in the network, as shown in Tables 1 and 2, respectively.

Figure 6 shows the BER versus SNR for the decentralized processing cases using w-LLR-Ref and LLR-Censoring. It is clear that the case using LLR refinement has a lower BER than the one using LLR censoring. The UE can only be decoded at the AP, achieving the highest mean absolute value when LLR censoring is used. In contrast, LLR refinement enhances performance by performing a linear combination of the LLRs from all APs. Nonetheless, censoring LLRs prevents the redundant processing of the LLRs. The only difficulty that might arise is a slight increase in the hardware complexity of the receiver design since the CPU must constantly scan all the APs to identify the one that offers the largest absolute value of LLRs to a specific UE. However, the CPU is usually designed with strong computing power, so it can handle such complexity. One could interpret the

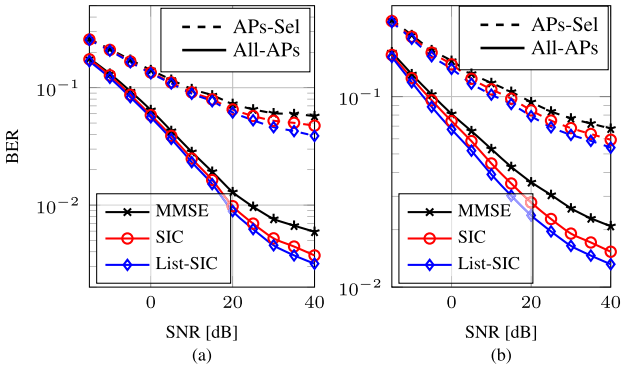



**FIGURE 7.** BER versus SNR for a case that uses All APs and a case that uses APs-Sel for $L = 4$, $N = 4$, $K = 4$: (a) Centralized Processing and (b) Decentralized Processing.

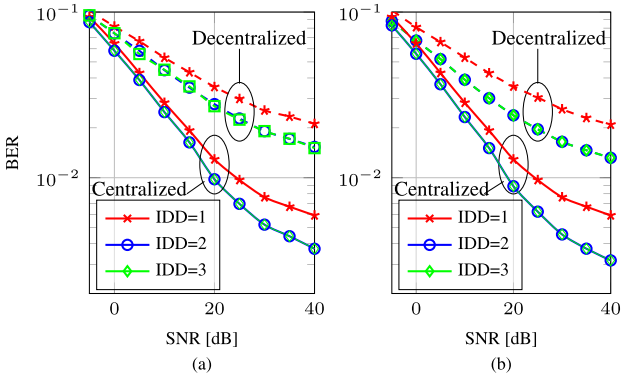



**FIGURE 8.** BER versus SNR while varying number of IDD iterations for $L = 4$, $N = 4$, $K = 4$: (a) SIC and (b) List-SIC.

proposed LLR censoring and refinement as analogous to the selection combining and maximal ratio combining used in diversity analysis. However, the former leverages the distributed computation of LLRs from each AP, and therefore it should not be confused with the latter schemes.

Figure 7 plots the BER versus SNR for the case that the detectors use all APs (All-APs) and the case that uses APs selection (APs-Sel) with (a) centralized and (b) decentralized processing schemes. It can be observed that for both processing levels, the system that uses All-APs achieves lower BER values as compared to the one that selects the APs. This is because selecting APs reduces the number of antennas in the network, distorting the performance. However, APs-Sel reduces the signaling load, making the network more scalable and practical. Moreover, the distributed location, the delay spread of the APs, and the associated signal propagation latency will limit the APs involved in cell-free MIMO systems. Therefore, APs-Sel techniques are key to reducing fronthaul signaling, computational costs, and latency.

The BER versus SNR of the detectors while varying the number of outer iterations for the detectors is presented in Figure 8. From the curves, it can be noticed that increasing the number of iterations reduces the BER. Specifically, for both centralized and decentralized (a) SIC and (b) List-SIC, there is a significant performance improvement when the number of iterations is increased from 1 to 2 iterations for both

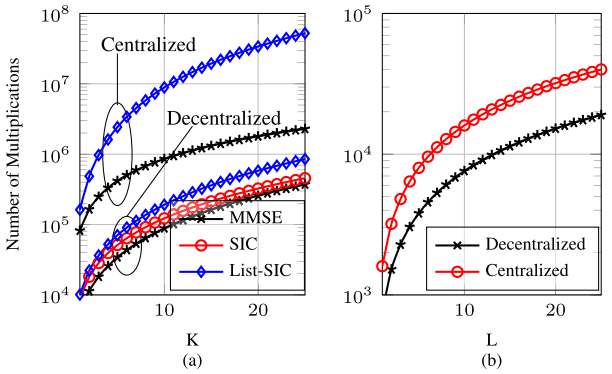



**FIGURE 9.** Number of multiplications versus number of UE $K$ and number of APs $L$ (a) Computational complexity for $L = 50$, $M_c = 2$, $N = 4$ and (b) Signaling load for $K = 4$, $N = 8$.

detectors. However, after the third iteration, the performance benefits are marginal.

Figure 9 (a) plots the computational complexity versus the number of UEs $K$ while considering the studied detectors for centralized and decentralized processing. It can be observed that the linear MMSE detectors have the lowest computation complexity for both cases. The list-based detectors have slightly higher computation complexity than the SIC-based detectors. It is worth mentioning that the differences in computational complexity between SIC and List-SIC are marginal, whereas centralized detectors have a higher computation cost than decentralized detectors. The signaling load is given by the number of complex scalars that must be exchanged in the network, which is shown in Figure 9 (b) for the centralized and decentralized CF-mMIMO setups. From the curves, it can be noticed that decentralized processing requires less signaling between the APs and CPU than the centralized setup. The decentralized processing requires knowledge of the $KC_{\text{leng}}L$-dimensional matrix of the LLRs from all APs to perform the final processing. Nevertheless, decentralized schemes greatly reduce the required signaling in the network and can achieve close performance to that of centralized processing while using LLR refinement.

## VIII. CONCLUSION

In this paper, an IDD scheme using LDPC codes has been devised with AP selection for centralized and decentralized CF-mMIMO architectures. In particular, we have proposed low-complexity interference mitigation techniques, including a list-based detector that uses an MMSE receive filter to improve performance. New closed-form expressions for the MMSE-soft-IC detectors have been derived for both the centralized and decentralized implementations, taking channel estimation errors and AP selection into account. The performance of the proposed list-based detector is compared with other baseline detectors, such as the soft linear MMSE and MMSE-SIC, and the results show that the list-based detector yields low BER values compared to the other detectors. We also proposed LLR refinement strategies based on combining and censoring LLRs. The results have shown that LLR refinement strategies obtained lower BER values

than standard processing. The proposed scalable APs-Sel scheme based on LLSF coefficients can reduce the signaling load between APs and the CPU, resulting in a trade-off between performance, signaling load, and network feasibility.

## APPENDIX A
## DERIVATION OF THE PROPOSED CENTRALIZED DETECTOR

We start our analysis by expressing the conditional expectation on the RHS of (13) as

$$F = \mathbb{E}\left\{ ||\tilde{s}_k - s_k||^2 \mid \hat{\mathbf{G}} \right\} = \mathbb{E}\left\{ (\tilde{s}_k - s_k)(\tilde{s}_k - s_k)^* \mid \hat{\mathbf{G}} \right\}. \tag{54}$$

Substituting (12) into (54) yields

$$\begin{aligned} F = {} & \mathbf{w}_k^H \mathbf{D}_k \mathbb{E}\left\{ \left( \mathbf{y} - \hat{\mathbf{G}}_{\mathrm{i}} \bar{\mathbf{s}}_{\mathrm{i}} \right) \left( \mathbf{y}^H - \bar{\mathbf{s}}_{\mathrm{i}}^H \hat{\mathbf{G}}_{\mathrm{i}}^H \right) \right\} \mathbf{D}_k^H \mathbf{w}_k \\ & - \mathbf{w}_k^H \mathbf{D}_k \mathbb{E}\left\{ \left( \mathbf{y} - \hat{\mathbf{G}}_{\mathrm{i}} \bar{\mathbf{s}}_{\mathrm{i}} \right) s_k^* \right\} - \mathbb{E}\left\{ s_k \left( \mathbf{y}^H - \bar{\mathbf{s}}_{\mathrm{i}}^H \hat{\mathbf{G}}_{\mathrm{i}}^H \right) \right\} \\ & \times \mathbf{D}_k^H \mathbf{w}_k + \mathbb{E}\left\{ s_k s_k^* \right\}. \end{aligned} \tag{55}$$

Further simplification of (55) can be done by letting $\mathbf{y}_R = \mathbf{y} - \hat{\mathbf{G}}_{\mathrm{i}} \bar{\mathbf{s}}_{\mathrm{i}}$. Thus, (55) can be re-written as

$$\begin{aligned} F = {} & \mathbf{w}_k^H \mathbf{D}_k \mathbb{E}\left\{ \mathbf{y}_R \mathbf{y}_R^H \right\} \mathbf{D}_k^H \mathbf{w}_k - \mathbf{w}_k^H \mathbf{D}_k \mathbb{E}\left\{ \mathbf{y}_R s_k^* \right\} \\ & - \mathbb{E}\left\{ s_k \mathbf{y}_R^H \right\} \mathbf{D}_k^H \mathbf{w}_k + \mathbb{E}\left\{ s_k s_k^* \right\}. \end{aligned} \tag{56}$$

Differentiating (56) with respect to w.r.t $\mathbf{w}_k^H$ we obtain

$$\frac{\partial F}{\partial \mathbf{w}_{\mathbf{k}}^{\mathbf{H}}} = \mathbf{D}_k \mathbb{E}\left\{ \mathbf{y}_R \mathbf{y}_R^H \right\} \mathbf{D}_k^H \mathbf{w}_k - \mathbf{D}_k \mathbb{E}\left\{ \mathbf{y}_R s_k^* \right\}. \tag{57}$$

The optimal MMSE filter is obtained by equating (57) to 0. Thus, the optimal MMSE filter $\mathbf{w}_k$ is given by

$$\mathbf{D}_k \mathbb{E}\left\{ \mathbf{y}_R \mathbf{y}_R^H \right\} \mathbf{D}_k^H \mathbf{w}_k - \mathbf{D}_k \mathbb{E}\left\{ \mathbf{y}_R s_k^* \right\} = 0. \tag{58}$$

The reader can confirm that (58) is the same as (14). By making $\mathbf{w}_k$ the subject of (58), we obtain (15). The terms $\mathbb{E}\{\mathbf{y}_R s_k^*\}$ and $\mathbb{E}\{\mathbf{y}_R \mathbf{y}_R^H\}$ are given by (16) and (17), where $\mathbb{E}\left\{ s_m s_m^* \right\} = |s_m|^2 + \sigma_m^2$, $\mathbb{E}\left\{ \tilde{\mathbf{g}}_m \tilde{\mathbf{g}}_m^H \right\} = \mathbf{C}_m$, $\mathbb{E}\left\{ \mathbf{n}\mathbf{n}^H \right\} = \sigma^2 \mathbf{I}_{NL}$, $\mathbb{E}\left\{ s_k s_k^H \right\} = \rho_k$, obtained after assuming statistical independence between each term in the RHS of (11) and using the orthogonality principle [35]. By substituting (16) and (17) into (15), we arrive at the centralized MMSE filter given by (18).

## APPENDIX B
## DERIVATION OF THE PROPOSED DECENTRALIZED DETECTOR

The derivation of the proposed local MMSE filter is similar to that of Appendix A. The expectation on the R.H.S of (26) can be expressed as

$$\begin{aligned} F_2 &= \mathbb{E}\left\{ ||\tilde{s}_{kl} - s_k||^2 \mid \hat{\mathbf{G}}_l \right\} \\ &= \mathbb{E}\left\{ (\tilde{s}_{kl} - s_k)(\tilde{s}_{kl} - s_k)^* \mid \hat{\mathbf{G}}_l \right\}. \end{aligned} \tag{59}$$

By substituting (25) into (59) we obtain

$$F_2 = \mathbb{E}\left\{ \left( \mathbf{w}_{kl}^H \mathbf{D}_{kl} \mathbf{y}_{Rl} - s_k \right) \left( \mathbf{w}_{kl}^H \mathbf{D}_{kl} \mathbf{y}_{Rl} - s_k \right)^* \right\}. \tag{60}$$

The term $\mathbf{y}_{Rl}$ is the residue signal obtained after soft-IC and substituting for $y_l$ in the term in brackets of (25), we get

$$\mathbf{y}_{Rl} = \hat{\mathbf{g}}_{kl} s_k + \hat{\mathbf{G}}_{\mathrm{i}l} \left( \mathbf{s}_{\mathrm{i}} - \bar{\mathbf{s}}_{\mathrm{i}} \right) + \sum_{m=1}^{K} \tilde{\mathbf{g}}_{ml} s_m + \mathbf{n}_l. \tag{61}$$

After some mathematical and algebraic manipulations, (60) can be re-written as

$$\begin{aligned} F_2 = {} & \mathbf{w}_{kl}^H \mathbf{D}_{kl} \mathbb{E}\left\{ \mathbf{y}_{Rl} \mathbf{y}_{Rl}^H \right\} \mathbf{D}_{kl}^H \mathbf{w}_{kl} - \mathbf{w}_{kl}^H \mathbf{D}_{kl} \mathbb{E}\left\{ \mathbf{y}_{Rl} s_k^* \right\} \\ & - \mathbb{E}\left\{ s_k \mathbf{y}_{Rl}^H \right\} \mathbf{D}_{kl}^H \mathbf{w}_{kl} + \mathbb{E}\left\{ s_k s_k^* \right\} \end{aligned} \tag{62}$$

We take the first derivative of (62) w.r.t $\mathbf{w}_{kl}^H$ to arrive at

$$\frac{\partial F_2}{\partial \mathbf{w}_{kl}^H} = \mathbf{D}_{kl} \mathbb{E}\left\{ \mathbf{y}_{Rl} \mathbf{y}_{Rl}^H \right\} \mathbf{D}_{kl}^H \mathbf{w}_{kl} - \mathbf{D}_{kl} \mathbb{E}\left\{ \mathbf{y}_{Rl} s_k^* \right\}. \tag{63}$$

After equating the resulting expression in (63) to 0, we obtain

$$\mathbf{D}_{kl} \mathbb{E}\left\{ \mathbf{y}_{Rl} \mathbf{y}_{Rl}^H \right\} \mathbf{D}_{kl}^H \mathbf{w}_{kl} - \mathbf{D}_{kl} \mathbb{E}\left\{ \mathbf{y}_{Rl} s_k^* \right\} = 0. \tag{64}$$

The optimal local MMSE filter $\mathbf{w}_{kl}$ can be obtained from (64). The terms $\mathbb{E}\left\{ \mathbf{y}_{Rl} \mathbf{y}_{Rl}^H \right\}$ and $\mathbb{E}\left\{ \mathbf{y}_{Rl} s_k^* \right\}$ can be obtained from (29) and (30), respectively, where the terms $\mathbb{E}\left\{ \tilde{\mathbf{g}}_{ml} \tilde{\mathbf{g}}_{ml}^H \right\} = \mathbf{C}_{ml}$ and $\mathbb{E}\left\{ \mathbf{n}_l \mathbf{n}_l^H \right\} = \sigma^2 \mathbf{I}_N$, by taking assumptions similar to those in Subsection A.

## APPENDIX C
## DERIVATION OF THE SOFT DEMAPPER PARAMETERS FOR CENTRALIZED PROCESSING

We start the proof by making some assumptions on the output of the MMSE filter to be a Gaussian approximation. The optimal soft bit metric which takes into account the channel estimation error and APs-Sel can be derived as below. Let $k$ denote the desired UE which minimizes the mean square error (MSE). Then, (12) can be expressed as

$$\tilde{s}_k = \mathbf{w}_k^H \mathbf{D}_k \mathbf{y} - \mathbf{w}_k^H \mathbf{D}_k \hat{\mathbf{G}}_i \bar{\mathbf{s}}_{\mathrm{i}}. \tag{65}$$

By substituting (11) into (65) we obtain

$$\begin{aligned} \tilde{s}_k = {} & \mathbf{w}_k^H \mathbf{D}_k \hat{\mathbf{g}}_k s_k + \mathbf{w}_k^H \mathbf{D}_k \hat{\mathbf{G}}_{\mathrm{i}} \left( \mathbf{s}_{\mathrm{i}} - \bar{\mathbf{s}}_{\mathrm{i}} \right) + \mathbf{w}_k^H \mathbf{D}_k \sum_{m=1}^{K} \tilde{\mathbf{g}}_m s_m \\ & + \mathbf{w}_k^H \mathbf{D}_k \mathbf{n}. \end{aligned} \tag{66}$$

By comparing (66) with (39), it can be observed that

$$\omega_k = \mathbf{w}_k^H \tilde{\mathbf{D}}_k \hat{\mathbf{g}}_k, \tag{67}$$

and the interference-plus-noise term is given by

$$z_k = \mathbf{w}_k^H \mathbf{D}_k \hat{\mathbf{G}}_{\mathrm{i}} \left( \mathbf{s}_{\mathrm{i}} - \bar{\mathbf{s}}_{\mathrm{i}} \right) + \mathbf{w}_k^H \mathbf{D}_k \sum_{m=1}^{K} \tilde{\mathbf{g}}_m s_m + \mathbf{w}_k^H \mathbf{D}_k \mathbf{n}. \tag{68}$$

By assuming that $z_k$ is a Gaussian random variable [13] and assuming statistical independence of each term of (68), the variance $\kappa^2 = \mathbb{E}\left\{\mid u_k - \omega_k s_k \mid^2\right\} = \mathbb{E}\left\{z_k z_k^*\right\}$ of $z_k$ is given by

$$\kappa^2 = \mathbf{w}_k^H \mathbf{D}_k \left( \hat{\mathbf{G}}_{\mathrm{i}} \Delta_{\mathrm{i}} \hat{\mathbf{G}}_{\mathrm{i}}^H + \sum_{m=1}^{K} \rho_m \mathbf{C}_m + \sigma^2 \mathbf{I}_{NL} \right) \mathbf{D}_k^H \mathbf{w}_k. \tag{69}$$

By substituting (67) and (69) into (43), the soft beliefs for the centralized processor can be obtained in each subsequent iteration.

## APPENDIX D
## DERIVATION OF THE SOFT DEMAPPER PARAMETER FOR DECENTRALIZED PROCESSING

We also start by assuming that the output of the MMSE-SIC filter is a Gaussian random variable. We then derive the local optimal soft bit metric which takes the APs-Sel matrix, and channel estimation error as follow Let $k$ denote the desired UE which minimizes the mean square error (MSE). Then, (12) can be expressed as

$$\tilde{s}_k = \mathbf{w}_{kl}^H \mathbf{D}_{kl} \mathbf{y}_l - \mathbf{w}_{kl}^H \mathbf{D}_{kl} \hat{\mathbf{G}}_{il} \bar{\mathbf{s}}_{\mathrm{i}}. \tag{70}$$

By substituting (24) into (70) we obtain

$$\tilde{s}_k = \mathbf{w}_{kl}^H \mathbf{D}_{kl} \hat{\mathbf{g}}_{kl} s_k + \mathbf{w}_{kl}^H \mathbf{D}_{kl} \hat{\mathbf{G}}_{\mathrm{il}} \left( \mathbf{s}_{\mathrm{i}} - \bar{\mathbf{s}}_{\mathrm{i}} \right) + \mathbf{w}_{kl}^H \mathbf{D}_{kl} \sum_{m=1}^{K} \tilde{\mathbf{g}}_{ml} s_m + \mathbf{w}_{kl}^H \mathbf{D}_{kl} \mathbf{n}_l. \tag{71}$$

By comparing (71) with (39), it can be observed that

$$\omega_{kl} = \mathbf{w}_{kl}^H \tilde{\mathbf{D}}_{kl} \hat{\mathbf{g}}_{kl}, \tag{72}$$

and the interference-plus-noise term is given by

$$z_{kl} = \mathbf{w}_{kl}^H \mathbf{D}_{kl} \hat{\mathbf{G}}_{\mathrm{il}} \left( \mathbf{s}_{\mathrm{i}} - \bar{\mathbf{s}}_{\mathrm{i}} \right) + \mathbf{w}_{kl}^H \mathbf{D}_{kl} \sum_{m=1}^{K} \tilde{\mathbf{g}}_{ml} s_m + \mathbf{w}_{kl}^H \mathbf{D}_{kl} \mathbf{n}_l. \tag{73}$$

By assuming that $z_{kl}$ is a Gaussian random variable [13] and assuming statistical independence of each term of (73), the variance $\kappa_l^2 = \mathbb{E}\left\{\mid u_{kl} - \omega_{kl} s_k \mid^2\right\} = \mathbb{E}\left\{z_{kl} z_{kl}^*\right\}$ of $z_{kl}$ is given by

$$\kappa_l^2 = \mathbf{w}_{kl}^H \mathbf{D}_{kl} \left( \hat{\mathbf{G}}_{\mathrm{il}} \Delta_{\mathrm{il}} \hat{\mathbf{G}}_{\mathrm{il}}^H + \sum_{m=1}^{K} \rho_m \mathbf{C}_{ml} + \sigma^2 \mathbf{I}_N \right) \mathbf{D}_{kl}^H \mathbf{w}_{kl}. \tag{74}$$

By substituting (72) and (74) into (43), the soft beliefs for the lcoal processors can be obtained in each subsequent iteration at each AP.

## REFERENCES


[1] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, Jan. 2020.

[2] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020.

[3] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[4] Z. Chen and E. Björnson, "Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5205–5219, Nov. 2018.

[5] H. T. Dao and S. Kim, "Effective channel gain-based access point selection in cell-free massive MIMO systems," *IEEE Access*, vol. 8, pp. 108127–108132, 2020.

[6] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 25–39, Mar. 2018.

[7] S. Mashdour, R. C. de Lamare, and J. P. S. H. Lima, "Enhanced subset greedy multiuser scheduling in clustered cell-free massive MIMO systems," *IEEE Commun. Lett.*, vol. 27, no. 2, pp. 610–614, Feb. 2023.

[8] S. Mashdour, R. C. de Lamare, A. Schmeink, and J. P. S. H. Lima, "MMSE-based resource allocation for clustered cell-free massive MIMO networks," in *Proc. 26th Int. ITG Workshop Smart Antennas 13th Conf. Syst., Commun.*, Braunschweig, Germany, 2023, pp. 1–6.

[9] I. L. Shakya and F. H. Ali, "Joint access point selection and interference cancellation for cell-free massive MIMO," *IEEE Commun. Lett.*, vol. 25, no. 4, pp. 1313–1317, Apr. 2021.

[10] Y. Li, Q. Lin, Y.-F. Liu, B. Ai, and Y.-C. Wu, "Asynchronous activity detection for cell-free massive MIMO: From centralized to distributed algorithms," *IEEE Trans. Wireless Commun.*, vol. 22, no. 4, pp. 2477–2492, Apr. 2023.

[11] X. Wang and H. V. Poor, "Iterative (turbo) soft interference cancellation and decoding for coded CDMA," *IEEE Trans. Commun.*, vol. 47, no. 7, pp. 1046–1061, Jul. 1999.

[12] B. Xiao, K. Xiao, S. Zhang, Z. Chen, B. Xia, and H. Liu, "Iterative detection and decoding for SCMA systems with LDPC codes," in *Proc. Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nanjing, China, Oct. 2015, pp. 1–5.

[13] A. Matache, C. Jones, and R. Wesel, "Reduced complexity MIMO detectors for LDPC coded systems," in *Proc. IEEE MILCOM Mil. Commun. Conf.*, Monterey, CA, USA, Nov. 2004, pp. 1073–1079.

[14] F. Sagheer, F. Lehmann, and A. O. Berthet, "Low-complexity dynamic channel estimation in multi-antenna grant-free NOMA," in *Proc. IEEE 95th Veh. Technol. Conf.*, Helsinki, Finland, Jun. 2022, pp. 1–7.

[15] R. De Lamare and R. Sampaio-Neto, "Minimum mean-squared error iterative successive parallel arbitrated decision feedback detectors for DS-CDMA systems," *IEEE Trans. Commun.*, vol. 56, no. 5, pp. 778–789, May 2008.

[16] R. B. Di Renna and R. C. de Lamare, "Joint channel estimation, activity detection and data decoding based on dynamic message-scheduling strategies for mMTC," *IEEE Trans. Commun.*, vol. 70, no. 4, pp. 2464–2479, Apr. 2022.

[17] H. Li, Y. Dong, C. Gong, X. Wang, and X. Dai, "Gaussian message passing detection with constant front-haul signaling for cell-free massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 72, no. 4, pp. 5395–5400, Apr. 2023.

[18] R. B. Di Renna and R. C. de Lamare, "Iterative list detection and decoding for massive machine-type communications," *IEEE Trans. Commun.*, vol. 68, no. 10, pp. 6276–6288, Oct. 2020.

[19] Z. Shao, L. T. N. Landau, and R. C. de Lamare, "Dynamic oversampling for 1-bit ADCs in large-scale multiple-antenna systems," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3423–3435, May 2021.

[20] H. T. Nguyen, D. A. Hoang, H. T. Bui, H. N. Dang, and T. V. Nguyen, "Large-scale MIMO communications with low-resolution ADCs using 16-ary QAM and protograph LDPC codes," in *Proc. IEEE 9th Int. Conf. Commun. Electron. (ICCE)*, Jul. 2022, pp. 43–47.

[21] S. Jing, C. Yang, J. Yang, X. You, and C. Zhang, "Joint detection and decoding of polar-coded SCMA systems," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Nanjing, China, Oct. 2017, pp. 1–6.

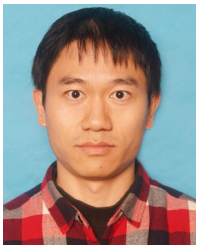
[22] P. Li, R. C. de Lamare, and R. Fa, "Multiple feedback successive interference cancellation detection for multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 10, no. 8, pp. 2434–2439, Aug. 2011.

[23] R. C. de Lamare, "Adaptive and iterative multi-branch MMSE decision feedback detection algorithms for multi-antenna systems," *IEEE Trans. Wireless Commun.*, vol. 12, no. 10, pp. 5294–5308, Oct. 2013.

[24] C. D'Andrea and E. G. Larsson, "Improving cell-free massive MIMO by local per-bit soft detection," *IEEE Commun. Lett.*, vol. 25, no. 7, pp. 2400–2404, Jul. 2021.

[25] H. He, H. Wang, X. Yu, J. Zhang, S. H. Song, and K. B. Letaief, "Distributed expectation propagation detection for cell-free massive MIMO," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Madrid, Spain, Dec. 2021, pp. 1–6.
[26] A. G. D. Uchoa, C. T. Healy, and R. C. de Lamare, "Iterative detection and decoding algorithms for MIMO systems in block-fading channels using LDPC codes," *IEEE Trans. Veh. Technol.*, vol. 65, no. 4, pp. 2735–2741, Apr. 2016.
[27] A. Krebs, M. Joham, and W. Utschick, "Comparative performance evaluation of error regularized turbo-MIMO MMSE-SIC detectors in Gaussian channels," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, South Brisbane, QLD, Australia, Apr. 2015, pp. 2984–2988.
[28] Z. H. Shaik, E. Björnson, and E. G. Larsson, "Distributed computation of a posteriori bit likelihood ratios in cell-free massive MIMO," in *Proc. 29th Eur. Signal Process. Conf. (EUSIPCO)*, Dublin, Ireland, Aug. 2021, pp. 935–939.
[29] H. Lee and I. Lee, "New approach for coded layered space-time OFDM systems," in *Proc. IEEE Int. Conf. Commun.*, Seoul, Korea (South), May 2005, pp. 608–612.
[30] H. Le, B. Lee, and I. Lee, "Iterative detection and decoding with an improved V-BLAST for MIMO-OFDM systems," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 504–513, Mar. 2006.
[31] Z. Shao, R. C. de Lamare, and L. T. N. Landau, "Iterative detection and decoding for large-scale multiple-antenna systems with 1-bit ADCs," *IEEE Wireless Commun. Lett.*, vol. 7, no. 3, pp. 476–479, Jun. 2018.
[32] T. Ssettumba, R. B. D. Renna, L. T. N. Landau, and R. C. D. Lamare, "Iterative detection and decoding for cell-free massive MIMO using LDPC codes," in *Proc. Anais do 40th Brazilian Symp. Telecommun. Signal Process.*, Sep. 2022, pp. 1–5.
[33] T. Ssettumba, R. B. Di Renna, L. T. N. Landau, and R. C. de Lamare, "List-based detector and access points selection in cell-free massive MIMO using LDPC codes," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Hangzhou, China, Oct. 2022, pp. 1–6.
[34] R. B. Di Renna and R. C. de Lamare, "Adaptive LLR-based APs selection for grant-free random access in cell-free massive MIMO," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Rio de Janeiro, Brazil, Dec. 2022, pp. 196–201, doi: 10.1109/GCWkshps56602.2022.10008736.
[35] J. G. Proakis, *Digital Communications* (Series in Electrical and Computer Engineering), 4th ed. New York, NY, USA: McGraw-Hill, 2001.
[36] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends Signal Process.*, vol. 11, nos. 3–4, pp. 154–655, 2017.

**TONNY SSETTUMBA** received the B.Sc. degree in telecommunications engineering from Makerere University, Kampala, Uganda, in 2016, and the M.Sc. degree in electronics and communications engineering from Egypt-Japan University of Science and Technology, New Borg El Arab, Egypt, in 2019. He is currently pursuing the Ph.D. degree in electrical engineering with the Pontifical Catholic University of Rio de Janeiro, Brazil. He was a Graduate Transmission Engineer with Africell Uganda (formerly Orange Uganda), Kampala, and an O&M Site Engineer with Huawei Technologies Ltd., Kampala. He is also a Lecturer with the Department of Electronics and Computer Engineering, Soroti University, Uganda. His current research interests include but not limited to iterative detection and decoding for cell-free MIMO networks and physical layer security of wireless communication systems.



**ZHICHAO SHAO** (Member, IEEE) received the B.Sc. degree in information engineering from Xidian University, China, in 2012, the M.Sc. degree in electrical engineering from Technische Universität Dresden, Germany, in 2016, and the Ph.D. degree in electrical engineering from the Pontifical Catholic University of Rio de Janeiro, Brazil, in 2020. From 2021 to 2023, he was a Postdoctoral Researcher with the National Key Laboratory of Wireless Communications, University of Electronic Science and Technology of China. He is currently an Associate Research Fellow with Yangtze Delta Region Institute (Quzhou), University of Electronic Science and Technology of China. His research interests include communications and signal processing.

**LUKAS T. N. LANDAU** (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees in electrical engineering and information technology from the Ilmenau University of Technology, Germany, in 2009 and 2011, respectively, and the Ph.D. degree in electrical engineering and information technology from Technische Universität Dresden, Germany, in 2016. Since 2016, he has been with the Pontifical Catholic University of Rio de Janeiro, Brazil, where he is currently an Adjunct Professor. His research interests include communications and signal processing. He also serves as an Associate Editor for IEEE Wireless Communications Letters and the *EURASIP Journal on Wireless Communications and Networking*.

**MICHELLE SOARES PEREIRA FACINA** received the B.S. and M.Sc. degrees in electrical engineering from the Federal University of Juiz de Fora (UFJF), Brazil, in 2013 and 2015, respectively, and the Ph.D. degree in electrical engineering from the University of Campinas (UNICAMP), Brazil, in 2020. She was a Visiting Ph.D. Student with CentraleSupelec (L2S Group-France) for one year. Since 2021, she has been involved on Brasil 6G Project, before as a Postdoctoral Researcher and then as a Researcher with Fundação Centro de Pesquisa e Desenvolvimento em Telecomunicações (CPQD).

**PAULO RICARDO BRANCO DA SILVA** received the B.Sc. degree in electrical engineering from Instituto Mauá de Tecnologia, Brazil, in 2012, and the M.Sc. and Ph.D. degrees in forward error correction from the Federal University of Santa Catarina, Brazil, in 2015 and 2019, respectively. He was a Postdoctoral Researcher of synthetic aperture radars (SAR) for ultrawideband images with the Instituto Tecnológico de Aeronáutica (ITA), Brazil, with an emphasis on change detection using information theoretical distances. He is currently with CPQD, Brazil, as a Machine Learning Researcher, with most of his work focused on the fields of intelligent networks with applied artificial intelligence, forecasting, and reinforcement learning.

**RODRIGO C. DE LAMARE** (Senior Member, IEEE) was born in Rio de Janeiro, Brazil, in 1975. He received the Diploma degree in electronic engineering from the Federal University of Rio de Janeiro, in 1998, and the M.Sc. and Ph.D. degrees in electrical engineering from the Pontifical Catholic University of Rio de Janeiro (PUC-Rio), in 2001 and 2004, respectively. Since January 2006, he has been with the Communications Group, School of Physics, Engineering and Technology, University of York, U.K., where he is currently a Professor. Since April 2013, he has been a Professor with PUC-Rio. His research interests include communications and signal processing, areas in which he has published over 500 papers in international journals and conferences. He is a an Elected Member of the IEEE Signal Processing for Communications and Networking Committee. He served as an Editor for IEEE Wireless Communications Letters and IEEE Transactions on Communications.