







Cell-Free Massive MIMO Symbiotic Radio for IoT: RIS or BD?

Feiyang Li , Qiang Sun , *Member, IEEE*, Xiaomin Chen , *Member, IEEE*, Bile Peng , *Member, IEEE*, Jiayi Zhang , *Senior Member, IEEE*, and Kai-Kit Wong , *Fellow, IEEE*

Abstract—Cell-free massive multiple-input multiple-output symbiotic radio (CF-mMIMO-SR) is a promising technology to address the requirements of high-rate and spectrum-efficient communication for the Internet of Things (IoT). However, in the conventional CF-mMIMO-SR system aided by backscatter devices (BDs), the backscatter link is impacted by double fading without any supplementary compensation, resulting in significantly low spectral efficiency (SE) on the backscatter link. To address this issue, we propose the usage of reconfigurable intelligent surfaces (RISs) instead of BD for symbol-level reflection on the backscatter link, leading to a novel RIS-aided CF-mMIMO-SR (RIS-CF-SR) system. In this paper, we conduct a comprehensive analysis of the RIS-CF-SR system considering different levels of cooperation among the access points (APs). Specifically, we analyze the uplink SEs of four different implementations with arbitrary linear processing on both the direct and backscatter links. Moreover, we investigate different signal cancellation schemes based on full or local channel state information (CSI) to improve the SE of the backscatter link. Through the simulation results, we find that RISs can significantly improve the SE of the backscatter link due to the large number of reflection elements, whereas additional appropriate signal processing schemes are required for the direct link. More specifically, from Level 1 to Level 3, RIS-CF-SR does not have significant advantage in SE over BD-CF-SR on the direct link. At Level 4, RIS-CF-SR can outperform BD-CF-SR on the direct link with the MMSE combining scheme.

Index Terms—Cell-free massive MIMO, symbiotic radio, large-scale fading decoding, spectral efficiency, energy efficiency.

I. INTRODUCTION

Nowadays, the emerging new Internet of Things (IoT) applications (such as remote robotic surgery and parking sensor) has raised demands for high-rate transmission, posing new challenges to the communications systems [1], [2]. To this end, cellular massive multiple-input multiple-output (MIMO) has been widely used to support IoT, whereas the inter-cell

interference can be one of the major obstacles for the future IoT requiring high-rate transmission [3]. Cell-free massive MIMO (CF-mMIMO) has been proposed as an effective solution to tackle this issue [4], [5], which is regarded as cell-free IoT (CF-IoT) [6]. In the CF-IoT, numerous access points (APs) are strategically dispersed throughout the coverage area, connecting to a central processing unit (CPU) via fronthaul links [8], [9], and provide coherent service to multiple user equipments (UEs) by spatially multiplexing [10], [11].

Different from the cellular massive MIMO, CF-mMIMO has no cell boundaries, where each UE is served by the most suitable set of APs. Thus, the inter-cell interference can be effectively suppressed, thus the transmission rate is increased [12], [13]. CF-IoT was initially explored in [6], in which the uplink closed-form signal-to-interference-and-noise ratio (SINR) expressions are derived. Given the considerable number of IoT devices, it is often impractical to allocate orthogonal pilots to each one. Consequently, the issue of pilot contamination becomes pronounced, particularly considering the low transmit power of these IoT devices [7]. In order to mitigate pilot contamination, a framework for the detection of IoT device activity and channel estimation was proposed [14]. However, the scarcity of radio spectrum resources remains as a challenge, especially when the CF-mMIMO is utilized to support massive number of IoT devices [15]. It has been reported that around 76 GHz spectrum resources are required if the exclusive spectrum is allocated to the IoT connections [16]. Therefore, a spectrum-efficient communication scheme is needed for the future CF-IoT.

The symbiotic radio (SR) was proposed in [17] for the first time to address the challenge of radio spectrum resources shortage in CF-IoT. The SR comprises three main components: a primary transmitter (PT), a primary receiver (PR), and a backscatter device (BD). The BD is designed to modulate information onto continuous wave signals emitted by specialized carrier emitters [18]. The SR can be categorized into two types depending on the symbol rate of the BD, i.e., parasitic symbiotic radio (PSR) and commensal symbiotic radio (CSR). While the equal symbol rates are required in PSR for both primary and BD transmissions, it can introduce additional interference to the primary transmission. In contrast, CSR employs a considerably lower symbol rate for the BD transmission. As a result, it is regarded as an additional multipath component to the primary transmission [17]. Moreover, the BD operates within the same radio spectrum as the PT, without the need for dedicated spectrum allocation [19], [20], offering significant improvements in spectrum utilization for CF-IoT.

This work was supported in part by National Natural Science Foundation of China under Grant 62371262, 62401297, and 62341131, in part by the Key Research and Development Program of Nantong under Grant GZ2024002, in part by the Qiluan Project of Jiangsu Province, in part by the Scientific Research Program of Nantong under Grant JC22022026, and in part by the Postgraduate Research & Practice Innovation Program of Jiangsu Province under KYCX23_3399. (Corresponding authors: Qiang Sun.)

F. Li, Q. Sun, and X. Chen are with the School of Information Science and Technology, Nantong University, Nantong 226019, China.

Bile Peng is with the Department of Information Theory and Communication Systems, Technische Universität Braunschweig, 38112 Braunschweig, Germany

J. Zhang is with the School of Electronic and Information Engineering and the Frontiers Science Center for Smart High-speed Railway System, Beijing Jiaotong University, Beijing 100044, China.

K.-K. Wong is with the Department of Electronic and Electrical Engineering, University College London, London WC1E 6BT, U.K.

TABLE I: Comparison of relevant works with this paper.

Ref.	CSR	RIS	EE	Imperfect CSI	LSFD	Fully centralized processing	MMSE processing	Closed-form
[9]	×	×	×	✓	✓	✓	✓	×
[21]	×	×	×	✓	×	×	×	×
[22]	✓	×	×	✓	×	×	×	✓
[23]	×	×	×	✓	×	×	✓	✓
[31]	✓	✓	✓	×	×	×	×	×
Proposed	✓	✓	✓	✓	✓	✓	✓	✓

In view of these advantages, researchers have been intensively investigating the combination of SR and CF-mMIMO. The cell-free massive MIMO symbiotic radio (CF-mMIMO-SR) was first considered in [21], in which the CF-mMIMO-SR under a single IoT device and perfect channel state information (CSI) is considered. Furthermore, the CF-mMIMO-SR under several IoT devices and imperfect CSI are investigated in [22] and [23]. However, all these papers focused on BD-aided CF-mMIMO to realize SR (refer to BD-CF-SR), where the backscatter link is affected by double fading without any supplementary compensation, resulting in a very low spectral efficiency (SE) on the backscatter link.

To enhance the SE of backscatter link for SR, the authors of [24] introduced the reconfigurable intelligent surface (RIS) as the BD into SR. In contrast to traditional RISs, these RISs designed for SR are essentially embedded environmental sensors. This distinctive characteristic enables them not only to reflect incoming signals but also to autonomously transmit their own data to receivers. Moreover, the SE of backscatter-link can be bolstered by deploying a large number of reflection elements in the RIS. Particularly, the RIS reflecting elements solely passively reflect incoming signals without necessitating sophisticated signal processing operations that typically require radio frequency transceiver hardware. Consequently, in contrast to conventional relays, RIS operation entails significantly lower costs in terms of hardware and power consumption [25], [26]. Specifically, the problem of minimizing transmit power has been explored in [27] within the context of RIS-assisted MIMO SR. Furthermore, [28], [29], and [30] have delved into minimizing bit error rate (BER), optimizing transmit power to enhance the energy efficiency (EE), and maximizing secondary transmission rate, respectively, in the scenario of RIS-assisted MISO SR. Nevertheless, few papers explored the integration of RIS and CF-mMIMO-SR. The combination of the two is first appeared in [31], but the model considers perfect CSI. Such overly idealized conditions might not yield practical insights for analyzing the real-world system performance. In addition, all the works about CF-mMIMO-SR [21]–[23], [31] are built on a simple processing scheme given in [5] (called “Level 2” in [9]) but ignore more complex signal processing schemes, e.g., “large-scale fading decoding (LSFD) scheme” and “fully centralized processing”.

Motivated by the above observations, we explore a RIS-aided CF-mMIMO-SR (RIS-CF-SR) with pilot contamination and imperfect CSI. We first introduce fully centralized process-

ing and LSFD scheme into CF-mMIMO-SR system. Besides, we utilize the MMSE processing to suppress the additional interference caused by SR. To provide a comprehensive understanding of our work, we present a brief comparison between this paper and relevant works, as summed up in Table I. The main contributions of this paper are summarized as follows:

- 1) A novel RIS-CF-SR system with pilot contamination is considered, where the channel from the device and RIS operates over Rician fading and the other channels operate over Rayleigh fading. Four different uplink implementations are investigated, including fully centralized processing, local processing with LSFD scheme, local processing with simple centralized decoding, and a small-cell network from fully centralized to fully distributed.
- 2) Achievable SE expressions are derived for four implementations with arbitrary combining schemes on both the direct and backscatter links. Based on full or local CSI, different signal cancellation schemes are also considered for four different implementations to enhance the backscatter-link SE. Moreover, by using the use-and-then-forget (UatF), the optimal LSFD scheme is proposed for Level 3 to maximize the direct-link SE and backscatter-link SE. In addition, novel closed-form SE expressions with MR combining schemes are calculated for Level 2 and Level 3.
- 3) The performance of RIS-CF-SR and BD-CF-SR are compared. We find that from Level 1 to Level 3, RIS-CF-SR has no significant advantage over BD-CF-SR on the direct link. Furthermore, at Level 4, with the MMSE combining scheme, RIS-CF-SR is superior to BD-CF-SR on both the direct and backscatter links.

Notation: We apply boldface uppercase letters to represent matrices. Boldface lowercase letters denote column vectors. Conjugate, transpose, and conjugate transpose operations are signified by the superscripts $(\cdot)^*$, $(\cdot)^T$, and $(\cdot)^H$, respectively. We use \triangleq to denote definitions. Moreover, the $N \times N$ identity matrix is represented as \mathbf{I}_N . The multivariate circularly symmetric complex Gaussian distribution with correlation matrix \mathbf{G} is denoted by $\mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{G})$. $\mathbb{E}\{\mathbf{h}\}$ represent the expected value of \mathbf{h} .

II. SYSTEM MODEL

The proposed RIS-CF-SR system is shown in Fig. 1, where several RIS are placed to support the transmission between the

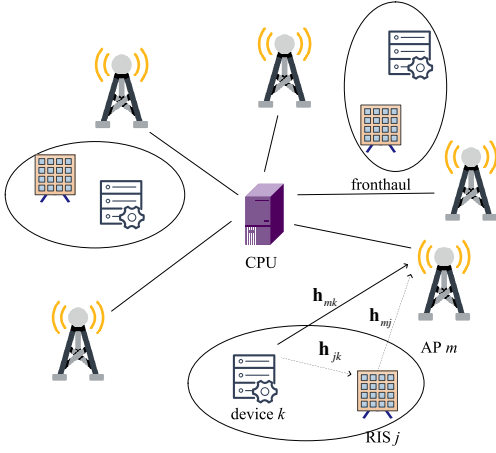


Fig. 1: RIS-CF-SR system model.

devices and the APs. Simultaneously, these RIS also function as extra devices, utilizing backscatter modulation to transmit their information to the APs. More specifically, J RISs, M APs, and K devices with single-antenna are proposed in the system. We assume that each RIS has N_{RIS} reflection elements ($N_{\text{RIS}} \geq 1$) and each AP is equipped with N antennas. Since the number of RISs is larger than the number of devices, each device can be paired with a different nearby RIS. For simplicity, the RIS near the device i is denoted as RIS j . During the uplink transmission, device i transmits the signal to all the APs through the direct-link channel, and the corresponding RIS j can assist the direct-link channel transmission. Moreover, at the same time, it also modulates its information onto the incident signals. Note that the distance between the RIS j and the other device is very long, thus the interfering links from RIS j and the other far devices can be ignored.

All the APs are linked to a CPU via error-free fronthaul links. The direct-link channel from device k to AP m is denoted by $\mathbf{h}_{mk} \in \mathbb{C}^{N \times 1}$, which can be modeled by Rayleigh fading as $\mathbf{h}_{mk} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R}_{mk})$. Note that the positive semi-infinite correlation matrix $\mathbf{R}_{mk} \in \mathbb{C}^{N \times N}$ captures the spatial properties of the channel, and $\beta_{mk} \triangleq \text{tr}(\mathbf{R}_{mk})/N$ symbolizes the large-scale fading coefficient related to geometric pathloss and shadowing. Moreover, the same setup is also used for the channel from RIS j to AP m . The channel can be represented as $\mathbf{h}_{mj} \in \mathbb{C}^{N \times N_{\text{RIS}}}$.

Particularly, $\mathbf{h}_{jk} \in \mathbb{C}^{N_{\text{RIS}} \times 1}$ are used to represent the channel from device k to RIS j . Since device k and RIS j are relatively close, the channel typically consists of a combination of a semideterministic line-of-sight (LoS) path and small-scale fading caused by multipath propagation, which can be modeled as Rician fading [24], [27]. More specifically, the channel between device k and RIS j is modeled as $\mathbf{h}_{jk} \sim \mathcal{N}_{\mathbb{C}}(\Psi_{jk}\bar{\mathbf{h}}_{jk}, \mathbf{R}_{jk})$, where $\bar{\mathbf{h}}_{jk} \in \mathbb{C}^{N_{\text{RIS}} \times 1}$ represents the deterministic LoS component and $\mathbf{R}_{jk} \in \mathbb{C}^{N_{\text{RIS}} \times N_{\text{RIS}}}$ is the spatial correlation matrix for the non-LoS (NLoS) propagation. Moreover, $\Psi_{jk} = \text{diag}(e^{j\psi_{jk1}}, \dots, e^{j\psi_{jkN_{\text{RIS}}}}) \in \mathbb{C}^{N_{\text{RIS}} \times N_{\text{RIS}}}$, where $\psi_{jkn} \sim \mathcal{U}[-\pi, \pi]$ is the additional phase-shift of the LoS component between the n -th elements of RIS j and device k . Furthermore, the reflection

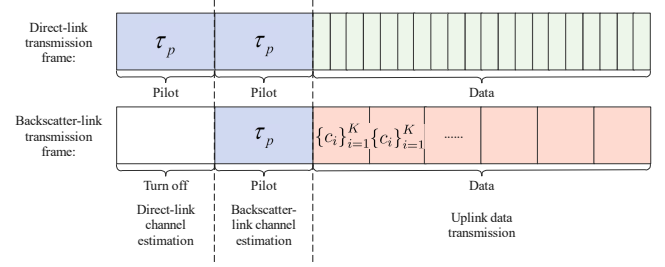


Fig. 2: Channel estimation and data transmission.

matrix at RIS j is $\Theta_j = \text{diag}(\Phi_j) \in \mathbb{C}^{N_{\text{RIS}} \times N_{\text{RIS}}}$, where $\Phi_j = \alpha_j \text{diag}(\exp(j\theta_{j1}), \exp(j\theta_{j2}), \dots, \exp(j\theta_{jN_{\text{RIS}}})) \in \mathbb{C}^{N_{\text{RIS}} \times 1}$ is the RIS reflect beamforming vector, with $\theta_{jn} \in [0, 2\pi]$ being the induced phase-shift and $\alpha_j \in [0, 1]$ describing the given amplitude reflection coefficient of RIS j . Based on the above descriptions, the total backscatter-link channel can be defined, i.e., the backscatter-link channel from device k to AP m via RIS j is drawn as $\mathbf{g}_{mk} = \mathbf{h}_{mj}\Phi_j\mathbf{h}_{jk} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{Q}_{mk})$, where $\mathbf{Q}_{mk} = \mathbb{E}\{\mathbf{g}_{mk}\mathbf{g}_{mk}^H\}$.

In addition, the BD-CF-SR system is also presented in this paper. Let $\mathbf{h}_{mj,\text{BD}} \in \mathbb{C}^{N \times 1}$ and $h_{jk,\text{BD}} \in \mathbb{C}$ be the channel from the AP m and BD j , and from the BD j to device k . The fading setups is the same as the RIS-CF-SR system. Thus, the backscatter-link channel from device k to AP m via BD j can be written as $\mathbf{g}_{mk,\text{BD}} = \alpha_{j,\text{BD}}\mathbf{h}_{mj,\text{BD}}h_{jk,\text{BD}} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{Q}_{mk,\text{BD}})$, where $\alpha_{j,\text{BD}}$ reflects the power reflection coefficient of BD j , and $\mathbf{Q}_{mk,\text{BD}} = \mathbb{E}\{\mathbf{g}_{mk,\text{BD}}\mathbf{g}_{mk,\text{BD}}^H\}$.

When the symbol rate of the backscatter-link signals transmitted by the RISs is much lower than that of the direct-link signals sent by devices, it enables enhanced collaboration between these two links [17], thereby contributing to improved overall performance. Hence, we make the assumption that whenever RIS j sends a backscatter-link symbol to the APs, device i transmits L direct-link symbols. Let T_{dl} and T_{bl} represent the symbol period of the device and RIS, respectively. In this case, we can express the relationship between the two as $T_{\text{bl}} = LT_{\text{dl}}$.

We employ the standard block fading model, wherein time-frequency blocks are composed of τ_c channel uses. As depicted in Fig. 2, the system operates in time division duplex (TDD) mode. Each coherence interval is divided into three phases: the channel estimation phase for the direct link taking τ_p channel uses, the channel estimation phase for the backscatter link taking τ_p channel uses, and finally, the uplink data transmission phase taking $\tau_c - 2\tau_p$ channel uses. It is demanded that $2\tau_p < \tau_c$.

A. Channel Estimation

A two-phase channel estimation approach is considered (see Fig. 2). First, we perform direct-link channel estimation while turning off the RISs. In this phase, this scheme is similar to the traditional CF-mMIMO channel estimation.

Let $\sqrt{\tau_p}\varphi_k^H \in \mathbb{C}^{\tau_p \times 1}$ represents the pilot sequence utilized by device k , and $\|\varphi_k\|^2 = 1$. All devices need to simultane-

ously transmit pilot sequences of length τ_p samples to all APs simultaneously. The signal $\mathbf{y}_m^{\text{dl}} \in \mathbb{C}^{N \times \tau_p}$ received at AP m is

$$\mathbf{y}_m^{\text{dl}} = \sqrt{\tau_p} \sum_{i=1}^K \sqrt{p_i^{\text{ce}}} \mathbf{h}_{mi} \boldsymbol{\varphi}_i^T + \mathbf{w}_m^{\text{dl}}, \quad (1)$$

where p_i^{ce} reflects power of the device i , $\mathbf{w}_m^{\text{dl}} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2)$ represents the noise received at AP m , and σ^2 denotes the noise power. Next, AP m needs to correlate the pilot $\boldsymbol{\varphi}_k^*$ with the received signal to achieve

$$\hat{\mathbf{y}}_{mk}^{\text{dl}} = \sqrt{\tau_p p_k^{\text{ce}}} \mathbf{h}_{mk} + \sqrt{\tau_p} \sum_{i \neq k} \sqrt{p_i^{\text{ce}}} \mathbf{h}_{mi} \boldsymbol{\varphi}_i^T \boldsymbol{\varphi}_k^* + \mathbf{w}_m^{\text{dl}} \boldsymbol{\varphi}_k^*. \quad (2)$$

Using the standard MMSE estimation [10], the channel estimation of the direct link $\hat{\mathbf{h}}_{mk}$ is

$$\hat{\mathbf{h}}_{mk} = \boldsymbol{\Omega}_{mk} \hat{\mathbf{y}}_{mk}^{\text{dl}}, \quad (3)$$

where

$$\boldsymbol{\Omega}_{mk} \triangleq \frac{\sqrt{\tau_p p_k^{\text{ce}}} \mathbf{R}_{mk}}{\tau_p \sum_{i=1}^K p_i^{\text{ce}} \mathbf{R}_{mi} |\boldsymbol{\varphi}_i^T \boldsymbol{\varphi}_k^*|^2 + \sigma^2 \mathbf{I}_N}. \quad (4)$$

The channel estimation error $\tilde{\mathbf{h}}_{mk}$ for the direct link is given by $\tilde{\mathbf{h}}_{mk} = \mathbf{h}_{mk} - \hat{\mathbf{h}}_{mk}$. Furthermore, the estimation errors and estimation are distributed as $\tilde{\mathbf{h}}_{mk} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \tilde{\mathbf{R}}_{mk})$, and $\hat{\mathbf{h}}_{mk} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \hat{\mathbf{R}}_{mk})$, respectively, where $\tilde{\mathbf{R}}_{mk} = \mathbf{R}_{mk} - \hat{\mathbf{R}}_{mk}$. Note that $\hat{\mathbf{R}}_{mk} = \sqrt{p_k^{\text{ce}} \tau_p} \mathbf{R}_{mk} \boldsymbol{\Omega}_{mk}$.

After the direct-link channel estimation $\hat{\mathbf{h}}_{mk}$ is achieved, to estimate the channel of the backscatter link, all devices need to retransmit the identical pilot signals while activating the RISs. Note that although the RISs are turned on, they do not send any signals during the channel estimation phase. Therefore, the symbol rate of the RISs and the devices can be regarded as the same at this point (see Fig. 2).

Subsequently, the pilot signals are transmitted independently via the backscatter link and the direct link. Thus, the received pilot signals \mathbf{y}_m^{bl} during the backscatter-link channel estimation phase are

$$\mathbf{y}_m^{\text{bl}} = \sqrt{\tau_p} \sum_{i=1}^K \sqrt{p_i^{\text{ce}}} \mathbf{h}_{mi} \boldsymbol{\varphi}_i^T + \sqrt{\tau_p} \sum_{i=1}^K \sqrt{p_i^{\text{ce}}} \mathbf{g}_{mi} \boldsymbol{\varphi}_i^T + \mathbf{w}_m^{\text{bl}}. \quad (5)$$

where $\mathbf{w}_m^{\text{bl}} \in \mathbb{C}^{N \times \tau_p}$ denotes the noise received at AP m .

Note that the direct-link channel estimation $\hat{\mathbf{h}}_{mk}$ is known, the components $\sqrt{\tau_p} \sum_{i=1}^K \sqrt{p_i^{\text{ce}}} \hat{\mathbf{h}}_{mi} \boldsymbol{\varphi}_i^T$ can be eliminated from the signal given in $\mathbf{y}_m^{\text{bl}} \in \mathbb{C}^{N \times \tau_p}$. The eliminated signal $\bar{\mathbf{y}}_m^{\text{bl}}$ is

$$\bar{\mathbf{y}}_m^{\text{bl}} = \sqrt{\tau_p} \sum_{i=1}^K \sqrt{p_i^{\text{ce}}} (\tilde{\mathbf{h}}_{mi} + \mathbf{g}_{mi}) \boldsymbol{\varphi}_i^T + \mathbf{w}_m^{\text{bl}}. \quad (6)$$

Then APs are required to relate the eliminated signals $\bar{\mathbf{y}}_m^{\text{bl}}$ to the relevant pilot signals $\boldsymbol{\varphi}_k^*$ as

$$\hat{\mathbf{y}}_{mk}^{\text{bl}} = \bar{\mathbf{y}}_m^{\text{bl}} \boldsymbol{\varphi}_k^* = \sqrt{\tau_p} \sum_{i=1}^K \sqrt{p_i^{\text{ce}}} (\tilde{\mathbf{h}}_{mi} + \mathbf{g}_{mi}) \boldsymbol{\varphi}_i^T \boldsymbol{\varphi}_k^* + \mathbf{w}_m^{\text{bl}} \boldsymbol{\varphi}_k^*. \quad (7)$$

From $\hat{\mathbf{y}}_{mk}^{\text{bl}}$, the MMSE estimation of the backscatter-link channel \mathbf{g}_{mk} is

$$\hat{\mathbf{g}}_{mk} = \frac{\mathbb{E} \left\{ \mathbf{g}_{mk} (\hat{\mathbf{y}}_{mk}^{\text{bl}})^H \right\}}{\mathbb{E} \left\{ \hat{\mathbf{y}}_{mk}^{\text{bl}} (\hat{\mathbf{y}}_{mk}^{\text{bl}})^H \right\}} \hat{\mathbf{y}}_{mk}^{\text{bl}} = \boldsymbol{\Psi}_{mk} \hat{\mathbf{y}}_{mk}^{\text{bl}}, \quad (8)$$

where

$$\boldsymbol{\Psi}_{mk} \triangleq \frac{\sqrt{p_k^{\text{ce}} \tau_p} \mathbf{Q}_{mk}}{\tau_p \sum_{i=1}^K p_i^{\text{ce}} (\tilde{\mathbf{R}}_{mi} + \mathbf{Q}_{mi}) |\boldsymbol{\varphi}_i^T \boldsymbol{\varphi}_k^*|^2 + \sigma^2 \mathbf{I}_N}. \quad (9)$$

The backscatter-link channel estimation $\hat{\mathbf{g}}_{mk}$ and estimation error $\tilde{\mathbf{g}}_{mk}$ are independent distributed as $\hat{\mathbf{g}}_{mk} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \hat{\mathbf{Q}}_{mk})$ with $\hat{\mathbf{Q}}_{mk} = \sqrt{p_k^{\text{ce}} \tau_p} \mathbf{Q}_{mk} \boldsymbol{\Psi}_{mk}$ and $\tilde{\mathbf{g}}_{mk} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{Q}_{mk} - \hat{\mathbf{Q}}_{mk})$.

Remark 1. It is worth emphasizing that the outstanding performance brought by RIS relies on the phase modulation. In this paper, all the elements on the RISs are treated as a whole for channel estimation, which means we use random phases without the phase modulation. This almost represents the worst-case scenario for RISs. Hence, this paper can be viewed as a comparison between the lower bound of RIS-CF-SR and the general case of BD-CF-SR.

B. Uplink Data Transmission

Let c_i be the symbol sent by RIS j . We assume c_i applies binary phase-shift keying (BPSK) modulation, i.e., $c_i \in \{1, -1\}$. The l -th received signal at AP m during the uplink data transmission is

$$\mathbf{y}_m^{\text{up}}(l) = \sum_{i=1}^K \sqrt{p_i} \mathbf{h}_{mi} s_i(l) + \sum_{i=1}^K \sqrt{p_i} \mathbf{g}_{mi} s_i(l) c_i + \mathbf{w}_m, \quad (10)$$

where p_i denotes the transmit power of device i , $s_i(l) \sim \mathcal{N}_{\mathbb{C}}(0, 1)$ is the l -th information-bearing signal sent by device i , and $\mathbf{w}_m \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ reflects the receiver noise at AP m .

III. FOUR LEVELS OF SIGNAL PROCESSING SCHEMES

In RIS-CF-SR, all APs are connected to the CPU through fronthaul, enabling them to transmit information to the CPU for designing better signal processing schemes and improving performance. However, overly complex signal processing schemes may place higher demands on the computing power. Therefore, it is essential to balance the complexity of signal processing with system performance. According to the above observations, four different implementation levels of processing schemes are proposed in this section.

A. Level 4: Fully Centralized Processing

The most advanced processing scheme is that all APs send their pilots and data signals to the CPU for centralized processing. In this case, all the APs act as relays while the CPU is required to perform all the channel estimation and signal detection.

The l -th received signals at the CPU is

$$\mathbf{y}^{\text{up}}(l) = \sum_{i=1}^K \sqrt{p_i} \mathbf{h}_i s_i(l) + \sum_{i=1}^K \sqrt{p_i} \mathbf{g}_i s_i(l) c_i + \mathbf{w}, \quad (11)$$

where $\mathbf{w} = [\mathbf{w}_1, \dots, \mathbf{w}_M]$. Moreover, the collective channel \mathbf{h}_i and \mathbf{g}_i can be written as $\mathbf{h}_i = [\mathbf{h}_{1i}, \dots, \mathbf{h}_{Mi}] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{R}_k)$ and $\mathbf{g}_i = [\mathbf{g}_{1i}, \dots, \mathbf{g}_{Mi}] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \mathbf{Q}_k)$ where $\mathbf{R}_k = \text{diag}(\mathbf{R}_{1k}, \dots, \mathbf{R}_{Mk}) \in \mathbb{C}^{MN \times MN}$ and $\mathbf{Q}_k = \text{diag}(\mathbf{Q}_{1k}, \dots, \mathbf{Q}_{Mk}) \in \mathbb{C}^{MN \times MN}$, respectively.

Since all the pilot signals and channel statistics is transmitted to the CPU, the MMSE channel estimation can be computed jointly. For device k , the direct-link channel estimation is given by

$$\hat{\mathbf{h}}_k \triangleq \begin{bmatrix} \hat{\mathbf{h}}_{1k} \\ \vdots \\ \hat{\mathbf{h}}_{Mk} \end{bmatrix} \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sqrt{p_k^{\text{ce}} \tau_p} \mathbf{R}_k \mathbf{\Omega}_k), \quad (12)$$

where $\mathbf{\Omega}_k = \text{diag}(\mathbf{\Omega}_{1k}, \dots, \mathbf{\Omega}_{Mk}) \in \mathbb{C}^{MN \times MN}$. Moreover, the direct-link channel estimation errors are $\tilde{\mathbf{h}}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \tilde{\mathbf{R}}_k)$ with $\tilde{\mathbf{R}}_k = \text{diag}(\tilde{\mathbf{R}}_{1k}, \dots, \tilde{\mathbf{R}}_{Mk}) \in \mathbb{C}^{MN \times MN}$. Similarly, the backscatter-link channel estimation for device k is given by $\hat{\mathbf{g}}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \sqrt{p_k^{\text{ce}} \tau_p} \mathbf{Q}_k \mathbf{\Psi}_k)$ with $\mathbf{\Psi}_k = \text{diag}(\mathbf{\Psi}_{1k}, \dots, \mathbf{\Psi}_{Mk}) \in \mathbb{C}^{MN \times MN}$. And the backscatter-link channel estimation error is $\tilde{\mathbf{g}}_k \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \tilde{\mathbf{Q}}_k)$, where $\tilde{\mathbf{Q}}_k = \text{diag}(\tilde{\mathbf{Q}}_{1k}, \dots, \tilde{\mathbf{Q}}_{Mk}) \in \mathbb{C}^{MN \times MN}$. Next, based on the received signal (11), the CPU aspires to decode both direct-link channel signals $s_i(l)$ and backscatter-link signals c_i by selecting an arbitrary receive combining vector $\mathbf{v}_k \in \mathbb{C}^{MN \times 1}$. Furthermore, centralizing the signal processing at a CPU offers an additional advantage: it provides access ample computational resources, enabling the potential utilization of more advanced decoding methods. Different from the most existing papers about SR [17]–[20], we consider the successive interference cancellation (SIC) technique to assist in decoding with imperfect CSI. Specifically, the signals are decoded individually by the CPU, with interference caused by each decoded signal sequentially eliminated from the remaining signals. In this paper, we apply the smaller index to represent the set of $s_i(l)$ and c_i that are decoded earlier. It means that the signal $s_i(l)$ is decoded first and signal $s_K(l)$ is decoded last when the CPU decodes direct-link channel signals. To this end, the received signals is given by

$$\begin{aligned} \mathbf{y}^{\text{up}}(l) = & \sqrt{p_k} \mathbf{v}_k^H (\mathbf{h}_k s_k(l) + \mathbf{g}_k s_k(l) c_k) + \sum_{i \neq k}^K \sqrt{p_i} \mathbf{v}_k^H \mathbf{h}_i s_i(l) \\ & - \sum_{i=0}^{k-1} \sqrt{p_i} \mathbf{v}_k^H \hat{\mathbf{h}}_i s_i(l) + \sum_{i \neq k}^K \sqrt{p_i} \mathbf{v}_k^H \mathbf{g}_i s_i(l) c_i + \mathbf{w}. \end{aligned} \quad (13)$$

Take into account that the symbol period of the direct link is significantly lower than that for the backscatter link, thus the second term in (13) can convert into the output of the direct-link signal s_k transmitting through a slowly varying channel $\mathbf{v}_k^H \mathbf{g}_k c_k$ [17]. Meanwhile the other signal comes from the backscatter link is treated as interference, and, therefore, the

average power of which is $\mathbb{E} \left\{ p_i |\mathbf{v}_i^H|^2 |\mathbf{g}_i|^2 |s_i(l)|^2 |c_i|^2 \right\} = p_i |\mathbf{v}_i^H \mathbf{g}_i|^2$.

Note that the capacity of SR with perfect CSI has been extensively investigated [17]–[20], [24], [27]–[31]. However, the perfect CSI is generally unknown. Fortunately, we can use the standard capacity lower bounds to rigorously analyze the performance [9], which can be viewed as the so-called “achievable SEs”.

Proposition 1. *For direct link, the achievable SE of device k at Level 4 is*

$$\text{SE}_k^{(4), \text{dl}} = \left(1 - \frac{2\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_k^{(4), \text{dl}} \right) \right\}, \quad (14)$$

where the instantaneous SINR is given in (15) (see top of this page).

Proof: See Proof B. ■

It is worth noting that any combining vectors \mathbf{v}_k can be applied in equation (15), one option is the MR combining with $\mathbf{v}_k = \hat{\mathbf{h}}_k + \hat{\mathbf{g}}_k$, referred to as “hybrid MR combining” in [23]. This approach maximizes the desired signal while overlooking interference. However, with the introduction of RISs, additional interference may very large. Thus, we employ the MMSE combining, which is computed by minimizing the mean squared error (MSE), and $\text{MSE}_k = \mathbb{E} \left\{ |s_k - \mathbf{v}_k^H \mathbf{y}^{\text{up}}|^2 \mid \hat{\mathbf{h}}_i, \hat{\mathbf{g}}_i \right\}$. The MMSE combining vector \mathbf{v}_k is

$$\mathbf{v}_k = p_k \left(\sum_{i=1}^K p_i (\mathbf{A}_i + \tilde{\mathbf{R}}_i + \tilde{\mathbf{Q}}_i) + \sigma^2 \mathbf{I}_{MN} \right)^{-1} (\hat{\mathbf{h}}_k + \hat{\mathbf{g}}_k), \quad (16)$$

where $\mathbf{A}_i \triangleq \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^H + \hat{\mathbf{g}}_i \hat{\mathbf{g}}_i^H \in \mathbb{C}^{N \times N}$.

Proof: The proof follows the similar steps as in [32] and is therefore omitted. ■

We have to take into consideration that (15) is a generalized Rayleigh quotient with respect to \mathbf{v}_k , and the same result can also be obtained by following the similar steps outlined in [9].

After decoding all the direct-link signals $\{s_i(l) : i = 1, 2, \dots, K\}$, CPU tends to decode the signals of backscatter link. Similarly, by applying SIC technique, (13) can be simplified as

$$\begin{aligned} \mathbf{y}^{\text{up}}(l) = & \sqrt{p_k} \mathbf{v}_k^H \mathbf{g}_k s_k(l) c_k + \sum_{i \neq k}^K \sqrt{p_i} \mathbf{v}_k^H \mathbf{g}_i s_i(l) c_i \\ & - \sum_{i=0}^{k-1} \sqrt{p_i} \mathbf{v}_k^H \hat{\mathbf{g}}_i s_i(l) c_i + \sum_{i=1}^K \sqrt{p_i} \mathbf{v}_k^H \tilde{\mathbf{h}}_i s_i(l) + \mathbf{w}. \end{aligned} \quad (17)$$

Subsequently, utilizing (17), the achievable SEs can be obtained. Since only one symbol from the backscatter link is transmitted during L successive symbol periods of the direct link, the direct-link signal $s_k(l)$ can be considered as a spread-spectrum code with a length of L for backscatter-link signal c_k [17]. Consequently, the SINR for decoding the backscatter-link signal increases by a factor of L , while the achievable SEs decrease by $1/L$. Therefore, the achievable SEs of the backscatter link are expressed as follows.

$$\text{SINR}_k^{(4),\text{dl}} = \frac{p_k \left| \mathbf{v}_k^H \hat{\mathbf{h}}_k + \mathbf{v}_k^H \hat{\mathbf{g}}_k \right|^2}{\sum_{i=k+1}^K p_i \left| \mathbf{v}_k^H \hat{\mathbf{h}}_i \right|^2 + \sum_{i \neq k}^K p_i \left| \mathbf{v}_k^H \hat{\mathbf{g}}_i \right|^2 + \mathbf{v}_k^H \left(\sum_{i=1}^K p_i (\tilde{\mathbf{R}}_i + \tilde{\mathbf{Q}}_i) + \sigma^2 \mathbf{I}_{MN} \right) \mathbf{v}_k} \quad (15)$$

$$\text{SINR}_k^{(4),\text{bl}} = \frac{L p_k \left| \mathbf{v}_k^H \hat{\mathbf{g}}_k \right|^2}{\sum_{i=k+1}^K p_i \left| \mathbf{v}_k^H \hat{\mathbf{g}}_i \right|^2 + \mathbf{v}_k^H \left(\sum_{i=1}^K p_i (\tilde{\mathbf{R}}_i + \tilde{\mathbf{Q}}_i) + \sigma^2 \mathbf{I}_{MN} \right) \mathbf{v}_k} \quad (19)$$

Proposition 2. For device k , the SE of the backscatter link at Level 4 is

$$\text{SE}_k^{(4),\text{bl}} = \left(\frac{\tau_c - 2\tau_p}{L\tau_c} \right) \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_k^{(4),\text{bl}} \right) \right\}, \quad (18)$$

where the instantaneous effective SINR can be written as (19) (see top of the next page).

Proof: The proof of (19) is similar to the proof of (15) and is therefore omitted. ■

Remark 2. We observe that the MMSE combining vector \mathbf{v}_k is designed for the SINR of the direct link. Therefore, it may not be the optimal combining vector for maximizing the SINR of the backscatter link. Nevertheless, it has been demonstrated in [23] that the suboptimal MMSE combining still offers improved performance compared to the simple MR combining.

B. Level 3: Local Processing & Large-Scale Fading Decoding

Instead of transmitting all pilot and data signals to the CPU for centralized processing, the APs have the capability to preprocess their signals locally. They can compute local estimates of the data and subsequently transmit them to the CPU for final decoding.

Similar to Level 4, the direct-link channel signals are assumed to be processed first. However, since the APs do not have ample computational resource like the CPU, the SIC technology is not available at Level 3. Let $\mathbf{v}_{mk} \in \mathbb{C}^{N \times 1}$ be the local combining vectors calculated by AP m , the l -th received signals with the local combining vector at AP m is

$$\begin{aligned} \hat{s}_{\text{dl},mk}(l) &\triangleq \mathbf{v}_{mk}^H \mathbf{y}_m^{\text{up}}(l) \\ &= \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{h}_{mk} s_k(l) + \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{g}_{mk} s_k(l) c_k \\ &\quad + \sum_{i \neq k}^K \sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{h}_{mi} s_i(l) + \sum_{i \neq k}^K \sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{g}_{mi} s_i(l) c_i \\ &\quad + \mathbf{v}_{mk}^H \mathbf{w}_m. \end{aligned} \quad (20)$$

We notice that the above expression allows for the adoption of any combining vectors. A simple solution is MR combining designed by the AP m as $\mathbf{v}_{mk} = \hat{\mathbf{h}}_{mk} + \hat{\mathbf{g}}_{mk}$ [23], but the better solution is the MMSE combining computed by the local CSI at AP m , which can be called as local MMSE (L-MMSE)

combining scheme. Similar to (16), the L-MMSE combining vector for Level 3 is given by

$$\mathbf{v}_{mk} = p_k \left(\sum_{i=1}^K p_i (\mathbf{A}_{mi} + \tilde{\mathbf{R}}_{mi} + \tilde{\mathbf{Q}}_{mi}) + \sigma^2 \mathbf{I}_N \right)^{-1} (\hat{\mathbf{h}}_{mk} + \hat{\mathbf{g}}_{mk}), \quad (21)$$

where $\mathbf{A}_{mi} \triangleq \hat{\mathbf{h}}_{mi} \hat{\mathbf{h}}_{mi}^H + \hat{\mathbf{g}}_{mi} \hat{\mathbf{g}}_{mi}^H \in \mathbb{C}^{N \times N}$. Then all these signals are sent to the CPU for the final decoding. The CPU sees

$$\begin{aligned} \hat{s}_k &= \sum_{m=1}^M \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{h}_{mk} s_k + \sum_{m=1}^M \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{g}_{mk} s_k c_k \\ &\quad + \sum_{m=1}^M \sum_{i \neq k}^K \sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{h}_{mi} s_i + \sum_{m=1}^M \sum_{i \neq k}^K \sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{g}_{mi} s_i c_i \\ &\quad + \sum_{m=1}^M \mathbf{v}_{mk}^H \mathbf{w}_m. \end{aligned} \quad (22)$$

A second layer decoding structure is applied, which is called “LSFD” in [9]. The CPU can select the weights $\{a_{mk} : m = 1, 2, \dots, M\}$ for all the signals from the APs to enhance the direct-link SE. Let $\mathbf{f}_{ki}^{\text{dl}} = [\mathbf{v}_{1k}^H \mathbf{h}_{1i} \dots \mathbf{v}_{Mk}^H \mathbf{h}_{Mi}]^T$, $\mathbf{f}_{ki}^{\text{bl}} = [\mathbf{v}_{1k}^H \mathbf{g}_{1i} \dots \mathbf{v}_{Mk}^H \mathbf{g}_{Mi}]^T$, and $\mathbf{w}'_k = [\mathbf{v}_{1k}^H \mathbf{w}_1 \dots \mathbf{v}_{Mk}^H \mathbf{w}_M]^T$ be the M -dimensional vector, (22) becomes

$$\begin{aligned} \hat{s}_k &= \mathbf{a}_k^H \mathbf{f}_{kk}^{\text{dl}} s_k + \mathbf{a}_k^H \mathbf{f}_{kk}^{\text{bl}} s_k c_k \\ &\quad + \sum_{i \neq k}^K \mathbf{a}_k^H \mathbf{f}_{ki}^{\text{dl}} s_i + \sum_{i \neq k}^K \mathbf{a}_k^H \mathbf{f}_{ki}^{\text{bl}} s_i c_i + \mathbf{a}_k^H \mathbf{w}'_k, \end{aligned} \quad (23)$$

where $\mathbf{a}_k = [a_{1k} \dots a_{Mk}]^T \in \mathbb{C}^{M \times 1}$ is the weighting coefficient vector. Note that at the CPU, only statistical CSI are accessible. Hence, the achievable SE for the LSFD scheme is obtained by employing the established use-and-then-forget (UatF) bound, as delineated in the subsequent proposition.

Proposition 3. The achievable SE of device k at Level 3 for direct link is

$$\text{SE}_k^{(3),\text{dl}} = \left(1 - \frac{2\tau_p}{\tau_c} \right) \log_2 \left(1 + \text{SINR}_k^{(3),\text{dl}} \right) \quad (24)$$

with the effective SINR as

$$\text{SINR}_k^{(3),\text{dl}} = \frac{|\mathbf{a}_k^H \boldsymbol{\Sigma}_k^{\text{dl}}|^2}{\mathbf{a}_k^H \mathbf{D}_k^{\text{dl}} \mathbf{a}_k}, \quad (25)$$

where $\Sigma_k^{\text{dl}} = \mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{dl}} \} + \mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{bl}} \}$, and $\mathbf{D}_k^{\text{dl}} = \sum_{i=1}^K \mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}}|^2 \} - |\mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{dl}} \}|^2 + \sum_{i=1}^K \mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{bl}}|^2 \} - |\mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{bl}} \}|^2 + \mathbb{E} \{ |\mathbf{w}'_k|^2 \}$.

Proof: The proof follows the similar steps as in [9] and is therefore omitted. ■

The structure of (25) enables the computation of the deterministic weighting vector \mathbf{a}_k that maximizes the SINR of the direct link. The expression for this weighting vector is as follows.

Corollary 1. *The effective SINR of the direct link for device k can be maximized by*

$$\mathbf{a}_k = (\mathbf{D}_k^{\text{dl}})^{-1} \Sigma_k^{\text{dl}}, \quad (26)$$

which leads to the maximum value as

$$\text{SINR}_k^{(3),\text{dl}} = (\Sigma_k^{\text{dl}})^H (\mathbf{D}_k^{\text{dl}})^{-1} \Sigma_k^{\text{dl}}. \quad (27)$$

Proof: The proof follows from [32, Lemma B.10] by noting that (25) is a generalized Rayleigh quotient with respect to \mathbf{a}_k . ■

After decoding the direct-link signals, the CPU aspires to decode the signals from the backscatter-link. In order to reduce interference from the direct link, CPU first needs to perform signal cancellation on (22). We observe that the average value is non-zero and deterministic, although the exact information at the CPU remains unknown. Consequently, we can treat it as a known quantity and use the average value $\sum_{i=1}^K \mathbb{E} \{ \sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}} \}$ to remove partial interference from the direct link. Thus, the resulting signal \hat{s}_k^{bl} reduces to

$$\begin{aligned} \hat{s}_k^{\text{bl}} = & \sqrt{p_k} \mathbf{f}_{kk}^{\text{bl}} s_k c_k + \sum_{i \neq k} \sqrt{p_i} \mathbf{f}_{ki}^{\text{bl}} s_i c_i \\ & + \left(\sum_{i=1}^K \sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}} - \sum_{i=1}^K \mathbb{E} \{ \sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}} \} \right) s_i + \mathbf{w}'_k. \end{aligned} \quad (28)$$

Similar to (23), LSFD can be utilized once more for the second layer decoding of the backscatter-link signal. Therefore, the weighting coefficient vector $\mathbf{b}_k = [b_{1k} \dots b_{Mk}]^T \in \mathbb{C}^{M \times 1}$ is selected to maximize the backscatter-link SE. The resulting signal \hat{s}_k^{bl} becomes

$$\begin{aligned} \hat{s}_k^{\text{bl}} = & \sqrt{p_k} \mathbf{b}_k^H \mathbf{f}_{kk}^{\text{bl}} s_k c_k + \sum_{i \neq k} \sqrt{p_i} \mathbf{b}_k^H \mathbf{f}_{ki}^{\text{bl}} s_i c_i \\ & + \left(\sum_{i=1}^K \sqrt{p_i} \mathbf{b}_k^H \mathbf{f}_{ki}^{\text{dl}} - \sum_{i=1}^K \mathbb{E} \{ \sqrt{p_i} \mathbf{b}_k^H \mathbf{f}_{ki}^{\text{dl}} \} \right) s_i + \mathbf{b}_k^H \mathbf{w}'_k. \end{aligned} \quad (29)$$

Then, the achievable SE for the backscatter link can be achieved as follows.

Proposition 4. *An achievable SE of device k at Level 3 for backscatter link is given by*

$$\text{SE}_k^{(3),\text{bl}} = \left(\frac{\tau_c - 2\tau_p}{L\tau_c} \right) \log_2 \left(1 + \text{SINR}_k^{(3),\text{bl}} \right) \quad (30)$$

with the effective SINR as

$$\text{SINR}_k^{(3),\text{bl}} = \frac{L |\mathbf{b}_k^H \Sigma_k^{\text{bl}}|^2}{\mathbf{b}_k^H \mathbf{D}_k^{\text{bl}} \mathbf{b}_k}, \quad (31)$$

where $\Sigma_k^{\text{bl}} = \mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{bl}} \}$ and $\mathbf{D}_k^{\text{bl}} = \sum_{i=1}^K \mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{bl}}|^2 \} - |\mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{bl}} \}|^2 + \sum_{i=1}^K \mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}}|^2 \} - |\mathbb{E} \{ \sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}} \}|^2 + \mathbb{E} \{ |\mathbf{w}'_k|^2 \}$.

Proof: The proof of (31) is similar to the proof of (25) and is therefore omitted. ■

Analogous to (25), the weighting vector \mathbf{b}_k can be computed to maximize the SINR of the backscatter link by noting that (31) is a generalized Rayleigh quotient with respect to \mathbf{b}_k .

Corollary 2. *The effective SINR of the backscatter link for device k is maximized by*

$$\mathbf{b}_k = (\mathbf{D}_k^{\text{bl}})^{-1} \Sigma_k^{\text{bl}}, \quad (32)$$

which leads to the maximum value as

$$\text{SINR}_k^{(3),\text{bl}} = (\Sigma_k^{\text{bl}})^H (\mathbf{D}_k^{\text{bl}})^{-1} \Sigma_k^{\text{bl}}. \quad (33)$$

When employing L-MMSE combining, the expectations in (25) and (31) cannot be calculated in closed form, but can be determined through Monte Carlo simulations. Nevertheless, employing MR combining $\mathbf{v}_{mk} = \hat{\mathbf{h}}_{mk} + \tilde{\mathbf{g}}_{mk}$ enables us to derive the subsequent closed-form expressions.

Corollary 3. *If MR combining is used, then the closed-form expectations in (25) and (31) can be derived as follows.*

$$\begin{aligned} \Sigma_k^{\text{dl}} = & \sqrt{p_k} \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{R}}_{mk} \right) + \sqrt{p_k} \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{Q}}_{mk} \right) \\ & + \sqrt{p_k} \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{Q}}_{mk}}{\mathbf{Q}_{mk}} \tilde{\mathbf{R}}_{mk} \right), \end{aligned} \quad (34)$$

$$\begin{aligned} \mathbf{D}_k^{\text{dl}} = & \sum_{i=1}^K p_i \left(\sum_{m=1}^M \text{tr} \left(\hat{\mathbf{R}}_{mk} \mathbf{R}_{mi} \right) + \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{Q}}_{mk} \mathbf{R}_{mi} \right) \right) \\ & + \sum_{i=1}^K p_i \left(\sum_{m=1}^M \text{tr} \left(\hat{\mathbf{R}}_{mk} \mathbf{Q}_{mi} \right) + \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{Q}}_{mk} \mathbf{Q}_{mi} \right) \right) \\ & + \sum_{i \neq k} p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{R}}_{mk}}{\mathbf{R}_{mk}} \mathbf{R}_{mi} \right) \right|^2 + \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{R}}_{mk} \right) \sigma^2 \\ & + \sum_{i \neq k} p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{Q}}_{mk}}{\mathbf{Q}_{mk}} \mathbf{Q}_{mi} \right) \right|^2 + \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{Q}}_{mk} \right) \sigma^2 \\ & + \sum_{i \neq k} p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{Q}}_{mk}}{\mathbf{Q}_{mk}} \tilde{\mathbf{R}}_{mi} \right) \right|^2, \end{aligned} \quad (35)$$

$$\Sigma_k^{\text{bl}} = \sqrt{p_k} \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{Q}}_{mk} \right), \quad (36)$$

and

$$\mathbf{D}_k^{\text{bl}} = \mathbf{D}_k^{\text{dl}} - p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{R}}_{mk} \mathbf{R}_{mi}}{\mathbf{R}_{mk}} \right) \right|^2 - p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{Q}}_{mk} \tilde{\mathbf{R}}_{mi}}{\mathbf{Q}_{mk}} \right) \right|^2. \quad (37)$$

Finally, $\mathbf{a}_k = (\mathbf{D}_k^{\text{dl}})^{-1} \Sigma_k^{\text{dl}}$ and $\mathbf{b}_k = (\mathbf{D}_k^{\text{bl}})^{-1} \Sigma_k^{\text{bl}}$.

Proof: See Appendix A. ■

C. Level 2: Local Processing & Simple Centralized Decoding

While LSFD scheme can enhance system performance, this processing scheme demands the extensive knowledge of large-scale fading parameter, which might be excessively high in CF-mMIMO-SR. To address this concern, the CPU can generate the signal \hat{s}_k from device k by simply averaging the local estimates, as widely applied in early papers on this topic [21]–[23], [31]. Based on (20), the decoding signals can be written as

$$\hat{s}_{\text{dl},k} = \frac{1}{M} \sum_{m=1}^M \hat{s}_{\text{dl},mk}. \quad (38)$$

And the achievable SEs for both the direct and backscatter links are derived as follows.

Proposition 5. *An achievable SE of device k at Level 2 for direct link can be written as*

$$\text{SE}_k^{(2),\text{dl}} = \left(1 - \frac{2\tau_p}{\tau_c} \right) \log_2 \left(1 + \frac{|\Lambda_k^{\text{dl}}|^2}{\Xi_k^{\text{dl}}} \right), \quad (39)$$

where $\Lambda_k^{\text{dl}} = \sum_{m=1}^M (\mathbb{E} \{ \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{h}_{mk} \} + \mathbb{E} \{ \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{g}_{mk} \})$, and $\Xi_k^{\text{dl}} = \sum_{m=1}^M \left(\sum_{i=1}^K \mathbb{E} \{ |\sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{h}_{mi}|^2 \} - |\mathbb{E} \{ \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{h}_{mk} \}|^2 \right) + \sum_{m=1}^M \left(\sum_{i=1}^K \mathbb{E} \{ |\sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{g}_{mi}|^2 \} - |\mathbb{E} \{ \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{g}_{mk} \}|^2 \right) + \sum_{m=1}^M \mathbb{E} \{ |\mathbf{w}'_k|^2 \}$

Similarly, the backscatter-link signal does not use LSFD but only applies signal cancellation. The resulting signal is drawn as

$$\hat{s}_k^{\text{bl}} = \sum_{m=1}^M \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{g}_{mk} s_k c_k + \sum_{m=1}^M \sum_{i \neq k} \sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{g}_{mi} s_i c_i + \sum_{m=1}^M \sum_{i=1}^K (\sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{h}_{mi} - \mathbb{E} \{ \sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{h}_{mi} \}) s_i + \mathbf{w}'_k. \quad (40)$$

Proposition 6. *An achievable SE of device k at Level 2 for backscatter link is*

$$\text{SE}_k^{(2),\text{bl}} = \left(\frac{\tau_c - 2\tau_p}{L\tau_c} \right) \log_2 \left(1 + \frac{L|\Lambda_k^{\text{bl}}|^2}{\Xi_k^{\text{bl}}} \right), \quad (41)$$

where $\Lambda_k^{\text{bl}} = \sum_{m=1}^M \mathbb{E} \{ \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{g}_{mk} \}$ and $\Xi_k^{\text{bl}} = \sum_{m=1}^M \left(\sum_{i=1}^K \mathbb{E} \{ |\sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{g}_{mi}|^2 \} - |\mathbb{E} \{ \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{g}_{mk} \}|^2 \right) + \sum_{m=1}^M \sum_{i=1}^K (\mathbb{E} \{ |\sqrt{p_i} \mathbf{v}_{mk}^H \mathbf{h}_{mi}|^2 \} - |\mathbb{E} \{ \sqrt{p_k} \mathbf{v}_{mk}^H \mathbf{h}_{mk} \}|^2) + \sum_{m=1}^M \mathbb{E} \{ |\mathbf{w}'_k|^2 \}$.

Resembling Level 3, the closed-form expressions in (39) and (41) also can be derived when MR combining $\mathbf{v}_{mk} = \hat{\mathbf{h}}_{mk} + \hat{\mathbf{g}}_{mk}$ is used. However, due to its high similarity with the closed-form expressions in Level 3, we decided to omit it and only verify it through simulation results in Section V.

D. Level 1: Small-Cell Network

The last level is that the signals are only decoded by AP, which means there are no signals exchanged between APs and the CPU. Taking this into account, the RIS-CF-SR network becomes truly distributed, transforming into a small-cell network. As referenced in [9], the macro-diversity is achieved by choosing the best AP that provides the highest SE to a given device. To this end, the achievable SEs are given as follows.

Proposition 7. *At Level 1, an achievable SE of device k for direct link is*

$$\text{SE}_k^{(1),\text{dl}} = \left(1 - \frac{\tau_p}{\tau_c} \right) \max_{m \in \{1, \dots, M\}} \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_{mk}^{\text{dl}} \right) \right\}, \quad (42)$$

where the instantaneous effective SINR is given by (43) (see top of this page).

Then, the backscatter-link signals are designed to be decoded. Similar to (42), the best AP is selected to decode the backscatter-link signal. However, unlike the previous level, there are no signals exchanged between APs and the CPU. Therefore, if the AP that decodes the backscatter-link is different from the AP that previously decodes the direct-link signal, the AP cannot perform signal cancellation. Considering the large number of APs in the system, this is a high-probability occurrence. Thus, the backscatter-link SE without signal cancellation is shown as follows.

Proposition 8. *An achievable SE of device k at Level 1 for backscatter link is*

$$\text{SE}_k^{(1),\text{bl}} = \left(\frac{\tau_c - \tau_p}{L\tau_c} \right) \max_{m \in \{1, \dots, M\}} \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_{mk}^{\text{bl}} \right) \right\}, \quad (44)$$

where the instantaneous effective SINR is given by (45) (see top of the next page).

Remark 3. *Note that the main difference between RIS-CF-SR and BD-CF-SR lies in the backscatter link, and the considered signal processing also can be applied in BD-CF-SR. Thus, the achievable SEs for BD-CF-SR are presented by changing the backscatter link, e.g., turning \mathbf{g}_{mk} into $\mathbf{g}_{mk,\text{BD}}$.*

$$\text{SINR}_{mk}^{(1),\text{dl}} = \frac{p_k |\mathbf{v}_{mk}^H \hat{\mathbf{h}}_{mk} + \mathbf{v}_{mk}^H \hat{\mathbf{g}}_{mk}|^2}{\sum_{i \neq k}^K p_i |\mathbf{v}_{mk}^H \hat{\mathbf{h}}_{mi}|^2 + \sum_{i \neq k}^K p_i |\mathbf{v}_{mk}^H \hat{\mathbf{g}}_{mi}|^2 + \mathbf{v}_{mk}^H \left(\sum_{i=1}^K p_i (\tilde{\mathbf{R}}_{mi} + \tilde{\mathbf{Q}}_{mi}) + \sigma^2 \mathbf{I}_N \right) \mathbf{v}_{mk}}. \quad (43)$$

$$\text{SINR}_k^{(1),\text{bl}} = \frac{L p_k |\mathbf{v}_{mk}^H \hat{\mathbf{g}}_{mk}|^2}{\sum_{i \neq k}^K p_i |\mathbf{v}_{mk}^H \hat{\mathbf{g}}_{mi}|^2 + \sum_{i \neq k}^K p_i |\mathbf{v}_{mk}^H \hat{\mathbf{h}}_{mi}|^2 + \mathbf{v}_{mk}^H \left(\sum_{i=1}^K p_i (\tilde{\mathbf{R}}_{mi} + \tilde{\mathbf{Q}}_{mi}) + \sigma^2 \mathbf{I}_N \right) \mathbf{v}_{mk}}. \quad (45)$$

TABLE II: Simulation Parameters

Parameter	Value
Communication bandwidth	20 MHz
Noise power, σ^2	-94 dBm
Shadow fading with the standard deviation σ_{sh}	8 dB
Distance, D, d_0, d_1, d_2	1000 m, 50 m, 10 m, 10 m
Coherence time, τ_c	200 msec
Uplink training duration, τ_p	5 msec
Power reflection coefficient, α	1
Total number of antennas, $M \times N$	120
Number of RISs and BDs, J	10
Number of devices, K	10
Number of RIS elements, N_{RIS}	64
Transmission period, L	20
Uplink transmit power per device, p_k^{ce} and p_k	0.1 W

IV. NUMERICAL RESULTS

In this section, the comparison of the uplink performance of RIS-CF-SR is shown, with the different cooperation implementation levels and either MMSE/L-MMSE or MR combining scheme, and BD-CF-SR.

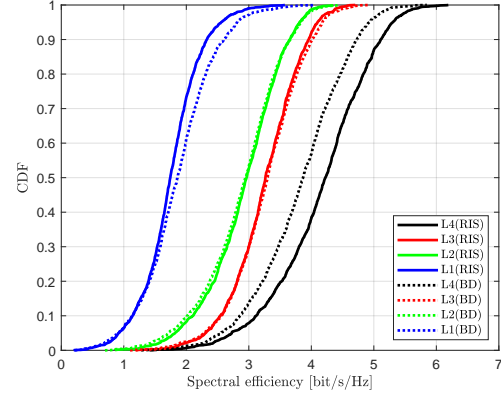
A. Simulation Setup and Propagation Model

We consider M APs with N antennas and K devices that are uniformly distributed in a $D \times D$ km² area with a wrap-around scheme. Furthermore, we assume each RIS is located at the area of a circle centered at its associated device with a radius of d_{jk} . To be fair, the BD-CF-SR follows the same setup. In addition, the phase of each RIS is random without phase modulation. Distance measurements are wraps around the edges of the simulation area to avoid boundary effects when simulating the CF-mMIMO network. The large-scale fading β_{mk} is modeled as

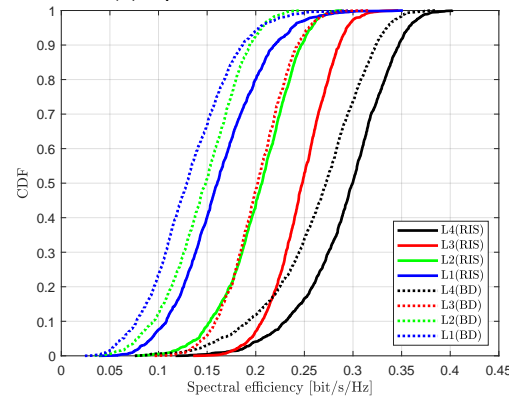
$$\beta_{mk} = \text{PL}_{mk} \cdot 10^{\frac{\sigma_{\text{sh}} z_{mk}}{10}}, \quad (46)$$

where PL_{mk} represents the path-loss, $z_{mk} \sim \mathcal{N}(0, 1)$, and $10^{\frac{\sigma_{\text{sh}} z_{mk}}{10}}$ is the shadow fading with the standard deviation $\sigma_{\text{sh}} = 8$ dB. Moreover, a three-sloped model is considered for the path-loss model. [5]. Let d_{tw} denotes the distance between the w -th vertex and the t -th vertex, where $t, w \in [m, j, k]$ for $m = 1 \dots M, j = 1 \dots J$, and $k = 1 \dots K$. In the case when $d_{tw} > d_0$, the Hata-COST231 propagation model is applied, where the path-loss exponent equals 3.5 for \mathbf{h}_{mk} , 2.8 for \mathbf{h}_{mj} and 2.0 for \mathbf{h}_{jk} , equals 2 if $d_0 < d_{tw} \leq d_1$, and equals 0 if $d_{tw} \leq d_1$ [33]. Note that when $d_{tw} \leq d_1$ there is no shadowing. Moreover, the large-scale coefficients of the channel \mathbf{h}_{jk} between the RIS j and device k is

$$\beta_{jk}^{\text{LoS}} = \frac{\kappa_{jk}}{\kappa_{jk} + 1} \beta_{jk}, \beta_{jk}^{\text{NLoS}} = \frac{1}{\kappa_{jk} + 1} \beta_{jk}, \quad (47)$$



(a) Uplink SEs of direct link.



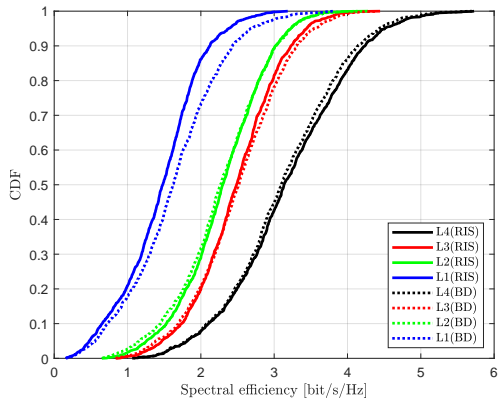
(b) Uplink SEs of backscatter link.

Fig. 3: Uplink SEs per device with MMSE/L-MMSE combining when $M = 120$ APs with $N = 1$ antenna.

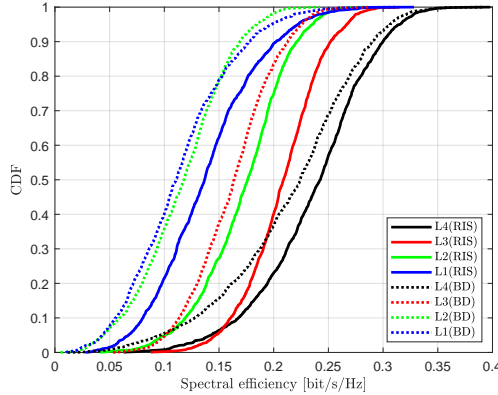
where $\kappa_{jk} = 10^{1.3-0.003d_{jk}}$ reflects the Rician κ -factor. Note that the n -th element of the deterministic LoS component $\bar{\mathbf{h}}_{j,k}$ can be written as $[\bar{\mathbf{h}}_{mk}]_n = \sqrt{\beta_{mk}^{\text{LoS}}} e^{j2\pi d_H(n-1) \sin(\theta_{mk})}$, where θ_{jk} is the angle of arrival to the device k seen from RIS j and d_H denotes the elements spacing parameter. All the simulation parameters are shown in Table II.

B. SE of Different Implementation Levels

Fig. 3 compares RIS-CF-SR and BD-CF-SR with MMSE/L-MMSE combining when $M = 120$ and $N = 1$. For the direct-link, at the 95% likely SE points, Level 4 provides the highest SEs, while Level 1 achieves the lowest SEs. Level 2 and Level 3 are in the middle of the two. Since RIS-CF-SR introduces a lot of interference, this actually reflects the



(a) Uplink SEs of direct link.



(b) Uplink SEs of backscatter link.

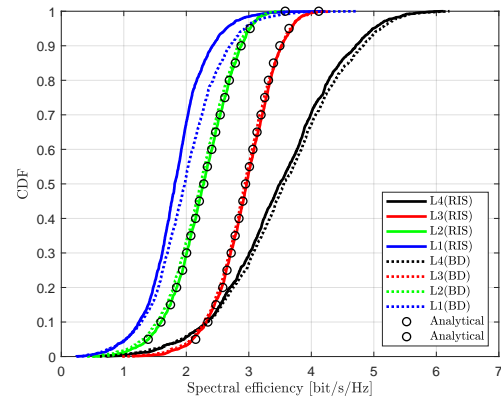
Fig. 4: Uplink SEs per device with MMSE/L-MMSE combining when $M = 60$ APs with $N = 2$ antennas.

ability to suppress the interference under different levels. In addition, It is observed that RIS-CF-SR is superior to BD-CF-SR only at Level 4. The reasons for this are as follows: RIS offers higher signal gains compared to BD, enhancing the signal strength. Nevertheless, this increased signal strength also leads to increased interference among devices. Hence, on the direct-link, RIS-CF-SR has the significant advantage only at Level 4, where the interference suppression is the strongest.

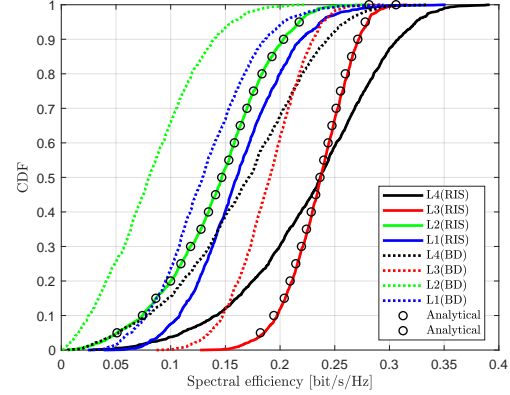
Fig. 3(b) shows the cumulative distribution function (CDF) of SEs for the backscatter link. Due to the proposed signal cancellation schemes, numerous interference has been eliminated. Thus, on the backscatter-link, RIS-CF-SR can outperform BD-CF-SR across varying levels at the 95% likely SE points.

Fig. 4 examines the same setups but with fewer APs that are equipped with multi-antennas: $M = 60$ and $N = 2$. The general trends are the same as in Fig. 3 but all the levels lose in SEs due to the reduced macro diversity. In particular, the advantages of RIS-CF-SR over BD-CF-SR at Level 4 are significantly smaller 4 on the direct-link owing to the weakened interference suppression capability. Nevertheless, the backscatter-link SE of RIS-CF-SR remains better than that of BD-CF-SR.

Fig. 5 draws the CDF of SEs with MR combining when $M = 120$ and $N = 1$. On the direct link, since MR combining cannot effectively suppress the interference, the poor combining scheme results in a large SE loss except for level 1. Due to the challenges in interference suppression, the



(a) Uplink SEs of direct link.



(b) Uplink SEs of backscatter link.

Fig. 5: Uplink SEs per device with MR combining when $M = 120$ APs with $N = 1$ antenna.

performance of RIS-CF-SR and BD-CF-SR is almost the same.

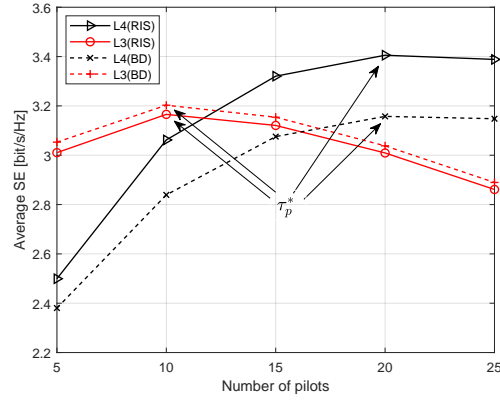
Taking into account the backscatter link, it is discovered that the backscatter-link SE loss is larger than the direct-link SE. Since the MR combining is not optimal, the impact of the poor combining scheme is more severe. It can be found that Level 3 outperforms other levels in this scenario by using the LSFD to suppress the interference. Furthermore, for levels 2 and 3, the “o” markers from the analytic results coincide with the simulated curves. This is a confirmation of the accuracy of the closed-form SE expressions inferred by us.

C. Effects of Uplink Training Duration

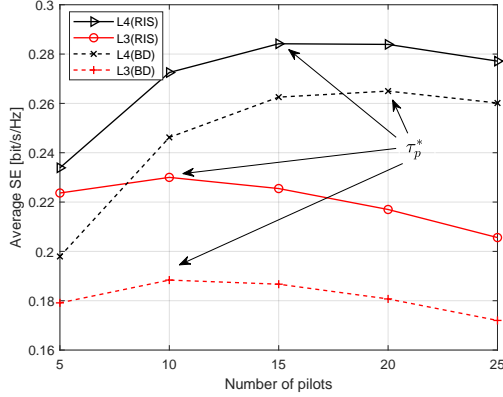
Fig. 6 depicts the average SE with MMSE/L-MMSE combining for Level 3 and Level 4. The parameters are set as: $M = 360$, $N = 1$, $K = 30$, and $\tau_c = 200$. In Fig. 6(a), we find that Level 4 relies more on the accuracy of channel estimation compared to Level 3, hence the optimal point τ_p^* is higher than that of Level 3. Furthermore, when the number pilots is very small (i.e., $\tau_p = 5$), Level 4 underperforms Level 3. The general trends of Fig. 6(b) are the same as in Fig. 6(a). The major difference is that Level 4 reaches the maximum SE value with τ_p^* earlier. The reason for this is that the suboptimal combining reduces the benefits of orthogonal pilots.

D. Effects of RIS Elements

Fig. 7 presents the average SE when varying the number of RIS elements. At level 4, the SE of the RIS-CF-SR is



(a) Average SE of direct link.



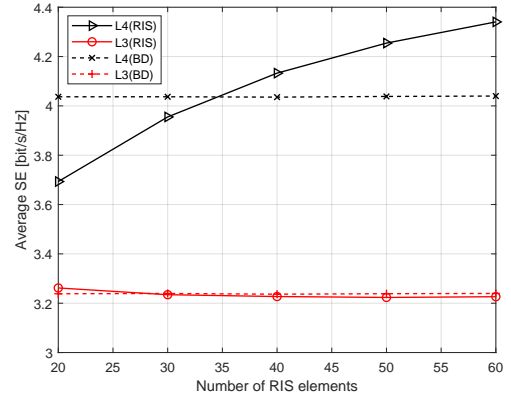
(b) Average SE of backscatter link.

Fig. 6: Average SE with MMSE/L-MMSE combining for Level 3 and Level 4 when $M = 360$ APs with $N = 1$ antenna and $K = 30$ devices.

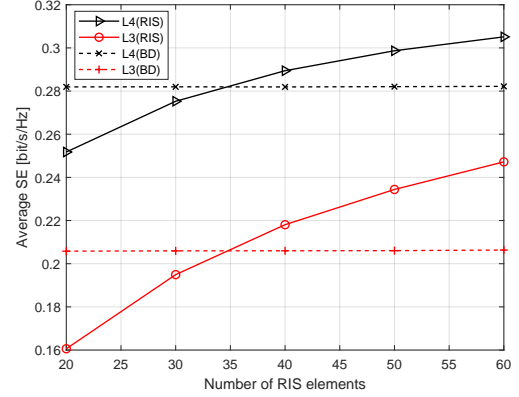
improved with the number of RIS elements on both the direct and backscatter links. However, due to insufficient ability to suppress the interference, increasing the number of RIS elements can lead to the direct link SE loss at level 3.

V. CONCLUSION

In this study, we have conducted an investigation into the uplink performance of a RIS-CF-SR system. We have considered four different implementations, ranging from fully centralized to fully distributed, and based on different CSI, we have designed distinct signal cancellation schemes to improve the SE of the backscatter link. Furthermore, we have derived novel closed-form SE expressions for Level 2 and Level 3 using the MR combining scheme. We compared the uplink performance of RIS-CF-SR and BD-CF-SR systems. Simulation results illustrated that RISs can significantly improve the SE of the backscatter link due to the large number of reflection elements, whereas additional appropriate signal processing schemes are required for the direct link. Notably, RIS-CF-SR is observed to outperform BD-CF-SR at Level 4 in SE on the direct link with the MMSE combining scheme. However, from Level 1 to Level 3, RIS-CF-SR has no significant advantage on the direct link. Moreover, increasing the number of RIS elements may lead to the direct-link SE loss without the appropriate signal processing scheme. Therefore, we recommend using fully centralized processing in RIS-CF-SR to suppress



(a) Average SE of direct link.



(b) Average SE of backscatter link.

Fig. 7: Average SE with MMSE/L-MMSE combining for Level 3 and Level 4 when $M = 120$ APs with $N = 1$ antenna and $J = 10$ RISs/BDs.

the additional interference caused by RIS, thereby improving SE.

APPENDIX

A. Proof of Corollary 3

1) Compute $\mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{dl}} \}$:

$$\begin{aligned} & \mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{dl}} \} \\ &= \sqrt{p_k} \mathbb{E} \left\{ \sum_{m=1}^M \hat{\mathbf{h}}_{mk}^H \hat{\mathbf{h}}_{mk} + \sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H (\hat{\mathbf{h}}_{mk} + \tilde{\mathbf{h}}_{mk}) \right\} \\ &= \sqrt{p_k} \sum_{m=1}^M \text{tr}(\hat{\mathbf{R}}_{mk}) + \sqrt{p_k} \mathbb{E} \left\{ \sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H \tilde{\mathbf{h}}_{mk} \right\}, \end{aligned} \quad (48)$$

and

$$\mathbb{E} \left\{ \sum_{m=1}^M \tilde{\mathbf{h}}_{mk}^H \hat{\mathbf{g}}_{mk} \right\} = \sqrt{p_k} \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{Q}}_{mk}}{\mathbf{Q}_{mk}} \tilde{\mathbf{R}}_{mk} \right). \quad (49)$$

Substitution of (49) into (48) yields

$$\mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{dl}} \} = \sqrt{p_k} \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{Q}}_{mk}}{\mathbf{Q}_{mk}} \tilde{\mathbf{R}}_{mk} \right) + \sqrt{p_k} \sum_{m=1}^M \text{tr}(\hat{\mathbf{R}}_{mk}), \quad (50)$$

Similarly, $\sum_{i=1}^K |\mathbb{E} \{ \sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}} \}|^2$ can be easily derived and therefore be omitted.

2) Compute $\mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{bl}} \}$: Similar to (48), $\mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{bl}} \}$ can be easily computed by

$$\mathbb{E} \{ \sqrt{p_k} \mathbf{f}_{kk}^{\text{bl}} \} = \mathbb{E} \left\{ \sum_{m=1}^M \sqrt{p_k} \hat{\mathbf{g}}_{mk}^H (\hat{\mathbf{g}}_{mk} + \tilde{\mathbf{g}}_{mk}) \right\} = \sqrt{p_k} \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{Q}}_{mk} \right). \quad (51)$$

3) Compute $\mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}}|^2 \}$:

$$\begin{aligned} \mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}}|^2 \} &= p_i \mathbb{E} \left\{ \left| \sum_{m=1}^M \mathbf{h}_{mi}^H (\hat{\mathbf{g}}_{mk} + \tilde{\mathbf{h}}_{mk}) \right|^2 \right\} \\ &\stackrel{(c)}{=} p_i \mathbb{E} \left\{ \left| \sum_{m=1}^M \mathbf{h}_{mi}^H \hat{\mathbf{g}}_{mk} \right|^2 \right\} + p_i \mathbb{E} \left\{ \left| \sum_{m=1}^M \mathbf{h}_{mi}^H \tilde{\mathbf{h}}_{mk} \right|^2 \right\} \\ &= p_i \mathbb{E} \left\{ \left| \sum_{m=1}^M \mathbf{h}_{mi}^H \hat{\mathbf{g}}_{mk} \right|^2 \right\} + p_i \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{R}}_{mk} \mathbf{R}_{mi} \right) \\ &\quad + p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{R}}_{mk}}{\mathbf{R}_{mk}} \mathbf{R}_{mi} \right) \right|^2, \end{aligned} \quad (52)$$

in which case the cross-point in (c) is very small and can therefore be omitted, and similar cases are treated in the same manner in the subsequent proofs. Then

$$\begin{aligned} &\mathbb{E} \left\{ \left| \sum_{m=1}^M \mathbf{h}_{mi}^H \hat{\mathbf{g}}_{mk} \right|^2 \right\} \\ &= \mathbb{E} \left\{ \left| \sqrt{\tau_p} \sum_{m=1}^M \sum_{k'=1}^K \sqrt{p_{k'}^{\text{ce}}} \mathbf{h}_{mi}^H \Psi_{mk} \tilde{\mathbf{h}}_{mk'} \varphi_{k'}^T \varphi_k^* \right|^2 \right\} \\ &\quad + \tau_p \sum_{m=1}^M \sum_{k'=1}^K \text{tr} \left(p_{k'}^{\text{ce}} \mathbf{R}_{mi} \Psi_{mk}^2 \mathbf{Q}_{mk'} |\varphi_{k'}^T \varphi_k^*|^2 \right) \\ &\quad + \sigma^2 \sum_{m=1}^M \text{tr} \left(\mathbf{R}_{mi} \Psi_{mk}^2 \right), \end{aligned} \quad (53)$$

and

$$\mathbb{E} \left\{ \left| \sqrt{\tau_p} \sum_{m=1}^M \sum_{k'=1}^K \sqrt{p_{k'}^{\text{ce}}} \mathbf{h}_{mi}^H \Psi_{mk} \tilde{\mathbf{h}}_{mk'} \varphi_{k'}^T \varphi_k^* \right|^2 \right\} \quad (54)$$

$$= p_i^{\text{ce}} \tau_p |\varphi_i^T \varphi_k^*|^2 \mathbb{E} \left\{ \left| \sum_{m=1}^M \mathbf{h}_{mi}^H \Psi_{mk} \tilde{\mathbf{h}}_{mi} \right|^2 \right\} \quad (55)$$

$$+ \tau_p \mathbb{E} \left\{ \left| \sum_{m=1}^M \sum_{k' \neq i}^K \sqrt{p_{k'}^{\text{ce}}} \mathbf{h}_{mi}^H \Psi_{mk} \tilde{\mathbf{h}}_{mk'} \varphi_{k'}^T \varphi_k^* \right|^2 \right\}, \quad (56)$$

where

$$\begin{aligned} &\mathbb{E} \left\{ \left| \sum_{m=1}^M \left(\mathbf{h}_{mi}^H \Psi_{mk} \tilde{\mathbf{h}}_{mi} \right) \right|^2 \right\} \\ &= \sum_{m=1}^M \text{tr} \left(\mathbf{R}_{mk} \Psi_{mk}^2 \tilde{\mathbf{R}}_{mi} \right) + 2 \sum_{m=1}^M \text{tr} \left(\Psi_{mk}^2 \tilde{\mathbf{R}}_{mi}^2 \right) \end{aligned}$$

$$+ \sum_{m=1}^M \sum_{n \neq m}^M \text{tr} \left(\tilde{\mathbf{R}}_{mi} \Psi_{mk} \Psi_{nk} \tilde{\mathbf{R}}_{ni} \right). \quad (57)$$

Similarly, we have

$$\begin{aligned} &\mathbb{E} \left\{ \left| \sum_{m=1}^M \sum_{k' \neq i}^K \sqrt{p_{k'}^{\text{ce}}} \mathbf{h}_{mi}^H \Psi_{mk} \tilde{\mathbf{h}}_{mk'} \varphi_{k'}^T \varphi_k^* \right|^2 \right\} \\ &= \sum_{m=1}^M \sum_{k' \neq i}^K \text{tr} \left(p_{k'}^{\text{ce}} \mathbf{R}_{mi} \Psi_{mk}^2 \tilde{\mathbf{R}}_{mk'} |\varphi_{k'}^T \varphi_k^*|^2 \right). \end{aligned} \quad (58)$$

Substitution of (53), (54), (57) and (58) into (52) yields

$$\begin{aligned} &\mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{dl}}|^2 \} \\ &= p_i \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{R}}_{mk} \mathbf{R}_{mi} \right) + p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{R}}_{mk}}{\mathbf{R}_{mk}} \mathbf{R}_{mi} \right) \right|^2 \\ &\quad + p_i \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{Q}}_{mk} \tilde{\mathbf{R}}_{mi} \right) + p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{Q}}_{mk}}{\mathbf{Q}_{mk}} \tilde{\mathbf{R}}_{mi} \right) \right|^2, \end{aligned} \quad (59)$$

4) Compute $\mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{bl}}|^2 \}$:

$$\begin{aligned} &\mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{bl}}|^2 \} \\ &= \mathbb{E} \left\{ \left| \sum_{m=1}^M \sqrt{p_i} \mathbf{g}_{mi}^H \hat{\mathbf{g}}_{mk} \right|^2 \right\} + p_i \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{R}}_{mk} \mathbf{Q}_{mi} \right). \end{aligned} \quad (60)$$

Follow the similar steps as in (53), we obtain

$$\begin{aligned} &\mathbb{E} \left\{ \left| \sum_{m=1}^M \sqrt{p_i} \mathbf{g}_{mi}^H \hat{\mathbf{g}}_{mk} \right|^2 \right\} \\ &= p_i \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{Q}}_{mk} \mathbf{Q}_{mi} \right) + p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{Q}}_{mk}}{\mathbf{Q}_{mk}} \mathbf{Q}_{mi} \right) \right|^2. \end{aligned} \quad (61)$$

Based on the (60) and (61), $\mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{bl}}|^2 \}$ can be computed as follows.

$$\begin{aligned} &\mathbb{E} \{ |\sqrt{p_i} \mathbf{f}_{ki}^{\text{bl}}|^2 \} \\ &= p_i \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{R}}_{mk} \mathbf{Q}_{mi} \right) + p_i \sum_{m=1}^M \text{tr} \left(\hat{\mathbf{Q}}_{mk} \mathbf{Q}_{mi} \right) \\ &\quad + p_i |\varphi_k^H \varphi_i|^2 \left| \sum_{m=1}^M \text{tr} \left(\frac{\hat{\mathbf{Q}}_{mk}}{\mathbf{Q}_{mk}} \mathbf{Q}_{mi} \right) \right|^2. \end{aligned} \quad (62)$$

B. Proof of Proposition 2

The interference term of (13) is

$$\begin{aligned} \mathbf{z} &= \sqrt{p_k} \mathbf{v}_k^H \left(\tilde{\mathbf{h}}_k s_k(l) + \tilde{\mathbf{g}}_k s_k(l) c_k \right) + \sum_{i \neq k}^K \sqrt{p_i} \mathbf{v}_k^H \mathbf{h}_i s_i(l) \\ &\quad - \sum_{i=0}^{k-1} \sqrt{p_i} \mathbf{v}_k^H \tilde{\mathbf{h}}_i s_i(l) + \sum_{i \neq k}^K \sqrt{p_i} \mathbf{v}_k^H \mathbf{g}_i s_i(l) c_i + \mathbf{w}. \end{aligned} \quad (63)$$

Thus, the conditional variance of the zero-mean interfering signal is

$$\begin{aligned}
 p(h, g, z) &= \mathbb{E} \left\{ |z|^2 \mid \{\hat{\mathbf{h}}_k, \hat{\mathbf{g}}_k\} \right\} \\
 &\stackrel{(a)}{=} \mathbb{E} \left\{ |s_k|^2 \right\} \mathbb{E} \left\{ |p_k \mathbf{v}_k^H \tilde{\mathbf{h}}_k|^2 \mid \{\hat{\mathbf{h}}_k, \hat{\mathbf{g}}_k\} \right\} \\
 &\quad + \mathbb{E} \left\{ |s_k|^2 \right\} \mathbb{E} \left\{ |c_k|^2 \right\} \mathbb{E} \left\{ |p_k \mathbf{v}_k^H \tilde{\mathbf{g}}_k|^2 \mid \{\hat{\mathbf{h}}_k, \hat{\mathbf{g}}_k\} \right\} \\
 &\quad + \sum_{i \neq k}^K \mathbb{E} \left\{ |s_i|^2 \right\} \mathbb{E} \left\{ |p_i \mathbf{v}_k^H \hat{\mathbf{h}}_i|^2 \mid \{\hat{\mathbf{h}}_i, \hat{\mathbf{g}}_i\} \right\} \\
 &\quad - \sum_{i=0}^{k-1} \mathbb{E} \left\{ |s_i|^2 \right\} \mathbb{E} \left\{ |p_i \mathbf{v}_k^H \hat{\mathbf{h}}_i|^2 \mid \{\hat{\mathbf{h}}_i, \hat{\mathbf{g}}_i\} \right\} \\
 &\quad + \sum_{i \neq k}^K \mathbb{E} \left\{ |s_i|^2 \right\} \mathbb{E} \left\{ |c_i|^2 \right\} \mathbb{E} \left\{ |p_i \mathbf{v}_k^H \hat{\mathbf{g}}_i|^2 \mid \{\hat{\mathbf{h}}_i, \hat{\mathbf{g}}_i\} \right\} \\
 &\quad + \mathbb{E} \left\{ |\mathbf{w}|^2 \mid \{\hat{\mathbf{h}}_i, \hat{\mathbf{g}}_i\} \right\} \\
 &= \sum_{i=k+1}^K p_i |\mathbf{v}_k^H \hat{\mathbf{h}}_i|^2 + \sum_{i \neq k}^K p_i |\mathbf{v}_k^H \hat{\mathbf{g}}_i|^2 \\
 &\quad + \mathbf{v}_k^H \left(\sum_{i=1}^K p_i (\mathbf{R}_i + \mathbf{Q}_i) + \sigma^2 \mathbf{I}_{MN} \right) \mathbf{v}_k,
 \end{aligned} \tag{64}$$

where (a) follows from the independence among each of the zero-mean signals s_i and c_i , as well as the independence between signals and channels (since the estimation error is independent of the channel estimate). Similarly, the desired term also can be calculated as shown in (15).

REFERENCES

- [1] D. C. Nguyen et al., "6G Internet of Things: A comprehensive survey," *IEEE Internet Things J.*, vol. 9, no. 1, pp. 359-383, Jan. 1, 2022.
- [2] X. Li, D. Li, J. Wan, C. Liu, and M. Imran, "Adaptive transmission optimization in SDN-based industrial Internet of Things with edge computing," *IEEE Internet Things J.*, vol. 5, no. 3, pp. 1351-1360, Jun. 2018.
- [3] L. Chettri and R. Bera, "A comprehensive survey on Internet of Things (IoT) toward 5G wireless systems," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 16-32, Jan. 2020.
- [4] A.-S. Bana et al., "Massive MIMO for Internet of Things (IoT) connectivity," *Phys. Commun.*, vol. 37, no. 100859, pp. 1-17, Sep. 2019.
- [5] 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.
- [6] S. Rao, A. Ashikhmin, and H. Yang, "Internet of Things based on cell-free massive MIMO," *Proc. 53rd Asilomar Conf. Signals Syst. Comput.*, Pacific Grove, CA, USA, 2019, pp. 1946-1950.
- [7] T. Zhao, X. Chen, Q. Sun, J. Zhang, "Energy-efficient federated learning over cell-free IoT networks: Modeling and optimization," *IEEE Internet Things J.*, early access, 2023, DOI: 10.1109/JIOT.2023.3273619.
- [8] S. Chen, J. Zhang, E. Björnson, J. Zhang, and B. Ai, "Structured massive access for scalable cell-free massive MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 4, pp. 1086-1100, Apr. 2021.
- [9] 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. 2019.
- [10] J. Zhang, J. Zhang, E. Björnson, and B. Ai, "Local partial zero-forcing combining for cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8459-8473, Dec. 2021.
- [11] Z. Wang, J. Zhang, B. Ai, C. Yuen, and M. Debbah, "Uplink performance of cell-free massive MIMO with multi-antenna users over jointly correlated Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 21, no. 9, pp. 7391-7406, Sep. 2022.
- [12] J. Zhang, E. Björnson, M. Matthaiou, D. W. K. Ng, H. Yang, and D. J. Love, "Prospective multiple antenna technologies for beyond 5G," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1637-1660, Aug. 2020.
- [13] J. Zhang, J. Zhang, D. W. K. Ng, S. Jin, and B. Ai, "Improving sum-rate of cell-free massive MIMO with expanded compute-and-forward," *IEEE Trans. Signal Process.*, vol. 70, no. 12, pp. 202-215, Dec. 2021.
- [14] M. Guo and M. C. Gursoy, "Joint activity detection and channel estimation in cell-free massive MIMO networks with massive connectivity," *IEEE Trans. Commun.*, vol. 70, no. 1, pp. 317-331, Jan. 2022.
- [15] J. Zhang, J. Fan, J. Zhang, D. W. K. Ng, Q. Sun, and B. Ai, "Performance analysis and optimization of NOMA-based cell-free massive MIMO for IoT," *IEEE Internet Things J.*, vol. 9, no. 12, pp. 9625-9639, 2022.
- [16] *Identification and Quantification of Key Socio-Economic Data to Support Strategic Planning for the Introduction of 5G in Europe*, Europe Union, Brussels, Belgium, 2016.
- [17] R. Long, Y.-C. Liang, H. Guo, G. Yang, and R. Zhang, "Symbiotic radio: A new communication paradigm for passive Internet of Things," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1350-1363, Feb. 2020.
- [18] J. Kimionis, A. Bletsas, and J. N. Sahalos, "Increased range bistatic scatter radio," *IEEE Trans. Commun.*, vol. 62, no. 3, pp. 1091-1104, Mar. 2014.
- [19] C. Zhang, H. Zhou and, Y. -C. Liang, "Interference-free MU-MISO symbiotic radios via RIS partitioning design" in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Kuala Lumpur, Malaysia, 2023, pp. 3264-3269.
- [20] Y.-C. Liang, Q. Zhang, E. G. Larsson, and G. Y. Li, "Symbiotic radio: Cognitive backscattering communications for future wireless networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 6, no. 4, pp. 1242-1255, Dec. 2020.
- [21] Z. Dai, R. Li, J. Xu, Y. Zeng, and S. Jin, "Rate-Region characterization and channel estimation for cell-free symbiotic radio communications," *IEEE Trans. Commun.*, vol. 71, no. 2, pp. 674-687, Feb. 2023.
- [22] M. Ataeshojai, R. C. Elliott, W. A. Krzymień, C. Tellambura, and I. Maljević, "Symbiotic backscatter communication underlying a cell-free massive MIMO system," *IEEE Internet Things J.*, vol. 10, no. 19, pp. 16758-16777, 1 Oct. 1, 2023.
- [23] F. Li, Q. Sun, X. Chen, and J. Zhang, "Spectral efficiency analysis of uplink cell-free massive MIMO symbiotic radio," *IEEE Internet Things J.*, vol. 11, no. 2, pp. 3614-3627, 15 Jan. 15, 2024.
- [24] J. Hu, Y.-C. Liang, and Y. Pei, "Reconfigurable intelligent surface enhanced multi-user MISO symbiotic radio system," *IEEE Trans. Commun.*, vol. 69, no. 4, pp. 2359-2371, 2021.
- [25] M. Di Renzo et al., "Reconfigurable Intelligent Surfaces vs. Relaying: Differences, Similarities, and Performance Comparison," *IEEE Open J. Commun. Soc.*, vol. 1, pp. 798-807, 2020.
- [26] J. Ye, A. Kammoun and M. -S. Alouini, "Spatially-Distributed RISs vs Relay-Assisted Systems: A Fair Comparison," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 799-817, 2021.
- [27] Q. Zhang, Y.-C. Liang, and H. V. Poor, "Reconfigurable intelligent surface assisted MIMO symbiotic radio networks," *IEEE Trans. Commun.*, vol. 69, no. 7, pp. 4832-4846, 2021.
- [28] M. Hua, Q. Wu, L. Yang, R. Schober and H. V. Poor, "A novel wireless communication paradigm for intelligent reflecting surface based symbiotic radio systems," *IEEE Trans. Signal Proc.*, vol. 70, pp. 550-565, 2022.
- [29] M. Hua, L. Yang, Q. Wu, C. Pan, C. Li and A. L. Swindlehurst, "UAV-assisted intelligent reflecting surface symbiotic radio system," *IEEE Wireless Commun.*, vol. 20, no. 9, pp. 5769-5785, Sept. 2021.
- [30] J. Hu, Y.-C. Liang, Y. Pei, S. Sun, and R. Liu, "Reconfigurable intelligent surface based uplink MU-MIMO symbiotic radio system," *IEEE Trans. Wireless Commun.*, vol. 22, no. 1, pp. 423-438, Aug. 2022.
- [31] X. Peng, Q. Tao, X. Gan, and C. Zhong, "Intelligent reflecting surface enhanced cell-free symbiotic radio systems," *IEEE Internet Things J.*, early access, 2023, DOI: 10.1109/JIOT.2023.3267658.
- [32] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," in *Foundations and Trends in Signal Processing*, vol. 11, nos. 3-4. 2017, pp. 154-655.
- [33] B. Al-Nahhas, M. Obeed, A. Chaaban and M. J. Hossain, "RIS-aided cell-free massive MIMO: performance analysis and competitiveness," in *Proc. IEEE Int. Conf. Commun. Workshops (ICC Workshops)*, 2021, pp. 1-6.



Feiyang Li received the B.S. degree from the College of Information Science and Technology, Nantong University, Nantong, China, in 2021. He is currently pursuing the Ph.D. degree with Nantong University. His research interests include massive MIMO systems, backscatter communication, and performance analysis of wireless communication systems.



Qiang Sun (Member, IEEE) received the Ph.D. degree in communications and information systems from Southeast University, Nanjing, China, in 2014. He was a Visiting Scholar with the University of Delaware, Newark, DE, USA, in 2016. He is currently a Professor with the School of Information Science and Technology, Nantong, China. His research interests include deep learning and wireless communications. He was a member of Technical Program Committee and a reviewer for a number of IEEE conferences/journals.



Xiaomin Chen (Member, IEEE) received the B.S. and the M.S. degrees in communications engineering from Tongji University, Shanghai, China, in 2005 and 2008, respectively, and the Ph.D. degree (Hons.) in electronic engineering from Technische Universität Braunschweig, Germany, in 2014. She is currently with the School of Information Science and Technology, Nantong University, Nantong, China. Her research interests include error-control coding, wireless networks, and machine learning in communication systems.



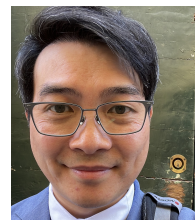
Bile Peng (Member, IEEE) received the B.S. degree from Tongji University, Shanghai, China, in 2009, the M.S. degree from the Technische Universität Braunschweig, Braunschweig, Germany, in 2012, and the Ph.D. degree with distinction from the Institut für Nachrichtentechnik, Technische Universität Braunschweig in 2018. In 2019, he was a Marie Skłodowska-Curie Research Fellow with the Chalmers University of Technology, Gothenburg, Sweden. He is currently a Researcher with the Institut für Nachrichtentechnik, Technische Universität Braunschweig. His research interests include wireless channel measurement, modeling and estimation, massive MIMO signal processing, Bayesian inference, and machine learning algorithms, in particular deep reinforcement learning, for wireless communication and localization.

Dr. Peng is a major contributor to the IEEE Standard for High Data Rate Wireless Multi-Media Networks Amendment 2: 100 Gb/s Wireless Switched Point-to-Point Physical Layer (IEEE Std 802.15.3d-2017) and was the recipient of the IEEE Vehicular Technology Society 2019 Neal Shepherd Memorial Best Propagation Paper Award.



Jiayi Zhang (Senior Member, IEEE) received the B.Sc. and Ph.D. degree of Communication Engineering from Beijing Jiaotong University, China in 2007 and 2014, respectively.

Since 2016, he has been a Professor with School of Electronic and Information Engineering, Beijing Jiaotong University, China. From 2014 to 2016, he was a Postdoctoral Research Associate with the Department of Electronic Engineering, Tsinghua University, China. From 2014 to 2015, he was also a Humboldt Research Fellow in Institute for Digital Communications, Friedrich-Alexander-University Erlangen-Nürnberg (FAU), Germany. From 2012 to 2013, he was a visiting scholar at the Wireless Group, University of Southampton, United Kingdom. His current research interests include cell-free massive MIMO, reconfigurable intelligent surface (RIS), communication theory and applied mathematics. Dr. Zhang received the Best Paper Awards at the WCSP 2017 and IEEE APCC 2017, the URSI Young Scientist Award in 2020, and the IEEE ComSoc Asia-Pacific Outstanding Young Researcher Award in 2020. He was recognized as an exemplary reviewer of the IEEE COMMUNICATIONS LETTERS in 2015-2017. He was also recognized as an exemplary reviewer of the IEEE TRANSACTIONS ON COMMUNICATIONS in 2017-2019. He was the Lead Guest Editor of the special issue on "Multiple Antenna Technologies for Beyond 5G" of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS and an Editor for IEEE COMMUNICATIONS LETTERS from 2016-2021. He currently serves as an Associate Editor for IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE ACCESS and IET COMMUNICATIONS.



Kai-Kit Wong (Fellow, IEEE) (M'01-SM'08-F'16) received the BEng, the MPhil, and the PhD degrees, all in Electrical and Electronic Engineering, from the Hong Kong University of Science and Technology, Hong Kong, in 1996, 1998, and 2001, respectively. After graduation, he took up academic and research positions at the University of Hong Kong, Lucent Technologies, Bell-Labs, Holmdel, the Smart Antennas Research Group of Stanford University, and the University of Hull, UK. He is Chair in Wireless Communications at the Department of Electronic and Electrical Engineering, University College London, UK. His current research centers around 6G and beyond mobile communications. He is Fellow of IEEE and IET. He served as the Editor-in-Chief for IEEE Wireless Communications Letters between 2020 and 2023.

Dr. Wong is a major contributor to the IEEE Standard for High Data Rate Wireless Multi-Media Networks Amendment 2: 100 Gb/s Wireless Switched Point-to-Point Physical Layer (IEEE Std 802.15.3d-2017) and was the recipient of the IEEE Vehicular Technology Society 2019 Neal Shepherd Memorial Best Propagation Paper Award.