Cell-free massive MIMO with multiple active eavesdroppers

Atiya, Y. S., Mobini, Z., Ngo, H.-Q., & Matthaiou, M. (2025). Cell-free massive MIMO with multiple active eavesdroppers. *IEEE Open Journal of the Communications Society*. Advance online publication. https://doi.org/10.1109/OJCOMS.2025.3534640

Published in:
IEEE Open Journal of the Communications Society

Document Version:
Publisher's PDF, also known as Version of record

Received XX Month, XXXX; revised XX Month, XXXX; accepted XX Month, XXXX; Date of publication XX Month, XXXX; date of current version 11 January, 2024.

Digital Object Identifier 10.1109/OJCOMS.2024.011100

Cell-Free Massive MIMO with Multiple Active Eavesdroppers

Yasseen Sadoon Atiya, Zahra Mobini, *Member, IEEE,* Hien Quoc Ngo, *Fellow, IEEE,* and Michail Matthaiou, *Fellow, IEEE*

CORRESPONDING AUTHOR: Yasseen Sadoon Atiya (e-mail: yhimiari01@qub.ac.uk).

The authors are with the Centre for Wireless Innovation (CWI), Queen's University Belfast, BT3 9DT Belfast, U.K. email:{yhimiari01, zahra.mobini, hien.ngo, m.matthaiou}@qub.ac.uk. Yasseen Sadoon Atiya is also a lecturer at Imam Alkadhim University College, Iraq. H. Q. Ngo and M. Matthaiou are also affiliated with the Department of Electronic Engineering, Kyung Hee University, Yongin-si, Gyeonggi-do 17104, South Korea. This work is a contribution by Project REASON, a UK Government funded project under the Future Open Networks Research Challenge (FONRC) sponsored by the Department of Science Innovation and Technology (DSIT). The work of Z. Mobini and H. Q. Ngo was supported by the U.K. Research and Innovation Future Leaders Fellowships under Grant MR/X010635/1, and a research grant from the Department for the Economy Northern Ireland under the US-Ireland R&D Partnership Programme. The work of M. Matthaiou has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 101001331). Parts of this paper were presented at the 2024 IEEE WCNC conference [1].

ABSTRACT This paper investigates the secrecy performance of cell-free massive MIMO (CF-mMIMO) systems in the presence of active spoofing attacks by multiple eavesdroppers (Eves). Each Eve conducts a spoofing attack on a different legitimate user during the uplink training phase, aiming to intercept the information intended for that user during the downlink transmission phase. To counter these attacks, we propose a joint access point (AP) selection and power optimization strategy to enhance the security performance of the CF-mMIMO system. Specifically, we formulate an optimization problem that seeks to maximize the sum-spectral efficiency (SE) of the legitimate users while ensuring a positive secrecy spectral efficiency (SSE) for all the attacked users. A sub-optimal solution to this mixed-integer non-convex problem is obtained using an efficient, low-complexity accelerated projected gradient (APG)-based algorithm. Moreover, for system design purposes, we introduce two simple and efficient methods: *i)* detecting the presence of multiple active Eves within the system and identifying which users are under attack, and *ii)* estimating the large-scale fading coefficients between the APs and the Eves. Our findings show that the proposed approach achieves a median sum-SE performance that is 62% better than that of equal power allocation without AP selection scheme. Furthermore, the results demonstrate that the proposed strategy significantly improves the sum-SE while ensuring a positive secrecy rate, thereby safeguarding the confidentiality of all transmitted information signals to all users, even in the presence of a relatively large number of Eves.

INDEX TERMS Accelerated projected gradient, access point selection, cell-free massive multiple-input multiple-output, multiple active Eves, physical layer security, power control, sum spectral efficiency.

I. INTRODUCTION

CELL-FREE massive MIMO (CF-mMIMO) has recently attracted significant research attention for the next generations of wireless networks. In a CF-mMIMO network, a large number of access points (APs) is deployed, each equipped with multiple antennas to serve a much smaller number of users at the same time/frequency resources [2]–[4]. The APs cooperate by connecting to central processing units (CPUs) via fronthaul links. CF-mMIMO networks leverage this architecture to provide high spectral efficiency (SE), reliability, and energy efficiency, all of which are achieved with relatively simple signal processing. Furthermore, CF-mMIMO leverages high macro diversity provided by distributed antennas to ensure high connectivity for all users in the networks [5], [6].

The architecture of a CF-mMIMO network introduces additional challenges in ensuring secure communication against eavesdropping attacks due to the short distances between the APs and users [7]. An eavesdropper, referred to as Eve, aims to intercept the legitimate information intended for the

legitimate users and can operate either passively or actively. In passive eavesdropping, Eve can only access the information intended for a legitimate user [8]. Conversely, in active eavesdropping, Eve can interfere with the communication channel between the APs and the targeted user by sending the same pilot signal as the attacked user or by transmitting a jamming signal [9]. Numerous studies have investigated the physical-layer security aspects of massive MIMO, proposing various strategies to mitigate these vulnerabilities and enhance the overall communication security. More specifically, several works on the detection of active pilot spoofing attacks have been proposed in [10]–[13] among others. Moreover, enhancing the secrecy performance of massive MIMO using a multiple antenna cooperative jammer was proposed in [14]. Instead of jamming signals, the authors in [15]–[18] utilized artificial noise signals to mitigate the effects of eavesdropping. Other scenarios were considered in [19]–[23] to reduce the rate of eavesdropping by employing resource allocation strategies and beamforming designs.

To date, there have been a few studies addressing the physical-layer security aspect of CF-mMIMO [1], [24]–[26], [28]–[36]. In [29], the authors investigated the secrecy performance of CF-mMIMO networks under the attack of a single-antenna Eve by deriving the secrecy spectral efficiency (SSE). They also provided a comparative analysis with co-located massive MIMO systems, highlighting the differences in SSE between the two architectures. Then, the authors in [24] considered the same setup as in [29] and proposed a method to detect the presence of an active Eve within the network. Also, they conceived power control schemes using the successive convex approximation (SCA)-based optimization approach to maximize the SSE of the user under attack. In [25], the downlink performance of CF-mMIMO was examined under pilot spoofing attacks, and a novel downlink transmission protocol was proposed to mitigate these threats. In addition, Alageli *et al.* in [30] presented an optimization problem aimed at enhancing the non-linear power control during the downlink phase of simultaneous wireless information and power transfer (SWIPT) in CF-mMIMO systems, considering the presence of active eavesdropping. The authors in [31] examined the impact of hardware impairments on the secrecy performance of a CF-mMIMO system, while the effect of low-resolution analog-to-digital converters and radio frequency (RF) impairments on the secrecy performance of a CF-mMIMO system experiencing active eavesdropping was investigated in [32]. Park *et al.* in [33] proposed a power control scheme aided by artificial noise for CF-mMIMO experiencing the attack of passive Eves. The proposed scheme aims at enhancing the SSE of the system. The problem of collusive eavesdropping was investigated in [26]. In [34], the physical-layer security in an intelligent reflecting surface-unmanned aerial vehicle (IRS-UAV) CF-mMIMO, in the presence of passive Eves, was investigated, while the authors in [35] studied the pilot assignment and power control issues for secure UAV communications in a CF-mMIMO architecture, where the pilot spoofing attack is performed by one malicious UAV.

However, it is important to point out that most studies focus on simplified scenarios when considering Eve's attacks. Specifically, the above literature [1], [24], [25], [29]–[32], [34], [35] assumes a single Eve, which is somewhat idealistic. In reality, practical CF-mMIMO systems are likely to face multiple Eves simultaneously targeting different users. Addressing the challenges posed by multiple Eves is crucial for improving the security resilience of CF-mMIMO networks in real-world environments. In addition, a common assumption in the existing literature is that all APs transmit to all users within their coverage area. In reality, some APs may be situated far from legitimate users and, hence, do not contribute much on the performance of these users. Therefore, an efficient selection of APs to serve users during an active spoofing attack is crucial in practice. In this space, our recent work in [28] proposed an AP selection approach to improve the SSE and showed that the AP selection and power optimization approaches provide significant SE gains. However, [28] considered a single-Eve scenario and a simple heuristic AP selection scheme. Moreover, there were no joint AP selection and power optimization designs to improve the secrecy performance.

Motivated by the aforementioned considerations, this paper exploits a CF-mMIMO system facing multiple-Eves' attacks. We propose a novel approach that combines AP selection and power optimization using the accelerated proximal gradient (APG) method to maximize the sum-SE while ensuring positive SSE for all attacked users. It is worth mentioning that simultaneously optimizing the power allocation and selecting an optimal subset of APs for serving each user in CF-mMIMO networks is essential for enhancing the energy efficiency and improving the SE. Equally importantly, utilizing the APG algorithm offers significantly higher computational efficiency, particularly for large-scale CF-mMIMO systems. This enhanced efficiency makes it more suitable for practical architectures [37]. In particular, compared to SCA-based methods used in [24] and [28] and machine learning techniques, the APG method provides a scalable and efficient solution for optimization. Each iteration of the APG algorithm is computationally lightweight and memory-efficient, leveraging closed-form expressions and gradient computations that can be executed quickly. These features make APG particularly suitable for large-scale CF-mMIMO systems, offering effective optimization with low computational complexity while maintaining robust system performance.

We also note that this work builds upon our previous study as presented in [1]. In [1], we focused on minimizing the SE at Eve while meeting specific SE requirements for the legitimate users, considering a simplified scenario with a single Eve and without exploring the SSE condition for all attacked users. In contrast, our current study not only extends the scope to multiple-Eves scenarios but also introduces an

TABLE 1. Comparison of Our Contributions to Existing Literature on Secure CF-mMIMO

Feature	[1]	[24]	[25]	[26]	[27]	[28]	**our work**
Multiple-antenna APs	✓			✓	✓	✓	✓
Multiple Eves-Multiple attacked users							✓
Single Eve's detection		✓				✓	✓
Multiple Eves' detection							✓
AP selection	✓		✓	✓		✓	✓
Joint power optimization and AP selection	✓						✓

efficient scheme for detecting the presence of multiple Eves and identifying which users are under attack. Table I clearly and explicitly highlights our contributions and benchmarks them against the state-of-the-art. We further detail the main contributions of this work in a point-by-point format as follows:

- We develop a framework for a CF-mMIMO system operating under active spoofing attacks orchestrated by multiple Eves targeting various users during the uplink training phase.
- We propose a joint AP selection and power optimization approach to enhance the security of the CF-mMIMO system. Our approach involves formulating a mixed-integer non-convex optimization problem aimed at maximizing the sum-SE of legitimate users while ensuring a positive SSE for all attacked users. The challenging formulated problem is transformed into a tractable form and an efficient algorithm is proposed to solve it using an APG-based scheme.
- We present an effective and simple method for detecting the presence of multiple active Eves in the system and identifying the specific users under attack. This approach can be implemented distributively at each AP and is based on the sample average power of the received pilot signals.
- Numerical results show that our proposed optimization approach significantly enhances the sum-SE and ensures a positive secrecy rate, effectively safeguarding the confidentiality of transmitted signals, even with a high number of Eves.

Note that while our analysis assumes orthogonal pilot sequences, it can also be extended to scenarios involving non-orthogonal pilot sequences. Addressing the non-orthogonal pilot scenario requires modifications to the SE equations and derivations to accurately reflect the effects of pilot non-orthogonality, although the joint power allocation and AP selection optimization approach remains unchanged. Furthermore, the Eves' detection method must be redesigned in the non-orthogonal pilot scenario to account for the interference arising not only from the Eves but also from legitimate users sharing the same pilot as the targeted user.

The remainder of the paper is as follows: In Section II, we present a system model for CF-mMIMO in the presence of multiple Eves and analyze the achievable SE of the users and Eves. Section III introduces a joint AP selection and power optimization problem formulation, along with an APG-based approach to solve it. In Section IV, we propose a method for detecting multiple active Eves in the system and identifying the specific users under attack. Moreover, extensive simulations are conducted to demonstrate the effectiveness of the proposed optimization problem and Eve's detection scheme. Finally, Section VI concludes the paper.

Notation: Matrices are denoted by bold upper case letters, while bold lower case letters indicate vectors. The superscript $(\cdot)^{\mathrm{H}}$ stands for the conjugate-transpose (Hermitian); we use $\mathbb{C}^{L\times N}$ to denote a $L\times N$ matrix; an $M\times M$ identity matrix is represented by $\mathbf{I}_M$; $(\cdot)$ refers to the trace operation. The statistical expectation is denoted by $\mathbb{E}\{\cdot\}$. Finally, a zero mean circular symmetric complex Gaussian distribution with variance σ^2 is denoted by $\mathcal{CN}(0,\sigma^2)$.

II. System Model and Secrecy Performance Analysis

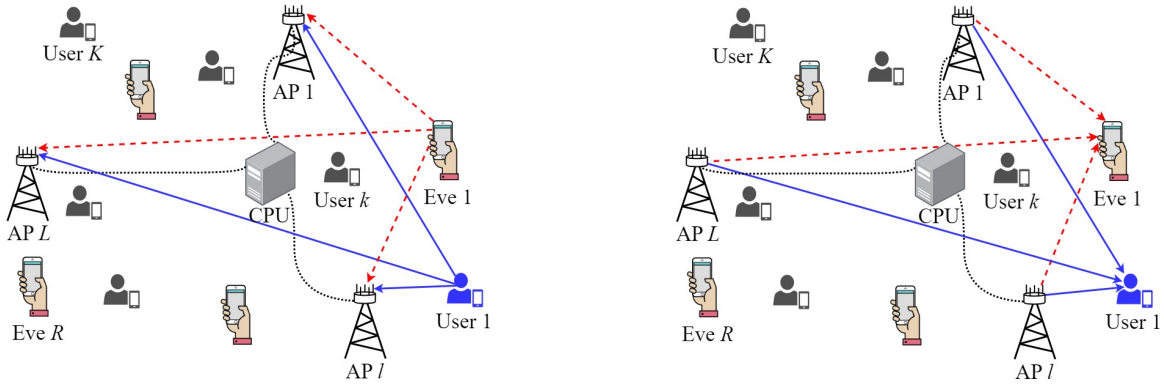
We consider a CF-mMIMO system, where L APs serve K single-antenna users in a time division duplex (TDD) mode, while the uplink and downlink transmissions use the same frequency but occur at different times. Each AP l, $l\in\mathcal{L}=\{1,\ldots,L\}$, is equipped with M antennas such that $LM > K$. In addition, there are R single-antenna active Eves in the system, as illustrated in Fig. 1, that attempt to intercept the information signals transmitted from the APs to some users (referred to as attacked/targeted users). We assume that different Eves overhear signals transmitted to different users. This assumption does not compromise generality, since multiple Eves attacking a given user can be treated as a single multi-antenna Eve. The users and Eves are randomly deployed over a large area. The sets of all users, Eves, and attacked users are, respectively, denoted by $\mathcal{K}=\{1,\ldots,K\}$, $\mathcal{R}=\{1,\ldots,R\}$, and $\hat{\mathcal{T}}\subset\mathcal{K}$. Let $\mathbf{h}_{l,k}\in\mathbb{C}^{M\times 1}$ and $\mathbf{h}^E_{l,r}\in\mathbb{C}^{M\times 1}$ be the channel vectors from the l-th AP to the k-th user and the r-th Eve, respectively. They can be modeled as

$$\mathbf{h}_{l,k}=\sqrt{\beta_{l,k}}\mathbf{g}_{l,k}, \tag{1}$$

and

$$\mathbf{h}^E_{l,r}=\sqrt{\beta^E_{l,r}}\mathbf{g}^E_{l,r}, \tag{2}$$

where $\beta_{l,k}$ and $\beta^E_{l,r}$ are the corresponding large-scale fading coefficients, while $\mathbf{g}_{l,k}\sim\mathcal{CN}(\mathbf{0},\mathbf{I}_M)$ and $\mathbf{g}_{l,r}\sim\mathcal{CN}(\mathbf{0},\mathbf{I}_M)$ represent the small-scale fading coefficient vectors. We consider Rayleigh fading channels for the following reasons: (i)



(a) Uplink training phase

(b) Downlink data transmission phase

FIGURE 1. CF-mMIMO with *L* APs, each equipped with *M* antennas, and *K* legitimate users facing multiple active Eves.

they simplify the analysis, enabling us to obtain important insights and achieve simple system designs; (ii) they are reasonable in many scenarios with rich scatterers, such as in dense urban environments; and (iii) it has been shown in [38] that the performance gap between CF-mMIMO under Rayleigh and Ricean fading channels is small. Here, the targeted users, the number of Eves, and the large-scale fading coefficients are assumed to be known a priori. The method to obtain these details is provided in Section IV.

We assume that the downlink transmission includes two main phases: the uplink training phase and the downlink data transmission phase. During the uplink training phase, all K users send their pilot sequences to the APs so that the APs can estimate the channels to all users. These channel estimates will be used to precode the signals before beamforming to all users in the downlink data transmission phase. To overhear the information transmitted to the users, the Eves conduct a pilot spoofing attack, i.e., they send the same pilot sequences as the pilots used by the users they want to intercept [1], [24]. By doing so, the channel estimates from the APs to the attacked users will be contaminated by the channels from the APs to the Eves. As a result, when the APs beam information to the attacked users, they also inadvertently beam to the Eves. Hence, the Eves can effectively overhear the signals transmitted to their targeted users.

A. Uplink Training Transmission

All users transmit their pilot sequences to all APs in each coherence block in the uplink training phase. Let τ_p denote the length of pilot sequences. Assume that $\tau_p \geq K$ so that the pilot sequences are pairwisely orthogonal. The pilot sequence sent by the k-th user is denoted by $\boldsymbol{\phi}_k \in \mathbb{C}^{\tau_p \times 1}$, where $\|\boldsymbol{\phi}_k\|^2 = 1$. Assume that the r-th Eve wishes to overhear the signals transmitted to the $\hat{t}$-th user; therefore, during the uplink training phase it sends a pilot sequence $\boldsymbol{\phi}_r^E = \boldsymbol{\phi}_{\hat{t}}$ to all the APs. Then, the received pilot matrix at the l-th AP is given by

$$\mathbf{Y}_{p,l} = \sqrt{\tau_p \rho_u} \sum_{k \in \mathcal{K}} \mathbf{h}_{l,k} \boldsymbol{\phi}_k^{\mathrm{H}} + \sqrt{\tau_p \rho_e} \sum_{r \in \mathcal{R}} \mathbf{h}_{l,r}^E (\boldsymbol{\phi}_r^E)^{\mathrm{H}} + \mathbf{N}_l, \tag{3}$$

where $\mathbf{N}_l$ is the additive noise with independent and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$ elements, $\rho_u \triangleq P_u/N_0$ and $\rho_e \triangleq P_e/N_0$, where N_0 is the average noise power at the APs, while P_u and P_e are the transmit powers of each user and Eve, respectively.

Using the minimum-mean-square-error (MMSE) estimation method at the l-th AP, the estimate of the channel between the l-th AP and k-th user is given by

$$\hat{\mathbf{h}}_{l,k} = \frac{\sqrt{\tau_p \rho_u} \beta_{l,k}}{\tau_p \rho_u \beta_{l,k} + \tau_p \rho_e \sum\limits_{r \in \mathcal{R}} \lambda_{k,r} \beta_{l,r}^E + 1} \mathbf{y}_{l,k}, \tag{4}$$

where $\mathbf{y}_{l,k} = \mathbf{Y}_{p,l} \boldsymbol{\phi}_k$ and $\lambda_{k,r}$ is a binary parameter defined as

$$\lambda_{k,r} \triangleq \begin{cases} 1, & \text{if the } r\text{-th Eve does attack the } k\text{-th user,} \\ 0, & \text{othewise.} \end{cases} \tag{5}$$

Since a different Eve would attack different user, we have $\sum\limits_{r \in \mathcal{R}} \lambda_{k,r} = 1$. Hence, $\hat{\mathbf{h}}_{l,k} \sim \mathcal{CN}(\mathbf{0}, \gamma_{l,k} \mathbf{I}_M)$, where

$$\gamma_{l,k} = \frac{\tau_p \rho_u \beta_{l,k}^2}{\tau_p \rho_u \beta_{l,k} + \tau_p \rho_e \sum\limits_{r \in \mathcal{R}} \lambda_{k,r} \beta_{l,r}^E + 1}. \tag{6}$$

Additionally, assuming that Eve r conducts a spoofing attack on user $\hat{t}$, then the estimate of $\mathbf{h}_{l,E}$ is given by

$$\hat{\mathbf{h}}_{l,r}^E = \sqrt{\alpha_{l,r}} \hat{\mathbf{h}}_{l,\hat{t}}, \tag{7}$$

whose elements are independent and identically distributed (i.i.d.) Gaussian random variables (RVs) with zero mean and variance $\gamma_{l,r}^E$, given by

$$\gamma_{l,r}^E = \alpha_{l,r} \gamma_{l,\hat{t}}, \tag{8}$$

where $\alpha_{l,r} \triangleq \frac{\rho_e \left(\beta_{l,r}^E\right)^2}{\rho_u \beta_{l,\hat{t}}^2}$.

B. Downlink Data Transmission

During this phase, all APs use the channel estimates obtained from the uplink training phase to transmit signals to the users. We consider a CF-mMIMO with AP selection (referred as the AP-user association), where each user is served by a subset of APs rather than all APs. Specifically, we define the binary variables as follows to represent the AP-user association:

$$a_{l,k} \triangleq \begin{cases} 1, & \text{if user } k \text{ is served by } l\text{-th AP}, \\ 0, & \text{othewise}. \end{cases} \tag{9}$$

The data symbol transmitted to the k-th user is denoted by s_k, where $\mathbb{E}\{|s_k|^2\} = 1$. Then, the precoded data signal sent to all users by the l-th AP can be expressed as follows

$$\mathbf{x}_l = \sum_{k \in \mathcal{K}} \sqrt{\rho_d} \theta_{l,k} \mathbf{w}_{l,k} s_k, \tag{10}$$

where ρ_d and $\theta_{l,k}$ is the maximum normalized transmit power at each AP and the power control coefficient, respectively. The vector $\mathbf{w}_{l,k} \in \mathbb{C}^{M \times 1}$, where $\mathbb{E}\{\|\mathbf{w}_{l,k}\|^2\} = 1$, is the precoding vector associated with user k. For all users, we enforce $\theta_{l,k} = 0$ whenever $a_{l,k} = 0$. Hence, the signals received at the k-th user and the r-th Eve are given by

$$z_k = \sum_{l \in \mathcal{L}} \mathbf{h}_{l,k}^{\mathrm{H}} \mathbf{x}_l + n_k, \tag{11}$$

and

$$z_r^E = \sum_{l \in \mathcal{L}} (\mathbf{h}_{l,r}^E)^{\mathrm{H}} \mathbf{x}_l + n_r^E, \tag{12}$$

respectively, where $n_k \sim \mathcal{CN}(0,1)$ and $n_r^E \sim \mathcal{CN}(0,1)$.

In our work, we consider the local protective partial zero-forcing (PPZF) precoder because it offers very good balance between performance and implemenation complexity [39]. Specifically, PPZF precoding enhances the performance of CF-mMIMO by splitting users into two subgroups: weak and strong groups $\mathcal{W}_l$, $\mathcal{S}_l$, respectively, with $\mathcal{S}_l \cap \mathcal{W}_l = \varnothing$ and $|\mathcal{S}_l| + |\mathcal{W}_l| = K$. The grouping process is done based on the channel gains $\beta_{l,k}$, $l \in \mathcal{L}, k \in \mathcal{K}$. In PPZF design, the APs use a partial zero forcing (PZF) scheme for strong users and a protective maximum ratio transmission (PMRT) scheme for weak users [39]. Since PZF is used for the strong group, we must have $M \geq |\mathcal{S}_l|$. Let us denote the collective matrix of the channel estimates from AP l to all users by $\hat{\mathbf{H}}_l = [\hat{\mathbf{h}}_{l,1}, \ldots, \hat{\mathbf{h}}_{l,K}] \in \mathbb{C}^{M \times K}$, while $\mathbf{E}_{\mathcal{S}_l} = \left[\mathbf{e}_i : i \in \mathcal{S}_l\right] \in \mathbb{C}^{K \times |\mathcal{S}_l|}$, where $\mathbf{e}_i$ is the i-th column of $\mathbf{I}_K$. Furthermore, let the j-th column of $\mathbf{I}|\mathcal{S}_l|$ be denoted by $\boldsymbol{\pi}_k$, where user k corresponds to the j-th element of set $\mathcal{S}_l$, with $j \in 1, \ldots, |\mathcal{S}_l|$. Thus, the precoding vector in (10) can be written as

$$\mathbf{w}_{l,k} = \begin{cases} \mathbf{w}_{l,k}^{\mathtt{PZF}}, & \text{if } k \in \mathcal{S}_l, \\ \mathbf{w}_{l,k}^{\mathtt{PMRT}}, & \text{if } k \in \mathcal{W}_l, \end{cases} \tag{13}$$

where $\mathbf{w}_{l,k}^{\mathtt{PZF}}$ and $\mathbf{w}_{l,k}^{\mathtt{PMRT}}$ are calculated as

$$\mathbf{w}_{l,k}^{\mathtt{PZF}} = \sqrt{(M - |\mathcal{S}_l|)\gamma_{l,j}} \hat{\mathbf{H}}_l \mathbf{E}_{\mathcal{S}_l} \left(\mathbf{E}_{\mathcal{S}_l}^{\mathrm{H}} \hat{\mathbf{H}}_l^{\mathrm{H}} \hat{\mathbf{H}}_l \mathbf{E}_{\mathcal{S}_l} \right)^{-1} \boldsymbol{\pi}_k, \tag{14}$$

$$\mathbf{w}_{l,j}^{\mathtt{PMRT}} = \frac{\mathbf{B}_l \hat{\mathbf{H}}_l \mathbf{e}_j}{\sqrt{(M - |\mathcal{S}_l|)\gamma_{l,j}}}, \tag{15}$$

respectively, with

$$\mathbf{B}_l = \mathbf{I}_M - \hat{\mathbf{H}}_l \mathbf{E}_{\mathcal{S}_l} \left(\mathbf{E}_{\mathcal{S}_l}^{\mathrm{H}} \hat{\mathbf{H}}_l^{\mathrm{H}} \hat{\mathbf{H}}_l \mathbf{E}_{\mathcal{S}_l} \right)^{-1} \mathbf{E}_{\mathcal{S}_l}^{\mathrm{H}} \hat{\mathbf{H}}_l^{\mathrm{H}}. \tag{16}$$

Thus, (10) can be re-written as

$$\mathbf{x}_l = \sum_{k \in \mathcal{S}_l} \sqrt{\rho_d} \theta_{l,k} \mathbf{w}_{l,k}^{\mathtt{PZF}} s_k + \sum_{j \in \mathcal{W}_l} \sqrt{\rho_d} \theta_{l,j} \mathbf{w}_{l,j}^{\mathtt{PMRT}} s_j. \tag{17}$$

C. Achievable SE at the Users and Eves

Similar to user grouping, the APs are divided into two subsets: weak and strong, based on the precoding vector utilized in the downlink phase. Let $\mathcal{Z}_k$ and $\mathcal{M}_k$ denote the indices associated with the APs that employ PZF and PMRT, respectively, which are given by

$$\mathcal{Z}_k \triangleq \{l : k \in \mathcal{S}_l, l \in \mathcal{L}\}, \tag{18}$$

$$\mathcal{M}_k \triangleq \{l : k \in \mathcal{W}_l, l \in \mathcal{L}\}, \tag{19}$$

with $\mathcal{Z}_k \cap \mathcal{M}_k = \varnothing$, and $|\mathcal{Z}_k| + |\mathcal{M}_k| = L$.

Then, the received signal at the k-th user in (11) can be re-expressed as

$$z_k = \left(\sum_{l \in \mathcal{Z}_k} \sqrt{\rho_d} \theta_{l,k} \mathbf{h}_{l,k}^{\mathrm{H}} \mathbf{w}_{l,k}^{\mathtt{PZF}} + \sum_{p \in \mathcal{M}_k} \sqrt{\rho_d} \theta_{p,k} \mathbf{h}_{p,k}^{\mathrm{H}} \mathbf{w}_{p,k}^{\mathtt{PMRT}} \right) s_k$$
$$+ \sum_{\substack{t \in \mathcal{K} \\ t \neq k}} \left(\sum_{l \in \mathcal{Z}_t} \sqrt{\rho_d} \theta_{l,t} \mathbf{h}_{l,k}^{\mathrm{H}} \mathbf{w}_{l,t}^{\mathtt{PZF}} + \sum_{p \in \mathcal{M}_t} \sqrt{\rho_d} \theta_{p,t} \mathbf{h}_{p,k}^{\mathrm{H}} \mathbf{w}_{p,t}^{\mathtt{PMRT}} \right) s_t + n_k.$$

Therefore, the achievable SE of the k-th user, using the widely-used use-and-then-forget bounding method [2], is given by $\mathrm{SE}_k = \log_2(1 + \mathrm{SINR}_k)$, where

$$\mathrm{SINR}_k = \frac{\left(\sum_{l \in \mathcal{L}} \sqrt{\rho_d (M - |\mathcal{S}_l|) \gamma_{l,k}} \theta_{l,k} \right)^2}{\sum_{t \in \mathcal{K}} \sum_{l \in \mathcal{L}} \rho_d \theta_{l,t}^2 (\beta_{l,k} - \delta_{l,k} \gamma_{l,k}) + 1}, \tag{20}$$

with $\delta_{l,k} = 1$ if $k \in \mathcal{S}_l$ and $\delta_{l,k} = 0$ if $k \in \mathcal{W}_l$. Note that the derivation of (20) can follow a similar methodology used in [27] where a single Eve was considered.

Furthermore, the received signal at the r-th Eve in (12) that overhears the signal transmitted to the $\hat{t}$-th user, can be rewritten as

$$z_r^E = \sum_{l \in \mathcal{L}} \sqrt{\rho_d} \theta_{l,\hat{t}} (\mathbf{h}_{l,r}^E)^{\mathrm{H}} \mathbf{w}_{l,\hat{t}} s_{\hat{t}}$$
$$+ \sum_{\substack{t \in \mathcal{K} \\ t \neq \hat{t}}} \sum_{l \in \mathcal{L}} \sqrt{\rho_d} \theta_{l,t} (\mathbf{h}_{l,r}^E)^{\mathrm{H}} \mathbf{w}_{l,\hat{t}} s_{\hat{t}} + n_r^E. \tag{21}$$

As a result, the SE of the r-th Eve can be approximated as[1]

$$\mathrm{SE}_r^E \approx \log_2\left(1 + \mathrm{SINR}_r^E\right), \tag{22}$$

[1]The approximation in calculating (22) follows [40, Lemma 1].

where

$$\mathrm{SINR}_r^E = \frac{\left(\sum\limits_{l\in\mathcal{L}}\sqrt{\rho_d(M-|\mathcal{S}_l|)\gamma_{l,r}^E}\theta_{l,\hat{t}}\right)^2 + \sum\limits_{l\in\mathcal{L}}\rho_d\theta_{l,\hat{t}}^2(\beta_{l,r}^E - \delta_{l,\hat{t}}\gamma_{l,r}^E)}{\sum\limits_{\substack{t\in\mathcal{K}\\ t\neq\hat{t}}}\sum\limits_{l\in\mathcal{L}}\rho_d\theta_{l,t}^2(\beta_{l,r}^E - \delta_{l,\hat{t}}\gamma_{l,r}^E) + 1}. \tag{23}$$

The derivation of (23) can also follow a similar method used in [27]. As a result, the SSE of the $\hat{t}$-th targeted user can be calculated as follows:

$$\mathrm{SSE}_{\hat{t}} = \left[\mathrm{SE}_{\hat{t}} - \sum_{r\in\mathcal{R}}\lambda_{\hat{t},r}\mathrm{SE}_r^E\right]^+, \tag{24}$$

where $[x]^+ = \max\{0, x\}$.

III. Joint Power Allocation and AP Selection

In this section, our objective is to jointly optimize the power control coefficients $\boldsymbol{\theta} \triangleq \{\theta_{l,k}\}$ and AP selection coefficients $\mathbf{a} \triangleq \{a_{l,k}\}$ in order to maximize the sum-SE of all the users under a transmit power constraint at each AP, positive SSE requirement for all the attacked users, and fronthaul constraint. In practical CF-mMIMO systems, the fronthaul links between the CPU and the APs have limited capacity. Consequently, it is necessary to limit the fronthaul signal load to each AP. To achieve this, the number of users served by each AP should be restricted. As a result, each user is served only by a subset of associated APs, rather than all the APs in the network.

The average normalized power constraint at AP l can be expressed as

$$\sum_{k\in\mathcal{K}}\theta_{l,k}^2 \leq 1, \forall l. \tag{25}$$

Now, the joint power allocation and AP selection problem is formulated as

$$\max_{\mathbf{a},\boldsymbol{\theta}} \quad \sum_{k\in\mathcal{K}}\mathrm{SE}_k(\boldsymbol{\theta}), \tag{26a}$$

$$\text{s.t.} \quad \theta_{l,k} \geq 0, \forall l, k, \tag{26b}$$

$$\sum_{k\in\mathcal{K}}\theta_{l,k}^2 \leq 1, \forall l, \tag{26c}$$

$$\mathrm{SE}_{\hat{t}}(\boldsymbol{\theta}) - \sum_{r\in\mathcal{R}}\lambda_{\hat{t},r}\mathrm{SE}_r^E(\boldsymbol{\theta}) \geq 0, \qquad \forall\hat{t}\in\hat{\mathcal{T}} \tag{26d}$$

$$\left(\theta_{l,k}^2 = 0, \forall k, \text{ if } a_{l,k} = 0\right), \quad \forall l, \tag{26e}$$

$$\sum_{l\in\mathcal{L}}a_{l,k} \geq 1, \forall k, \tag{26f}$$

$$\sum_{k\in\mathcal{K}}a_{l,k} \leq \widehat{K}_l, \forall l, \tag{26g}$$

The constraint in (26f) ensures that each user associates with at least one AP, while $\widehat{K}_l$ in (26g) represents the maximum number of users served by each AP l to satisfy the fronthaul constraint. We highlight that the sufficient number of APs chosen by the optimization problem sustains the key characteristics of CF-mMIMO, even under the constraints in (26f) and (26g). Deploying APs with multiple antennas (M) ensures that each user is served by at least M antennas, enhancing signal strength, spatial diversity, and overall network performance. In particular, our proposed optimization approach selects a suboptimal set of APs, ensuring that at least one AP is assigned to each user. This allocation maintains positive SSE for attacked users while achieving high sum-SE across the network. APs not directly serving specific users can still assist others, while distant APs have minimal impact on SE. This helps preserve the quality of service (QoS) even when some APs are inactive. However, nearby APs to Eves can compromise security, making effective AP selection crucial for robust performance and secrecy.

To solve the optimization problem in (26), we first need to tackle the non-convexity issue posed by the binary constraint (9). To this end, we utilize the fact that $x \in \{0,1\} \Leftrightarrow x \in [0,1],\ \&\ x - x^2 \leq 0$ [41]. Thus, we substitute (9) with

$$Q(\mathbf{a}) \triangleq \sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}}\left(a_{l,k} - a_{l,k}^2\right) \leq 0, \tag{27}$$

$$0 \leq a_{l,k}, \forall l, k, \tag{28}$$

$$a_{l,k} \leq 1, \forall l, k. \tag{29}$$

In accordance with equation (26c), (26e) can be re-written as

$$\theta_{l,k}^2 \leq a_{l,k}, \forall l, k. \tag{30}$$

Now, problem (26) can be reformulated as

$$\min_{\mathbf{x}\in\mathcal{C}} \quad -\sum_{k\in\mathcal{K}}\mathrm{SE}_k(\boldsymbol{\theta}), \tag{31}$$

where $\mathbf{x} \triangleq \{\mathbf{a}, \boldsymbol{\theta}\}, \mathcal{C} \triangleq \{(26\text{b}), (26\text{c}), (26\text{d}), (26\text{f}), (26\text{g}), (27) - (30)\}$ is a feasible set. The optimization problem (31), involving AP selection and power allocation, is non-convex, making it impractical to find a globally optimal solution in a reasonable timeframe for large-scale CF-mMIMO systems. Instead, the APG-based method is employed to obtain a locally optimal solution. The degree of sub-optimality is influenced by factors, such as network size and configuration, as well as the approximations used in key equations. The APG-based method has demonstrated high effectiveness for resource allocation in wireless systems [42], [43]. It offers excellent performance and can substantially reduce the complexity compared to conventional SCA algorithms, especially as the network size increases[2]

APG Approach

To begin, we modify the optimization problem (31) through a change in variables, enabling a more efficient computation

[2]While SCA-based methods scale poorly with large problems, the APG-based algorithm, despite requiring more iterations, is computationally more efficient. The APG method is memory-efficient, relying on quick closed-form expressions and gradient computations. Additionally, as our proposed optimization approach uses statistical CSI, optimization results can be reused across different subcarriers and frame durations, ensuring practicality and efficiency in dynamic scenarios.

of the function's gradient and the subsequent projection. To this end, we first introduce $z_{l,k}^2 \triangleq a_{l,k}, \forall l,k$, where

$$0 \leq z_{l,k} \leq 1, \tag{32}$$

and define the new variables as follows

- $U_k(\boldsymbol{\theta}) = \left(\sum_{l\in\mathcal{L}} \sqrt{\rho_d(M-|\mathcal{S}_l|)\gamma_{l,k}}\theta_{l,k}\right)^2$,
- $V_k(\boldsymbol{\theta}) = \sum_{t\in\mathcal{K}}\sum_{l\in\mathcal{L}} \rho_d\theta_{l,t}^2(\beta_{l,k}-\delta_{l,k}\gamma_{l,k})+1$,
- $U_r^E(\boldsymbol{\theta}) = \left(\sum_{l\in\mathcal{L}} \sqrt{\rho_d(M-|\mathcal{S}_l|)\gamma_{l,r}^E}\theta_{l,\hat{t}}\right)^2 + \sum_{l\in\mathcal{L}} \rho_d\theta_{l,\hat{t}}^2(\beta_{l,r}^E-\delta_{l,\hat{t}}\gamma_{l,r}^E)$,
- $V_r^E(\boldsymbol{\theta}) = \sum_{\substack{t\in\mathcal{K}\\ t\neq\hat{t}}}\sum_{l\in\mathcal{L}} \rho_d\theta_{l,t}^2(\beta_{l,r}^E-\delta_{l,\hat{t}}\gamma_{l,r}^E)+1$.

Now, a key challenge in developing an efficient APG algorithm for solving problem (31) is the positive SSE constraint (26d) as well as the AP association constraints (26f), (27), and (30). One possible way to address this problem is to incorporate these constraints into the objective function by introducing a penalty parameter, resulting in the formulation of the penalized problem [42]. The APG technique is subsequently utilized to address the penalized problem, while this process is iterated until a predefined stopping criterion is met. To this end, for each constraint, we introduce a quadratic loss function as:

- $\Psi_1(\boldsymbol{\theta}) \triangleq \sum_{\hat{t}\in\hat{\mathcal{T}}}\left[\max\left(0, \sum_{r\in\mathcal{R}}\lambda_{\hat{t},r}\mathrm{SE}_r^E(\boldsymbol{\theta}) - \mathrm{SE}_{\hat{t}}(\boldsymbol{\theta})\right)\right]^2 \leq 0$, for constraint (26d).
- $\Psi_2(\mathbf{z}) \triangleq \sum_{k\in\mathcal{K}}\sum_{l\in\mathcal{L}}\left(z_{l,k}^2 - z_{l,k}^4\right)$, where $\mathbf{z} \triangleq \left[\mathbf{z}_1^T,\ldots,\mathbf{z}_L^T\right]^T$ with $\mathbf{z}_l \triangleq [z_{l,1},\ldots,z_{l,K}]^T$ for constraint (27).
- $\Psi_3(\boldsymbol{\theta},\mathbf{z}) \triangleq \sum_{k\in\mathcal{K}}([\max(0,1 - \sum_{l\in\mathcal{L}} z_{l,k}^2)]^2 + \sum_{l\in\mathcal{L}}[\max(0,\theta_{l,k}^2 - z_{l,k}^2]^2)$ for constraints (26f) and (30).

Now, the constraint in (26f) can be rewritten as

$$\sum_{k\in\mathcal{K}} z_{l,k}^2 \leq \widehat{K}_l, \forall l. \tag{33}$$

Therefore, for the given penalty coefficients μ_1, μ_2, and μ_3, the penalized objective function of problem (26), can be expressed as

$$\begin{aligned} f(\mathbf{v}) \triangleq -\sum_{k\in\mathcal{K}} \mathrm{SE}_k(\boldsymbol{\theta}) + \varrho[\mu_1\Psi_1(\boldsymbol{\theta}) \\ + \mu_2\Psi_2(\mathbf{z}) + \mu_3\Psi_3(\boldsymbol{\theta},\mathbf{z})], \end{aligned} \tag{34}$$

with $\mathbf{v} \triangleq \left[\boldsymbol{\theta}^T, \mathbf{z}^T\right]^T$ and $\varrho > 0$. Accordingly, at each iteration of the iterative process, the regularized optimization problem

$$\min_{\mathbf{v}\in\widehat{\mathcal{C}}} f(\mathbf{v}), \tag{35}$$

Algorithm 1 The Proposed Algorithm for Solving (35)

1: **Initialization**: $\varrho, \varsigma > 1$, $v > 0$, $n = 1, q^{(0)} = 0, q^{(1)} = 1, \mathbf{v}^{(0)}, \overline{\mathbf{v}}^{(0)} \in \widehat{\mathcal{C}}, \alpha_{\overline{\mathbf{v}}} > 0, \alpha_{\mathbf{v}} > 0, \tilde{\mathbf{v}}^{(1)} = \mathbf{v}^{(1)} = \mathbf{v}^{(0)}, \zeta \in [0,1), b^{(1)} = 1, c^{(1)} = f\left(\mathbf{v}^{(1)}\right)$
2: **repeat** *(outer loop: penalty method)*
3: **repeat** *(inner loop: APG method)*
4: Update $\overline{\mathbf{v}}^{(n)}$ as $\overline{\mathbf{v}}^{(n)} = \mathbf{v}^{(n)} + \frac{q^{(n-1)}}{q^{(n)}}(\tilde{\mathbf{v}}^{(n)} - \mathbf{v}^{(n)}) + \frac{q^{(n-1)}-1}{q^{(n)}}(\mathbf{v}^{(n)} - \mathbf{v}^{(n-1)})$,
5: Set $\tilde{\mathbf{v}}^{(n+1)} = \mathcal{P}_{\widehat{\mathcal{C}}}\left(\overline{\mathbf{v}}^{(n)} - \alpha_{\mathbf{v}}\nabla f\left(\overline{\mathbf{v}}^{(n)}\right)\right)$,
6: **if** $f\left(\tilde{\mathbf{v}}^{(n+1)}\right) \leq c^{(n)} - \zeta\left\|\tilde{\mathbf{v}}^{(n+1)} - \overline{\mathbf{v}}^{(n)}\right\|^2$ **then**
7: $\mathbf{v}^{(n+1)} = \tilde{\mathbf{v}}^{(n+1)}$
8: **else**
9: Update $\hat{\mathbf{v}}^{(n+1)}$ as (55) and $\mathbf{v}^{(n+1)}$ as (56)
10: **end if**
11: Set $q^{(n+1)} = \frac{1+\sqrt{4\left(q^{(n)}\right)^2+1}}{2}$.
12: Update $b^{(n+1)}$ as (54), and $c^{(n+1)}$ as (53)
13: Set $n = n+1$
14: **until** $\left|\frac{f\left(\mathbf{v}^{(n)}\right)-f\left(\mathbf{v}^{(n-10)}\right)}{f\left(\mathbf{v}^{(n)}\right)}\right| \leq \epsilon$ or $\left|\frac{h\left(\boldsymbol{\theta}^{(n)}\right)-h\left(\boldsymbol{\theta}^{(n-1)}\right)}{h\left(\boldsymbol{\theta}^{(n)}\right)}\right| \leq \epsilon$
15: Increase $\varrho = \varrho \times \varsigma$
16: **until** Convergence.

for a given ϱ is solved, where $\widehat{\mathcal{C}}$ represents the convex feasible set of (35) (i.e., (26b), (26c), (32), (33)). We outline our proposed approach to address problem (35), which integrates the penalty method with the APG method, in **Algorithm 1**. We note that $\mathcal{P}_{\widehat{\mathcal{C}}}(\mathbf{x})$ in **Algorithm 1** shows the Euclidean projection operator which is defined as

$$\mathcal{P}_{\widehat{\mathcal{C}}}(\mathbf{x}) = \mathrm{argmin}_{\mathbf{u}\in\mathcal{C}}\|\mathbf{x}-\mathbf{u}\|. \tag{36}$$

It is clear that the two main operations in the implementation of **Algorithm 1** are the computation of the gradient of the objective function and the projections.

1) Gradient of the objective function: The gradients $\frac{\partial}{\partial\theta_{l,k}}f(\mathbf{v})$ and $\frac{\partial}{\partial z_{l,k}}f(\mathbf{v})$ can be calculated as

$$\frac{\partial}{\partial\theta_{l,k}}f(\mathbf{v}) = -\sum_{i\in\mathcal{K}}\frac{\partial}{\partial\theta_{l,k}}\mathrm{SE}_i(\mathbf{v}) + \varrho\frac{\partial}{\partial\theta_{l,k}}\widetilde{\Psi}(\mathbf{v}), \tag{37}$$

$$\frac{\partial}{\partial z_{l,k}}f(\mathbf{v}) = -\sum_{i\in\mathcal{K}}\frac{\partial}{\partial z_{l,k}}\mathrm{SE}_i(\mathbf{v}) + \varrho\frac{\partial}{\partial z_{l,k}}\widetilde{\Psi}(\mathbf{v}), \tag{38}$$

where

$$\begin{aligned} &\frac{\partial}{\partial\theta_{l,k}}\mathrm{SE}_i(\mathbf{v}) \\ &= \frac{1}{\log 2}\left[\frac{\frac{\partial}{\partial\theta_{l,k}}(U_i(\mathbf{v})+V_i(\mathbf{v}))}{(U_i(\mathbf{v})+V_i(\mathbf{v}))} - \frac{\frac{\partial}{\partial\theta_{l,k}}V_i(\mathbf{v})}{V_i(\mathbf{v})}\right], \end{aligned} \tag{39}$$

where $\frac{\partial}{\partial\theta_{l,k}}U_i(\mathbf{v})$ and $\frac{\partial}{\partial\theta_{l,k}}V_i(\mathbf{v})$ are given by

$$\frac{\partial}{\partial\theta_{l,k}}U_i(\mathbf{v})=\begin{cases}2(\sum\limits_{l\in\mathcal{L}}\sqrt{\rho_d\left(M-|\mathcal{S}_l|\right)\gamma_{l,k}}\theta_{l,k})\\ \times\sqrt{\rho_d\left(M-|\mathcal{S}_l|\right)\gamma_{l,k}}, & i=k\backslash\{\hat{t}\},\\ 0, & i\neq k,\end{cases} \tag{40}$$

$$\frac{\partial}{\partial\theta_{l,k}}V_i(\mathbf{v})=\begin{cases}2\rho_d\theta_{l,k}\left(\beta_{l,k}-\delta_{l,k}\gamma_{l,k}\right), & i=k\backslash\{\hat{t}\},\\ 2\rho_d\theta_{l,k}\left(\beta_{l,i}-\delta_{l,i}\gamma_{l,i}\right), & i\neq k.\end{cases} \tag{41}$$

Additionally, $\frac{\partial}{\partial\theta_{l,k}}\widetilde{\Psi}(\mathbf{v})$ and $\frac{\partial}{\partial z_{l,k}}\widetilde{\Psi}(\mathbf{v})$ in (37) and (38) are respectively given by

$$\begin{aligned}\frac{\partial}{\partial\theta_{l,k}}\widetilde{\Psi}(\mathbf{v})&=4\mu_3\max\left(0,\theta_{l,k}^2-z_{lk}^2\right)\theta_{l,k}\\&+\mu_2\sum_{\hat{t}\in\hat{\mathcal{T}}}2\max\left(0,\sum_{r\in\mathcal{R}}\lambda_{\hat{t},r}\mathrm{SE}_r^E(\boldsymbol{\theta})-\mathrm{SE}_{\hat{t}}(\boldsymbol{\theta})\right)\\&\frac{\partial}{\partial\theta_{l,k}}\left(\sum_{r\in\mathcal{R}}\lambda_{\hat{t},r}\mathrm{SE}_r^E(\boldsymbol{\theta})-\mathrm{SE}_{\hat{t}}(\boldsymbol{\theta})\right),\end{aligned} \tag{42}$$

$$\begin{aligned}\frac{\partial}{\partial z_{l,k}}\widetilde{\Psi}(\mathbf{v})=&\mu_1\left(2z_{l,k}-4z_{l,k}^3\right)-4\mu_3\max\left(0,\theta_{l,k}^2-z_{l,k}^2\right)z_{l,k}\\&-4\mu_3\max\left(0,1-\sum_{l\in\mathcal{L}}z_{l,k}^2\right)z_{l,k}.\end{aligned} \tag{43}$$

where

$$\begin{aligned}&\frac{\partial}{\partial\theta_{l,k}}\mathrm{SE}_r^E(\theta)\\&=\frac{1}{\log 2}\left[\frac{\frac{\partial}{\partial\theta_{l,k}}\left(U_r^E(\theta)+V_r^E(\theta)\right)}{\left(U_r^E(\theta)+V_r^E(\theta)\right)}-\frac{\frac{\partial}{\partial\theta_{l,k}}V_r^E(\theta)}{V_r^E(\theta)}\right],\end{aligned} \tag{44}$$

with

$$\frac{\partial}{\partial\theta_{l,k}}U_r^E(\theta)=\begin{cases}2\left(\sum\limits_{l\in\mathcal{L}}\sqrt{\rho_d\left(M-|\mathcal{S}_l|\right)\gamma_{l,r}^E}\theta_{l,k}\right)\\ \times\sqrt{\rho_d\left(M-|\mathcal{S}_l|\right)\gamma_{l,r}^E}\\ +2\rho_d\theta_{l,k}\left(\beta_{l,r}^E-\delta_{l,\hat{t}}\gamma_{l,r}^E\right), & \hat{t}=k,\\ 0, & \hat{t}\neq k,\end{cases} \tag{45a}$$

$$\frac{\partial}{\partial\theta_{l,k}}V_r^E(\mathbf{v})=\begin{cases}2\rho_d\theta_{l,k}\left(\beta_{l,r}^E-\delta_{l,\hat{t}}\gamma_{l,r}^E\right), & t=k\backslash\{\hat{t}\},\\ 0, & t\neq k.\end{cases} \tag{45b}$$

2) Projection onto feasible set $\widehat{\mathcal{C}}$: The projection onto the feasible set $\widehat{\mathcal{C}}$ in **Algorithm 1** can be done by solving the problem

$$\begin{aligned}\mathcal{P}_{\widehat{\mathcal{C}}}(\mathbf{v}):&\min_{\mathbf{v}\in\mathbb{R}^{2LK\times1}}\|\mathbf{v}-\mathbf{r}\|^2\\&\text{s.t.}\quad(26\text{b}),(26\text{c}),(32),\end{aligned} \tag{46}$$

where $\mathbf{r}=\left[\mathbf{r}_1^T,\mathbf{r}_2^T\right]^T\in\mathbb{R}^{2LK\times1}$ with $\mathbf{r}_1\triangleq\left[\mathbf{r}_{1,1}^T,\ldots,\mathbf{r}_{1,L}^T\right]^T$ and $\mathbf{r}_{1,l}\triangleq[r_{1,l1},\ldots,r_{1,lK}]^T$. Problem (46) can be split into two distinct subproblems as

$$\begin{aligned}&\min_{\boldsymbol{\theta}_l\in\mathbb{R}^{LK\times1}}\|\boldsymbol{\theta}_l-\mathbf{r}_{1,l}\|^2\\&\text{s.t. }\|\boldsymbol{\theta}_l\|^2\leq1,\boldsymbol{\theta}_l\geq\mathbf{0},\end{aligned} \tag{47}$$

and

$$\begin{aligned}&\min_{\mathbf{z}_l\in\mathbb{R}^{LK\times1}}\|\mathbf{z}_l-\mathbf{r}_{2,l}\|^2\\&\text{s.t. }\|\mathbf{z}_l\|^2\leq\widehat{K}_l,\mathbf{z}_l\geq\mathbf{0},\mathbf{z}_l\leq\mathbf{1},\end{aligned} \tag{48}$$

where the constraints in the problems (47) and (48) adhere to the conditions outlined in (26b), (26c), and (32).[3]

Solving problem (47) involves projecting a given point onto the intersection of a Euclidean ball and the positive orthant, and this projection can be calculated using a closed-form expression [37] as

$$\boldsymbol{\theta}_l=\frac{1}{\max\left(1,\left\|\left[\mathbf{r}_{1,l}\right]_0^+\right\|\right)}\left[\mathbf{r}_{1,l}\right]_0^+, \tag{49}$$

where $[\mathbf{x}]_0^+\triangleq[\max(0,x_1),\ldots,\max(0,x_K)]^T,\forall\mathbf{x}\in\mathbb{R}^{K\times1}$. Following [44], the solution of (48) can be approximated as

$$\mathbf{z}_l=\left[\frac{\sqrt{\widehat{K}_l}}{\max\left(\sqrt{\widehat{K}_l},\left\|\left[\mathbf{r}_{2,l}\right]_{0+}\right\|\right)}\left[\mathbf{r}_{2,l}\right]_{0+}\right]_{1-}, \tag{50}$$

where $[\mathbf{x}]_{1-}\triangleq[\min(1,x_1),\ldots,\min(1,x_K)]^{\mathrm{T}},\forall\mathbf{x}\in\mathbb{R}^{K\times1}$.

We note that in **Algorithm 1**, starting from $\overline{\mathbf{v}}^{(n)}$, we move along the gradient of the objective function using a specific step size $\alpha_{\overline{\mathbf{v}}}$. By projecting the resulting point $(\overline{\mathbf{v}}-\alpha_{\overline{\mathbf{v}}}\nabla f(\overline{\mathbf{v}}))$ onto the feasible set $\widehat{\mathcal{C}}$, we obtain

$$\tilde{\mathbf{v}}^{(n+1)}=\mathcal{P}_{\widehat{\mathcal{C}}}\left(\overline{\mathbf{v}}^{(n)}-\alpha_{\mathbf{v}}\nabla f\left(\overline{\mathbf{v}}^{(n)}\right)\right). \tag{51}$$

However, $f\left(\tilde{\mathbf{v}}^{(n+1)}\right)$ may not improve the objective sequence because $f(\mathbf{v})$ is not convex. In this case,$\mathbf{v}^{(n+1)}=$

[3]In this work, we consider active eavesdroppers with pilot spoofing attacks. In CF-mMIMO, active eavesdroppers pose a significantly greater threat than passive eavesdroppers who remain silent in the uplink training phase and try to only intercept data in the downlink data transmission phase. The main reason is, in CF-mMIMO, the use of a massive number of antennas enables the system to beam signals precisely to the user locations. Thus, the information leaked to the passive eavesdropper is negligible. However, to address the passive eavesdropping scenario, modifications are required, including updates to (20) and (22) to reflect the absence of pilot contamination. Furthermore, (7) would no longer be applicable and must be revised. Once these equations are updated, the optimization problem can be reformulated accordingly. Specifically, (40), (41) and (45) in Section III-C must be adjusted based on the new SE expressions. After these modifications, the same optimization methodology can be applied to solve the updated problem.

$\tilde{\mathbf{v}}^{(n+1)}$ is accepted if and only if the objective value $f\left(\tilde{\mathbf{v}}^{(n+1)}\right)$ is below $c^{(n)}$, which represents a relaxation of $f\left(\mathbf{v}^{(n)}\right)$, but relatively close to $f\left(\mathbf{v}^{(n)}\right)$. Following [45], $c^{(n)}$ is computed as follows:

$$c^{(n)} = \frac{\sum_{n=1}^{\kappa} \zeta^{(\kappa-n)} f\left(\mathbf{v}^{(n)}\right)}{\sum_{n=1}^{\kappa} \zeta^{(\kappa-n)}}, \tag{52}$$

where $\zeta \in [0,1)$ used to control the non-monotonicity degree. After each iteration, $c^{(n)}$ can be iteratively updated as follows:

$$c^{(n+1)} = \frac{\zeta b^{(n)} c^{(n)} + f\left(\mathbf{v}^{(n)}\right)}{b^{(n+1)}}, \tag{53}$$

where $c^{(1)}{=}f\left(\mathbf{v}^{(1)}\right)$ and $b^{(1)}{=}1$, and $b^{(n+1)}$ is calculated as

$$b^{(n+1)} = \zeta b^{(n)} + 1. \tag{54}$$

In case the condition $f\left(\tilde{\mathbf{v}}^{(n+1)}\right) \leq c^{(n)} - \zeta\|\tilde{\mathbf{v}}^{(n+1)} - \overline{\mathbf{v}}^{(n)}\|^2$ is not satisfied, extra correction steps are employed to avoid this situation, where $\|\mathbf{x}\|$ denotes the Euclidean norm of vector $\mathbf{x}$. Specifically, an additional point is calculated with a dedicated step size $\alpha_{\mathbf{v}}$ as

$$\hat{\mathbf{v}}^{(n+1)} = \mathcal{P}_{\widehat{\mathcal{C}}}\left(\mathbf{v}^{(n)} - \alpha_{\mathbf{v}} \nabla f\left(\mathbf{v}^{(n)}\right)\right). \tag{55}$$

Then, $\mathbf{v}^{(n+1)}$ is updated by comparing the objective values at $\tilde{\mathbf{v}}^{(n+1)}$ and $\hat{\mathbf{v}}^{(n+1)}$ as

$$\mathbf{v}^{(n+1)} \triangleq \begin{cases} \tilde{\mathbf{v}}^{(n+1)}, & \text{if } f\left(\tilde{\mathbf{v}}^{(n+1)}\right) \leq f\left(\hat{\mathbf{v}}^{(n+1)}\right), \\ \hat{\mathbf{v}}^{(n+1)}, & \text{otherwise.} \end{cases} \tag{56}$$

Since the feasible set $\widehat{\mathcal{C}}$ has bounds, it is valid to assert that $\nabla f(\mathbf{v})$ is Lipschitz continuous, with a known constant value of J, i.e.,

$$\|\nabla f(\mathbf{x}) - \nabla f(\mathbf{y})\| \leq J\|\mathbf{x} - \mathbf{y}\|, \forall \mathbf{x}, \mathbf{y} \in \widehat{\mathcal{C}}. \tag{57}$$

IV. Detection of Eavesdropping Attacks

Our proposed joint power control and AP selection method in Section III requires knowledge of the eavesdropping attacks, including the number of Eves, which users are being attacked, and the large-scale fading coefficients between the APs and the Eves. In this section, we propose a method to address these issues. The proposed method builds upon the approach in [28] but is specifically designed to accommodate scenarios with multiple eavesdroppers. Additionally, it demonstrates the capability to estimate the large-scale fading coefficients of the eavesdroppers, enhancing its applicability in complex network environments.

A. Eves' Detection

To detect which users are being overheard and how many Eves are in the system, we will utilize the pilot signals received at all APs across the entire system bandwidth. It is worth noting that in previous sections, we considered transmission in a specific coherence block, and the index regarding the coherence bandwidth was neglected. In this section, for the sake of completeness, coherence bandwidth indexes are included. In the n-th coherence interval, the received pilot matrix at the l-th AP, given by (3), can be rewritten as

$$\mathbf{Y}_{p,l}(n) = \sqrt{\tau_p \rho_u} \sum_{k=1}^{K} \mathbf{h}_{l,k}(n) \boldsymbol{\phi}_k^{\mathrm{H}} + \sqrt{\tau_p \rho_e} \sum_{r \in R} \mathbf{h}_{l,r}^{E}(n) (\boldsymbol{\phi}_r^{E})^{\mathrm{H}} + \mathbf{N}_l(n). \tag{58}$$

Note that the small-scale fading parts of $\mathbf{h}_{l,k}(n)$ and $\mathbf{h}_{l,r}^{E}(n)$ vary independently over n, but their large-scale fading parts remain unchanged, as large-scale fading is consistent across frequencies.

By using $\mathbf{Y}_{p,l}(n)$, we can check if user k is overheard by an Eve (say Eve r) or not. Let us consider two hypotheses: $\mathcal{H}_{k,1}$ if user k is overheard by Eve r; and $\mathcal{H}_{k,0}$ otherwise. We define $\mathbf{y}_{l,k}(n) = \mathbf{Y}_{p,l}(n)\boldsymbol{\phi}_k$. Then, the observation vector $\mathbf{y}_{l,k}(n)$ under hypotheses $\mathcal{H}_{k,0}$ and $\mathcal{H}_{k,1}$ is formulated as

$$\mathbf{y}_{l,k}(n) = \begin{cases} \sqrt{\tau_p \rho_u} \mathbf{h}_{l,k}(n) + \mathbf{N}_l(n)\boldsymbol{\phi}_k, & \mathcal{H}_{k,0}, \\ \sqrt{\tau_p \rho_u} \mathbf{h}_{l,k}(n) + \sqrt{\tau_p \rho_e} \mathbf{h}_{l,r}^{E}(n) + \mathbf{N}_l(n)\boldsymbol{\phi}_k, & \mathcal{H}_{k,1}. \end{cases} \tag{59}$$

Let N_{cb} be the number of coherence bandwidth intervals within a whole system bandwidth. Then, we can stack $\mathbf{y}_{l,k}(n)$ from all coherence bandwidth intervals to obtain an $N_{\mathrm{cb}}M \times 1$ vector as

$$\bar{\mathbf{y}}_{l,k} = [\mathbf{y}_{l,k}^{T}(1) \ \ldots \ \mathbf{y}_{l,k}^{T}(N_{\mathrm{cb}})]^{T}. \tag{60}$$

The collective observation vector $\bar{\mathbf{y}}_{l,k}$ under hypotheses $\mathcal{H}_{k,0}$ and $\mathcal{H}_{k,1}$ is

$$\bar{\mathbf{y}}_{l,k} = \begin{cases} \sqrt{\tau_p \rho_u} \bar{\mathbf{h}}_{l,k} + \bar{\mathbf{n}}_{l,k}, & \mathcal{H}_{k,0}, \\ \sqrt{\tau_p \rho_u} \bar{\mathbf{h}}_{l,k} + \sqrt{\tau_p \rho_e} \bar{\mathbf{h}}_{l,r}^{E} + \bar{\mathbf{n}}_{l,k}, & \mathcal{H}_{k,1}. \end{cases} \tag{61}$$

where $\bar{\mathbf{h}}_{l,k} = [\mathbf{h}_{l,k}^{T}(1) \ \ldots \ \mathbf{h}_{l,k}^{T}(N_{\mathrm{cb}})]^{T}$, $\bar{\mathbf{h}}_{l,r}^{E} = [(\mathbf{h}_{l,r}^{E})^{T}(1) \ \ldots \ (\mathbf{h}_{l,r}^{E})^{T}(N_{\mathrm{cb}})]^{T}$, and $\bar{\mathbf{n}}_{l,k} = [\mathbf{N}_l^{T}(1) \ \ldots \ \mathbf{N}_l^{T}(N_{\mathrm{cb}})]^{T}\boldsymbol{\phi}_k$.

By using the law of large numbers, as $N_{\mathrm{cb}}M \to \infty$, we have

$$\frac{\|\bar{\mathbf{y}}_{l,k}\|^2}{N_{\mathrm{cb}}M} \to \begin{cases} \tau_p \rho_u \beta_{l,k} + 1, & \mathcal{H}_{k,0}, \\ \tau_p \rho_u \beta_{l,k} + \tau_p \rho_e \beta_{l,r}^{E} + 1, & \mathcal{H}_{k,1}. \end{cases} \tag{62}$$

The derivation of (62) is provided in Appendix A. The above result implies that, when $N_{\mathrm{cb}}M$ is large enough, if $\frac{\|\bar{\mathbf{y}}_{l,k}\|^2}{N_{\mathrm{cb}}M}$ is close to $\tau_p \rho_u \beta_{l,k} + 1$, then user k is not being overheard by any Eves. Otherwise, we can conclude that there is an Eve attempting to overhear signals transmitted to user k. In other words, we can use the following normalized gap:

$$\epsilon_{l,k} = \frac{\left|\frac{\|\bar{\mathbf{y}}_{l,k}\|^2}{N_{\mathrm{cb}}M} - (\tau_p \rho_u \beta_{l,k} + 1)\right|}{\tau_p \rho_u \beta_{l,k} + 1}, \tag{63}$$

to determine if user k is a targeted user or not. More precisely, the user is targeted if $\epsilon_{l,k}$ is large, and not otherwise.

Based on the above observation, we propose a method to identify the targeted users, as summarized in **Algorithm 2.**

Algorithm 2 Eves' Detection Method

1: For each coherence time, at AP l, using received pilot signals over N_{cb} coherence bandwidth intervals $\mathbf{Y}_{p,l}(n)$, for $n = 1, \ldots, N_{\text{cb}}$, to compute $\mathbf{y}_{l,k}(n) = \mathbf{Y}_{p,l}(n)\boldsymbol{\phi}_k$, for $k = 1, \ldots, K$.
2: Compute $\frac{\|\bar{\mathbf{y}}_{l,k}\|^2}{N_{\text{cb}}M}$, where $\bar{\mathbf{y}}_{l,k}$ is given by (60). Then, compute the normalized gap $\epsilon_{l,k}$ using (63).
3: Choose a threshold $\epsilon_{\text{threshold}}$:
- If $\epsilon_{l,k} \geq \epsilon_{\text{threshold}}$, then AP l decides that user k is overheard by an Eve. Set $\text{count}(l,k) = 1$.
- Otherwise, there is no Eve overhearing signals transmitted to user k. $\text{count}(l,k) = 0$
4: Compute the total of APs that conclude that user k is targeted, i.e. compute $\sum_{l=1}^{L} \text{count}(l,k)$. If $\sum_{l=1}^{L} \text{count}(l,k) \geq L/2$, then, we can conclude there is an Eve intercepting information sent to user k. Otherwise, user k is not targeted.

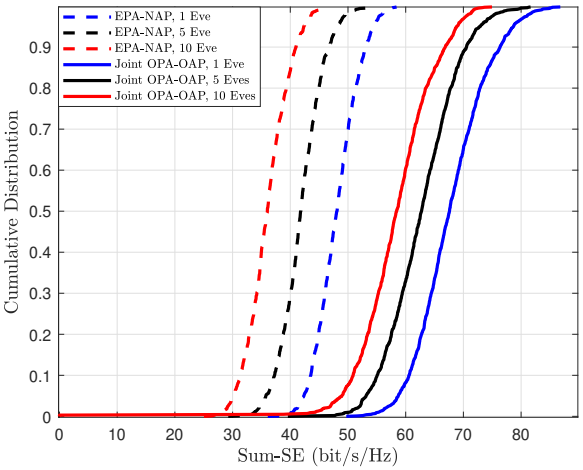


FIGURE 2. CDFs of the sum-SE for EPA-NAP and joint OPA-OAP schemes when the network is attacked by: 1) 1 Eve, 2) 5 Eves, and 3) 10 Eves, 8 of them attacking the same user. Here, $L = 50$, $M = 4$, $r = 0.2$ km, and $K = 10$.

The total number of Eves can then be determined, i.e., it equals the total number of targeted users.

B. Estimation of Eves' Large-Scale Fading Coefficients

Large-scale fading coefficients are required at the APs for system design purposes (e.g., power allocation and AP selection). These coefficients associated with the users (i.e. $\beta_{l,k}$) can be easily estimated because they change slowly with time and belong to legitimate systems. However, estimating the large-scale fading coefficients between the APs and the Eves (i.e. $\beta^E_{l,r}$) is challenging. In this section, we propose a simple method to estimate these coefficients based on the observations in Section A.

Let us assume that Eve r wants to overhear signals transmitted to user k, and that AP l can detect this Eve using

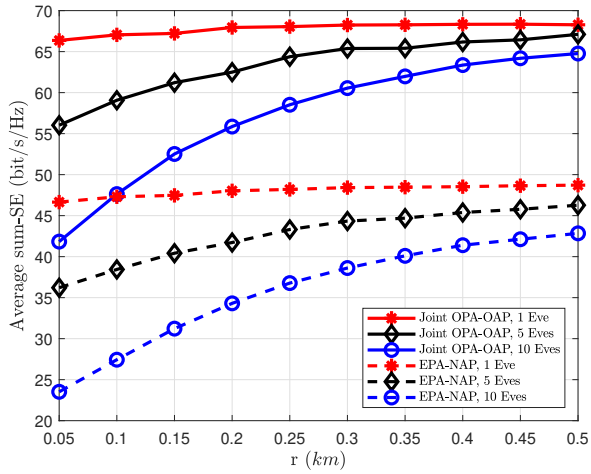


FIGURE 3. Average sum-SE versus r for EPA-NAP and joint OPA-OAP schemes when the network is attacked by: 1) 1 Eve, 2) 5 Eves, and 3) 10 Eves. Here, $L = 50$, $M = 4$, and $K = 10$.

Algorithm 2. Next, from (62), for large $N_{\text{cb}}M$, we have

$$\frac{\|\bar{\mathbf{y}}_{l,k}\|^2}{N_{\text{cb}}M} \approx \tau_p \rho_u \beta_{l,k} + \tau_p \rho_e \beta^E_{l,r} + 1. \tag{64}$$

Note that the above approximation is derived from the law of large numbers. In practical systems (e.g. 5G NR), the number of coherence bandwidth intervals, N_{cb}, is around 300, while each AP can be equipped with several antennas, M. Thus $N_{\text{cb}}M$ is large enough to make the above approximation tight.

Based on (64), the estimate of $\beta^E_{l,r}$ can be obtained be using $\bar{\mathbf{y}}_{l,k}$ as

$$\hat{\beta}^E_{l,r} = \left(\frac{\|\bar{\mathbf{y}}_{l,k}\|^2}{N_{\text{cb}}M} - \tau_p \rho_u \beta_{l,k} - 1 \right) \Big/ \tau_p \rho_e. \tag{65}$$

Remark 1:
The assumption of orthogonal pilots is feasible in scenarios with a long coherence interval and/or a moderate number of users, as typically observed in low- or medium-mobility environments [24], [30], [31]. In contrast, high-mobility or dense networks necessitate the use of non-orthogonal pilots, which demand more advanced analyses and designs, such as pilot contamination cancellation techniques. To extend the analysis of this paper to scenarios involving non-orthogonal pilots, specific modifications are required. In particular, additional terms must be incorporated into the denominator of (20) and (23) to account for the interference caused by pilot contamination. Furthermore, (3)–(8) in the uplink channel estimation process also require adjustments. These changes will consequently affect the SEs of both users and Eves. It is worth noting that the proposed optimization methodology in **Algorithm 1** remains applicable to the non-orthogonal pilot case. However, the Eves' detection method in **Algorithm 2** requires modification. In the non-orthogonal pilot scenario,

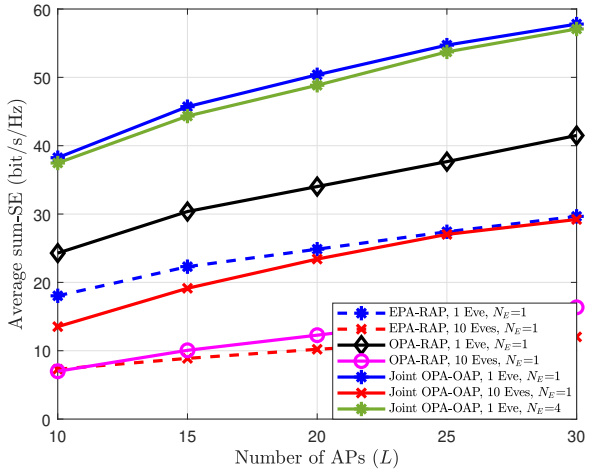


FIGURE 4. **Average sum-SE versus L for EPA-RAP, OPA-RAP, and joint OPA-OAP schemes including the scenario of a multiple-antenna Eve when the network is attacked by: 1) 1 Eve, 2) 10 Eves. Here, $M = 4$, $K = 10$, and $r = 0.05$ km.**

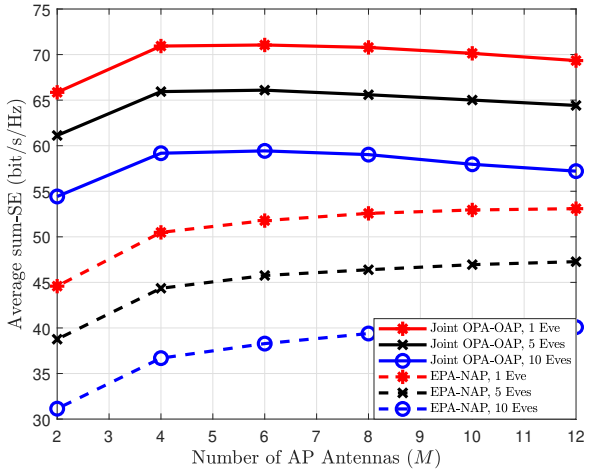


FIGURE 5. **Average sum-SE versus M when LM is fixed ($LM = 240$) for EPA-NAP and joint OPA-OAP schemes when the network is attacked by: 1) 1 Eve, 2) 5 Eves, and 3) 10 Eves. Here, $K = 10$, and $r = 0.2$ km.**

interference arises not only from the Eve but also from other legitimate users sharing the same pilot as the attacked user. Thus, **Algorithm 2** must be redesigned to account for this additional complexity.

V. Numerical Results

In this section, we present numerical results to investigate the security performance of CF-mMIMO under multiple active Eves using the proposed joint power control and AP selection approach.

A. Network Setup and Parameter Setting

We consider a CF-mMIMO network with multiple active Eves, where all L APs and K users are randomly distributed within a square area of 1×1 km^2. Additionally, each Eve is randomly located inside a circle of radius r_E around its targeted user. To mitigate boundary effects, we employ a wrapped-around technique. The large-scale fading $\beta_{l,k}$ is formulated as follows:

$$\beta_{l,k} = 10^{\frac{\text{PL}_{l,k}^{d}}{10}} 10^{\frac{F_{l,k}}{10}}, \tag{66}$$

where $10^{\frac{\text{PL}_{l,k}^{d}}{10}}$ is the path loss, and $10^{\frac{F_{l,k}}{10}}$ denotes the shadowing effect with $F_{l,k} \in \mathcal{N}(0, 4^2)$ (in dB). However, $\text{PL}_{l,k}^{d}$ (in dB) can be calculated as

$$\text{PL}_{l,k}^{d} = -30.5 - 36.7 \log_{10}\left(\frac{d_{l,k}}{1\,\text{m}}\right), \tag{67}$$

while the correlation among the shadowing terms from the l-th AP to different k users can be given by

$$\mathbb{E}\{F_{l,k} F_{j,k'}\} \triangleq \begin{cases} 4^2 2^{-\vartheta_{k,k'}/d_c}, & j = l, \\ 0, & \text{otherwise}, \end{cases} \tag{68}$$

where $\vartheta_{k,k'}$ is the physical distance between users k and k', and $d_c = 9$ m represents the decorrelation distance [39]. Similarly, the large-scale fading coefficient between the APs and Eves and between the users and Eves can be modelled by a proper change of indices. In addition, we choose the bandwidth $B = 20$ MHz, a noise power equal to -92 dBm, and the maximum transmit power for each AP and each user as 1 W and 100 mW, respectively. Finally, $\tau_p = K$.

B. Performance Evaluation

1) Benefits of the Proposed Joint Power Allocation and AP Selection, as well as Effect of the Number of Eves: In Fig. 2, we examine the effectiveness of the joint optimized power allocation and AP selection (**OPA-OAP**) approach presented in **Algorithm 1**. The cumulative distribution function (CDF) of the sum-SE is depicted for two scenarios: equal power allocation with no AP selection (**EPA-NAP**) and our proposed APG-based joint **OPA-OAP**. The results indicate that our proposed joint optimization approach yields 50% likely sum-SE enhancement of up to 62% compared to the **EPA-NAP** scheme. Also, as expected, from Fig. 2 it can be seen that the sum-SE reduces with an increase in the number of Eves. More specifically, for the proposed joint **OPA-OAP** scheme, the sum-SE decreases by up to 7.5% and 21% when the number of Eves increases from 1 to 5 and from 1 to 10, respectively. Nevertheless, the CF-mMIMO system employing the proposed **OPA-OAP** still yields an excellent sum-SE performance while yielding the positive SSE for all the attacked users.

2) Effect of Eves' Locations: In Fig. 3, we investigate the impact of Eves' locations on the SSE performance by plotting the average sum-SE against the radius of a circle surrounding attacked users. The average is taken over large-scale fading realizations. Our results show that the sum-SE significantly improves as the Eves move further away

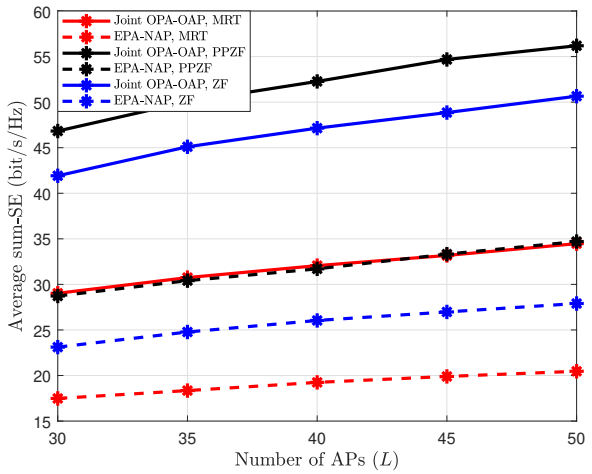
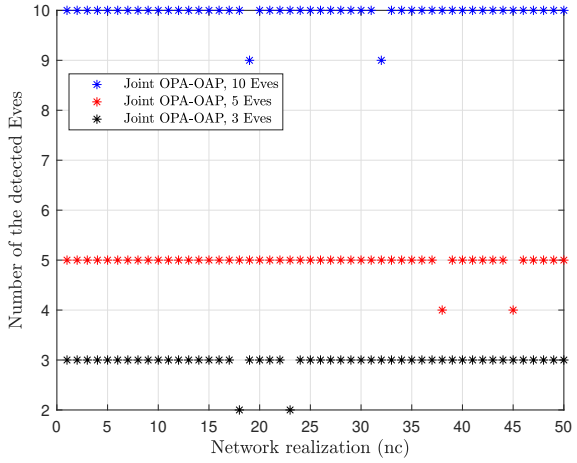
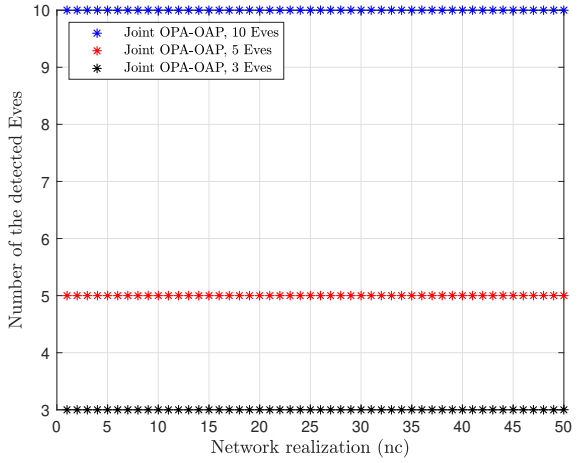


FIGURE 6. **Average sum-SE versus L when the network is attacked by 10 Eves for MRT, PPZF, and ZF precoding schemes. Here, $M = 4$, $K = 10$, and $r = 0.2$ km.**



(a) $\epsilon_{\text{threshold}} = 0.025$



(b) $\epsilon_{\text{threshold}} = 0.01$

FIGURE 7. **Number of the detected Eves using the proposed Eves' detection method for cell-free mMIMO under multiple active eavesdropping attacks for two cases: 1) $\epsilon_{\text{threshold}} = 0.025$, and 2) $\epsilon_{\text{threshold}} = 0.01$. Here, $L = 30$, $M = 4$, $K = 10$, $r = 0.2$ km, and $N_{\text{cb}} = 100$.**

from their targeted users. Notably, the performance gap between the joint optimization approach and baseline method becomes more pronounced in challenging scenarios where either the Eves are closer to the target users or the number of Eves are higher. Specifically, as r decreases from 500 m to 50 m, the average sum-SE performance gap between the joint optimization approach and the **EPA-NAP** benchmark increases from 40% to 42% and from 51% to 77.8% when there is 1 and 10 Eves in the system, respectively. These findings highlight the superiority of our joint optimization method over heuristic approaches and further confirm that effective power allocation and AP selection are essential for securing CF-mMIMO systems when multiple active Eves are located in close proximity to the targeted users.

3) Effect of the Number of APs (L): In Fig. 4, we investigate the impact of the number of APs (L) on the sum-SE. As expected, when the number of APs increases, the system has more degrees of freedom, and hence, the sum-SE increases. Particularly, the sum-SE increases by up to 39%, 41%, and 28% when L is doubled for equal power allocation with random AP selection (**EPA-RAP**), optimized power allocation with random AP selection (**OPA-RAP**), and joint **OPA-OAP** schemes, respectively. In addition, as the number of APs increases, the relative performance gaps for the cases with 1 and 10 Eves decrease, especially with the joint **OPA-OAP** scheme. This is because, at high L, more degrees of freedom can be effectively leveraged to mitigate the effects of attacks from multiple Eves.

In Fig. 4, we also study the impact of the number of antennas at Eve (N_E) on the sum-SE. To calculate the SE at a multiple-antenna Eve, we follow the same strategy as [28, Section V-B]. It is observed from Fig. 4 that the sum-SE remains largely unaffected as the number of antennas, N_E, increases from 1 to 4.

4) Effect of the Number of Antennas Per AP (M): In Fig. 5, we study the sum-SE of CF-mMIMO as a function of the number of antennas per AP, M, while maintaining a constant total number of service antennas $LM = 240$. It is evident that, up to a certain threshold, increasing M improves the sum-SE for both the **EPA-NAP** and **OPA-OAP** schemes. However, the rate of SE improvement diminishes as M increases further. This degradation occurs because, at larger M values, while array gain increases, the number of APs, L, decreases, resulting in higher path losses and reduced macro-diversity gain. This observation allows us to design a system where the optimal number of antennas per AP is selected. In addition, we observe that decreasing the number of AP antennas increases the gain of **OPA-OAP** over **EPA-NAP**.

6) Effect of the Precoding Schemes: Figure 6 illustrates the sum-SE of the CF-mMIMO system with multiple active

Eves relying on the proposed optimization approach as a function of the number of APs for various precoding schemes: maximum ratio transmission (MRT), zero-forcing (ZF), and PPZF. We can observe that the PPZF scheme outperforms the MRT and ZF schemes in terms of the sum-SE performance, primarily due to its excellent capability to mitigate inter-user interference and the effective balance it offers between canceling that interference and amplifying the desired signal.

*6) Performance of the Proposed Eves' Detection Method in **Algorithm 2**:* Figure 7 illustrates the scatter plots of the number of the detected Eves for 50 network realizations using the proposed Eves' detection method described in **Algorithm 2** when the network is under attack by multiple active Eves. The findings indicate that the proposed Eves' detection scheme effectively identifies the presence of active Eves in the network and the number of Eves. Furthermore, it shows that the performance of the proposed method strongly depends on the $\epsilon_{\text{threshold}}$ values.

VI. Conclusion

This paper addressed the critical challenge of ensuring secure communication in CF-mMIMO systems under the threat of active spoofing attacks by multiple Eves. To combat this, we proposed a joint AP selection and power optimization strategy designed to enhance the security and performance of CF-mMIMO systems. Our approach involves formulating a mixed-integer non-convex optimization problem aimed at maximizing the sum-SE of legitimate users while ensuring a positive SSE for all attacked users. The challenging formulated problem was transformed into a tractable form and an efficient algorithm was proposed to solve it using APG based path-following algorithm. Numerical results showed that our joint optimization approach significantly outperforms the heuristic approach in terms of sum-SE. More importantly, in the presence of a high number of Eves, e.g., 10 Eves, the scheme can provide high sum-SE gains, up to 72%, while achieving a positive secrecy spectral efficiency for all the attacked users. We also demonstrated that our proposed Eve detection algorithm is efficient and is able to provide a detection probability close to 1. While the proposed optimization scheme demonstrated significant improvements in both the sum-SE and secrecy performance, future research could explore the integration of jamming strategies to disrupt the Eves and further mitigate their impact on targeted users. Additionally, investigating the pilot contamination scenario when $\tau_p < K$ and/or under Ricean fading channels, is recommended for future studies.

Appendix A
Proof of (62)

From (61), under hypothesis $\mathcal{H}_{k,0}$, we have

$$\begin{aligned}\frac{\|\bar{\mathbf{y}}_{l,k}\|^2}{N_{\text{cb}}M} &= \frac{1}{N_{\text{cb}}M}\left(\sqrt{\tau_p\rho_u}\bar{\mathbf{h}}_{l,k}+\bar{\mathbf{n}}_{l,k}\right)^H\left(\sqrt{\tau_p\rho_u}\bar{\mathbf{h}}_{l,k}+\bar{\mathbf{n}}_{l,k}\right)\\ &= \tau_p\rho_u\frac{\bar{\mathbf{h}}_{l,k}^H\bar{\mathbf{h}}_{l,k}}{N_{\text{cb}}M}+\sqrt{\tau_p\rho_u}\frac{\bar{\mathbf{h}}_{l,k}^H\bar{\mathbf{n}}_{l,k}}{N_{\text{cb}}M}+\sqrt{\tau_p\rho_u}\frac{\bar{\mathbf{n}}_{l,k}^H\bar{\mathbf{h}}_{l,k}}{N_{\text{cb}}M}\\ &\quad+\frac{\bar{\mathbf{n}}_{l,k}^H\bar{\mathbf{n}}_{l,k}}{N_{\text{cb}}M}. \end{aligned} \tag{69}$$

Since the large-scale fading is independent of frequencies, the elements of $\bar{\mathbf{h}}_{l,k}$ are i.i.d. $\mathcal{CN}(0,\beta_{l,k})$ RVs. Thus the law of large numbers gives $\frac{\bar{\mathbf{h}}_{l,k}^H\bar{\mathbf{h}}_{l,k}}{N_{\text{cb}}M}\to\beta_{l,k}$, as $N_{\text{cb}}M$ goes to infinity. Similarly, $\frac{\bar{\mathbf{n}}_{l,k}^H\bar{\mathbf{n}}_{l,k}}{N_{\text{cb}}M}\to 1$. The other two terms of (69) converge to 0 when $N_{\text{cb}}M$ approaches infinity, as $\bar{\mathbf{h}}_{l,k}$ and $\bar{\mathbf{n}}_{l,k}$ are independent. By substituting these results into (69), we obtain, as $N_{\text{cb}}M\to\infty$,

$$\frac{\|\bar{\mathbf{y}}_{l,k}\|^2}{N_{\text{cb}}M}\to\tau_p\rho_u\beta_{l,k}+1. \tag{70}$$

A similar method can be applied to the case under hypothesis $\mathcal{H}_{k,1}$, and hence, we can arrive at (62).

REFERENCES

[1] Y. S. Atiya, Z. Mobini, H. Q. Ngo, and M. Matthaiou, "Joint power optimization and AP selection for secure cell-free massive MIMO," in *Proc. IEEE WCNC*, Apr. 2024.

[2] 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.

[3] 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.

[4] M. Mohammadi, Z. Mobini, H. Q. Ngo, and M. Matthaiou, "Next generation multiple access with cell-free massive MIMO," *Proc. IEEE*, vol. 119, no. 9, p. 1372–1420, Sep. 2024.

[5] 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.

[6] M. Matthaiou, O. Yurduseven, H. Q. Ngo, D. Morales-Jimenez, S. L. Cotton, and V. F. Fusco, "The road to 6G: Ten physical layer challenges for communications engineers," *IEEE Commun. Mag.*, vol. 59, no. 1, pp. 64–69, Jan. 2021.

[7] Z. Mobini, H. Q. Ngo, M. Matthaiou, and L. Hanzo, "Cell-free massive MIMO surveillance of multiple untrusted communication links," *IEEE Internet Things J.*, vol. 11, no. 20, pp. 33 010–33 026, Jul. 2024.

[8] Z. Mobini, M. Mohammadi, and C. Tellambura, "Wireless-powered full-duplex relay and friendly jamming for secure cooperative communications," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 3, pp. 621–634, Mar. 2019.

[9] D. Kapetanovic, G. Zheng, and F. Rusek, "Physical layer security for massive MIMO: An overview on passive eavesdropping and active attacks," *IEEE Commun. Mag.*, vol. 53, no. 6, pp. 21–27, Jun. 2015.

[10] Q. Xiong, Y.-C. Liang, K. H. Li, and Y. Gong, "An energy-ratio-based approach for detecting pilot spoofing attack in multiple-antenna systems," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 5, pp. 932–940, May 2015.

[11] J. Xie, Y.-C. Liang, J. Fang, and X. Kang, "Two-stage uplink training for pilot spoofing attack detection and secure transmission," in *Proc. IEEE ICC*, May 2017.

[12] X. Zhang, D. Guo, K. An, Z. Ding, and B. Zhang, "Secrecy analysis and active pilot spoofing attack detection for multigroup multicasting cell-free massive MIMO systems," *IEEE Access*, vol. 7, pp. 57 332–57 340, Apr. 2019.

[13] N. Li, Y. Gao, K. Xu, M. Guo, N. Sha, X. Wang, and G. Wang, "Spatial sparsity-based pilot attack detection and transmission countermeasure for cell-free massive MIMO system," *IEEE Syst. J.*, vol. 17, no. 2, pp. 2065–2076, Jun. 2023.

[14] Y. Fan, X. Wang, and X. Liao, "On the secure degrees of freedom for two-user MIMO interference channel with a cooperative jammer," *IEEE Trans. Commun.*, vol. 67, no. 8, pp. 5390–5402, Aug. 2019.
[15] J. Zhu, R. Schober, and V. K. Bhargava, "Secure transmission in multicell massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4766–4781, Sep. 2014.
[16] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission with an active eavesdropper," *IEEE Trans. Inf. Forensics Security*, vol. 62, no. 7, pp. 3880–3900, Jul. 2016.
[17] N.-P. Nguyen, H. Q. Ngo, T. Q. Duong, H. D. Tuan, and K. Tourki, "Secure massive MIMO with the artificial noise-aided downlink training," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 802–816, Apr. 2018.
[18] W. Xu, B. Li, L. Tao, and W. Xiang, "Artificial noise assisted secure transmission for uplink of massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 6750–6762, Jul. 2021.
[19] J. Chen, X. Chen, W. H. Gerstacker, and D. W. K. Ng, "Resource allocation for a massive MIMO relay aided secure communication," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 8, pp. 1700–1711, Aug. 2016.
[20] K. Guo, Y. Guo, and G. Ascheid, "Security-constrained power allocation in MU-massive-MIMO with distributed antennas," *IEEE Trans. Wireless Commun.*, vol. 15, no. 12, pp. 8139–8153, Dec. 2016.
[21] M. Li, G. Ti, and Q. Liu, "Secure beamformer designs in MU-MIMO systems with multiuser interference exploitation," *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8288–8301, Sep. 2018.
[22] F. Zhu, F. Gao, H. Lin, S. Jin, J. Zhao, and G. Qian, "Robust beamforming for physical layer security in BDMA massive MIMO," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 775–787, Apr. 2018.
[23] Z. Lin, M. Lin, B. Champagne, W.-P. Zhu, and N. Al-Dhahir, "Secure and energy efficient transmission for RSMA-based cognitive satellite-terrestrial networks," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 251–255, Feb. 2021.
[24] T. M. Hoang, H. Q. Ngo, T. Q. Duong, H. D. Tuan, and A. Marshall, "Cell-free massive MIMO networks: Optimal power control against active eavesdropping," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4724–4737, Oct. 2018.
[25] S. Elhoushy and W. Hamouda, "Nearest APs-based downlink pilot transmission for high secrecy rates in cell-free massive MIMO," in *Proc. GLOBECOM*, Dec. 2020.
[26] X. Gao, Y. Li, W. Cheng, L. Dong, and P. Liu, "Secure optimal precoding for user-centric cell-free massive MIMO system," *IEEE Wireless Commun. Lett.*, vol. 12, no. 1, pp. 31–35, Jan. 2023.
[27] Y. S. Atiya, Z. Mobini, H. Q. Ngo, and M. Matthaiou, "Cell-free massive MIMO with protective partial zero-forcing and active eavesdropping," in *Proc. IEEE VTC*, Jun. 2023.
[28] ——, "Secure transmission in cell-free massive MIMO under active eavesdropping," *IEEE Trans. Wireless Commun*, vol. 23, no. 12, pp. 18 036–18 052, Dec. 2024.
[29] S. Timilsina, D. Kudathanthirige, and G. Amarasuriya, "Physical layer security in cell-free massive MIMO," in *Proc. IEEE GLOBECOM*, Dec. 2018.
[30] M. Alageli, A. Ikhlef, F. Alsifiany, M. A. M. Abdullah, G. Chen, and J. Chambers, "Optimal downlink transmission for cell-free SWIPT massive MIMO systems with active eavesdropping," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 1983–1998, Nov. 2019.
[31] X. Zhang, D. Guo, K. An, and B. Zhang, "Secure communications over cell-free massive MIMO networks with hardware impairments," *IEEE Syst. J.*, vol. 14, no. 2, pp. 1909–1920, Jun. 2020.
[32] X. Zhang, T. Liang, K. An, G. Zheng, and S. Chatzinotas, "Secure transmission in cell-free massive MIMO with RF impairments and low-resolution ADCs/DACs," *IEEE Trans. Veh. Technol.*, vol. 70, no. 9, pp. 8937–8949, Sep. 2021.
[33] J. Park, S. Yun, and J. Ha, "Secure power control for downlink cell-free massive MIMO with passive eavesdroppers," *IEEE Trans. Veh. Technol.*, vol. 73, no. 6, pp. 9038–9043, Jun. 2024.
[34] X.-T. Dang, H. V. Nguyen, and O.-S. Shin, "Physical layer security for IRS-UAV-assisted cell-free massive MIMO systems," *IEEE Access*, vol. 12, pp. 89 520–89 537, Jun. 2024.
[35] Y. Chen, X. Zhang, F. Yao, K. An, G. Zheng, and S. Chatzinotas, "Pilot assignment and power control in secure UAV-enabled cell-free massive MIMO networks," *IEEE Internet Things J.*, vol. 11, no. 2, pp. 3377–3391, Jan. 2024.
[36] W. Xu, R. Wang, Y. Zhang, H. Quoc Ngo, and W. Xiang, "Pilot spoofing attack on the downlink of cell-free massive MIMO: From the perspective of adversaries," *IEEE Trans. Inf. Forensics Security*, vol. 19, pp. 5641–5654, May 2024.
[37] C. Hao, T. Vu, H.-Q. Ngo, M. Dao, X. Dang, and M. Matthaiou, "User association and power control in cell-free massive MIMO with the APG method," in *Proc. IEEE EUSIPCO*, Sep. 2023.
[38] H. Q. Ngo, H. Tataria, M. Matthaiou, S. Jin, and E. G. Larsson, "On the performance of cell-free massive MIMO in Ricean fading," in *Proc. IEEE ASILOMAR*, Nov. 2018, pp. 980–984.
[39] G. Interdonato, M. Karlsson, E. Björnson, and E. G. Larsson, "Local partial zero-forcing precoding for cell-free massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 19, no. 7, pp. 4758–4774, Jul. 2020.
[40] Q. Zhang, S. Jin, K.-K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.
[41] T. T. Vu, T. N. Duy, H. Q. Ngo, M. N. Dao, N. H. Tran, and R. H. Middleton, "Joint resource allocation to minimize execution time of federated learning in cell-free massive MIMO," *IEEE Internet Things J.*, vol. 9, no. 21, pp. 21 736–21 750, Nov. 2022.
[42] T. C. Mai, H. Q. Ngo, and L.-N. Tran, "Energy efficiency maximization in large-scale cell-free massive MIMO: A projected gradient approach," *IEEE Trans. Wireless Commun.*, vol. 21, no. 8, pp. 6357–6371, Aug. 2022.
[43] M. Farooq, H. Q. Ngo, E.-K. Hong, and L.-N. Tran, "Utility maximization for large-scale cell-free massive MIMO downlink," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 7050–7062, Nov. 2021.
[44] C. Hao, T. T. Vu, H. Q. Ngo, M. N. Dao, X. Dang, C. Wang, and M. Matthaiou, "Joint user association and power control for cell-free massive MIMO," *IEEE Internet Things J.*, vol. 11, no. 9, pp. 15 823–15 841, May 2024.
[45] H. Li and Z. Lin, "Accelerated proximal gradient methods for nonconvex programming," in *Proc. NeurIPS*, vol. 1, no. 9, Dec. 2015.

Yasseen Sadoon Atiya received the B.Sc degree in electrical engineering from the University of Babylon, Babylon, in 2007 and the M.Sc degree in electronics and communication from the University of Baghdad, Baghdad, in 2012. Currently, he is pursuing his Ph.D. at the Centre for Wireless Innovation (CWI), Queen's University Belfast. His research focuses on physical-layer security and cell-free massive MIMO systems. He has served as a reviewer for many journals, including the IEEE journals STSP and TIFS.

Zahra Mobini received the B.S. degree in electrical engineering from Isfahan University of Technology, Isfahan, Iran, in 2006, and the M.S and Ph.D. degrees, both in electrical engineering, from the M. A. University of Technology and K. N. Toosi University of Technology, Tehran, Iran, respectively. From November 2010 to November 2011, she was a Visiting Researcher at the Research School of Engineering, Australian National University, Canberra, ACT, Australia. She is currently a Post-Doctoral Research Fellow at the Centre for Wireless Innovation (CWI), Queen's University Belfast (QUB). Before joining QUB, she was an Assistant and then Associate Professor with the Faculty of Engineering, Shahrekord University, Shahrekord, Iran (2015-2021). Her research interests include physical-layer security, massive MIMO, cell-free massive MIMO, full-duplex communications, resource management, and optimization. She has co-authored many research papers in wireless communications. She has actively served as the reviewer for a variety of IEEE journals, such as TWC, TCOM, and TVT.

Hien Quoc Ngo is currently a Reader with Queen's University Belfast, U.K. His main research interests include massive MIMO systems, cell-free massive MIMO, reconfigurable intelligent surfaces, physical layer security, and cooperative communications. He has co-authored many research papers in wireless communications and co-authored the Cambridge University Press textbook *Fundamentals of Massive MIMO* (2016). He received the IEEE ComSoc Stephen O. Rice Prize in 2015, the IEEE ComSoc Leonard G. Abraham Prize in 2017, the Best Ph.D. Award from EURASIP in 2018, and the IEEE CTTC Early Achievement Award in 2023. He also received the IEEE Sweden VT-COM-IT Joint Chapter Best Student Journal Paper Award in 2015. He was awarded the UKRI Future Leaders Fellowship in 2019. He serves as the Editor for the IEEE Transactions on Wireless Communications, IEEE Transactions on Communications, the Digital Signal Processing, and the Physical Communication (Elsevier). He was a Guest Editor of IET Communications, and a Guest Editor of IEEE ACCESS in 2017.

Michail Matthaiou obtained his Ph.D. degree from the University of Edinburgh, U.K. in 2008. He is currently a Professor of Communications Engineering and Signal Processing and Deputy Director of the Centre for Wireless Innovation (CWI) at Queen's University Belfast, U.K. He is also an Eminent Scholar at the Kyung Hee University, South Korea. He has held research/faculty positions at Munich University of Technology (TUM), Germany and Chalmers University of Technology, Sweden. His research interests span signal processing for wireless communications, beyond massive MIMO, reflecting intelligent surfaces, mm-wave/THz systems and AI-empowered communications.

Dr. Matthaiou and his coauthors received the IEEE Communications Society (ComSoc) Leonard G. Abraham Prize in 2017. He currently holds the ERC Consolidator Grant BEATRICE (2021-2026) focused on the interface between information and electromagnetic theories. To date, he has received the prestigious 2023 Argo Network Innovation Award, the 2019 EURASIP Early Career Award and the 2018/2019 Royal Academy of Engineering/The Leverhulme Trust Senior Research Fellowship. His team was also the Grand Winner of the 2019 Mobile World Congress Challenge. He was the recipient of the 2011 IEEE ComSoc Best Young Researcher Award for the Europe, Middle East and Africa Region and a co-recipient of the 2006 IEEE Communications Chapter Project Prize for the best M.Sc. dissertation in the area of communications. He has co-authored papers that received best paper awards at the 2018 IEEE WCSP and 2014 IEEE ICC. In 2014, he received the Research Fund for International Young Scientists from the National Natural Science Foundation of China. He is currently the Editor-in-Chief of Elsevier Physical Communication, a Senior Editor for IEEE WIRELESS COMMUNICATIONS LETTERS and IEEE SIGNAL PROCESSING MAGAZINE, an Area Editor for IEEE TRANSACTIONS ON COMMUNICATIONS and Editor-in-Large for IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY. He is an IEEE and AAIA Fellow.