Exploiting Underlay Spectrum Sharing in Cell-Free Massive MIMO Systems

Diluka Loku Galappaththige, *Student Member, IEEE*,
and Gayan Amarasuriya Aruma Baduge, *Senior Member, IEEE*

Abstract— We investigate the coexistence of underlay spectrum sharing in cell-free massive multiple-input multiple-output (MIMO) systems. A primary system with geographically distributed primary access points (P-APs) serves a multitude of primary users (PUs), while a secondary system serves a large number of secondary users (SUs) in the same primary/licensed spectrum by exploiting the underlay spectrum sharing. To mitigate the secondary co-channel interference inflected at PUs, stringent secondary transmit power constraints are defined for the secondary access points (S-APs). A generalized pilots sharing scheme is used to locally estimate the uplink channels at P-APs/S-APs, and thereby, conjugate precoders are adopted to serve PUs/SUs in the same time-frequency resource element. Moreover, the effect of a user-centric AP clustering scheme is investigated by assigning a suitable set of APs to a particular user. The impact of estimated downlink (DL) channels at PUs/SUs via DL pilots beamformed by P-APs/S-APs is investigated. The achievable primary/secondary rates at PUs/SUs are derived for the statistical DL and estimated DL CSI cases. User-fairness for PUs/SUs is achieved by designing efficient transmit power control policies based on a multi-objective optimization problem formulation of joint underlay spectrum sharing and max-min criteria. The proposed orthogonal multiple-access based analytical framework is also extended to facilitate non-orthogonal multiple-access. Our analysis and numerical results manifest that the primary/secondary performance of underlay spectrum sharing can be boosted by virtue of the average reduction of transmit powers/path-losses, uniform coverage/service, and macro-diversity gains, which are inherent to distributed transmissions/receptions of cell-free massive MIMO.

Index Terms— Spectrum sharing, cell-free massive MIMO.

I. INTRODUCTION

MASSIVE multiple-input multiple-output (MIMO) operating in sub-6 GHz can simultaneously serve many users in the same time-frequency resource element by virtue of aggressive spatial multiplexing gains rendered by large base-station (BS) antenna arrays [1]–[4]. Co-located massive MIMO in which all BS antennas are packed into the same array is currently being deployed in the United States [5]. Thus, the co-located massive MIMO enabled with fully-digital beamforming has already become a reality [6].

Recently, a distributed/cell-free massive MIMO architecture, which deploys a large number of distributed access points (APs), is proposed to enhance the coverage probability and hence to provide a uniformly better service to users within a much larger geographical area [7]–[14]. Coordinated multi-point (CoMP) and network MIMO [15] are two related technologies, which exploit the notion of cooperation among BSs located within dedicated cells. On the contrary to these existing techniques, all APs are deployed in a cell-free architecture aiming to jointly serve all users for a given/large geographical area with no cell-boundaries. The uplink (UL) channel state information (CSI) is locally estimated at each AP via pilots send by users, and these APs are connected to a central processing unit (CPU) via a fronthaul/backhaul network [8]. By exploiting the channel reciprocity of time-division duplexing (TDD) mode of operation, APs acquire the downlink (DL) CSI via UL channel estimates and thereby design precoders for DL data transmission. The requirement for CSI exchange among APs via the CPU strictly depends on the AP precoder design. Owing to the fact that the distributed APs enable a user-centric architecture, the average transmission distances of a cell-free massive MIMO are inherently smaller than that of the co-located counterpart [16], [17]. This benefit translates into transmit power and path-loss reductions, which in turn lead to boosted energy efficiencies. The distributed APs circumvent the impediments caused by spatially correlated fading and shadowing due to large obstacles. The underlying macro-diversity gains can be exploited to boost the achievable rates [18], [19]. Joint beamfoming at a massive number of APs also enables unprecedented spatial multiplexing gains. Thus, cell-free massive MIMO enables a large number of concurrent connections with a guaranteed uniform service throughout a given/large geographical area.

Cognitive radio techniques based on the spectrum sharing are extensively explored to mitigate spectrum scarcity and underutilization/holes for the next-generation wireless communication systems [20]–[22]. In particular, many spectrum sharing techniques are evolved through three main paradigms, namely underlay, overlay, and interweave [20]. The overlay spectrum sharing involves sophisticated signal processing techniques and requires codebook knowledge of the

non-cognitive users. The interweave spectrum sharing adopts opportunistic frequency reuse over the available spectrum holes and hence requires stringent activity information of the non-cognitive users. The underlay spectrum sharing allows cognitive/secondary users (SUs) to simultaneously operate within the licensed spectrum if the secondary interference inflicted at the non-cognitive/primary users (PUs) is below a certain threshold. The main reasons for adopting the underlay spectrum sharing in this paper are its implementation simplicity and the achievable high spectrum utilization [20]–[22], compared to the significantly sophisticated overlay and interweave counterparts. Thus, in the underlay spectrum sharing, an underlaid secondary system can be simultaneously operated in the same licensed primary spectrum by defining stringent secondary transmit power constraints such that the secondary co-channel interference (CCI) caused to the PUs always falls below a predefined primary interference threshold (PIT) [20].

A. Related Prior Research on Underlay Spectrum Sharing With Massive MIMO

In [22], an initial foundation for investigating the feasibility of underlay spectrum sharing in massive MIMO with co-located antenna arrays at BSs is established. The BSs in [22] use the maximum ratio transmission (MRT) for signal transmission, and the DL achievable rate of the secondary system which is underlaid in a primary massive MIMO system is investigated. Reference [21] investigates the impact of inherent intra/inter-cell pilot contamination in multi-cell multi-user massive MIMO system with underlay spectrum sharing. In [23], a dual-hop enabled spectrum sharing system is analyzed, and the achievable rates are derived. Thereby, the detrimental effects of inter/intra-cell pilot contamination are investigated. The fundamental performance limits for relay selection strategies in massive MIMO two-way relaying are explored for perfect CSI in [24]. Moreover, in [24], the asymptotic achievable rates are derived for the best relay selection by deriving the asymptotic signal-to-interference-plus-noise ratio (SINR). In [25], the achievable rates of reserve-TDD based underlay spectrum sharing are presented. In [13], the achievable rates of underlay spectrum sharing co-located massive MIMO non-orthogonal multiple-access (NOMA) are presented. In [26], a low-complexity sub-optimal user-clustering technique for NOMA based underlay spectrum sharing in cell-free massive MIMO is proposed, and thereby, the achievable rates are derived for fixed transmit power allocation. In [27], [28], the performance bounds of spectrum sharing for massive MIMO with stochastic BS/user locations are derived. Moreover, by investigating pilot contamination, path-loss inversion power control, and receiver association policies, the secondary interference for a random cognitive massive MIMO system is characterized in [27], [28]. In [29], the quality-of-service aware power allocation and user selection schemes are studied for cognitive massive MIMO systems. Pilot decontamination techniques are proposed to asymptotically mitigate the residual interference in an underlaid single user massive MIMO cognitive radio system in [30]. In [18], the macro-multiplexing gain achieved from

optimization of antenna locations is characterized in terms of the ambient dimension of the cell and the path-loss exponent. Reference [19] derives the upper and lower bounds of the achievable rate with the perfect/imperfect CSI for cell-free massive MIMO systems. Thereby, [19] shows that the bounds of the achievable rate converge to a common lower bound owing to the extra distance-diversity or macro-diversity gain offered by distributed antennas in cell-free massive MIMO. In [31], the effects of finite capacity of fronthauls in the presence of residual hardware impairments at the users and APs are investigated by deriving the achievable rates for the compress-forward-estimate, estimate-compress-forward, and estimate-multiply-compress-forward strategies. Reference [32] proposes a low-complexity power allocation scheme to maximize the energy efficiency for a cell-free massive MIMO system with user centric approach operating at millimeter-wave (mmWave) frequencies. In [33], two pilot assignments, namely the user-group and interference-based K-means schemes are proposed for the structured massive access.

B. Our Motivation

The aforementioned related prior references [21]–[25] [27], [30] have investigated the coexistence of massive MIMO and underlay spectrum sharing with only co-located antenna arrays at the BSs. The closely related references [13] and [26], respectively, investigate the rate performance of co-located and cell-free massive MIMO NOMA with underlay spectrum sharing by only considering fixed power allocation. To the best of our knowledge, multi-objective max-min fairness-based transmit power allocation, impact of beamformed DL pilots by APs, and the achievable rates with DL estimated CSI at users have not yet been investigated. Thus, in our paper, we fill this gap by exploring multi-objective power allocation, impact of DL pilots, and practically realizable performance bounds of underlay spectrum sharing in cell-free massive MIMO with imperfectly estimated UL/DL CSI. Since the APs are spatially distributed over a given geographical area, the distributed transmissions of cell-free massive MIMO architecture can be beneficial in providing uniformly better average rate to users by virtue of mitigating the near-far effect via max-min power control than that of the co-located massive MIMO counterparts. The average amount of secondary CCI inflicted at a particular PU from cell-free/distributed massive MIMO transmissions can be better constrained as specified by the PIT with low-power distributed APs. Moreover, cell-free massive MIMO has more robustness against the detrimental effects of correlated small/large scale fading than the co-located counterpart [8]. Both primary and secondary systems can provide a higher coverage probability because there are no cell boundaries, and users are much closer to the APs in cell-free setting. Thus, the cell-free massive MIMO can significantly boost the performance of underlay spectrum sharing.

C. Our Contribution and Its Difference Relative to the Existing Literature

Our main contribution is to investigate different UL/DL CSI cases at APs/users, and transmit power control and their
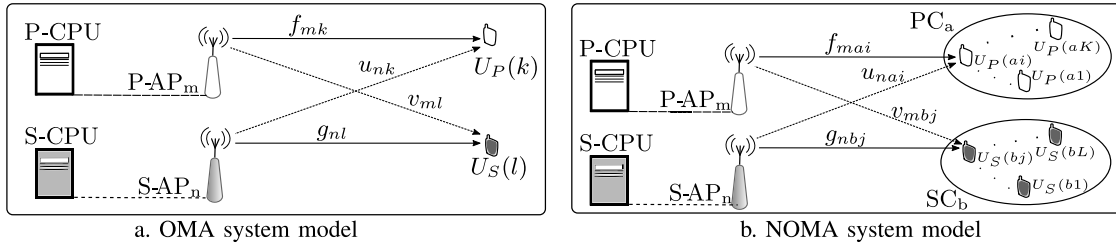
Fig. 1. A cell-free massive MIMO system models with underlay spectrum sharing.

effects on the achievable rates of underlay spectrum sharing in orthogonal multiple-access (OMA)/NOMA-based cell-free massive MIMO. Specifically, we derive the performance metrics by proposing max-min based multi-objective transmit power control algorithms and by exploring the deleterious impact of imperfectly estimated UL CSI at APs, availability of estimated DL CSI at the users, and effects of using statistical DL CSI with imperfect successive interference cancellation (SIC) signal decoding. Both the primary system and the secondary system, which is underlaid within the primary licensed spectrum, adopt a generalized pilot sharing scheme to minimize the training overhead. Thus, the pilot sequences sent by PUs and SUs are used to locally estimate the UL channels at primary access points (P-APs)/secondary access points (S-APs), respectively. Thereby, the impact of imperfectly estimated CSI is considered for our analysis. Our performance metrics for cell-free underlay spectrum sharing are categorized by taking into account the availability of long-term/statistical CSI and estimated DL CSI at PUs/SUs. When only UL channel estimation is used, the users rely on statistical CSI for signal decoding. Nonetheless, when APs beamform DL pilots, the users adopt estimated DL CSI to decode signals. Thus, the performance bounds are established for these two user CSI cases. Moreover, the effect of user-centric achieved through clustering the APs that serve for a particular user in both primary and secondary systems is investigated for cell-free underlay spectrum sharing.

To mitigate the secondary CCI inflected at PUs due to simultaneous transmission, the stringent secondary transmit power constraints are introduced at S-APs. Thus, by defining a PIT at PUs, the secondary transmit power is constrained such that the total secondary CCI at any PU falls below the predefined PIT. Then, in the presence of imperfectly estimated UL CSI at APs and DL CSI at users with intra-system pilot contamination and estimation errors, the achievable rates for PUs/SUs are derived. A multi-objective transmit power control algorithm is designed based on the max-min fairness criterion to guarantee a uniform quality-of-service among all the users. Moreover, the above OMA-based system model is extended to facilitate NOMA transmissions, and the corresponding performance bounds are established. The practical viability of underlay spectrum sharing in cell-free massive MIMO is explored by using numerical results through our analysis and Monte-Carlo simulations.

This paper goes well beyond our related conference papers [1], [2] by presenting a multi-objective optimization of

max-min fairness-based transmit power allocation for the OMA-based primary system, user-centric AP clustering aspects, beamforming of DL pilots, impact of estimated DL channel estimates, adoption of estimated DL CSI for signal decoding, and the corresponding achievable rates with estimated DL CSI for underlay spectrum sharing within cell-free massive MIMO. All numerical results/comparisons, except for Fig. 8, and their descriptions are distinctive from those of [1], [2], and the corresponding figures have been regenerated for different system parameters with respect to [1], [2].

Notation: $\mathbf{z}^T$ denotes the transpose of $\mathbf{z}$. The conjugate of z is denoted by z^*. The notation $z \sim \mathcal{CN}(\cdot, \cdot)$ denotes that z is a complex-valued circularly symmetric Gaussian distributed random variable. The operators $\mathbb{E}[\cdot]$ and $\mathbb{Var}[\cdot]$ are the expectation and variance, respectively.

II. SYSTEM, CHANNEL AND SIGNAL MODELS

A. System and Channel Model for OMA/NOMA

We consider a TDD cell-free massive MIMO system with underlay spectrum sharing (see Fig. 1a). A secondary system is underlaid within a primary licensed spectrum in order to enhance the overall spectrum efficiency by eliminating spectrum holes. The primary system having M single-antenna P-APs serves K single-antenna PUs, while the secondary system with N single-antenna S-APs uses the same time-frequency spectrum to serve L single-antenna SUs. We introduce the secondary transmit power constraints for all S-APs to ensure that the performance of the primary system is not hindered by the simultaneous secondary transmission in the same primary spectrum. Thus, the CCI caused by the secondary system at PUs falls below a predefined PIT, which defines a upper limit for the CCI endurance ability of PUs. The PIT constraint mitigates excessive secondary CCI at PUs. A synchronized operation between the primary and the secondary systems is assumed such that all P-APs and S-APs simultaneously serve all PUs and SUs by adopting spatial multiplexing rendered by cell-free massive MIMO [8]. Moreover, P-APs/S-APs are connected to their respective primary/secondary central processing units (P-CPU/S-CPU). The kth PU and the lth SU are denoted by $U_P(k)$ and $U_S(l)$, respectively.

In Fig. 1a, f_{mk}, g_{nl}, v_{ml}, and u_{nk} are the channel coefficients between the mth P-AP and $U_P(k)$, the nth S-AP and $U_S(l)$, the mth P-AP and $U_S(l)$, and the nth S-AP and $U_P(k)$, respectively, where $m \in \{1, \cdots, M\}$, $k \in \{1, \cdots, K\}$,

$n \in \{1, \cdots, N\}$, and $l \in \{1, \cdots, L\}$. The above channels can be modeled in a unified manner as

$$h_{ab} = \tilde{h}_{ab}\zeta_{h_{ab}}^{1/2}, \tag{1}$$

where $h \in \{f, g, u, v\}$, $a \in \{m, n\}$, and $b \in \{k, l\}$. Here, $\tilde{h}_{ab} \sim \mathcal{CN}(0, 1)$ captures the independent quasi-static Rayleigh fading and stays fixed during the coherence interval, while $\zeta_{h_{ab}}$ accounts for the large-scale fading, including path-loss and shadow fading. Since the large-scale coefficients stay fixed for several coherence intervals, it is assumed that the large-scale coefficients are known a-prior at both P-APs and S-APs [8]. Thus, the estimation of these large-scale fading coefficients can be done in once about tens/hundreds of coherence intervals [34].

Next, we extend our cell-free massive MIMO underlay spectrum sharing techniques to facilitate NOMA. To this end, we consider a system setup with AK number of PUs and BL number of SUs, where A and B are the numbers of primary clusters (PCs) and secondary clusters (SCs), respectively (see Fig. 1b). Based on the spatial directions of the users [25], [35], K and L number of PUs and SUs are assigned to each PC and SC, respectively. The kth PU in the ath PC and the lth SU in the bth SC are denoted by $U_P(a, k)$ and $U_S(b, l)$, respectively. The channel between the mth P-AP and $U_P(a, k)$ is denoted by f_{mak}, where $m \in \{1, \cdots, M\}$, $a \in \{1, \cdots, A\}$, and $k \in \{1, \cdots, K\}$. The channel between the nth S-AP and $U_S(b, l)$ is represented by g_{nbl}, where $n \in \{1, \cdots, N\}$, $b \in \{1, \cdots, B\}$, and $l \in \{1, \cdots, L\}$. Further, v_{mbl} and u_{nak} denote the interference channels between the mth P-AP and $U_S(b, l)$, and the nth S-AP and $U_P(a, k)$, respectively. These channels are modeled similar to (1) as

$$h_{qrs} = \tilde{h}_{qrs}\zeta_{h_{qrs}}^{1/2}, \tag{2}$$

where $h \in \{f, g, u, v\}$, $q \in \{m, n\}$, $r \in \{a, b\}$, and $s \in \{k, l\}$. Moreover, $\tilde{h}_{qrs} \sim \mathcal{CN}(0, 1)$ captures the small-scale fading, while $\zeta_{h_{qrs}}$ captures the large-scale fading, including path-loss and shadowing.

B. The UL Channel Estimation for OMA/NOMA

For OMA, the UL channels are estimated locally at P-APs and S-APs by using respective user pilots [8]. During UL channel estimation period, τ_p symbols out of the coherence interval having τ_c symbols are used to transmit the UL pilot. Then, by using these pilot sequences, the channels f_{mk} and g_{nl} are estimated at P-APs and S-APs, respectively. In practice, the number of orthogonal pilot sequences is limited as it is defined by the channel coherence interval [34]. To reduce the pilot overhead and to increase the number of served PUs/SUs, in this work, the pilot sequences are shared among PUs and SUs by adopting the following pilot sharing strategy. It is assumed that $Q \leq \min(K, L)$ number of pilot sequences having a length of τ_p symbol duration is shared among PUs and SUs. Then, we define the complete pilot sets used by PUs ($\boldsymbol{\Phi}_P$) and SUs ($\boldsymbol{\Phi}_S$) as

$$\boldsymbol{\Phi}_P = \begin{bmatrix} \boldsymbol{\Phi}; \tilde{\boldsymbol{\Phi}}_P \end{bmatrix} \quad \text{and} \quad \boldsymbol{\Phi}_S = \begin{bmatrix} \boldsymbol{\Phi}; \tilde{\boldsymbol{\Phi}}_S \end{bmatrix}, \tag{3}$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{Q \times \tau_p}$ denotes the pilots shared by Q PUs/SUs having a length of τ_p symbol duration. $\tilde{\boldsymbol{\Phi}}_P \in \mathbb{C}^{(K-Q) \times \tau_p}$ and $\tilde{\boldsymbol{\Phi}}_S \in \mathbb{C}^{(L-Q) \times \tau_p}$ are the pilots assigned for the remaining $(K-Q)$ PUs or $(L-Q)$ SUs, respectively. We can define the orthogonal properties among these pilots as $\boldsymbol{\Phi}^H \tilde{\boldsymbol{\Phi}}_P = 0$, $\boldsymbol{\Phi}^H \tilde{\boldsymbol{\Phi}}_S = 0$, and $\tilde{\boldsymbol{\Phi}}_P^H \tilde{\boldsymbol{\Phi}}_S = 0$. We define $\boldsymbol{\Phi}_P = \begin{bmatrix} \phi_{P_1}^T, \cdots, \phi_{P_k}^T, \cdots, \phi_{P_K}^T \end{bmatrix}^T$ and $\boldsymbol{\Phi}_S = \begin{bmatrix} \phi_{S_1}^T, \cdots, \phi_{S_l}^T, \cdots, \phi_{S_L}^T \end{bmatrix}^T$, where $\phi_{P_k} \in \mathbb{C}^{1 \times \tau_p}$ and $\phi_{S_l} \in \mathbb{C}^{1 \times \tau_p}$ are the pilot sequences sent by $U_P(k)$ and $U_S(l)$, respectively, and $\|\phi_{P_k}\|^2 = 1$ and $\|\phi_{S_l}\|^2 = 1$ for $k \in \{1, \cdots, K\}$ and $l \in \{1, \cdots, L\}$. Then, we can write the pilot signal received at the mth P-AP and the nth S-AP as

$$\mathbf{y}'_{P_m} = \sqrt{P_p}\sum_{k=1}^{K} f_{mk}\phi_{P_k} + \sqrt{P_p}\sum_{l=1}^{L} v_{ml}\phi_{S_l} + \mathbf{n}'_{P_m}, \tag{4a}$$

$$\mathbf{y}'_{S_n} = \sqrt{P_p}\sum_{l=1}^{L} g_{nl}\phi_{S_l} + \sqrt{P_p}\sum_{k=1}^{K} u_{nk}\phi_{P_k} + \mathbf{n}'_{S_n}, \tag{4b}$$

where $P_p = \tau_p P$ and P denotes the average transmitted pilot power at each PU/SU. Moreover, $\mathbf{n}'_{P_m}$ and $\mathbf{n}'_{S_n}$ are additive white Gaussian noise (AWGN) vectors, having independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ elements at the mth P-AP and the nth S-AP, respectively. The sufficient statistics for estimating f_{mk} and g_{nl} can be obtained by projecting $\phi_{P_k}^H$ and $\phi_{S_l}^H$ onto (4a) and (4b), respectively, as

$$y_{P_{mk}} = \phi_{P_k}^H \mathbf{y}'_{P_m} = \sqrt{P_p}f_{mk} + \sqrt{P_p}v_{mk} + n_{P_m}, \tag{5a}$$

$$y_{S_{nl}} = \phi_{S_l}^H \mathbf{y}'_{S_n} = \sqrt{P_p}g_{nl} + \sqrt{P_p}u_{nl} + n_{S_n}, \tag{5b}$$

where $\{\phi_{P_k}, \phi_{S_l}\} \in \boldsymbol{\Phi}$, $n_{P_m} = \phi_{P_k}^H \mathbf{n}'_{P_m} \sim \mathcal{CN}(0, 1)$, and $n_{S_n} = \phi_{S_l}^H \mathbf{n}'_{S_n} \sim \mathcal{CN}(0, 1)$ as ϕ_{P_k} and ϕ_{S_l} are unitary vectors.

Proposition 1: The minimum mean square error (MMSE) estimates of f_{mk} and g_{nl} are given by

$$\hat{f}_{mk} = \left(\mathbb{E}\left[y_{P_{mk}}^* f_{mk}\right] / \mathbb{E}\left[|y_{P_{mk}}|^2\right]\right) y_{P_{mk}} = c_{P_{mk}} y_{P_{mk}}, \tag{6a}$$

$$\hat{g}_{nl} = \left(\mathbb{E}\left[y_{S_{nl}}^* g_{nl}\right] / \mathbb{E}\left[|y_{S_{nl}}|^2\right]\right) y_{S_{nl}} = c_{S_{nl}} y_{S_{nl}}, \tag{6b}$$

where $c_{P_{mk}}$ and $c_{S_{nl}}$ are given by

$$c_{P_{mk}} = \frac{\sqrt{\tau_p P}\zeta_{f_{mk}}}{\tau_p P \left(\zeta_{f_{mk}} + \zeta_{v_{mk}}\right) + 1}, \tag{7a}$$

$$c_{S_{nl}} = \frac{\sqrt{\tau_p P}\zeta_{g_{nl}}}{\tau_p P \left(\zeta_{g_{nl}} + \zeta_{u_{nl}}\right) + 1}. \tag{7b}$$

Proof: Appendix A-A. ■

Remark 1: The MMSE channel estimates in (6a) and (6b) are valid for the coexistence of PUs/SUs with shared pilot sequences defined by (3). The MMSE channel estimation with the conventional orthogonal pilots for cell-free massive MIMO is reported in [8].

Due to TDD channel reciprocity, P-APs and S-APs utilize locally estimated $\hat{f}_{mk}$ and $\hat{g}_{nl}$ as DL CSI to construct their precoders [34]. Furthermore, the actual channels can be written as

$$f_{mk} = \hat{f}_{mk} + \epsilon_{f_{mk}} \quad \text{and} \quad g_{nl} = \hat{g}_{nl} + \epsilon_{g_{nl}}, \tag{8}$$

where $\epsilon_{f_{mk}}$ and $\epsilon_{g_{nl}}$ are the estimation errors, which are independent of the corresponding channel estimates yielded from orthogonality property of MMSE criterion [36].

For NOMA, the UL channel (f_{mak}) is also estimated locally at P-APs from the pilots sent by PUs within PCs. Again, we assume that $Q \leq \min(A, B)$ number of pilots is shared among PCs and SCs as

$$\boldsymbol{\Phi}_{PA} = \left[\boldsymbol{\Phi}; \tilde{\boldsymbol{\Phi}}_{PA}\right] \quad \text{and} \quad \boldsymbol{\Phi}_{SB} = \left[\boldsymbol{\Phi}; \tilde{\boldsymbol{\Phi}}_{SB}\right], \qquad (9)$$

where $\boldsymbol{\Phi} \in \mathbb{C}^{Q \times \tau_p}$ denotes the shared Q-pilot sequence among PCs and SCs. Then, the remaining $(A - Q)$ PCs or $(B - Q)$ SCs are assigned with the pilot sequences $\tilde{\boldsymbol{\Phi}}_{PA} \in \mathbb{C}^{(Q-A) \times \tau_p}$ and $\tilde{\boldsymbol{\Phi}}_{SB} \in \mathbb{C}^{(Q-B) \times \tau_p}$, respectively. Furthermore, $\boldsymbol{\Phi}_{PA} = \left[\boldsymbol{\phi}_{P_1}^T, \cdots, \boldsymbol{\phi}_{P_a}^T, \cdots, \boldsymbol{\phi}_{P_A}^T\right]^T$ and $\boldsymbol{\Phi}_{SB} = \left[\boldsymbol{\phi}_{S_1}^T, \cdots, \boldsymbol{\phi}_{S_b}^T, \cdots, \boldsymbol{\phi}_{S_B}^T\right]^T$, where $\boldsymbol{\phi}_{P_a} \in \mathbb{C}^{1 \times \tau_p}$ and $\boldsymbol{\phi}_{S_b} \in \mathbb{C}^{1 \times \tau_p}$ are the pilot sequences assigned to the ath PC and the bth SC, satisfying $\|\boldsymbol{\phi}_{P_a}\|^2 = 1$ and $\|\boldsymbol{\phi}_{S_b}\|^2 = 1$ for $a \in \{1, \cdots, A\}$ and $b \in \{1, \cdots, B\}$. The received pilot vector at the mth P-AP can be written as

$$\mathbf{y}_{P_m} = \sqrt{P_p} \sum_{a=1}^{A} \sum_{k=1}^{K} f_{mak} \boldsymbol{\phi}_{P_a}$$
$$+ \sqrt{P_p} \sum_{b=1}^{B} \sum_{l=1}^{L} v_{mbl} \boldsymbol{\phi}_{S_b} + \mathbf{n}_{P_m}, \quad (10)$$

where $\mathbf{n}_{P_m} \sim \mathcal{CN}(0, 1)$ is the AWGN vector at the mth P-AP. Then, by projecting $\boldsymbol{\phi}_{P_a}^H \in \boldsymbol{\Phi}$ into (10), we obtain a sufficient statistic to estimate f_{mak} as

$$y_{P_{ma}} = \boldsymbol{\phi}_{P_a}^H \mathbf{y}_{P_m}$$
$$= \sqrt{P_p} \sum_{i=1}^{K} f_{mak} + \sqrt{P_p} \sum_{j=1}^{L} v_{maj} + n_{P_m}. \quad (11)$$

By following steps similar to those in Appendix A-A, the MMSE estimate of f_{mak} can be derived as

$$\hat{f}_{mak} = \frac{\mathbb{E}\left[y_{P_{ma}}^* f_{mak}\right]}{\mathbb{E}\left[|y_{P_{ma}}|^2\right]} y_{P_{ma}}$$
$$= \frac{\sqrt{\tau_p P} \zeta_{f_{mak}}}{\tau_p P \left(\sum_{i=1}^{K} \zeta_{f_{mai}} + \sum_{j=1}^{L} \zeta_{v_{maj}}\right) + 1} y_{P_{ma}}. \quad (12)$$

Since $y_{P_{ma}}$ is Gaussian distributed, we have $\hat{f}_{mak} \sim \mathcal{CN}(0, \alpha_{f_{mak}})$, where $\alpha_{f_{mak}}$ is given as

$$\alpha_{f_{mak}} = \mathbb{E}\left[|\hat{f}_{mak}|^2\right] = \frac{\tau_p P \zeta_{f_{mak}}^2}{\tau_p P \left(\sum_{i=1}^{K} \zeta_{f_{mai}} + \sum_{j=1}^{L} \zeta_{v_{maj}}\right) + 1}. \quad (13)$$

Then, the channel estimation error of f_{mak} is defined as $\epsilon_{f_{mak}} = f_{mak} - \hat{f}_{mak} \sim \mathcal{CN}(0, \zeta_{f_{mak}} - \alpha_{f_{mak}})$.

Remark 2: The parameters $y_{S_{nb}}$, $\hat{g}_{nbl}$, and $\alpha_{g_{nbl}}$ correspond ing to the secondary system can be obtained by replacing subscripts $\{P, m, a, K, k, i\}$ of (11), (12), and (13), respectively, by $\{S, n, b, L, l, j\}$. Hence, the explicit presentation of UL channel estimates at S-APs is omitted for the sake of brevity.

C. AP Clustering for OMA

The P-APs/S-APs are clustered to serve a particular set of users based on the locally estimated channel gains. We assume that only $\mathcal{M}_P = \{1, \cdots, M_P\}$ and $\mathcal{N}_S = \{1, \cdots, N_S\}$ sets of P-APs/S-APs are assigned to $U_P(k)$ and $U_S(l)$, respectively [37]. At each P-AP/S-AP, a set of transmission determination coefficients can be introduced based on the estimated channels such that $\delta_{P_{mk}} = \delta_{S_{nl}} = 1$ for $m \in \mathcal{M}_P$, $n \in \mathcal{N}_S$, and $\delta_{P_{mk}} = \delta_{S_{nl}} = 0$, otherwise.

Remark 3: When the aforementioned AP clustering is adopted for the proposed cell-free underlay spectrum sharing, the primary/secondary CCI can be reduced, and hence, the achievable rates of both primary/secondary systems can be boosted with respect to a conventional unclustered system as depicted in Fig. 5 in our numerical results in Section VIII.

D. Signal Model for OMA

Due to implementation simplicity and near optimal performance in large AP regime [34], conjugate precoding is used at P-APs and S-APs to transmit signals towards their respective users via the channel estimates in (6a) and (6b). We can write the transmitted signals at the mth P-AP and the nth S-AP as

$$x_{P_m} = \sqrt{P_P} \sum_{k=1}^{K} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} \hat{f}_{mk}^* q_{P_k}, \qquad (14a)$$

$$x_{S_n} = \sqrt{P_S} \sum_{l=1}^{L} \delta_{S_{nl}} \eta_{S_{nl}}^{1/2} \hat{g}_{nl}^* q_{S_l}, \qquad (14b)$$

where $\eta_{P_{mk}}$ and $\eta_{S_{nl}}$ are the power allocation coefficients at the mth P-AP and the nth S-AP, respectively. Moreover, P_P and P_S denote the maximum allowable transmit powers at each P-AP and S-AP, respectively. Here, $\eta_{S_{nl}}$ is selected to satisfy the total transmit power constraints given in (28). The signals intended to $U_P(k)$ and $U_S(l)$ are denoted by q_{P_k} and q_{S_l}, respectively, and they satisfy $\mathbb{E}\left[|q_{P_k}|^2\right] = 1$ and $\mathbb{E}\left[|q_{S_l}|^2\right] = 1$. Then, the received signals at $U_P(k)$ and $U_S(l)$ can be written as

$$r_{P_k} = \sum_{m=1}^{M} f_{mk} x_{P_m} + \sum_{n=1}^{N} u_{nk} x_{S_n} + n_{P_k}, \quad (15a)$$

$$r_{S_l} = \sum_{n=1}^{N} g_{nl} x_{S_n} + \sum_{m=1}^{M} v_{ml} x_{P_m} + n_{S_l}, \qquad (15b)$$

where $n_{P_k} \sim \mathcal{CN}(0, 1)$ and $n_{S_l} \sim \mathcal{CN}(0, 1)$ are AWGN at $U_P(k)$ and $U_S(l)$, respectively. Here, x_{P_m} and x_{S_n} are given in (14a). We can rearrange the received signal at $U_P(k)$ in (15a) as

$$r_{P_k} = \sqrt{P_P} \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} f_{mk} \hat{f}_{mk}^* q_{P_k}$$
$$+ \sqrt{P_P} \sum_{m=1}^{M} \sum_{i \neq k}^{K} \delta_{P_{mi}} \eta_{P_{mi}}^{1/2} f_{mk} \hat{f}_{mi}^* q_{P_i}$$
$$+ \sqrt{P_S} \sum_{n=1}^{N} \sum_{j=1}^{L} \delta_{S_{nj}} \eta_{S_{nj}}^{1/2} u_{nk} \hat{g}_{nj}^* q_{S_j} + n_{P_k}, \quad (16)$$

where the first term represents the desired signal component at $U_P(k)$, while the inter-system interference caused by beamforming uncertainty of conjugate precoding is captured

by the second term. The third term accounts for the intra-system interference yielded from the secondary CCI. Similarly, the received signal at $U_S(l)$ in (15b) is rewritten as

$$r_{S_l} = \sqrt{P_S} \sum_{n=1}^{N} \delta_{S_{nl}} \eta_{S_{nl}}^{1/2} g_{nl} \hat{g}_{nl}^* q_{S_l}$$
$$+ \sqrt{P_S} \sum_{n=1}^{N} \sum_{j \neq l}^{L} \delta_{S_{nj}} \eta_{S_{nj}}^{1/2} g_{nl} \hat{g}_{nj}^* q_{S_j}$$
$$+ \sqrt{P_P} \sum_{m=1}^{M} \sum_{i=1}^{K} \delta_{P_{mi}} \eta_{P_{mi}}^{1/2} v_{ml} \hat{f}_{mi}^* q_{P_i} + n_{S_l}. \quad (17)$$

E. Signal Model for NOMA

Again, P-APs employ conjugate precoding by using the channel estimates in (12). The transmitted signal at the mth P-AP can be written as

$$x_{P_m} = \sqrt{P_P} \sum_{a=1}^{A} \sum_{i=1}^{K} \eta_{P_{mai}}^{1/2} \hat{f}_{mai}^* q_{P_{ai}}, \quad (18)$$

where $\eta_{P_{mai}}$ is the transmit power control coefficient at the mth P-AP. Here, $\mathbb{E}\left[|q_{P_{ai}}|^2\right] = 1$, where $q_{P_{ai}}$ is the signal intended for $U_P(a,i)$. Thus, the received signal at $U_P(a,k)$ can be written as

$$r_{P_{ak}} = \sqrt{P_P} \sum_{m=1}^{M} \eta_{P_{mak}}^{1/2} f_{mak} \hat{f}_{mak}^* q_{P_{ak}}$$
$$+ \sqrt{P_P} \sum_{m=1}^{M} \sum_{i \neq k}^{K} \eta_{P_{mai}}^{1/2} f_{mak} \hat{f}_{mai}^* q_{P_{ai}}$$
$$+ \sqrt{P_P} \sum_{m=1}^{M} \sum_{a' \neq a}^{A} \sum_{i=1}^{K} \eta_{P_{ma'i}}^{1/2} f_{mak} \hat{f}_{ma'i}^* q_{P_{a'i}}$$
$$+ \sqrt{P_S} \sum_{n=1}^{N} \sum_{b=1}^{B} \sum_{j=1}^{L} \eta_{S_{nbj}}^{1/2} u_{nak} \hat{g}_{nbj}^* q_{S_{bj}}$$
$$+ n_{P_{ak}}, \quad (19)$$

where $n_{P_{ak}} \sim \mathcal{CN}(0,1)$ is the AWGN at $U_P(a,k)$. To apply the power-domain NOMA, we assume that the users in the ath PC are ordered based on the effective channel strength as [12], [25], [35]

$$\mathbb{E}\left[\left|\sum_{m=1}^{M} \hat{f}_{ma1}\right|^2\right] \geq \cdots \geq \mathbb{E}\left[\left|\sum_{m=1}^{M} \hat{f}_{mak}\right|^2\right]$$
$$\geq \cdots \geq \mathbb{E}\left[\left|\sum_{m=1}^{M} \hat{f}_{maK}\right|^2\right]. \quad (20)$$

In power-domain NOMA, higher transmit powers are allocated for the users with weaker channel conditions. Thus, the transmit powers are ordered as [12], [25], [35]

$$P_{P_{a1}} \leq \cdots \leq P_{P_{ak}} \leq \cdots \leq P_{P_{aK}}, \quad (21)$$

where $P_{P_{ak}} = P_P \eta_{P_{mak}}$. Consequently, $U_P(a,k)$ aims to decode the signal intended for $U_P(a,i)$ for $\forall i \geq k$ provided

that $U_P(a,k)$ can decode its own signal. Hence, $U_P(a,k)$ may successively cancel the intra-cluster interference from $U_P(a,i)$ before decoding its own signal for $\forall i \geq k$, and the residual interference due to SIC error propagation must also captured. Moreover, $U_P(a,k)$ treats the signals for users $\forall i < k$ as interference [12], [25], [35]. To this end, the received signal at $U_P(a,k)$ upon imperfect SIC with error propagation can be written as (22), shown at the bottom of the page.

Remark 4: In our proposed cell-free NOMA-aided underlay spectrum sharing, the perfect SIC is not feasible due to intra-cluster pilot contamination, intra-system interference, inter-system interference, channel estimation errors, and statistical CSI knowledge at the users. Thus, the residual interference caused by imperfect SIC needs to be modeled. The third term in (22) captures the error propagation due to imperfect SIC in which $\hat{q}_{P_{ai}}$ is the estimate of $q_{P_{ai}}$. Since $q_{P_{ai}}$ is Gaussian distributed, $\hat{q}_{P_{ai}}$ and $q_{P_{ai}}$ are assumed to be jointly Gaussian distributed with a normalized correlation coefficient $\vartheta_{P_{ai}}$ as [12]

$$q_{P_{ai}} = \vartheta_{P_{ai}} \hat{q}_{P_{ai}} + e_{P_{ai}}, \quad (23)$$

where $\hat{q}_{P_{ai}} \sim \mathcal{CN}(0,1)$, $e_{P_{ai}} \sim \mathcal{CN}(0, \sigma_{e_{P_{ai}}}^2 / (1 + \sigma_{e_{P_{ai}}}^2))$, and $\vartheta_{P_{ai}} = 1/\sqrt{1 + \sigma_{e_{P_{ai}}}^2}$. Furthermore, $\hat{q}_{P_{ai}}$ and $e_{P_{ai}}$ are statistically independent. Thus, the third term in (22) can be used to capture the residual interference caused by error propagation of imperfect SIC when evaluating the SINR and achievable rate.

Remark 5: The received signal at $U_S(b,l)$ for the secondary system $(r_{S_{bl}})$ and the signal intended for $(U_S(b,j), q_{S_{bj}})$ can be obtained by replacing the subscripts $\{P, M, m, A, a, K, k, i\}$ of (22) and (23), respectively, by $\{S, N, n, B, b, L, l, j\}$.

F. Secondary Transmit Power Constraints for OMA

We constrain the transmit power at each S-AP to guarantee that the secondary CCI inflected at PUs falls below the PIT of each PU. Thus, we define the total transmit power at the nth S-AP as

$$P_{S_n} \triangleq \sum_{l=1}^{L} P_S \delta_{S_{nl}} \eta_{S_{nl}}, \quad (24)$$

where $n \in \{1, \cdots, N\}$, $l \in \{1, \cdots, L\}$, and $\sum_{l=1}^{L} \delta_{S_{nl}} \eta_{S_{nl}} \leq 1$. Moreover, the transmit power allocation coefficient at the nth S-AP for $U_S(l)$ is represented by $\eta_{S_{nl}}$. Thus, the CCI received at $U_P(k)$ from all S-APs can be

---

$$r_{P_{ak}} = \underbrace{\sqrt{P_P} \sum_{m=1}^{M} \eta_{P_{mak}}^{1/2} f_{mak} \hat{f}_{mak}^* q_{P_{ak}}}_{\text{Desired signal}} + \underbrace{\sqrt{P_P} \sum_{m=1}^{M} \sum_{i=1}^{k-1} \eta_{P_{mai}}^{1/2} f_{mak} \hat{f}_{mai}^* q_{P_{ai}}}_{\text{Intra-cluster interference after SIC}}$$

$$+ \underbrace{\sqrt{P_P} \sum_{m=1}^{M} \sum_{i=k+1}^{K} \eta_{P_{mai}}^{1/2} \left(f_{mak} \hat{f}_{mai}^* q_{P_{ai}} - \mathbb{E}\left[f_{mak} \hat{f}_{mai}^*\right] \hat{q}_{P_{ai}} \right)}_{\text{Error propagation due to imperfect SIC}}$$

$$+ \underbrace{\sqrt{P_P} \sum_{m=1}^{M} \sum_{a' \neq a}^{A} \sum_{i=1}^{K} \eta_{P_{ma'i}}^{1/2} f_{mak} \hat{f}_{ma'i}^* q_{P_{a'i}}}_{\text{Intra-system interference}} + \underbrace{\sqrt{P_S} \sum_{n=1}^{N} \sum_{b=1}^{B} \sum_{j=1}^{L} \eta_{S_{nbj}}^{1/2} u_{nak} \hat{g}_{nbj}^* q_{S_{bj}}}_{\text{Inter-system interference}} + \underbrace{n_{P_{ak}}}_{\text{AWGN}} \quad (22)$$

written as

$$y_k = \sum_{n=1}^{N} u_{nk} x_{S_n} = \sqrt{P_S} \sum_{n=1}^{N} \sum_{l=1}^{L} \delta_{S_{nl}} \eta_{S_{nl}}^{1/2} u_{nk} \hat{g}_{nl}^* q_{S_l}, \quad (25)$$

where x_{S_n} is the transmitted signal at the nth S-AP and defined in (14b).

Proposition 2: The total average secondary CCI power (P_{I_k}) *inflicted at* $U_P(k)$ *is given by*

$$P_{I_k} = \mathbb{E}\left[|y_k|^2\right] = P_S \underbrace{\sum_{n=1}^{N} \sum_{l=1}^{L} \delta_{S_{nl}} \eta_{S_{nl}} \mathbb{E}\left[|u_{nk}\hat{g}_{nl}^*|^2\right]}_{Z_k},$$
$$(26)$$

where Z_k *is defined as*

$$Z_k = \sum_{n=1}^{N} \sum_{l=1}^{L} \delta_{S_{nl}} \eta_{S_{nl}} \rho_{g_{nl}} \zeta_{u_{nk}} + \sum_{n=1}^{N} \delta_{S_{nk}} \eta_{S_{nk}} \rho_{u_{nk}}^2, \quad (27)$$

where $\rho_{g_{nl}} \triangleq \sqrt{\tau_p P} c_{S_{nl}} \zeta_{g_{nl}}$ *and* $\rho_{u_{nk}} \triangleq \sqrt{\tau_p P} c_{S_{nk}} \zeta_{u_{nk}}$.

Proof: Appendix A-B. ∎

Thus, we can give the secondary transmit power constraint at the nth S-AP as follows:

$$P_{S_n} = \min\left(P_S, I_{T_1}/Z_1, \cdots, I_{T_k}/Z_k, \cdots, I_{T_K}/Z_K\right), \quad (28)$$

where I_{T_k} is the interference threshold of $U_P(k)$.

G. Secondary Transmit Power Control for NOMA

As per Section II-F, to ensure that the performance of primary system is not hindered by the secondary system, we constrain the transmit power of S-APs. The secondary CCI received at $U_P(a,k)$ is given by

$$y_{ak} = \sum_{n=1}^{N} u_{nak} x_{S_{nb}}$$
$$= \sqrt{P_S} \sum_{n=1}^{N} \sum_{b=1}^{B} \sum_{l=1}^{L} \eta_{S_{nbl}}^{1/2} u_{nak} \hat{g}_{nbl}^* q_{S_{bl}}, \quad (29)$$

where $x_{S_{nb}}$ is the transmit signal intended for L users in the bth SC. Then, the total average secondary CCI $(P_{I_{ak}})$ inflicted at $U_P(a,k)$ can be derived as

$$P_{I_{ak}} = \mathbb{E}\left[|y_{ak}|^2\right]$$
$$= P_S \underbrace{\sum_{n=1}^{N} \sum_{b=1}^{B} \sum_{l=1}^{L} \eta_{S_{nl}} \mathbb{E}\left[|u_{nak}\hat{g}_{nbl}^*|^2\right]}_{Z_{ak}}, \quad (30)$$

where Z_{ak} can be derived by following steps similar to those in Appendix A-B as

$$Z_{ak} = \sum_{n=1}^{N} \sum_{b=1}^{B} \sum_{l=1}^{L} \eta_{S_{nbl}} \alpha_{g_{nbl}} \zeta_{u_{nak}}$$
$$+ \sum_{n=1}^{N} \sum_{l=1}^{L} \eta_{S_{nal}} \alpha_{f_{mak}}^2 \left(\frac{\zeta_{u_{uak}} \zeta_{g_{nal}}}{\zeta_{f_{mak}}}\right)^2. \quad (31)$$

Thus, we derive the secondary transmit power constraint at the nth S-AP as

$$P_{S_n} = \min\left(P_S, I_{T_{11}}/Z_{11}, \cdots, I_{T_{ak}}/Z_{ak}, \cdots, I_{T_{AK}}/Z_{AK}\right), \quad (32)$$

where $I_{T_{ak}}$ is the interference threshold of $U_P(a,k)$.

III. ACHIEVABLE RATE ANALYSIS OF OMA-AIDED UNDERLAY SPECTRUM SHARING

A. Achievable Rate Analysis for the Primary System

When the P-APs/S-APs do not beamform DL pilots for the acquisition of DL CSI, the PUs/SUs are unaware of the instantaneous DL channel coefficients. Thus, the PUs/SUs must rely on the long-term/statistical DL channel coefficients for signal detection [34]. This is a typical scenario in TDD-based co-located massive MIMO in which only UL pilots are used to estimate UL channels at the P-APs/S-APs, and the instantaneous DL channel coefficients can be approximated by their statistical counterparts thanks to channel hardening property. To this end, the signal received at $U_P(k)$ can be rearranged as (33), shown at the bottom of the page. In (33), the effective noise consists of the sum of (i) interference caused by detection uncertainty, (ii) intra-system interference due to beamforming uncertainty, (iii) inter-system interference due to secondary CCI, and (iv) AWGN. The desired signal term and effective noise are uncorrelated. Due to the law of large number, the latter can be treated as worst-case independently distributed Gaussian noise [34], [38].

Theorem 1: The effective SINR at $U_P(k)$ *is given by*

$$\gamma_{P_k} = \frac{P_P \left|\mathbb{E}\left[\sum_{m=1}^{M} \delta_{P_{mik}} \eta_{P_{mk}}^{1/2} f_{mk} \hat{f}_{mk}^*\right]\right|^2}{P_P \mathrm{Var}\left[\sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} f_{mk} \hat{f}_{mk}^*\right] + \sum_{j=1}^{2} I_{Pj} + \mathbb{E}\left[|n_{P_k}|^2\right]}, \quad (34)$$

where I_{Pj} *for* $j \in \{1, 2\}$ *can be defined as*

$$I_{P1} = P_P \mathbb{E}\left[\left|\sum_{i \neq k}^{K} \sum_{m=1}^{M} \delta_{P_{mi}} \eta_{P_{mi}}^{1/2} f_{mk} \hat{f}_{mi}^*\right|^2\right], \quad (35a)$$

$$I_{P2} = P_S \mathbb{E}\left[\left|\sum_{j=1}^{L} \sum_{n=1}^{N} \delta_{S_{nj}} \eta_{S_{nj}}^{1/2} u_{nk} \hat{g}_{nj}^*\right|^2\right]. \quad (35b)$$

By evaluating the expectation and variance terms in (34) and (35), the SINR is given as (36), shown at the bottom of the next page, where $\rho_{f_{mk}} \triangleq \sqrt{\tau_p P} c_{P_{mk}} \zeta_{f_{mk}}$.

Proof: Appendix B. ∎

$$r_{P_k} = \underbrace{\sqrt{P_P}\mathbb{E}\left[\sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} f_{mk} \hat{f}_{mk}^*\right] q_{P_k}}_{\text{Desired signal}} + \underbrace{\sqrt{P_P}\left(\sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} f_{mk} \hat{f}_{mk}^* - \mathbb{E}\left[\sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} f_{mk} \hat{f}_{mk}^*\right]\right) q_{P_k}}_{\text{Detection uncertainty}}$$
$$+ \underbrace{\sqrt{P_P}\sum_{i \neq k}^{K} \sum_{m=1}^{M} \delta_{P_{mi}} \eta_{P_{mi}}^{1/2} f_{mk} \hat{f}_{mi}^* q_{P_i}}_{\text{Inter-system interference caused by beamforminmg uncertainty}} + \underbrace{\sqrt{P_S}\sum_{j=1}^{L} \sum_{n=1}^{N} \delta_{S_{nj}} \eta_{S_{nj}}^{1/2} u_{nk} \hat{g}_{nj}^* q_{S_j}}_{\text{Inter-system interference caused by secondary CCI}} + \underbrace{n_{P_k}}_{\text{AWGN}} \quad (33)$$

Then, we define the achievable rate of $U_P(k)$ as follows:

$$R_{P_k} = ((\tau_c - \tau_p)/\tau_c) \log_2 (1 + \gamma_{P_k}), \qquad (37)$$

where the effective portion of coherence interval for payload data transmission is captured by the pre-log factor $((\tau_c - \tau_p)/\tau_c)$ and γ_{P_k} is defined in (36).

B. Achievable Rate Definition for the Secondary System

By following **Theorem** 1, we derive the achievable rate of $U_S(l)$ as follows:

$$R_{S_l} = ((\tau_c - \tau_p)/\tau_c) \log_2 (1 + \gamma_{S_l}), \qquad (38)$$

$$\gamma_{S_l} = \frac{P_S \left| \mathbb{E} \left[\sum_{n=1}^{N} \delta_{S_{nl}} \eta_{S_{nl}}^{1/2} g_{nl} \hat{g}_{nl}^* \right] \right|^2}{P_S \mathrm{Var} \left[\sum_{n=1}^{N} \delta_{S_{nl}} \eta_{S_{nl}}^{1/2} g_{nl} \hat{g}_{nl}^* \right] + \sum_{j=1}^{2} I_{Sj} + \mathbb{E} \left[|w_{S_l}|^2 \right]}, \qquad (39)$$

where I_{Sj} for $j \in \{1, 2\}$ can be defined as

$$I_{S_1} = P_S \mathbb{E} \left[\left| \sum_{j \neq l}^{L} \sum_{n=1}^{N} \delta_{S_{nj}} \eta_{S_{nj}}^{1/2} g_{nl} \hat{g}_{nj}^* \right|^2 \right], \qquad (40a)$$

$$I_{S_2} = P_P \mathbb{E} \left[\left| \sum_{i=1}^{K} \sum_{m=1}^{M} \delta_{P_{mi}} \eta_{P_{mi}}^{1/2} v_{ml} \hat{f}_{mi}^* \right|^2 \right]. \qquad (40b)$$

Thus, by deriving the expectation and variance term in (39) in similar manner to (34), the SINR of $U_S(l)$ can be computed as (41), shown at the bottom of the page.

IV. TRANSMIT POWER CONTROL

In cell-free massive MIMO, the user-fairness must be guaranteed in terms of the achievable rate in order to provide a uniform quality-of-service to all users. To this end, max-min power control algorithms have been shown to be optimal in the sense of user-fairness in presence of near-far effects [34], [39]–[41]. For our proposed system model, a multi-objective optimization problem (MOOP) [42] is most suited as both primary and secondary systems operate simultaneously. Furthermore, the secondary transmit power constraints in (28) must also be considered when formulating this MOOP.

We compute the optimal power allocation coefficients of P-APs/S-APs to maximize the minimum achievable DL rate among all PUs/SUs by invoking the max-min optimization criterion [34]. Since the rates in (37) and (38) are monotonically increasing functions of their arguments, we can equivalently replace R_{P_k} and R_{S_l} by γ_{P_k} and γ_{S_l}, which are given in (36) and (41), respectively. Thus, we formulate a max-min transmit

power control problem by introducing a common SINR (λ_γ) for primary/secondary systems and by defining slack variables $\beta_{P_{mk}} \triangleq \eta_{P_{mk}}^{1/2}$ and $\beta_{S_{nl}} \triangleq \eta_{S_{nl}}^{1/2}$ as

$$\underset{\beta_{P_{mk}}, \beta_{S_{nl}} \forall m, k, n, l}{\text{maximize}} \quad (\gamma_{P_k})^{w_P} (\gamma_{S_l})^{w_S} = \lambda_\gamma, \qquad (42a)$$

$$\text{subject to } C_1 : \gamma_{P_k} \geq \lambda_\gamma \quad \text{and} \quad \gamma_{S_l} \geq \lambda_\gamma, \qquad (42b)$$

$$C_2 : \sum_{k=1}^{K} \delta_{P_{mk}} \beta_{P_{mk}}^2 \rho_{f_{mk}} \leq 1$$

$$\text{and} \quad \sum_{l=1}^{L} \delta_{S_{nl}} \beta_{S_{nl}}^2 \rho_{g_{nl}} \leq 1, \qquad (42c)$$

$$C_3 : P_{S_n} \leq I_{T_k}/Z_k, \qquad (42d)$$

$$C_4 : 0 \leq \beta_{P_{mk}} \quad \text{and} \quad 0 \leq \beta_{S_{nl}}, \qquad (42e)$$

where w_P and w_S are the priorities assigned for primary and secondary achievable SINR, respectively. Moreover, C_2 is obtained by invoking the maximum allowable transmit power constraints at the mth P-AP and the lth S-AP as follows:

$$\mathbb{E} \left[|x_{P_m}|^2 \right] \leq P_P \Rightarrow \mathbb{E} \left[\left| \sqrt{P_P} \sum_{k=1}^{K} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} \hat{f}_{mk}^* q_{P_k} \right|^2 \right] \leq P_P$$

$$\sum_{k=1}^{K} \delta_{P_{mk}} \eta_{P_{mk}} \mathbb{E} \left[|\hat{f}_{mk}^* q_{P_k}|^2 \right] \leq 1 \Rightarrow \sum_{k=1}^{K} \delta_{P_{mk}} \eta_{P_{mk}} \rho_{f_{mk}} \leq 1. \qquad (43)$$

It can be shown that $\sum_{l=1}^{L} \delta_{S_{nl}} \eta_{S_{nl}} \rho_{g_{nl}} \leq 1$. We adopt the secondary transmit power constraint in (28) to obtain C_3. Since the objective functions in (42) are quasi-concave functions, we can show that the underlaying optimization problem is also quasi-concave [8]. Thus, an optimal solution can be found by using the Bisection method as shown in Algorithm 1.

V. THE IMPLICATION OF DL PILOT TRANSMISSION

In our achievable rate analysis in Section III-A, the PUs/SUs are assumed to be unaware of DL channel estimates, and this is a typical assumption in co-located massive MIMO literature [34]. It is aimed at minimizing the pilot overhead to preserve system scalability. Thus, the users adopt long-term DL channel statistics to decode the received signals as the DL channel coefficients can be tightly approximated by their average counterparts by virtue of channel hardening [38]. Nevertheless, it has been shown in [43] that channel hardening in cell-free massive MIMO occurs only when a large number of APs is distributed in close-vicinity, and hence, the adopting statistical DL CSI for signal decoding at users may hinder the system performance. To circumvent this, DL pilots can be beamformed to estimate DL channels at the users, and this approach ensures that the DL pilot sequence length does not

$$\gamma_{P_k} = \frac{P_P \left(\sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} \rho_{f_{mk}} \right)^2}{P_P \sum_{m=1}^{M} \sum_{i=1}^{K} \delta_{P_{mi}} \eta_{P_{mi}} \rho_{f_{mi}} \zeta_{f_{mk}} + P_S \sum_{n=1}^{N} \sum_{j=1}^{L} \delta_{S_{nj}} \eta_{S_{nj}} \rho_{g_{nj}} \zeta_{u_{nk}} + P_S \sum_{n=1}^{N} \delta_{S_{nk}} \eta_{S_{nk}} \rho_{u_{nk}}^2 + 1} \qquad (36)$$

$$\gamma_{S_l} = \frac{P_S \left(\sum_{n=1}^{N} \delta_{S_{nl}} \eta_{S_{nl}}^{1/2} \rho_{g_{nl}} \right)^2}{P_S \sum_{n=1}^{N} \sum_{j=1}^{L} \delta_{S_{nj}} \eta_{S_{nj}} \rho_{g_{nj}} \zeta_{g_{nl}} + P_P \sum_{m=1}^{M} \sum_{i=1}^{K} \delta_{P_{mi}} \eta_{P_{mi}} \rho_{f_{mi}} \zeta_{v_{ml}} + P_P \sum_{m=1}^{M} \delta_{P_{ml}} \eta_{P_{ml}} \rho_{v_{ml}}^2 + 1} \qquad (41)$$

Algorithm 1 Bisection Algorithm

Input: Path-losses between all P-APs/S-APs and PUs/SUs, and the average transmit powers of PUs/SUs.

Output: The power control coefficients $\eta_{P_{mk}}$ for $m \in \{1, \cdots, M\}$, $k \in \{1, \cdots, K\}$, and $\eta_{S_{nl}}$ for $n \in \{1, \cdots, N\}$, $l \in \{1, \cdots, L\}$.

Initialization: Define an initial region for the objective functions by choosing appropriate values for λ_{min} and λ_{max}. Choose a tolerance $\epsilon > 0$.

1: **while** $\lambda_{max} - \lambda_{min} > \epsilon$ **do**
2:    Calculate, $\lambda_\gamma = (\lambda_{max} + \lambda_{min})/2$.
3:    Solve the convex feasibility problem, which can be formulated as

$$\|\boldsymbol{V}_{\gamma_{P_k}}\| \le \frac{1}{\sqrt{\lambda}}\left(\sum_{m=1}^{M}\delta_{P_{mk}}\beta_{P_{mk}}\rho_{f_{mk}}\right), \quad (44a)$$

$$\|\boldsymbol{V}_{\gamma_{S_l}}\| \le \frac{1}{\sqrt{\lambda}}\left(\sum_{n=1}^{N}\delta_{S_{nl}}\beta_{S_{nl}}\rho_{g_{nl}}\right), \quad (44b)$$

which is subjected to C_2, C_3, and C_4 given in (42c), (42d), and (42e), respectively. Moreover, $\boldsymbol{V}_{\gamma_{P_k}} \triangleq \left[\mathbf{v}_{P_1}^T, \frac{P_S}{P_P}\mathbf{v}_{P_2}^T, \frac{P_S}{P_P}\mathbf{v}_{P_3}^T, \frac{1}{\sqrt{P_P}}\right]$, where

$$\mathbf{v}_{P_1} \triangleq \left[\frac{\sqrt{\rho_{f_{11}}\zeta_{f_{1k}}}}{\delta_{P_{11}}^{-1}\beta_{P_{11}}^{-1}}, \cdots, \frac{\sqrt{\rho_{f_{MK}}\zeta_{f_{Mk}}}}{\delta_{P_{MK}}^{-1}\beta_{P_{MK}}^{-1}}\right]^T, \quad (45a)$$

$$\mathbf{v}_{P_2} \triangleq \left[\frac{\sqrt{\rho_{g_{11}}\zeta_{u_{1k}}}}{\delta_{S_{11}}^{-1}\beta_{S_{11}}^{-1}}, \cdots, \frac{\sqrt{\rho_{g_{NL}}\zeta_{u_{Nk}}}}{\delta_{S_{NL}}^{-1}\beta_{S_{NL}}^{-1}}\right]^T, \quad (45b)$$

$$\mathbf{v}_{P_3} \triangleq \left[\delta_{S_{1k}}\beta_{S_{1k}}c_{P_{1k}}\zeta_{u_{1k}}, \cdots, \delta_{S_{Nk}}\beta_{S_{Nk}}c_{P_{Nk}}\zeta_{u_{Nk}}\right]^T, \quad (45c)$$

and $\boldsymbol{V}_{\gamma_{S_l}} \triangleq \left[\mathbf{v}_{S_1}^T, \frac{P_P}{P_S}\mathbf{v}_{S_2}^T, \frac{P_P}{P_S}\mathbf{v}_{S_3}^T, \frac{1}{\sqrt{P_S}}\right]$, where

$$\mathbf{v}_{S_1} \triangleq \left[\frac{\sqrt{\rho_{g_{11}}\zeta_{g_{1l}}}}{\delta_{S_{11}}^{-1}\beta_{S_{11}}^{-1}}, \cdots, \frac{\sqrt{\rho_{g_{NL}}\zeta_{g_{Nl}}}}{\delta_{S_{NL}}^{-1}\beta_{S_{NL}}^{-1}}\right]^T, \quad (46a)$$

$$\mathbf{v}_{S_2} \triangleq \left[\frac{\sqrt{\rho_{f_{11}}\zeta_{v_{1l}}}}{\delta_{P_{11}}^{-1}\beta_{P_{11}}^{-1}}, \cdots, \frac{\sqrt{\rho_{f_{MK}}\zeta_{v_{Ml}}}}{\delta_{P_{MK}}^{-1}\beta_{P_{MK}}^{-1}}\right]^T, \quad (46b)$$

$$\mathbf{v}_{S_3} \triangleq \left[\delta_{P_{1l}}\beta_{P_{1l}}c_{S_{1l}}\zeta_{v_{1l}}, \cdots, \delta_{P_{Ml}}\beta_{P_{Ml}}c_{S_{Ml}}\zeta_{v_{Ml}}\right]^T. \quad (46c)$$

4:    If the status of the problem is feasible, then set $\lambda_{min} = \lambda_\gamma$, otherwise set $\lambda_{max} = \lambda_\gamma$.
5:    Stop if $\lambda_{max} - \lambda_{min} < \epsilon$. Otherwise go to Step 2.
6: **end while**
7: **return** $\eta_{P_{mk}} = \beta_{P_{mk}}^2$ and $\eta_{S_{nl}} = \beta_{S_{nl}}^2$ for $m \in \{1, \cdots, M\}$, $k \in \{1, \cdots, K\}$, $n \in \{1, \cdots, N\}$, and $l \in \{1, \cdots, L\}$.

---

scale with the number of APs. Next, we investigate the impact of DL pilots for the proposed cell-free massive MIMO with underlay spectrum sharing.

A. DL Channel Estimation

We define the effective DL desired and interference channel coefficients of the primary system based on the signal received at $U_P(k)$ in (16) as follows:

$$\mu_{P_{ki}} \triangleq \sum_{m=1}^{M}\delta_{P_{mi}}\eta_{P_{mi}}^{1/2}f_{mk}\hat{f}_{mi}^*, \quad (47a)$$

$$\lambda_{P_{kj}} \triangleq \sum_{n=1}^{N}\delta_{S_{nj}}\eta_{S_{nj}}^{1/2}u_{nk}\hat{g}_{nj}^*. \quad (47b)$$

To estimate DL channels, we need an additional $\tau_{p,d}$ symbol duration for transmitting DL pilots towards PUs/SUs. We consider the same pilot sharing technique that was used for UL pilot transmission. The same pilot sequences will be used between P-APs/S-APs and PUs/SUs. For the sake of exposition, we denote the primary and secondary DL pilot sequences by $\phi_{P_k,d}$ for $k \in \{1, \cdots, K\}$ and $\phi_{S_l,d}$ for $l \in \{1, \cdots, L\}$, respectively. The pilot signal sent by the mth P-AP can be written as

$$\mathbf{x}_{P_m,d} = \sqrt{P_{p,d}}\sum_{i=1}^{K}\delta_{P_{mi}}\eta_{P_{mi}}^{1/2}\hat{f}_{mi}^*\phi_{P_i,d}, \quad (48)$$

where $P_{p,d} \triangleq \tau_{p,d}P_d$ and P_d is the average DL pilot transmit power at each P-AP/S-AP. Moreover, $\hat{f}_{mi}$ is defined in (6a). Next, the pilot vector received at $U_P(k)$ can be written as

$$\mathbf{y}'_{P_k,d} = \sum_{m=1}^{M}f_{mk}\mathbf{x}_{P_m,d} + \sum_{n=1}^{N}u_{nk}\mathbf{x}_{S_n,d} + \mathbf{n}'_{P_k,d}, \quad (49)$$

where $\mathbf{n}'_{P_k,d}$ is the AWGN vector with i.i.d. $\mathcal{CN}(0,1)$ elements, at $U_P(k)$. Then, we rewrite the received pilot vector at $U_P(k)$ by substituting (48) into (49) as

$$\mathbf{y}'_{P_k,d} = \sqrt{P_{p,d}}\left(\sum_{i=1}^{K}\mu_{P_{ki}}\phi_{P_i,d} + \sum_{j=1}^{L}\lambda_{P_{kj}}\phi_{S_j,d}\right) + \mathbf{n}'_{P_k,d}, \quad (50)$$

where $\mu_{P_{ki}}$ and $\lambda_{P_{kj}}$ are the effective desired and interference DL channels (47). To estimate the effective primary desired DL channel, the sufficient statistics can be obtained by projecting $\phi_{P_k,d}^H$ into (50) as

$$y_{P_k,d} = \phi_{P_k,d}^H\mathbf{y}'_{P_k,d} = \sqrt{P_{p,d}}\left(\mu_{P_{kk}} + \lambda_{P_{kk}}\right) + n_{P_k,d}, \quad (51)$$

where $n_{P_k,d} = \phi_{P_k,d}^H\mathbf{n}'_{P_k,d} \sim \mathcal{CN}(0,1)$ is the AWGN at $U_P(k)$.

Proposition 3: With beamformed DL pilots, the MMSE estimate of $\mu_{P_{kk}}$ (47a) is given by

$$\hat{\mu}_{P_{kk}} = \mathbb{E}\left[\mu_{P_{kk}}\right] + \frac{\mathbb{Cov}\left(\mu_{P_{kk}}y_{P_k,d}^*\right)}{\mathbb{Cov}\left(y_{P_k,d}y_{P_k,d}^*\right)}\left(y_{P_k,d} - \mathbb{E}\left[y_{P_k,d}\right]\right)$$

$$= \mathbb{E}\left[\mu_{P_{kk}}\right] + \frac{\sqrt{P_{p,d}}\mathbb{Var}[\mu_{P_{kk}}]}{P_{p,d}\left(\mathbb{Var}[\mu_{P_{kk}}] + \mathbb{Var}[\lambda_{P_{kk}}]\right) + 1}$$

$$\times \left(y_{P_k,d} - \sqrt{P_{p,d}}\left(\mathbb{E}\left[\mu_{P_{kk}}\right] + \mathbb{E}\left[\lambda_{P_{kk}}\right]\right)\right). \quad (52)$$

Then, the MMSE estimate of $\mu_{P_{kk}}$ is given by evaluating (52) *as* (53)*, shown at the bottom of the next page. In* (53)*, $v_{P_{kk}}$ and $u_{P_{kk}}$ are defined as*

$$v_{P_{kk}} \triangleq \sum_{m=1}^{M}\delta_{P_{mk}}\eta_{P_{mk}}\zeta_{f_{mk}}\rho_{f_{mk}}, \quad (54a)$$

$$u_{P_{kk}} \triangleq \sum_{n=1}^{N}\delta_{S_{nk}}\eta_{S_{nk}}\zeta_{u_{nk}}\rho_{u_{nk}}. \quad (54b)$$

The actual effective/desired DL channel coefficient is given by $\mu_{P_{kk}} = \hat{\mu}_{P_{kk}} + \epsilon_{P_{kk}}^\mu$, where $\epsilon_{P_{kk}}^\mu$ is estimation error, which is independent of respective channel estimate.

Proof: Appendix A-C.   ∎

Remark 6: It is worth noting that the parameters $\mu_{S_{lj}}$, $\lambda_{S_{li}}$, $\hat{\mu}_{S_{ll}}$, $v_{S_{ll}}$, and $u_{S_{ll}}$ corresponding to the secondary system can be obtained by replacing subscripts $\{P, m, M, k, i\}$ in (47), (53), and (54), respectively, by $\{S, n, N, l, j\}$.

B. Primary Achievable DL Rate With DL Pilots

We can rewrite the received signal at $U_P(k)$ in (16) via the effective desired and interference DL channel coefficients as follows:

$$r_{P_k,d} = \sqrt{P_P}\mu_{P_{kk}}q_{P_k} + \sqrt{P_P}\sum_{i\neq k}^{K}\mu_{P_{ki}}q_{P_i}$$
$$+\sqrt{P_S}\sum_{j=1}^{L}\lambda_{P_{kj}}q_{S_j} + n_{P_k,d}, \quad (55)$$

where $n_{P_k,d} \sim \mathcal{CN}(0,1)$ is the AWGN at $U_P(k)$. Then, by using the DL channel estimate at $U_P(k)$, (55) can be rearranged to facilitate decoding the desired signal as (56), shown at the bottom of the page. By using (56), the SINR at $U_P(k)$ can be given as (57), shown at the bottom of the page. From the facts that (i) $\mu_{P_{kk}}$ is Gaussian distributed, (ii) $\hat{\mu}_{P_{kk}}$ and $\epsilon^{\mu}_{P_{kk}}$ are independent, and (iii) $\mu_{P_{kk}}$, $\mu_{P_{ki}}$, and $\lambda_{P_{kj}}$ are independent for $i, j \neq k$ [43], we rewrite the SINR in (57) as given in (58), shown at the bottom of the page. Thereby, we define the achievable DL rate as

$$R_{P_k,d} = (\tau_{d,d}/\tau_c)\,\mathbb{E}_{\hat{\mu}_{P_{kk}}}\left[\log_2\left(1 + \gamma_{P_k,d}\right)\right], \quad (59)$$

where $\tau_{d,d} = \tau_c - (\tau_p + \tau_{p,d})$. From Jensen's inequality, an upper bound for DL rate at $U_P(k)$ is derived as

$$R^{ub}_{P_k,d} = (\tau_{d,d}/\tau_c)\log_2\left(1 + \mathbb{E}_{\hat{\mu}_{P_{kk}}}\left[\gamma_{P_k,d}\right]\right), \quad (60)$$

where $\mathbb{E}_{\hat{\mu}_{kk}}\left[\gamma_{P_k,d}\right]$ is given in (61), as shown at the bottom of the page.

Proposition 4: By evaluating expectation terms in (61), the effective SINR at $U_P(k)$ in the case of beamformed DL pilots

by the P-APs can be derived as in (62), *as shown at the bottom of the page, where* $\kappa^{\epsilon}_{P_{kk}}$ *and* $v_{P_{ki}}$ *are defined as*

$$\kappa^{\epsilon}_{P_{kk}} \triangleq \frac{\left(1+P_{p,d}u_{P_{kk}}\right)^2 v_{P_{kk}} + \left(P_{p,d}v_{P_{kk}}\right)^2 u_{P_{kk}} + P_{p,d}v^2_{P_{kk}}}{\left(P_{p,d}\left(v_{P_{kk}} + u_{P_{kk}}\right) + 1\right)^2}, \quad (63a)$$

$$v_{P_{ki}} \triangleq \sum_{m=1}^{M}\delta_{P_{mi}}\eta_{P_{mi}}\rho_{f_{mi}}\zeta_{f_{mk}}, \quad (63b)$$

where $P_{p,d} \triangleq \tau_{p,d}P_d$.
  Proof: Appendix C. ∎

C. Secondary Achievable DL Rate With DL Pilots

We define an upper bound of the DL achievable rate at $U_S(l)$ via steps similar to (55)-(60) as

$$R^{ub}_{S_l,d} = (\tau_{d,d}/\tau_c)\log_2\left(1 + \mathbb{E}_{\hat{\mu}_{S_{ll}}}\left[\gamma_{S_l,d}\right]\right), \quad (64)$$

where $\mathbb{E}_{\hat{\mu}_{S_{ll}}}\left[\gamma_{S_l,d}\right]$ is given as (65), shown at the bottom of the next page. Then, by following steps similar to (62) to evaluate the expectation terms in (65), the effective SINR at $U_S(l)$ can be written as (66), as shown at the bottom of the next page, where $\kappa^{\epsilon}_{S_{ll}}$ and $v_{S_{lj}}$ are defined as

$$\kappa^{\epsilon}_{S_{ll}} \triangleq \frac{\left(1+P_{p,d}u_{S_{ll}}\right)^2 v_{S_{ll}} + \left(P_{p,d}v_{S_{ll}}\right)^2 u_{S_{ll}} + P_{p,d}v^2_{S_{ll}}}{\left(P_{p,d}\left(v_{S_{ll}} + u_{S_{ll}}\right) + 1\right)^2}, \quad (67a)$$

$$v_{S_{lj}} \triangleq \sum_{n=1}^{N}\delta_{S_{nj}}\eta_{S_{nj}}\rho_{g_{nj}}\zeta_{g_{nl}}. \quad (67b)$$

Remark 7: We reveal through our numerical results in Section VIII that the adoption of estimated DL CSI at the PUs/SUs for signal decoding can be exploited to boost the achievable rates by circumventing the less prevalent channel hardening property in cell-free massive MIMO compared to that of co-located counterpart. The underlying implication is

---

$$\hat{\mu}_{P_{kk}} = \frac{\sqrt{P_{p,d}}v_{P_{kk}}y_{P_k,d} + \sum_{m=1}^{M}\delta_{P_{mk}}\eta^{1/2}_{P_{mk}}\rho_{f_{mk}} + P_{p,d}\left(u_{P_{kk}}\sum_{m=1}^{M}\delta_{P_{mk}}\eta^{1/2}_{P_{mk}}\rho_{f_{mk}} - v_{P_{kk}}\sum_{n=1}^{N}\delta_{S_{nk}}\eta^{1/2}_{S_{nk}}\rho_{u_{nk}}\right)}{P_{p,d}\left(v_{P_{kk}} + u_{P_{kk}}\right) + 1} \quad (53)$$

$$r_{P_k,d} = \underbrace{\sqrt{P_P}\mathbb{E}\left[(\mu_{P_{kk}}|\hat{\mu}_{P_{kk}})\right]q_{P_k}}_{\text{Desired signal}} + \underbrace{\sqrt{P_P}\left((\mu_{P_{kk}}|\hat{\mu}_{P_{kk}})q_{P_k} - \mathbb{E}\left[(\mu_{P_{kk}}|\hat{\mu}_{P_{kk}})\right]q_{P_k}\right)}_{\text{Detection uncertainty}}$$
$$+ \underbrace{\sqrt{P_P}\sum_{i\neq k}^{K}(\mu_{P_{ki}}|\hat{\mu}_{P_{kk}})q_{P_i}}_{\text{Inter-system interference}} + \underbrace{\sqrt{P_S}\sum_{j=1}^{L}\left(\lambda_{P_{kj}}|\hat{\mu}_{P_{kk}}\right)q_{S_j}}_{\text{Intra-system interference}} + \underbrace{n_{P_k,d}}_{\text{AWGN}} \quad (56)$$

$$\gamma_{P_k,d} = P_P\left|\mathbb{E}\left[\mu_{P_{kk}}|\hat{\mu}_{P_{kk}}\right]\right|^2 \bigg/ \left(P_P\sum_{i=1}^{K}\mathbb{E}\left[|\mu_{P_{ki}}|^2|\hat{\mu}_{P_{kk}}\right] - P_P\left|\mathbb{E}\left[\mu_{P_{kk}}|\hat{\mu}_{P_{kk}}\right]\right|^2 + P_S\sum_{j=1}^{L}\mathbb{E}\left[|\lambda_{P_{kj}}|^2|\hat{\mu}_{P_{kk}}\right] + 1\right) \quad (57)$$

$$\gamma_{P_k,d} = P_P|\hat{\mu}_{P_{kk}}|^2 \bigg/ \left(P_P\left(\sum_{i\neq k}^{K}\mathbb{E}\left[|\mu_{P_{ki}}|^2\right] + \mathbb{E}\left[|\epsilon^{\mu}_{P_{kk}}|^2\right]\right) + P_S\sum_{j=1}^{L}\mathbb{E}\left[|\lambda_{P_{kj}}|^2\right] + 1\right) \quad (58)$$

$$\mathbb{E}_{\hat{\mu}_{P_{kk}}}\left[\gamma_{P_k,d}\right] = P_P\mathbb{E}\left[|\hat{\mu}_{P_{kk}}|^2\right] \bigg/ \left(P_P\left(\sum_{i\neq k}^{K}\mathbb{E}\left[|\mu_{P_{ki}}|^2\right] + \mathbb{E}\left[|\epsilon^{\mu}_{P_{kk}}|^2\right]\right) + P_S\sum_{j=1}^{L}\mathbb{E}\left[|\lambda_{P_{kj}}|^2\right] + 1\right) \quad (61)$$

$$\mathbb{E}_{\hat{\mu}_{P_{kk}}}\left[\gamma_{P_k,d}\right] = \frac{P_P\left(\sum_{m=1}^{M}\sum_{m'=1}^{M}\delta_{P_{mk}}\eta^{1/2}_{P_{mk}}\delta_{P_{m'k}}\eta^{1/2}_{P_{m'k}}\rho_{f_{mk}}\rho_{f_{m'k}} + v_{P_{kk}} - \kappa^{\epsilon}_{P_{kk}}\right)}{P_P\left(\sum_{i\neq k}^{K}v_{P_{ki}} + \kappa^{\epsilon}_{P_{kk}}\right) + P_S\left(\sum_{n=1}^{N}\sum_{j=1}^{L}\delta_{S_{nj}}\eta_{S_{nj}}\rho_{g_{nj}}\zeta_{u_{nk}} + \sum_{n=1}^{N}\delta_{S_{nk}}\eta_{S_{nk}}\rho^2_{u_{nk}}\right) + 1} \quad (62)$$

that the assumption of statistical DL channels are approximately equal to the instantaneous counterparts may not be accurate for cell-free massive MIMO.

VI. ACHIEVABLE RATE ANALYSIS OF NOMA-AIDED UNDERLAY SPECTRUM SHARING

A. Primary Achievable Rate With NOMA

We rearrange the received signal at $U_P(a,k)$ (22) to decode the desired signal as (68), shown at the bottom of the page, where $w_{P_{ak}}$ denotes the effective noise at $U_P(a,k)$ containing intra-cluster interference after SIC, error propagation due to imperfect SIC, intra-system interference, inter-system interference, and AWGN given in (22). From (68), we write the SINR at $U_P(a,k)$ as

$$\gamma_{P_{ak}} = \frac{P_P \left| \mathbb{E}\left[\sum_{m=1}^{M} \eta_{P_{mak}}^{1/2} f_{mak}\hat{f}_{mak}^* \right] \right|^2}{P_P \mathbb{V}\mathrm{ar}\left[\sum_{m=1}^{M} \eta_{P_{mak}}^{1/2} f_{mak}\hat{f}_{mak}^* \right] + P_P \sum_{i=1}^{4} I_{P_{ai}} + \mathbb{E}\left[\left| n_{P_{ak}} \right|^2 \right]}, \tag{69}$$

where $I_{P_{ai}}$ for $i \in \{1,2,3,4\}$ can be defined as

$$I_{P_{a1}} = \mathbb{E}\left[\left| \sum_{m=1}^{M} \sum_{i=1}^{k-1} \eta_{P_{mai}}^{1/2} f_{mak}\hat{f}_{mak}^* \right|^2 \right], \tag{70a}$$

$$I_{P_{a2}} = \mathbb{E}\left[\left| \sum_{m=1}^{M} \sum_{i=k+1}^{K} \eta_{P_{mai}}^{1/2} \left(f_{mak}\hat{f}_{mai}^* q_{P_{ai}} - \mathbb{E}\left[f_{mak}\hat{f}_{mai}^* \right]\hat{q}_{P_{ai}} \right) \right|^2 \right], \tag{70b}$$

$$I_{P_{a3}} = \mathbb{E}\left[\left| \sum_{m=1}^{M} \sum_{a' \neq a}^{A} \sum_{i=1}^{K} \eta_{P_{ma'i}}^{1/2} f_{mak}\hat{f}_{ma'i}^* \right|^2 \right], \tag{70c}$$

$$I_{P_{a4}} = \frac{P_S}{P_P}\mathbb{E}\left[\left| \sum_{n=1}^{N} \sum_{b=1}^{B} \sum_{j=1}^{L} \eta_{S_{nbj}}^{1/2} u_{nak}\hat{g}_{nbj}^* \right|^2 \right]. \tag{70d}$$

Then, we compute the SINR by evaluating the expectation and variance terms in (69) as given in (71), as shown at the bottom of the page, where $I'_{P_{ai}}$ for $i \in \{1,2,3\}$ is defined as

$$I'_{P_{a1}} = \sum_{m=1}^{M} \sum_{i=1}^{k-1} \eta_{P_{mai}} \alpha_{f_{mak}}^2 \left(\frac{\zeta_{f_{mai}}}{\zeta_{f_{mak}}} \right)^2, \tag{72a}$$

$$I'_{P_{a2}} = 2\sum_{m=1}^{M} \sum_{i=k+1}^{K} (1 - \vartheta_{P_{ai}})\eta_{P_{mai}} \alpha_{f_{mak}}^2 \left(\frac{\zeta_{f_{mai}}}{\zeta_{f_{mak}}} \right)^2, \tag{72b}$$

$$I'_{P_{a3}} = \frac{P_S}{P_P} \sum_{n=1}^{N} \sum_{j=1}^{L} \eta_{S_{naj}} \alpha_{f_{mak}}^2 \left(\frac{\zeta_{u_{nak}}\zeta_{g_{naj}}}{\zeta_{f_{mak}}} \right)^2. \tag{72c}$$

The derivation of (71) follows steps similar to those in Appendix B, and hence, it is omitted for the sake of brevity. Next, the achievable rate of $U_P(a,k)$ and the sum rate of primary system are given by

$$R_{P_{ak}} = ((\tau_c - \tau_p)/\tau_c) \log_2(1 + \gamma_{P_{ak}}), \tag{73}$$

$$R_P = \sum_{a=1}^{A} \sum_{k=1}^{K} R_{P_{ak}}, \tag{74}$$

where $\gamma_{P_{ak}}$ is defined in (71).

B. Secondary Achievable Rate With NOMA

We follow a similar analysis to Section VI-A for deriving the sum rate of the secondary system as

$$R_S = \sum_{b=1}^{B} \sum_{l=1}^{L} R_{S_{bl}}, \tag{75}$$

where $R_{S_{bl}}$ is the achievable rate of $U_S(b,l)$ and given by

$$R_{S_{bl}} = ((\tau_c - \tau_p)/\tau_c) \log_2(1 + \gamma_{S_{bl}}). \tag{76}$$

In (76), we obtain the SINR at $U_S(b,l)$ denoted by $\gamma_{S_{bl}}$ by following (71) and replacing the primary system variables with respective secondary system variables in (71) as given in (77), as shown at the bottom of the next page, where $I'_{S_{bi}}$ for $i \in \{1,2,3\}$ is given by

$$I'_{S_{b1}} = \sum_{n=1}^{N} \sum_{j=1}^{l-1} \eta_{S_{nbj}} \alpha_{g_{nbl}}^2 \left(\frac{\zeta_{g_{nbj}}}{\zeta_{g_{nbl}}} \right)^2, \tag{78a}$$

$$I'_{S_{b2}} = 2\sum_{n=1}^{N} \sum_{j=l+1}^{L} (1 - \vartheta_{S_{bj}})\eta_{S_{nbj}} \alpha_{g_{nbl}}^2 \left(\frac{\zeta_{g_{nbj}}}{\zeta_{g_{nbl}}} \right)^2, \tag{78b}$$

$$I'_{S_{b3}} = \frac{P_P}{P_S} \sum_{m=1}^{M} \sum_{i=1}^{K} \eta_{P_{mbi}} \alpha_{g_{nbl}}^2 \left(\frac{\zeta_{v_{mbl}}\zeta_{f_{mbi}}}{\zeta_{g_{nbl}}} \right)^2. \tag{78c}$$

$$\mathbb{E}_{\hat{\mu}_{S_{ll}}}[\gamma_{S_l,d}] = P_S \mathbb{E}\left[|\hat{\mu}_{S_{ll}}|^2 \right] \Bigg/ \left(P_S \left(\sum_{j \neq l}^{L} \mathbb{E}\left[|\mu_{S_{lj}}|^2 \right] + \mathbb{E}\left[|\epsilon_{S_{li}}^{\mu}|^2 \right] \right) + P_P \sum_{i=1}^{K} \mathbb{E}\left[|\lambda_{S_{li}}|^2 \right] + 1 \right) \tag{65}$$

$$\mathbb{E}_{\hat{\mu}_{S_{ll}}}[\gamma_{S_l,d}] = \frac{P_S \left(\sum_{n=1}^{N} \sum_{n'=1}^{N} \delta_{S_{nl}}\eta_{S_{nl}}^{1/2}\delta_{S_{n'l}}\eta_{S_{n'l}}^{1/2}\rho_{g_{nl}}\rho_{g_{n'l}} + v_{S_{ll}} - \kappa_{S_{ll}}^{\epsilon} \right)}{P_S \left(\sum_{j \neq l}^{L} v_{S_{lj}} + \kappa_{S_{ll}}^{\epsilon} \right) + P_P \left(\sum_{m=1}^{M} \sum_{i=1}^{K} \delta_{P_{mi}}\eta_{P_{mi}}\rho_{f_{mi}}\zeta_{v_{ml}} + \sum_{m=1}^{M} \delta_{P_{ml}}\eta_{P_{ml}}\rho_{v_{ml}}^2 \right) + 1} \tag{66}$$

$$r_{P_{ak}} = \underbrace{\sqrt{P_P}\mathbb{E}\left[\sum_{m=1}^{M} \eta_{P_{mak}}^{1/2} f_{mak}\hat{f}_{mak}^* \right] q_{P_{ak}}}_{\text{Desired signal}} + \underbrace{\sqrt{P_P}\sum_{m=1}^{M} \eta_{P_{mak}}^{1/2} f_{mak}\hat{f}_{mak}^* q_{P_{ak}} - \sqrt{P_P}\mathbb{E}\left[\sum_{m=1}^{M} \eta_{P_{mak}}^{1/2} f_{mak}\hat{f}_{mak}^* \right] q_{P_{ak}}}_{\text{Detection uncertinity}} + w_{P_{ak}} \tag{68}$$

$$\gamma_{P_{ak}} = \frac{P_P \left(\sum_{m=1}^{M} \eta_{P_{mak}}^{1/2} \alpha_{f_{mak}} \right)^2}{P_P \sum_{m=1}^{M} \sum_{a=1}^{A} \sum_{i=1}^{K} \eta_{P_{mai}}\alpha_{f_{mai}}\zeta_{f_{mak}} + P_P \sum_{i=1}^{3} I'_{P_{ai}} + \sum_{n=1}^{N} \sum_{b=1}^{B} \sum_{j=1}^{L} \eta_{S_{nbj}}\alpha_{g_{nbj}}\zeta_{u_{nak}} + 1} \tag{71}$$

VII. THE IMPLICATION OF DL PILOT TRANSMISSION WITH NOMA

The effective DL desired and interference channel coefficients at $U_P(a,k)$ are defined from (19) as

$$\mu_{P_{a'i}^{ak}} \triangleq \sum_{m=1}^{M} \eta_{P_{ma'i}}^{1/2} f_{mak} \hat{f}_{ma'i}^*, \tag{79a}$$

$$\lambda_{P_{bj}^{ak}} \triangleq \sum_{n=1}^{N} \eta_{S_{nbj}}^{1/2} u_{nak} \hat{g}_{nbj}^*. \tag{79b}$$

The P-APs/S-APs again use the same pilot sequences in (9) to beamform DL pilots toward PUs/SUs. Then, a sufficient statistic to estimate the desired effective DL channel at $U_P(a,k)$ is given by

$$y_{P_{ak},d} = \sqrt{P_{p,d}} \left(\sum_{i=1}^{K} \mu_{P_{ai}^{ak}} + \sum_{j=1}^{L} \lambda_{P_{aj}^{ak}} \right) + n_{P_{ak},d}, \tag{80}$$

where $n_{P_{ak},d} \sim \mathcal{CN}(0,1)$ is the AWGN at $U_P(a,k)$. Thereby, the MMSE estimate of $\mu_{P_{at}^{ak}}$ can be derived as [36], [43]

$$\hat{\mu}_{P_{at}^{ak}} = \mathbb{E}\left[\mu_{P_{at}^{ak}}\right] + \frac{\mathbb{Cov}\left(\mu_{P_{at}^{ak}} y_{P_{ak},d}^*\right)}{\mathbb{Cov}\left(y_{P_{ak},d} y_{P_{ak},d}^*\right)} \left(y_{P_{ak},d} - \mathbb{E}[y_{P_{ak},d}]\right)$$

$$= \mathbb{E}\left[\mu_{P_{at}^{ak}}\right] + \frac{\sqrt{P_{p,d}} \mathbb{Var}\left[\mu_{P_{at}^{ak}}\right] \times \left(y_{P_k,d} - \mathbb{E}[y_{P_{ak},d}]\right)}{P_{p,d}\left(\sum_{i=1}^{K} \mathbb{Var}\left[\mu_{P_{ai}^{ak}}\right] + \sum_{j=1}^{L} \mathbb{Var}\left[\lambda_{P_{aj}^{ak}}\right]\right) + 1}. \tag{81}$$

By evaluating (81) via steps similar to those in Appendix A-C, the MMSE estimate of $\mu_{P_{at}^{ak}}$ can be derived as

$$\hat{\mu}_{P_{at}^{ak}} = \theta_{P_{at}^{ak}} + \Omega_{P_{at}^{ak}} \left(y_{P_k,d} - \sqrt{P_{p,d}} \left(\sum_{i=1}^{K} \theta_{P_{ai}^{ak}} + \sum_{j=1}^{L} \psi_{P_{aj}^{ak}} \right) \right), \tag{82}$$

where $\theta_{P_{ai}^{ak}} = \sum_{m=1}^{M} \eta_{P_{mai}}^{1/2} \alpha_{f_{mak}} \zeta_{f_{mai}}/\zeta_{f_{mak}}$, $\psi_{P_{aj}^{ak}} = \sum_{n=1}^{N} \eta_{S_{naj}}^{1/2} \alpha_{u_{nak}} \zeta_{g_{nak}}/\zeta_{g_{naj}}$, and

$$\Omega_{P_{at}^{ak}} = \sqrt{P_{p,d}} \left(\varrho_{P_{at}^{ak}}^{\mu} - \theta_{P_{at}^{ak}}^2 \right) \Big/ \left(P_{p,d} \sum_{i=1}^{K} \left(\varrho_{P_{ai}^{ak}}^{\mu} - \theta_{P_{ai}^{ak}}^2 \right) \right.$$

$$\left. + P_{p,d} \sum_{j=1}^{L} \left(\varrho_{P_{bj}^{ak}}^{\lambda} - \psi_{P_{aj}^{ak}}^2 \right) + 1 \right), \tag{83a}$$

$$\varrho_{P_{ai}^{ak}}^{\mu} = \sum_{m=1}^{M} \eta_{P_{mai}} \alpha_{f_{mai}} \left(\zeta_{f_{mak}} + \alpha_{f_{mak}} \right), \tag{83b}$$

$$\varrho_{P_{bj}^{ak}}^{\lambda} = \sum_{n=1}^{N} \eta_{S_{nbj}} \left(\zeta_{u_{nak}} \alpha_{g_{nbj}} + \right). \tag{83c}$$

The actual effective DL channel gain is given by $\mu_{P_{at}^{ak}} = \hat{\mu}_{P_{at}^{ak}} + \epsilon_{P_{at}^{ak}}^{\mu}$, where $\epsilon_{P_{at}^{ak}}^{\mu}$ is an estimation error, which is independent of the channel estimate $\hat{\mu}_{P_{at}^{ak}}$.

A. Primary Achievable DL Rate With DL Pilots in NOMA

It is assumed that the users in same cluster are ordered based on the effective channel gains as per (20), and hence, SIC can be adopted to decode the power-domain NOMA signals at the users [35]. Thereby, the post-processed signal after an

$$\gamma_{S_{bl}} = \frac{P_S \left(\sum_{n=1}^{N} \eta_{S_{nbl}}^{1/2} \alpha_{g_{nbl}} \right)^2}{P_S \sum_{n=1}^{N} \sum_{b=1}^{B} \sum_{j=1}^{L} \eta_{S_{nbj}} \alpha_{g_{nbj}} \zeta_{g_{nbl}} + P_S \sum_{j=1}^{3} I'_{S_{bj}} + \sum_{m=1}^{M} \sum_{a=1}^{A} \sum_{j=1}^{K} \eta_{P_{mai}} \alpha_{f_{mai}} \zeta_{v_{mbl}} + 1} \tag{77}$$

$$r_{P_{ak},d} = \underbrace{\sqrt{P_P} \mu_{P_{ak}^{ak}} q_{P_{ak}}}_{\text{Desired signal}} + \underbrace{\sqrt{P_P} \sum_{i=1}^{k-1} \mu_{P_{ai}^{ak}} q_{P_{ai}}}_{\text{Intra-cluster interference after SIC}} + \underbrace{\sqrt{P_P} \sum_{i=k+1}^{K} \epsilon_{P_{ai}^{ak}}^{\mu} q_{P_{ai}}}_{\text{Error propagation due to imperfect SIC}}$$

$$+ \underbrace{\sqrt{P_P} \sum_{a' \neq a}^{A} \sum_{i=1}^{K} \mu_{P_{a'i}^{ak}} q_{P_{a'i}}}_{\text{Intra-system interference}} + \underbrace{\sqrt{P_S} \sum_{b=1}^{B} \sum_{j=1}^{L} \lambda_{P_{bj}^{ak}} q_{S_{bj}}}_{\text{Inter-system interference}} + \underbrace{n_{P_{ak}}}_{\text{AWGN}} \tag{84}$$

$$\gamma_{P_{ak},d} = \frac{P_P \left| \mathbb{E}\left[\mu_{P_{ak}^{ak}} | \hat{\mu}_{P_{ak}^{ak}} \right] \right|^2}{P_P \sum_{a'=1}^{A} \sum_{i=1}^{K} \mathbb{E}\left[\left| \mu_{P_{a'i}^{ak}} | \hat{\mu}_{P_{ak}^{ak}} \right|^2 \right] - P_P \sum_{i=k}^{K} \left| \mathbb{E}\left[\mu_{P_{ak}^{ak}} | \hat{\mu}_{P_{ak}^{ak}} \right] \right|^2 + P_S \sum_{b=1}^{B} \sum_{j=1}^{L} \mathbb{E}\left[\left| \lambda_{P_{aj}^{ak}} | \hat{\mu}_{P_{ak}^{ak}} \right|^2 \right] + 1} \tag{85}$$

$$\mathbb{E}_{\hat{\mu}_{P_{ak}^{ak}}}[\gamma_{P_{ak},d}] = \frac{P_P \mathbb{E}\left[\left| \hat{\mu}_{P_{ak}^{ak}} \right|^2 \right]}{P_P \sum_{a' \neq a}^{A} \sum_{i=k}^{K} \mathbb{E}\left[\left| \mu_{P_{ai}^{ak}} \right|^2 \right] + \sum_{j=1}^{2} I_j^{\hat{\mu}} + P_S \sum_{b=1}^{B} \sum_{j=1}^{L} \mathbb{E}\left[\left| \lambda_{P_{bj}^{ak}} \right|^2 \right] + 1} \tag{86}$$

$$\mathbb{E}_{\hat{\mu}_{P_{ak}^{ak}}}[\gamma_{P_{ak},d}] = \frac{P_P \Phi_{P_{ak}^{ak}}^{\mu}}{P_P \sum_{a' \neq a}^{A} \sum_{i=1}^{K} \varrho_{P_{ai}^{ak}}^{\mu} + P_P \sum_{i=k+1}^{K} \Phi_{P_{ak}^{ak}}^{\mu} + P_P \sum_{i=1}^{K} \left(\varrho_{P_{ai}^{ak}}^{\mu} - \Phi_{P_{ak}^{ak}}^{\mu} \right) + P_S \sum_{b=1}^{B} \sum_{j=1}^{L} \varrho_{P_{bj}^{ak}}^{\lambda} + 1} \tag{88}$$

$$\mathbb{E}_{\hat{\mu}_{S_{bl}^{bl}}}[\gamma_{S_{bl},d}] = \frac{P_S \Phi_{S_{bl}^{bl}}^{\mu}}{P_S \sum_{b' \neq b}^{B} \sum_{j=1}^{L} \varrho_{S_{bj}^{bl}}^{\mu} + P_S \sum_{j=l+1}^{L} \Phi_{S_{bl}^{bl}}^{\mu} + P_S \sum_{j=1}^{L} \left(\varrho_{S_{bj}^{bl}}^{\mu} - \Phi_{S_{bj}^{bl}}^{\mu} \right) + P_P \sum_{a=1}^{A} \sum_{i=1}^{K} \varrho_{S_{ai}^{bl}}^{\lambda} + 1} \tag{90}$$

imperfect SIC operation at $U_P(a,k)$ can be written as (84), shown at the bottom of the previous page. By using (84), the corresponding SINR at $U_P(a,k)$ can be derived as (85), shown at the bottom of the previous page. By using techniques similar to those used in (58), the expected value of $\gamma_{P_{ak},d}$ can be written as (86), shown at the bottom of the previous page, where $I_j^{\hat{\mu}}$ for $j \in \{1, 2\}$ is given as

$$I_1^{\hat{\mu}} = P_P \sum_{i=k+1}^{K} \mathbb{E}\left[\left|\hat{\mu}_{P_{ak}^{ak}}\right|^2\right], \qquad (87a)$$

$$I_2^{\hat{\mu}} = P_P \sum_{i=1}^{K} \mathbb{E}\left[\left|\epsilon_{P_{ai}^{ak}}^{\mu}\right|^2\right]. \qquad (87b)$$

By evaluating the expectation terms in (86), the average SINR at $U_P(a,k)$ can be derived as given in (88), shown at the bottom of the previous page, where $\Phi_{P_{ai}^{ak}}^{\mu}$ is given by

$$\Phi_{P_{ai}^{ak}}^{\mu} \triangleq \theta_{P_{at}^{ak}}^2 + \Omega_{P_{at}^{ak}}^2 \left(P_{p,d} \left(\sum_{i=1}^{K} \varrho_{P_{ai}^{ak}}^{\mu} + \sum_{j=1}^{L} \varrho_{P_{bj}^{ak}}^{\lambda} \right) \right.$$

$$\left. + 1 - P_{p,d} \left(\sum_{i=1}^{K} \theta_{P_{ai}^{ak}} + \sum_{j=1}^{L} \psi_{P_{aj}^{ak}} \right)^2 \right). \qquad (89)$$

Then, an upper bound for the DL rate at $U_P(a,k)$ with estimated DL channels via the beamformed pilots can be derived as $R_{P_{ak},d}^{ub} = (\tau_{d,d}/\tau_c) \log_2\left(1 + \mathbb{E}_{\hat{\mu}_{P_{ak}^{ak}}}\left[\gamma_{P_{ak},d}\right]\right)$, where $\mathbb{E}_{\hat{\mu}_{P_{ak}^{ak}}}\left[\gamma_{P_{ak},d}\right]$ is defined in (88).

B. Secondary Achievable DL Rate With DL Pilots in NOMA

By following an analysis similar to Section VII-A, the achievable rate at $U_S(b,l)$ with estimated DL channels can be derived as $R_{S_{bl},d}^{ub} = (\tau_{d,d}/\tau_c) \log_2\left(1 + \mathbb{E}_{\hat{\mu}_{P_{bl}^{bl}}}\left[\gamma_{S_{bl},d}\right]\right)$, where $\mathbb{E}_{\hat{\mu}_{P_{bl}^{bl}}}\left[\gamma_{S_{bl},d}\right]$ is defined in (90), shown at the bottom of the previous page, where $\Phi_{S_{bl}^{bl}}^{\mu}$, $\varrho_{S_{bj}^{bl}}^{\mu}$, and $\varrho_{S_{ai}^{bl}}^{\lambda}$ can be obtained by replacing the primary subscripts $\{P, M, m, A, a, K, k, i\}$ with respective secondary subscripts $\{S, N, n, B, b, L, l, j\}$ in (89), (83b), and (83c), respectively.

VIII. NUMERICAL RESULTS

In this section, our numerical results are presented to obtained useful insights. The simulation parameters are as follows: $\tau_c = 196$, $\tau_p = \tau_{p,d} = \max(K, L)$, and $\zeta_{h_{ab}} = (d_0/d_{ab})^\nu \times 10^{\varphi_{ab}/10}$, where d_{ab} is transmission distance between the ath P-AP/S-AP and the bth PU/SU, d_0 is the reference distance, and ν is the path-loss exponent. Here, $10^{\varphi_{ab}/10}$ captures the shadow fading with $\varphi_{ab} \sim \mathcal{N}(0, 8)$. In an area of $800 \times 800\,m^2$, the P-APs/S-APs are uniformly distributed, while PUs/SUs are randomly placed.

In Fig. 2, the implications of our max-min based multi-objective transmit power allocation are investigated. To this end, the achievable user rates of the primary and secondary systems are plotted against the primary transmit power per P-AP (P_P). The maximum allowable secondary transmit power (P_S) at each S-AP is kept at $P_P/2$. Then, the primary and secondary transmit power control coefficients are
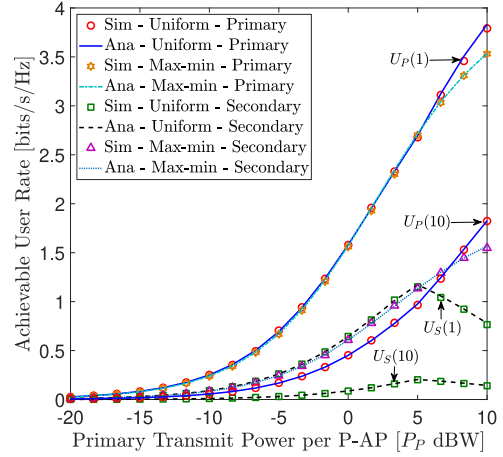

Fig. 2. User rate versus the primary transmit power for $K = L = 10$, $M = N = 32$, $I_T = 0$ dB, $\sigma_P^2 = \sigma_S^2 = 1$, $\nu = 2.4$, and $d_0 = 1\,\text{m}$. Moreover, $P_S = P_P/2$ and $P_{S_n} = \min(\check{P}_S, I_{T_1}/Z_1, \cdots, I_{T_k}/Z_k, \cdots, I_{T_K}/Z_K)$.

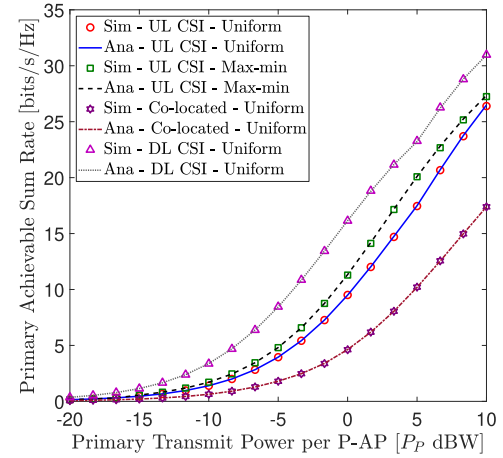
Fig. 3. The primary sum rate versus the primary transmit power for $K = L = 10$, $M = N = 32$, $I_T = 0$ dB, $\sigma_P^2 = \sigma_S^2 = 1$, $\nu = 2.4$, and $d_0 = 1\,\text{m}$.

computed by using the proposed max-min algorithm in Section IV. The rates with max-min power allocation are compared with those with uniform power allocation. The pair $\{U_P(1), U_S(1)\}$ is the users with strongest channels, while the pair $\{U_P(10), U_S(10)\}$ represents the users with weakest channel gains. Fig. 2 reveals that PUs/SUs experience distinct achievable rates when the uniform power allocation is adopted. Thus, the achievable rates are dependent on the detrimental near-far effects. When the proposed max-min power control is employed, all PUs/SUs achieve their respective common rates regardless of the near-far effects. For instance, at $P_P = 0\,\text{dBW}$, the weaker user $U_P(10)$ achieves a rate gain of 241.3% from our max-min power allocation over the uniform power allocation.

In Fig. 3 and Fig. 4, an achievable rate comparison is presented for cell-free/co-located, statistical/estimated DL CSI with max-min/uniform power allocations. In this context,
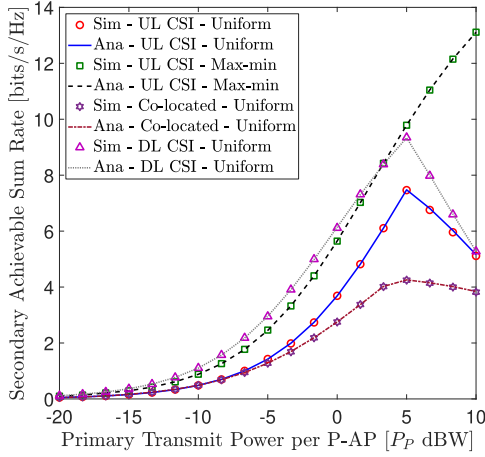
Fig. 4. The secondary sum rate versus the primary transmit power for $K = L = 10$, $M = N = 32$, $I_T = 0\,dB$, $\sigma_P^2 = \sigma_S^2 = 1$, $\nu = 2.4$, and $d_0 = 1\,\mathrm{m}$. Moreover, $P_S = P_P/2$ and $P_{S_n} = \min(P_S, I_{T_1}/Z_1, \cdots, I_{T_k}/Z_k, \cdots, I_{T_K}/Z_K)$.

the achievable sum rates of the primary and secondary systems, respectively, are plotted in Fig. 3 and Fig. 4 as a function of the primary transmit power per P-AP (P_P). In Fig. 4, P_S is set to $P_P/2$ at each S-AP. When the uniform power allocation is adopted, the achievable sum rate of the primary system increases monotonically with P_P. However, for the secondary system, the achievable sum rate gradually increases up to a maximum in the low P_P regime since P_S is proportional to P_P, and then it decreases as P_P grows without bound. The reason for this behavior is that when P_P increases, the respective secondary maximum allowable transmit power also increases since $P_S = P_P/2$. Thus, the secondary sum rate grows gradually until the secondary transmit power constraints in (28) are met. At this point S-APs transmit signals with their maximum allowable transmit power P_S, and the secondary sum rate reaches a maximum. Simultaneously, the primary transmit powers at the P-APs keep increasing, and this causes a high level of primary CCI at SUs. Consequently, the secondary sum rate decreases as P_P grows without bound. When max-min transmit power allocation algorithm is adopted for both systems, the primary and secondary achievable sum rates increase with the primary transmit power P_P. Furthermore, Fig. 3 and Fig. 4 reveal that the both systems achieve higher sum rates when DL CSI is adopted at the users for signal decoding over the statistical CSI case. In particular, the achieve rate performance of the proposed underlay spectrum sharing in cell-free massive MIMO is compared with that of the co-located counterpart in Fig. 3 and Fig. 4. This comparison shows that the cell-free version outperforms the co-located case in terms of the sum rate of the underlay spectrum sharing. For instance, in Fig. 4, the cell-free based secondary system achieves the rate gains of 75.6%, 119.6%, and 130.0% for the uniform power allocation, max-min power allocation, and DL CSI cases, respectively, compared to that of the co-located counterpart at $P_P = 5\,dBW$.

In Fig. 5, the effect of AP clustering is investigated. To this end the achievable sum rates of primary and secondary systems
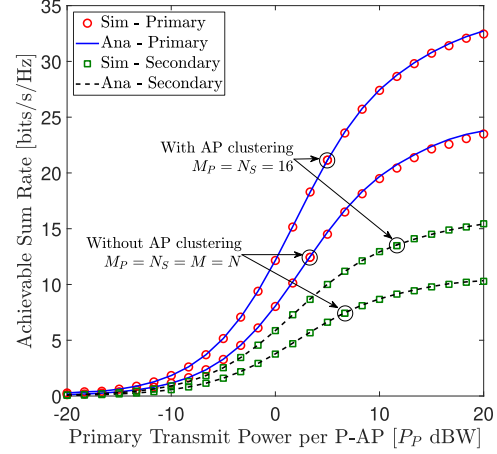

Fig. 5. The impact of user-centric AP clustering. The achievable sum rate versus the primary transmit power with/without AP clustering for $K = L = 10$, $M = N = 32$, $I_T = 0$ dB, $\sigma_P^2 = \sigma_S^2 = 1$, $\nu = 2.4$, and $d_0 = 1\,\mathrm{m}$.
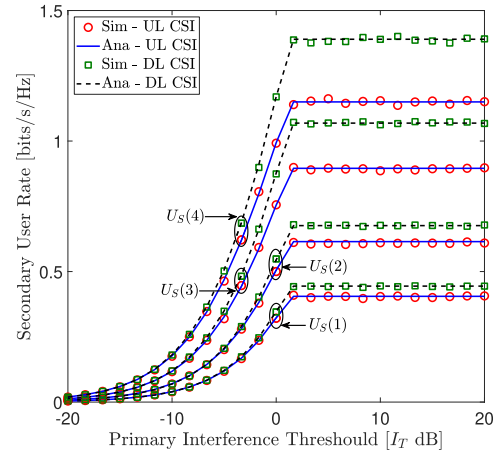

Fig. 6. The secondary user rate versus PIT (I_T) for $K = L = 4$, $M = N = 32$, $P_P = 0$ dBW, $\sigma_P^2 = \sigma_S^2 = 1$, $\nu = 2.4$, and $d_0 = 1\,\mathrm{m}$.

are plotted against the primary transmit power per P-AP with/without AP clustering by adopting max-min power allocation. Fig. 5 reveals that AP clustering boosts the achievable rates of both primary and secondary systems. For example, at P_P of $10\,dBW$, the primary and secondary systems achieve sum rate gains of 39.5% and 48.7%, respectively, when the user-centric AP clustering is adopted over the case of uniform AP deployment without a predefined AP clustering scheme. The reason for this behavior is that, when a certain number of APs is allocated for a particular user, the CCI from the remaining users of the own system and the other system is reduced. This reduced level of CCI translates into achievable rate gains.

In Fig. 6, the effects of secondary transmit power constraints on the achievable secondary user rates are explored. Two sets of rate curves are plotted as a function of PIT (I_T) for $M = N = 32$ and with/without DL CSI at SUs by keeping the maximum allowable transmit power $P_S = P_P/2$
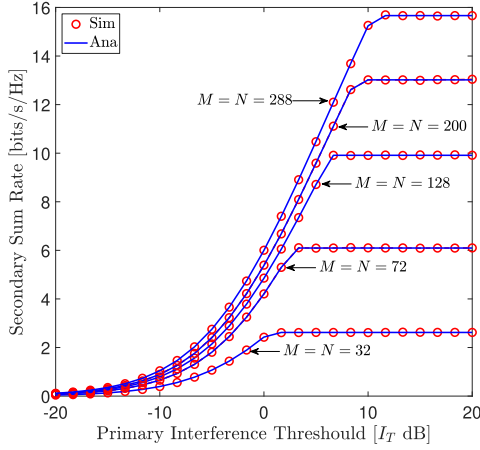
Fig. 7. The secondary sum rate versus PIT (I_T) for $K = L = 4$, $P_P = 0$ dBW, $\sigma_P^2 = \sigma_S^2 = 1$, $\nu = 2.4$, and $d_0 = 1$ m.

Fig. 8. A comparison between the achievable sum rate of NOMA and OMA for $M = N = 32$, $P_P = 0$ dBW, $\sigma_P^2 = \sigma_S^2 = 1$, $\nu = 2.4$, $\vartheta_{P_{ai}} = \vartheta_{S_{bj}} = 0.2$, and $d_0 = 1$ m.

at each S-AP. Fig. 6 clearly reveals that the secondary user rates grow exponentially in the low I_T regime in the both cases. In high I_T regime, the SU rates for the both CSI cases saturate to a maximum. The reason for this behavior is that in low I_T regime, a high amount of secondary CCI at PUs is allowed, whereas in high I_T regime, the SU rates saturate when the secondary transmit power meets the transmit power constraints in (28). Fig. 6 also reveals that the SU rates can be boosted when the estimated DL CSI from beamformed DL pilots is used over the statistical/long-term DL CSI counterpart for signal decoding at SUs. For example, $U_S(3)$ and $U_S(4)$ achieve rate gains of 20.9 % and 18.6 %, respectively, by using the DL CSI at $I_T = 10$ dB compared to those of without DL CSI.

In Fig. 7, the implication of the number of P-APs/S-APs is investigated by plotting the achievable secondary sum rate as a function of the PIT (I_T) with statistical DL CSI at the SUs. By varying the number of P-APs and S-APs as $M = N = 32$, $M = N = 72$, $M = N = 128$, $M = N = 200$, and $M = N = 288$, five sets of rate curves are plotted. Fig. 7 shows that the maximum saturation of the secondary sum rate is heavily depend on N/M. The reason for this behavior can be described as follows: In the low I_T regime, the secondary only can transmit smaller powers because PUs can withstand only to very small level of secondary CCI. Thus, those low secondary transmit powers result in smaller secondary achievable rates. Moreover, when I_T grows large, it allows to inflect a high amount of secondary CCI at the PUs, and thus, the S-APs can transmit high power and it exponentially increases the secondary sum rate. Once the secondary power constraints in (28) are met, the secondary sum rate saturates to a maximum.

In Fig. 8, a comparison of OMA versus NOMA in terms of the achievable sum rates of the primary and secondary systems is presented. To this end, the sum rates are plotted as a function of the number of users that can be served simultaneously. The analytical curves are plotted by using (74) and (75) for primary and secondary systems, respectively.
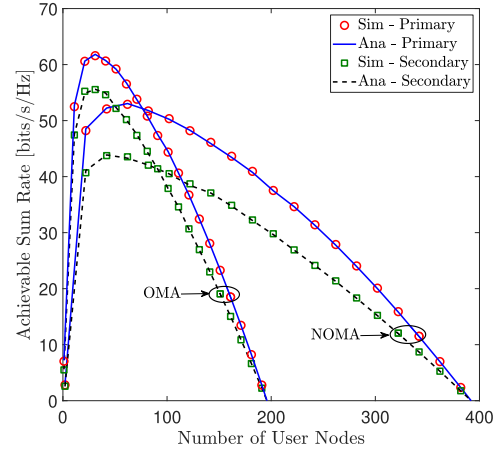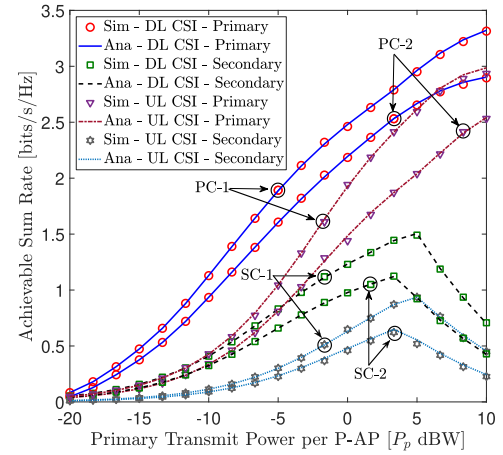

Fig. 9. The impact of DL pilots in NOMA. The achievable sum rate versus the primary transmit power for $M = N = 32$, $A = B = 2$, $K = L = 2$, $I_T = 0$ dB, $\sigma_P^2 = \sigma_S^2 = 1$, $\nu = 2.4$, and $d_0 = 1$ m. Moreover, $P_S = P_P/2$ and $P_{S_n} = \min\left(P_S, I_{T_1}/Z_1, \cdots, I_{T_k}/Z_k, \cdots, I_{T_K}/Z_K\right)$.

Since the coherence interval $\tau_c = 196$, the maximum number of users that can be served by OMA is limited to 196. On the other hand, NOMA can serve more number of users than that of OMA because of user clustering. However, Fig. 8 reveals that OMA outperforms NOMA in the regime of low number of users. The reason for this behavior is that the intra-cluster pilot contamination due to the shared pilots among NOMA clusters and the residual interference caused by error propagation from imperfect SIC hinder the achievable rates of NOMA in low user regime. However, NOMA-aided cell-free underlay spectrum sharing is beneficial in boosting the number of simultaneous served SUs and the achievable sum rates at high user rate regime.

In Fig. 9, the impact of DL pilots on the achievable sum rate of NOMA-aided underlay spectrum sharing for cell-free massive MIMO is studied. The analytical rate curves for NOMA with estimated DL CSI are plotted via our analysis in (88) and (90). Fig. 9 shows that the achievable sum

rate of the primary system with/without DL pilots grows monotonically with P_P. Nevertheless, the sum rate of the secondary system with/without DL channel training gradually grows until a maximum in the low power regime, and then, it starts to decrease as P_P continues to grow large. The reasons for this behavior are that the secondary transmit power is set to $P_S = P_P/2$, then it reaches the maximum allowable limit as per the secondary transmit power constraint in (28), and finally, the higher levels of primary CCI at SUs as P_P continues to grow large, while P_S stays fixed at its maximum level. Moreover, Fig. 9 reveals that the achievable rates can be boosted when the users adopt the estimated DL CSI via beamformed pilots for signal decoding at PUs/SUs. For instance, at a primary transmit power of $0\,\text{dBW}$, the primary system with DL pilots achieves a sum rate gain of 52.4% over a system that only relies on statistical DL CSI (without DL pilots) for signal decoding at the PUs.

IX. CONCLUSION

The practical feasibility of deploying underlay spectrum sharing in cell-free massive MIMO has been investigated by adopting UL/DL pilot-based channel estimations, max-min based MOOP, OMA/NOMA and the corresponding achievable rates for both primary and secondary systems. User fairness has been guaranteed by adopting an MOOP-based max-min fairness algorithm for P-APs/S-APs, while satisfying the secondary transmit power constraints, which are subjected to PIT. The achievable rates for both systems have been derived for locally estimated UL CSI at P-APs/S-APs and DL CSI at PUs/SUs. The impact of DL channel estimation at PUs/SUs via beamforming of DL pilots to boost the achievable rates has been studied. We reveal that our proposed MOOP-based max-min transmit power control algorithm can significantly boost the achievable rate of the both systems over the uniform power allocation by using carefully designed secondary transmit power constraints at each S-AP. The effect of user-centric cell-free massive MIMO in which the P-APs/S-APs are clustered to serve a particular PU/SU has been investigated. The trade-offs between OMA and NOMA in terms of the number of concurrently served PUs/SUs and the achievable rates have been explored, and thereby, it has been revealed that NOMA-aided underlay spectrum sharing can be beneficial in satisfying future massive access demands in a cell-free set-up. It has been shown that the achievable rates can be boosted when the PUs/SUs estimate DL channels from the beamformed DL pilots and adopt this estimated DL CSI to decode signals instead of solely relying on statistical DL CSI. Thus, our performance analysis establishes that a secondary system can be operated within the same primary/licensed spectrum without hindering the primary system performance in a cell-free massive MIMO set-up.

APPENDIX A

A. Derivation of the MMSE Estimates in (6a) and (6b)

By substituting (5a) into (6a), $\hat{f}_{mk}$ can be derived as [36]

$$\hat{f}_{mk} = \frac{\mathbb{E}\left[\left(\sqrt{\tau_p P}f^*_{mk} + \sqrt{\tau_p P}v^*_{mk} + n^*_{P_m}\right)f_{mk}\right]}{\mathbb{E}\left[\left|\sqrt{\tau_p P}f_{mk} + \sqrt{\tau_p P}v_{mk} + n_{P_m}\right|^2\right]}y_{P_{mk}}$$

$$= \frac{\sqrt{\tau_p P}\mathbb{E}\left[f^*_{mk}f_{mk}\right]y_{P_{mk}}}{\tau_p P\mathbb{E}\left[|f_{mk}|^2\right] + \tau_p P\mathbb{E}\left[|v_{mk}|^2\right] + \mathbb{E}\left[|n_{P_m}|^2\right]}$$

$$\overset{(a)}{=} c_{P_{mk}}y_{P_{mk}}, \tag{91}$$

where $c_{P_{mk}}$ is defined in (7a) and $n_{P_m} = \mathbf{n}_{P_m}\boldsymbol{\phi}^H_{P_k}$. The step (a) is written by using the fact that f_{mk}, v_{mk}, and n_{P_m} are Gaussian random variables with zero means and then, by evaluating the expectations of their squared norms [44]. Similarly, the MMSE estimate of g_{nl} can be derived as given in (6b).

B. Derivation of Z_k in (27)

By substituting (6b) and (5b) into (26), Z_k is computed as

$$Z_k$$
$$= \sum_{n=1}^{N}\delta_{S_{nk}}\eta_{S_{nk}}c^2_{S_{nk}}\mathbb{E}\left[\left|u_{nk}\left(\sqrt{P_p}g_{nk} + \sqrt{P_p}u_{nk} + n_{S_n}\right)^*\right|^2\right]$$
$$+ \sum_{n=1}^{N}\sum_{l\neq k}^{L}\delta_{S_{nl}}\eta_{S_{nl}}c^2_{S_{nl}}\mathbb{E}\left[\left|u_{nk}\left(\sqrt{P_p}g_{nl} + \sqrt{P_p}u_{nl} + n_{S_n}\right)^*\right|^2\right]$$
$$= \sum_{n=1}^{N}\sum_{l=1}^{L}\delta_{S_{nl}}\eta_{S_{nl}}\rho_{g_{nl}}\zeta_{u_{nk}} + \sum_{n=1}^{N}\delta_{S_{nk}}\eta_{S_{nk}}\rho^2_{u_{nk}}. \tag{92}$$

C. Derivation of the MMSE Estimate of $\mu_{P_{kk}}$ in (53)

By noting that $f_{mk} = \hat{f}_{mk} + \epsilon_{f_{mk}}$, $\mathbb{E}\left[\mu_{P_{kk}}\right]$ can be calculated as

$$\mathbb{E}\left[\mu_{P_{kk}}\right] = \mathbb{E}\left[\sum_{m=1}^{M}\delta_{P_{mk}}\eta^{1/2}_{P_{mk}}\hat{f}_{mk}\hat{f}^*_{mk}\right]$$
$$+ \mathbb{E}\left[\sum_{m=1}^{M}\delta_{P_{mk}}\eta^{1/2}_{P_{mk}}\epsilon_{f_{mk}}\hat{f}^*_{mk}\right]$$
$$= \sum_{m=1}^{M}\delta_{P_{mk}}\eta^{1/2}_{P_{mk}}\rho_{f_{mk}}, \tag{93}$$

where $\rho_{f_{mk}} \triangleq \sqrt{\tau_p P}c_{P_{mk}}\zeta_{f_{mk}}$. Then, $\mathbb{E}\left[\lambda_{P_{kk}}\right]$ can be evaluated as

$$\mathbb{E}\left[\lambda_{P_{kk}}\right]$$
$$= \mathbb{E}\left[\sum_{n=1}^{N}\delta_{S_{nk}}\eta^{1/2}_{S_{nk}}u_{nk}\hat{g}^*_{nk}\right]$$
$$\overset{(b)}{=} \mathbb{E}\left[\sum_{n=1}^{N}\delta_{S_{nk}}\eta^{1/2}_{S_{nk}}c_{S_{nk}}u_{nk}\left(\sqrt{\tau_p P}\left(\hat{g}^*_{nk} + \hat{u}^*_{nk}\right) + n_{S_n}\right)\right]$$
$$= \sum_{n=1}^{N}\delta_{S_{nk}}\eta^{1/2}_{S_{nk}}\rho_{u_{nk}}, \tag{94}$$

where the step (b) is written by using (6b) and (5b).

The variance term of $\mu_{P_{kk}}$ in (52) is given by

$$\mathbb{V}\text{ar}[\mu_{P_{kk}}] = \mathbb{E}\left[|\mu_{P_{kk}}|^2\right] - |\mathbb{E}[\mu_{P_{kk}}]|^2, \tag{95}$$

where the first expectation term in (95) can be derived as

$$\mathbb{E}\left[|\mu_{P_{kk}}|^2\right]$$
$$= \mathbb{E}\left[\left|\sum_{m=1}^{M}\delta_{P_{mk}}\eta^{1/2}_{P_{mk}}\hat{f}_{mk}\hat{f}^*_{mk}\right|^2\right]$$
$$+ \mathbb{E}\left[\left|\sum_{m=1}^{M}\delta_{P_{mk}}\eta^{1/2}_{P_{mk}}\epsilon_{f_{mk}}\hat{f}^*_{mk}\right|^2\right]$$

$$= \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}} \mathbb{E}\left[|\hat{f}_{mk}|^4\right]$$
$$+ \sum_{m=1}^{M} \sum_{m' \neq m}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} \delta_{P_{m'k}} \eta_{P_{m'k}}^{1/2} \mathbb{E}\left[|\hat{f}_{mk}|^2 |\hat{f}_{m'k}|^2\right]$$
$$+ \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}} (\zeta_{f_{mk}} - \rho_{f_{mk}}) \rho_{f_{mk}}$$
$$= 2 \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}} \rho_{f_{mk}}^2$$
$$+ \sum_{m=1}^{M} \sum_{m' \neq m}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} \delta_{P_{m'k}} \eta_{P_{m'k}}^{1/2} \rho_{f_{mk}} \rho_{f_{m'k}}$$
$$+ \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}} (\zeta_{f_{mk}} - \rho_{f_{mk}}) \rho_{f_{mk}}$$
$$= \sum_{m=1}^{M} \sum_{m'=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} \delta_{P_{m'k}} \eta_{P_{m'k}}^{1/2} \rho_{f_{mk}} \rho_{f_{m'k}}$$
$$+ \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}} \zeta_{f_{mk}} \rho_{f_{mk}}. \tag{96}$$

By substituting (93) and (96) into (95), we have

$$\mathbb{V}\mathrm{ar}[\mu_{P_{kk}}] = \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}} \zeta_{f_{mk}} \rho_{f_{mk}} \triangleq v_{P_{kk}}. \tag{97}$$

The variance of $\lambda_{P_{kk}}$ can be given as

$$\mathbb{V}\mathrm{ar}[\lambda_{P_{kk}}] = \mathbb{E}\left[|\lambda_{P_{kk}}|^2\right] - |\mathbb{E}[\lambda_{P_{kk}}]|^2. \tag{98}$$

Thus, the first expectation term in (98) can be derived as

$$\mathbb{E}\left[|\lambda_{P_{kk}}|^2\right]$$
$$= \mathbb{E}\left[\left|\sum_{n=1}^{N} \delta_{S_{nk}} \eta_{S_{nk}}^{1/2} u_{nk} \hat{g}_{nk}^*\right|^2\right]$$
$$\overset{(c)}{=} \sum_{n=1}^{N} \delta_{S_{nk}} \eta_{S_{nk}} c_{S_{nk}}^2 \zeta_{u_{nk}} (\tau_p P (\zeta_{g_{nk}} + 2\zeta_{u_{nk}}) + 1)$$
$$= \sum_{n=1}^{N} \delta_{S_{nk}} \eta_{S_{nk}} \rho_{g_{nk}} \zeta_{u_{nk}} + \sum_{n=1}^{N} \delta_{S_{nk}} \eta_{S_{nk}} \rho_{u_{nk}}^2, \tag{99}$$

where the step (c) is written by following steps similar to those used in the step (b) in (94). Therefore, by using the results in (94) and (99), we have

$$\mathbb{V}\mathrm{ar}[\lambda_{P_{kk}}] = \sum_{n=1}^{N} \delta_{S_{nk}} \eta_{S_{nk}} \zeta_{u_{nk}} \rho_{g_{nk}} \triangleq u_{P_{kk}}. \tag{100}$$

Then, by using (93), (94), (97), and (100), the effective channel estimate can be derived as (53) [43].

APPENDIX B
DERIVATION OF SINR IN (34)

The expectation in numerator of (34) can be derived as

$$\mathbb{E}\left[\sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} f_{mk} \hat{f}_{mk}^*\right] = \sum_{m=1}^{M} \eta_{P_{mk}}^{1/2} \mathbb{E}\left[\left(\hat{f}_{mk} + \epsilon_{f_{mk}}\right) \hat{f}_{mk}^*\right]$$
$$\overset{(d)}{=} \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} \rho_{f_{mk}}, \tag{101}$$

where $\epsilon_{f_{mk}}$ is an estimation error of f_{mk} such that $f_{mk} = \hat{f}_{mk} + \epsilon_{f_{mk}}$, satisfying $\mathbb{E}[\epsilon_{f_{mk}} \hat{f}_{mk}^*] = 0$. In (101), the step (d) is written by substituting (6a) and then evaluating expectation term as

$$\mathbb{E}\left[|\hat{f}_{mk}|^2\right] = c_{P_{mk}}^2 \mathbb{E}\left[|y_{P_{mk}}|^2\right] = \sqrt{\tau_p P} \zeta_{f_{mk}} c_{P_{mk}} = \rho_{f_{mk}}. \tag{102}$$

Then, the variance term in (34) can be derived as

$$\mathbb{V}\mathrm{ar}\left[\sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}}^{1/2} f_{mk} \hat{f}_{mk}^*\right]$$
$$= \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}} \left(\mathbb{E}\left[\left|\left(\hat{f}_{mk} + \epsilon_{f_{mk}}\right) \hat{f}_{mk}^*\right|^2\right] - \rho_{f_{mk}}^2\right)$$
$$= \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}} \left(2\rho_{f_{mk}}^2 + \rho_{f_{mk}} (\zeta_{f_{mk}} - \rho_{f_{mk}}) - \rho_{f_{mk}}^2\right)$$
$$= \sum_{m=1}^{M} \delta_{P_{mk}} \eta_{P_{mk}} \rho_{f_{mk}} \zeta_{f_{mk}}. \tag{103}$$

The expectation of the first term in (35) can be computed as

$$\mathbb{E}\left[\left|\sum_{i \neq k}^{K} \sum_{m=1}^{M} \delta_{P_{mi}} \eta_{P_{mi}}^{1/2} f_{mk} \hat{f}_{mi}^*\right|^2\right]$$
$$\overset{(e)}{=} \sum_{i \neq k}^{K} \sum_{m=1}^{M} \mathbb{E}\left[\left|\delta_{P_{mi}} \eta_{P_{mi}}^{1/2} f_{mk} c_{P_{mi}} y_{P_{mi}}\right|^2\right]$$
$$\overset{(f)}{=} \sum_{i \neq k}^{K} \sum_{m=1}^{M} \delta_{P_{mi}} \eta_{P_{mi}} c_{P_{mi}}^2 \zeta_{f_{mk}} (\tau_p P(\zeta_{f_{mi}} + \zeta_{v_{mi}}) + 1)$$
$$= \sum_{i \neq k}^{K} \sum_{m=1}^{M} \delta_{P_{mi}} \eta_{P_{mi}} \rho_{f_{mi}} \zeta_{f_{mk}}, \tag{104}$$

where the steps (e) and (f) are written by using (6a) and (5a), respectively. The expectation in the second term in (35) can be derived by following steps similar to those used in (92) as follows:

$$\mathbb{E}\left[\left|\sum_{j=1}^{L} \sum_{n=1}^{N} \delta_{S_{nj}} \eta_{S_{nj}}^{1/2} u_{nk} \hat{g}_{nj}^*\right|^2\right]$$
$$= \sum_{n=1}^{N} \delta_{S_{nk}} \eta_{S_{nk}} c_{S_{nk}}^2 \zeta_{u_{nk}} (\tau_p P (\zeta_{g_{nk}} + 2\zeta_{u_{nk}}) + 1)$$
$$+ \sum_{j \neq k}^{L} \sum_{n=1}^{N} \delta_{S_{nj}} \eta_{S_{nj}} c_{S_{nj}}^2 \zeta_{u_{nk}} (\tau_p P (\zeta_{g_{nj}} + \zeta_{u_{nj}}) + 1)$$
$$= \sum_{j=1}^{L} \sum_{n=1}^{N} \delta_{S_{nj}} \eta_{S_{nj}} \rho_{g_{nj}} \zeta_{u_{nk}} + \sum_{n=1}^{N} \delta_{S_{nk}} \eta_{S_{nk}} \rho_{u_{nk}}^2. \tag{105}$$

By substituting (101), (103), (104), and (105) into (34), the desired SINR can be derived as in (36).

APPENDIX C
DERIVATION OF SINR IN (62)

The expectation of magnitude squared error $(\epsilon_{P_{kk}}^\mu)$ in (61) can be calculated as

$$\mathbb{E}\left[|\epsilon_{P_{kk}}^\mu|^2\right]$$
$$= \mathbb{E}\left[|\mu_{P_{kk}} - \hat{\mu}_{P_{kk}}|^2\right]$$
$$\overset{(g)}{=} \frac{(1 + P_{p,d} u_{P_{kk}})^2 \mathbb{V}\mathrm{ar}[\mu_{P_{kk}}] + (P_{p,d} v_{P_{kk}})^2 \mathbb{V}\mathrm{ar}[\lambda_{P_{kk}}] + P_{p,d} v_{P_{kk}}^2}{(P_{p,d} (v_{P_{kk}} + u_{P_{kk}}) + 1)^2}$$
$$= \frac{(1 + P_{p,d} u_{P_{kk}})^2 v_{P_{kk}} + (P_{p,d} v_{P_{kk}})^2 u_{P_{kk}} + P_{p,d} v_{P_{kk}}^2}{(P_{p,d} (v_{P_{kk}} + u_{P_{kk}}) + 1)^2} \kappa_{P_{kk}}^\epsilon, \tag{106}$$

where $\kappa_{P_{kk}}^\epsilon$ is defined in the first term of (63) and the step (g) is written by substituting (53) and then, evaluating the expectation. The expectation term of $\mu_{P_{ki}}$ in the denominator of (61) can be derived as

$$\mathbb{E}\left[|\mu_{P_{ki}}|^2\right] = \mathbb{E}\left[\left|\sum_{m=1}^{M}\delta_{P_{mi}}\eta_{P_{mi}}^{1/2}\hat{f}_{mk}\hat{f}_{mi}^*\right|^2\right]$$

$$+ \mathbb{E}\left[\left|\sum_{m=1}^{M}\delta_{P_{mi}}\eta_{P_{mi}}^{1/2}\epsilon_{f_{mk}}\hat{f}_{mi}^*\right|^2\right]$$

$$= \sum_{m=1}^{M}\delta_{P_{mi}}\eta_{P_{mi}}\zeta_{f_{mk}}\rho_{f_{mi}} \triangleq v_{P_{ki}}. \quad (107)$$

Then, the expectation term in (61) with $\lambda_{P_{kj}}$ is evaluated similar to the steps those used in (105) as

$$\sum_{j=1}^{L}\mathbb{E}\left[|\lambda_{P_{kj}}|^2\right] = \sum_{j=1}^{L}\sum_{n=1}^{N}\delta_{S_{nj}}\eta_{S_{nj}}\rho_{g_{nj}}\zeta_{u_{nk}}$$

$$+ \sum_{n=1}^{N}\delta_{S_{nk}}\eta_{S_{nk}}\rho_{u_{nk}}^2. \quad (108)$$

The expectation term in the numerator of (61) is given by

$$\mathbb{E}\left[|\hat{\mu}_{P_{kk}}|^2\right] = \mathbb{E}\left[|\mu_{P_{kk}} - \epsilon_{P_{kk}}^\mu|^2\right] = \mathbb{E}\left[|\mu_{P_{kk}}|^2\right] - \mathbb{E}\left[|\epsilon_{P_{kk}}^\mu|^2\right]. \quad (109)$$

By substituting (96) and (106) into (109), we have

$$\mathbb{E}\left[|\hat{\mu}_{P_{kk}}|^2\right] = \sum_{m=1}^{M}\sum_{m'=1}^{M}\delta_{P_{mk}}\eta_{P_{mk}}^{1/2}\delta_{P_{m'k}}\eta_{P_{m'k}}^{1/2}\rho_{f_{mk}}\rho_{f_{m'k}}$$

$$+ v_{P_{kk}} - \kappa_{P_{kk}}^\epsilon. \quad (110)$$

By substituting (106), (107), (108), and (110) into (61), the effective DL SINR at $U_P(k)$ is evaluated as (62). By following step similar to (106)-(110), the DL SINR at $U_S(l)$ can be computed as (66).

REFERENCES

[1] D. L. Galappaththige and G. Amarasuriya, "Cell-free massive MIMO with underlay spectrum-sharing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–7.

[2] D. L. Galappaththige and G. Amarasuriya, "NOMA-aided cell-free massive MIMO with underlay spectrum-sharing," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2020, pp. 1–6.

[3] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[4] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[5] *Sprint Unveils six 5G-Ready Cities; Significant Milestone Toward Launching First 5G Mobile Network in the US*. Accessed: Aug. 5, 2020. [Online]. Available: https://newsroom.sprint.com/sprint-unveils-5G-ready-massive-MIMOmarkets. htm

[6] E. Björnson *et al.*, "Massive MIMO is a reality—What is next?: Five promising research directions for antenna arrays," *Digit. Signal Process.*, vol. 94, pp. 3–20, Nov. 2019.

[7] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO: Uniformly great service for everyone," in *Proc. IEEE 16th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2015, pp. 201–205.

[8] 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.

[9] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, Jul. 2017.

[10] T. C. Mai, H. Q. Ngo, M. Egan, and T. Q. Duong, "Pilot power control for cell-free massive MIMO," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 11264–11268, Nov. 2018.

[11] T. X. Doan, H. Q. Ngo, T. Q. Duong, and K. Tourki, "On the performance of multigroup multicast cell-free massive MIMO," *IEEE Commun. Lett.*, vol. 21, no. 12, pp. 2642–2645, Dec. 2017.

[12] Y. Li and G. A. A. Baduge, "NOMA-aided cell-free massive MIMO systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 6, pp. 950–953, Dec. 2018.

[13] Y. Li and G. A. A. Baduge, "Underlay spectrum-sharing massive MIMO NOMA," *IEEE Commun. Lett.*, vol. 23, no. 1, pp. 116–119, Jan. 2019.

[14] J. Zhang, E. Bjornson, 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.

[15] K. Hosseini, W. Yu, and R. S. Adve, "Large-scale MIMO versus network MIMO for multicell interference mitigation," in *Proc. IEEE 15th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jun. 2014, pp. 70–74.

[16] L. Dai, "A comparative study on uplink sum capacity with co-located and distributed antennas," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1200–1213, Jun. 2011.

[17] S. Chen, J. Zhang, J. Zhang, E. Björnson, and B. Ai, "A survey on user-centric cell-free massive MIMO systems," 2021, *arXiv:2104.13667*. [Online]. Available: https://arxiv.org/abs/2104.13667

[18] E. Koyuncu, "Performance gains of optimal antenna deployment in massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2633–2644, Apr. 2018.

[19] P. Liu, K. Luo, D. Chen, and T. Jiang, "Spectral efficiency analysis of cell-free massive MIMO systems with zero-forcing detector," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 795–807, Feb. 2020.

[20] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[21] H. Al-Hraishawi, G. A. A. Baduge, H. Q. Ngo, and E. G. Larsson, "Multi-cell massive MIMO uplink with underlay spectrum sharing," *IEEE Trans. Cognit. Commun. Netw.*, vol. 5, no. 1, pp. 119–137, Mar. 2019.

[22] L. Wang, H. Q. Ngo, M. Elkashlan, T. Q. Duong, and K.-K. Wong, "Massive MIMO in spectrum sharing networks: Achievable rate and power efficiency," *IEEE Syst. J.*, vol. 11, no. 1, pp. 20–31, Mar. 2017.

[23] Y. Li, D. Kudathanthirige, and G. A. A. Baduge, "Massive MIMO relay networks with underlay spectrum sharing," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 4, pp. 677–691, Dec. 2018.

[24] S. Silva, M. Ardakani, and C. Tellambura, "Interference suppression and energy efficiency improvement with massive MIMO and relay selection in cognitive two-way relay networks," *IEEE Trans. Green Commun. Netw.*, vol. 4, no. 2, pp. 326–339, Jun. 2020.

[25] D. Kudathanthirige and G. A. A. Baduge, "NOMA-aided multicell downlink massive MIMO," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 612–627, Jun. 2019.

[26] F. Rezaei, A. R. Heidarpour, C. Tellambura, and A. Tadaion, "Underlaid spectrum sharing for cell-free massive MIMO-NOMA," *IEEE Commun. Lett.*, vol. 24, no. 4, pp. 907–911, Apr. 2020.

[27] S. Kusaladharma and C. Tellambura, "Secondary user interference characterization for spatially random underlay networks with massive MIMO and power control," *IEEE Trans. Veh. Technol.*, vol. 66, no. 9, pp. 7897–7912, Sep. 2017.

[28] S. Kusaladharma and C. Tellambura, "Massive MIMO based underlay networks with power control," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.

[29] S. Chaudhari and D. Cabric, "QoS aware power allocation and user selection in massive MIMO underlay cognitive radio networks," *IEEE Trans. Cognit. Commun. Netw.*, vol. 4, no. 2, pp. 220–231, Jun. 2018.

[30] M. Filippou, D. Gesbert, and H. Yin, "Decontaminating pilots in cognitive massive MIMO networks," in *Proc. Int. Symp. Wireless Commun. Syst. (ISWCS)*, Aug. 2012, pp. 816–820.

[31] H. Masoumi and M. J. Emadi, "Performance analysis of cell-free massive MIMO system with limited fronthaul capacity and hardware impairments," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1038–1053, Feb. 2020.

[32] M. Alonzo, S. Buzzi, A. Zappone, and C. D'Elia, "Energy-efficient power control in cell-free and user-centric massive MIMO at millimeter wave," *IEEE Trans. Green Commun. Netw.*, vol. 3, no. 3, pp. 651–663, Sep. 2019.

[33] S. Chen, J. Zhang, E. Bjornson, 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.

[34] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge, U.K.: Cambridge Univ. Press, 2016.

[35] Z. Ding and H. V. Poor, "Design of massive-MIMO-NOMA with limited feedback," *IEEE Signal Process. Lett.*, vol. 23, no. 5, pp. 629–633, May 2016.

[36] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1993.

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

[38] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, May 2017.

[39] P. Marbach, "Priority service and max-min fairness," *IEEE/ACM Trans. Netw.*, vol. 11, no. 5, pp. 733–746, Oct. 2003.

[40] B. Radunovic and J.-Y. Le Boudec, "A unified framework for max-min and min-max fairness with applications," *IEEE/ACM Trans. Netw.*, vol. 15, no. 5, pp. 1073–1083, Oct. 2007.

[41] L. Zheng, D. W. H. Cai, and C. W. Tan, "Max-min fairness rate control in wireless networks: Optimality and algorithms by Perron-Frobenius theory," *IEEE Trans. Mobile Comput.*, vol. 17, no. 1, pp. 127–140, Jan. 2018.

[42] E. Bjornson, E. A. Jorswieck, M. Debbah, and B. Ottersten, "Multiobjective signal processing optimization: The way to balance conflicting metrics in 5G systems," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 14–23, Nov. 2014.

[43] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "How much do downlink pilots improve cell-free massive MIMO?" in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2016, pp. 1–7.

[44] H. Cramer, *Random Variables and Probability Distributions*. Cambridge, U.K.: Cambridge Univ. Press, 1970.

Gayan Amarasuriya Aruma Baduge (Senior Member, IEEE) received the B.Sc. degree (Hons.) in engineering from the Department of Electronics and Telecommunications Engineering, University of Moratuwa, Moratuwa, Sri Lanka, in 2006, and the Ph.D. degree in electrical engineering from the Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada, in 2013. He was a Post-Doctoral Research Fellow with the Department of Electrical Engineering, Princeton University, Princeton, NJ, USA, from 2014 to 2016. He is currently an Associate Professor with the School of Electrical, Computer, and Biomedical Engineering, Southern Illinois University, IL, USA. He is an Associate Editor of IEEE COMMUNICATIONS LETTERS, IEEE WIRELESS COMMUNICATIONS, and IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY.

Diluka Loku Galappaththige (Student Member, IEEE) received the B.Sc. degree (Hons.) from the Department of Electrical and Electronic Engineering, University of Peradeniya, Sri Lanka, in 2017. He is currently pursuing the Ph.D. degree with the School of Electrical, Computer, and Biomedical Engineering, Southern Illinois University, Carbondale, IL, USA. His current research interests include cell-free massive MIMO systems and intelligent reflective surfaces.