Rate Enhancement for Distributed Massive MIMO Systems with Underlay Spectrum Sharing

Fatemeh Rezaei[1], *Student Member, IEEE,* Chintha Tellambura[1], *Fellow, IEEE,* and Aliakbar Tadaion[2], *Senior Member, IEEE*
[1]Department of Electrical and Computer Engineering, University of Alberta, Edmonton, AB, Canada,
Email:{rezaeidi, ct4}@ualberta.ca, [2]Department of Electrical Engineering, University of Yazd, Yazd, Iran,
Email:tadaion@yazd.ac.ir

***Abstract*—We investigate the achievable rate of a single-cell distributed massive multi-input multi-output (MIMO) system underlaid a licensed primary multi-user co-located Massive MIMO network. We propose an access point (AP) selection procedure and power allocation (PA) method to improve the performance of the secondary network and derive the closed-form sum rate expression for Rayleigh fading channels by considering the effects of pilot contamination, inter-user interference and statistical downlink channel state information (CSI) at secondary users (SUs). Our results reveal that, the joint use of AP selection procedure and PA method significantly improve the achieved sum rate of the SUs while preserving the performance of the primary network.**

***Index Terms*—Distributed massive MIMO, underlay spectrum sharing, achievable sum rate.**

I. Introduction

Massive Multiple-Input Multiple-Output (MIMO) wireless has been identified as one of the key promising technologies for the fifth generation (5G) cellular networks, as it provides unprecedented spectral and energy efficiency gains with simple signal processing techniques [1], [2]. Thus, cognitive radios can also be integrated with massive MIMO to further enhance the spectral efficiency by opportunistic utilization of the spectrum resources [3]. In underlay spectrum sharing, secondary users (SUs) are allowed to simultaneously access the same licensed spectrum of the primary users (PUs). The caveat is that, the transmit power of the SUs must be curtailed to ensure that the co-channel interference (CCI) at the primary receivers is under a given threshold [3].

Massive MIMO techniques have thus been extensively investigated for cognitive radios with underlay spectrum sharing where both the primary and underlay networks employ massive MIMO at their base stations (BSs) [4], [5]. In [4], the SU interference for a cognitive massive MIMO system is specified by considering pilot contamination, path loss inversion power control and spatially random nodes. Furthermore, in [5], the uplink performance of a multi-cell, multi-user cognitive massive MIMO system underlaid a primary massive MIMO network is investigated and the achievable sum rates of the primary and secondary systems are derived. Importantly, this work assumes the availability of imperfect channel knowledge, a realistic assumption.

Since the underlay concept mandates that secondary transmit powers of the SUs must be below a certain peak level to limit interference on the primary network, flexibility and reconfigurability of the secondary system is a key design consideration. Distributed massive MIMO enables the secondary network to be dynamically adapted according to the network conditions; thereby leading to the performance improvement of PU/SU links [6], [7]. It uses distributed access points (APs) to provide better coverage and increase the macro diversity gain for the secondary network. It also has the flexibility to turn on or off some APs depending on the network load and traffic conditions. In [8], the performance of the secondary network is investigated where both the primary and underlay networks employ distributed massive MIMO. Thus, a set of distributed single-antenna APs serve the PUs/SUs simultaneously via conjugate precoding at the primary/secondary APs.

However, since massive MIMO is widely deployed for fifth generation (5G) cellular in the industry, e.g., Japanese SoftBank, China mobile, and Sprint in the United States [9], it makes more sense to consider massive MIMO as the primary BS in underlay networks. Following this idea, [10] investigated the performance of a secondary non-orthogonal multiple access (NOMA)-aided distributed massive MIMO system underlaid a primary massive MIMO network. The primary macro BS and secondary multi-antenna APs employ maximum ratio transmission (MRT) beamforming to transmit information to PUs and SUs, respectively. In order to maintain the performance of the PUs, a protected zone around the primary BS is enforced to satisfy the interference constraints.

Motivated by these considerations, in this paper, we investigate the achievable rate of a single-cell distributed massive MIMO system underlaid a licensed primary multi-user massive MIMO network. We consider a real scenario where unlike the previous works, the secondary APs and the SUs are spatially distributed in the given area without any constraints on their locations. We also consider an arbitrary number of active PUs/SUs in the network. The PUs are served by the primary macro BS, which uses zero forcing (ZF) beamforming. ZF yields better performance than MRT in terms of capacity in single-cell downlink massive MIMO systems [11]. In contrast, the secondary APs employ MRT beamforming to serve the SUs. This offers the advantage of having manageable fronthauling traffic for the secondary network.

Our contributions are summarized as follows; i) to improve the performance of the SUs while protecting the performance of the PUs, an AP selection procedure and power allocation (PA) strategy are proposed for the secondary network, ii) a closed-form sum rate expression of the SUs is derived by considering the effects of pilot contamination, inter-user interference and statistical downlink channel state information (CSI) at the SUs and iii) we show that the joint use of AP

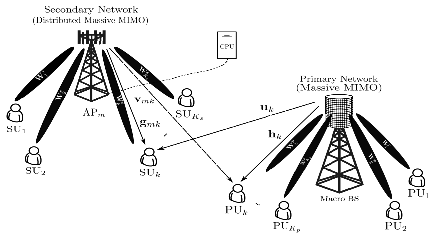

Fig. 1: System model of single-cell distributed massive MIMO system underlaid a licensed primary multi-user massive MIMO network.

selection procedure and PA method significantly improve the achieved sum rate of the SUs while preserving the performance of the primary network. Besides, by AP selection procedure and deactivating some of the secondary APs, the amount of overhead exchanged over the fronthaul network due to the payload data transmission decreases which is of great interest in practical scenarios. Numerical results are also presented to support our findings.

II. System Model and Preliminaries

A. System and Channel Models

We consider the downlink transmission of a single-cell distributed massive MIMO system, which is underlaid a licensed primary multi-user co-located massive MIMO system (Fig. 1). Underlaid refers to the fact that both the primary and secondary networks share the same licensed frequency spectrum. They both utilize time-division duplexing (TDD) in uplink and downlink and are synchronized perfectly to prevent undesired PU-SU interference [5]. The primary system consist of a macro BS equipped with L_p antennas and K_p spatially-distributed single-antenna users. In the secondary system, M APs, each with L_s antennas, are distributed uniformly in the cell and jointly serving K_s single-antenna SUs. Secondary APs are connected to a central processing unit (CPU) via an error-free fronthaul network to achieve coherent processing [12]. The information exchange between the secondary APs and the CPU is limited to only the payload data and large-scale parameters that change slowly. Due to the concurrent secondary transmissions, the transmit powers of the secondary APs must be constrained to manage the CCI inflicted at the PUs.

For the primary network, $\mathbf{h}_k \in \mathcal{C}^{L_p \times 1}$ represents the channel between the kth PU and the macro BS. In the secondary network, the channel between the kth SU and the mth secondary AP is $\mathbf{g}_{mk} \in \mathcal{C}^{L_s \times 1}$. The channel between the kth SU and the macro BS is $\mathbf{u}_k \in \mathcal{C}^{L_p \times 1}$. Moreover, $\mathbf{v}_{mk} \in \mathcal{C}^{L_s \times 1}$ is the channel between the mth secondary AP and the kth PU. All these four channels have the unified representation $\mathbf{a} = {\beta_{\mathbf{a}}}^{1/2}\bar{\mathbf{a}}$, where $\mathbf{a} \in \{\mathbf{h}_k, \mathbf{u}_k, \mathbf{g}_{mk}, \mathbf{v}_{mk}\}$ and $\beta_{\mathbf{a}}$ accounts for the large-scale pathloss and shadowing. This factor is constant for several coherence intervals. Whereas $\bar{\mathbf{a}} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{I}_{L_p|L_s}\right)$ captures Rayleigh fading.

B. Uplink Pilot Transmission

Due to the TDD mode in both primary and secondary networks, both uplink and downlink channels are the same. By exploiting this channel reciprocity, the primary macro BS and the secondary APs locally estimate the channel states by using uplink pilots transmitted by the PUs and SUs, respectively. In each coherence block of length τ_c, $\tau_p (< \tau_c)$ samples are used for uplink channel estimation. Here, we assume $\tau_p \geq \max(K_p, K_s)$. In this case, if $K_p < K_s$, the mutual orthogonal pilots assigned for K_p PUs are also shared among K_p SUs, while $K_s - K_p$ remaining SUs are assigned with mutual orthogonal pilots which are also orthogonal to those used by K_p PUs/SUs. On the other hand, if $K_s < K_p$, the mutual orthogonal pilots assigned for K_s SUs are also shared among K_s PUs, while $K_p - K_s$ remaining PUs are assigned with mutual orthogonal pilots which are also orthogonal to those used by K_s PUs/SUs.

Let the pilot sequence for an arbitrary user k be $\sqrt{\tau_p}\boldsymbol{\varphi}_k \in \mathcal{C}^{\tau_p \times 1}$ satisfying $\|\boldsymbol{\varphi}_k\|^2 = 1$. As the primary and secondary users transmit the pilot sequences in the uplink, the mth secondary AP estimates $\mathbf{g}_{mk}$ using minimum mean square error (MMSE) estimation [13]. The MMSE estimate of $\mathbf{g}_{mk}$ can be expressed as $\hat{\mathbf{g}}_{mk} = c_{mk}\tilde{\mathbf{y}}_{mk}^{s,\text{pilot}}$, where c_{mk} is given by [13] $c_{mk} = \sqrt{\tau_p p_p}\beta_{\mathbf{g}_{mk}}/(1 + \tau_p p_p(\beta_{\mathbf{g}_{mk}} + \mathcal{I}_k \beta_{\mathbf{v}_{mk}}))$ and $\tilde{\mathbf{y}}_{mk}^{s,\text{pilot}}$, the projected received pilot signal at the mth secondary AP onto $\boldsymbol{\varphi}_k$, is given as

$$\tilde{\mathbf{y}}_{mk}^{s,\text{pilot}} = \sqrt{\tau_p p_p}\mathbf{g}_{mk} + \mathcal{I}_k\sqrt{\tau_p p_p}\mathbf{v}_{mk} + \tilde{\mathbf{n}}_m, \qquad (1)$$

where p_p is the pilot transmit power and $\tilde{\mathbf{n}}_{mn}^s \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_L)$. Besides, $\mathcal{I}_k = 1$ for $1 \leq k \leq Q$ and $\mathcal{I}_k = 0$ otherwise, in which $Q = \min(K_p, K_s)$.

The channel can then be written as $\mathbf{g}_{mk} = \hat{\mathbf{g}}_{mk} + \tilde{\mathbf{g}}_{mk}$, in which the channel estimate and the estimation error, denoted by $\hat{\mathbf{g}}_{mk}$ and $\tilde{\mathbf{g}}_{mk}$, respectively, are distributed as $\hat{\mathbf{g}}_{mk} \sim \mathcal{CN}(\mathbf{0}, \theta_{\hat{\mathbf{g}}_{mk}}\mathbf{I}_{L_s})$, $\tilde{\mathbf{g}}_{mk} \sim \mathcal{CN}(\mathbf{0}, (\beta_{\mathbf{g}_{mk}} - \theta_{\hat{\mathbf{g}}_{mk}})\mathbf{I}_{L_s})$, where $\theta_{\hat{\mathbf{g}}_{mk}} = \tau_p p_p \beta_{\mathbf{g}_{mk}}^2/(1 + \tau_p p_p(\beta_{\mathbf{g}_{mk}} + \mathcal{I}_k\beta_{\mathbf{v}_{mk}}))$.

Similarly, for the primary network, the channel estimate and the estimation error are distributed as $\hat{\mathbf{h}}_k \sim \mathcal{CN}(\mathbf{0}, \theta_{\hat{\mathbf{h}}_k}\mathbf{I}_{L_p})$ and $\tilde{\mathbf{h}}_k \sim \mathcal{CN}(\mathbf{0}, (\beta_{\mathbf{h}_k} - \theta_{\hat{\mathbf{h}}_k})\mathbf{I}_{L_p})$ [13], where $\theta_{\hat{\mathbf{h}}_k} = \tau_p p_p \beta_{\mathbf{h}_k}^2/(1 + \tau_p p_p(\beta_{\mathbf{h}_k} + \mathcal{I}_k\beta_{\mathbf{u}_k}))$.

C. Downlink Data Transmission Model

The mth secondary AP transmits the signal $\mathbf{x}_m^s = \sqrt{p_t^s}\sum_{k=1}^{K_s}\sqrt{\lambda_{mk}^s}\mathbf{w}_{mk}^s s_k^s$, $\forall m$, where $\mathbf{w}_{mk}^s$ is the spatial directivity of the signal sent to the kth SU by the mth secondary AP. Symbols s_k^s, λ_{mk}^s and p_t^s denote the data signal, PA coefficient for the kth SU and the total transmitted power by each secondary AP, respectively. The set of PA coefficients satisfies $\sum_{k=1}^{K^s}\lambda_{mk}^s = 1$.

Similarly, the primary macro BS transmits the signal $\mathbf{x}_m^p = \sqrt{p_t^p}\sum_{k=1}^{K_p}\mathbf{w}_k^p\sqrt{\lambda_k^p}s_k^p$, where p_t^p, λ_k^p and $\mathbf{w}_k^p$ are respectively, the total transmit power, the power coefficients for the PUs ($\sum_{k=1}^{K^p}\lambda_k^p = 1$), and the precoding vector for the kth PU.

For the secondary network, to avoid sharing of channel state information between the secondary APs and to have manageable fronthauling traffic, we consider MRT beamforming which is given as [6] $\mathbf{w}_{mk}^s = \hat{\mathbf{g}}_{mk}/\sqrt{\mathbb{E}\{\|\hat{\mathbf{g}}_{mk}\|^2\}}$. The primary macro BS employs ZF beamforming to precode data signals and the precoding weight vector can be written as $\mathbf{w}_k^p = \hat{\mathbf{H}}(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}})^{-1}\mathbf{e}_k/\sqrt{\mathbb{E}\{\|\hat{\mathbf{H}}(\hat{\mathbf{H}}^{\mathrm{H}}\hat{\mathbf{H}})^{-1}\mathbf{e}_k\|^2\}}$, where $\hat{\mathbf{H}} = [\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \ldots, \hat{\mathbf{h}}_{K_p}]$ and $\mathbf{e}_k$ denotes the kth column of $\mathbf{I}_{K_p}$.

In order to enhance the performance of the SUs while protecting the primary network's communication, the CPU allows a set of secondary APs to serve the SUs and turns off the remaining secondary APs according to the criterion (3). Let the number of active secondary APs be M_0^s. Then, since these M_0^s secondary APs serve the K_s SUs simultaneously, using the statistical CSI knowledge of the effective channels at SUs i.e., $\mathbb{E}\left\{\mathbf{g}_{mk}^{\mathrm{H}}\mathbf{w}_{mk}^{s}\right\}$ $(\forall m,k)$, the received signal at the kth SU can be expressed as

$$
\begin{aligned}
y_k^s =& \underbrace{\sqrt{p_t^s}\sum_{m=1}^{M_0^s}\sqrt{\lambda_{mk}^s}\mathbb{E}\left\{\mathbf{g}_{mk}^{\mathrm{H}}\mathbf{w}_{mk}^{s}\right\}s_k^s}_{\text{desired signal}} \\
&+\underbrace{\sqrt{p_t^s}\sum_{m=1}^{M_0^s}\sqrt{\lambda_{mk}^s}\left(\mathbf{g}_{mk}^{\mathrm{H}}\mathbf{w}_{mk}^{s}-\mathbb{E}\left\{\mathbf{g}_{mk}^{\mathrm{H}}\mathbf{w}_{mk}^{s}\right\}\right)s_k^s}_{\text{beamforming gain uncertainty}} \\
&+\underbrace{\sum_{m=1}^{M_0^s}\mathbf{g}_{mk}^{\mathrm{H}}\sum_{j'=1,j'\neq k}^{K_s}\mathbf{w}_{mj}^{s}s_j^s}_{\text{inter-user interference}}+\underbrace{\sqrt{p_t^p}\mathbf{u}_k^{\mathrm{H}}\sum_{j=1}^{K_p}\mathbf{w}_j^p\sqrt{\lambda_j^p}s_j^p}_{\text{interference from the primary network}}+n_k^s, \quad (2)
\end{aligned}
$$

where $n_{nk}^s \sim \mathcal{CN}(0,1)$.

III. Performance analysis of the secondary network

In order to improve the performance of the underlaid distributed massive MIMO system while maintaining the performance of the primary massive MIMO network, we first investigate three important factors; i) AP selection procedure for the secondary network, ii) PA for the secondary network and iii) secondary transmit power constraints. After that, we derive the closed-form sum rate expression for the secondary network.

A. AP Selection for the Secondary Network

In order to enhance the spectral efficiency of the secondary network while preserving the primary network's performance, the CPU selects a set of secondary APs to serve the SUs and deactivates the remaining APs. This selection is based on the following idea. The CPU selects the secondary APs which have strong average effective channels with the SUs and the weak average interfering channels with the PUs. In particular, for each secondary AP, the CPU computes the ratio of its average effective local channels with the SUs and its average effective interfering channels with the PUs, $\mathcal{M}(m), \forall m$. So the secondary APs for which $\mathcal{M}(m)$ is higher than a predefined threshold δ, are selected to serve the SUs and the remaining secondary APs are turned off. The criterion is formally defined as follows: select the mth SU if $\mathcal{M}(m) > \delta$ where

$$
\mathcal{M}(m) = \frac{\sum_{k=1}^{K_s}\mathbb{E}\left\{|\mathbf{g}_{mk}^{\mathrm{H}}\mathbf{w}_{mk}^{s}|^2\right\}}{\sum_{k=1}^{K_p}\sum_{j=1}^{K_s}\mathbb{E}\left\{|\mathbf{v}_{mk}^{\mathrm{H}}\mathbf{w}_{mj}^{s}|^2\right\}},\ m=1,\ldots,M, \quad (3)
$$

which can be derived as (Appendix A)

$$
\mathcal{M}(m) = \frac{\sum_{k=1}^{K_s}\left(L_s\theta_{\hat{\mathbf{g}}_{mk}}+\beta_{\mathbf{g}_{mk}}\right)}{\left(\sum_{j=1}^{Q}\frac{L_s\tau_p p_p \beta_{\mathbf{v}_{mj}}^2}{\zeta_{mj}^s}+\sum_{j=1}^{K_p}\beta_{\mathbf{v}_{mj}}\right)}, \quad (4)
$$

where $\zeta_{mj}^s \triangleq 1+\tau_p p_p(\beta_{\mathbf{g}_{mj}}+\beta_{\mathbf{v}_{mj}})$.

As observed in (4), $\mathcal{M}(m), \forall m$ depends on the large scale parameters only. These parameters are easier to estimate and remain static for longer periods [14]. Thus, they are readily accessible by the CPU. Note that this secondary AP selection process for the secondary network has an added benefit. Since there are fewer active secondary APs, the amount of overhead exchanged over the fronthaul network due to the payload data transmission decreases.

B. Power Allocation for the Secondary Network

In the secondary network, each active secondary AP allocates its power to the SUs in proportion with the estimated channel strengths of the secondary links. In other words, the mth active secondary APs divides its power between the SUs such that, the SUs which have better channel strengths with respect to the mth secondary AP will receive a larger amounts of power than the SUs with bad channels condition. This way, each SU is dominantly served by the neighbouring APs with good channel strengths. The PA coefficients for the SUs can then be written as

$$
\lambda_{mk}^s = \frac{\theta_{\hat{\mathbf{g}}_{mk}}}{\sum_{j=1}^{K_s}\theta_{\hat{\mathbf{g}}_{mj}}}, \forall m,k. \quad (5)
$$

By employing this PA technique, not only the SUs are effectively served by the secondary APs, but also, the imposed interference on the primary network is decreased. Hence, this PA strategy is practically useful for the cognitive distributed massive MIMO networks.

C. Secondary Transmit Power Constraint

In underlay spectrum sharing, the transmit power of the secondary APs (p_t^s) are constrained to manage the interference imposed on the PUs due to the simultaneous secondary transmission

$$
p_{\mathrm{t}}^s = \min(p_{t,\max}^s, I_{p_1}/P(z_1),\ldots,I_{p_{K_p}}/P(z_{K_p})), \quad (6)
$$

where $p_{t,\max}^s$ is the maximum transmit power by each secondary AP and I_{p_k} is the interference temperature (maximum tolerable interference level) for the kth PU [8]. The interference power inflicted at the kth PU by the secondary network is scaled by $P(z_k)$ given as

$$
P(z_k) = \sum_{j=1}^{K_s}\mathbb{E}\left\{\left|\sum_{m=1}^{M_0^s}\sqrt{\lambda_{mj}^s}\mathbf{v}_{mk}^{\mathrm{H}}\mathbf{w}_{mj}\right|^2\right\},\ k=1,\ldots,K_p. \quad (7)
$$

which can be derived as (Appendix B)

$$
P(z_k) = \mathcal{I}_k\left(\sum_{m=1}^{M_0^s}\sqrt{\frac{L_s\lambda_{mk}^s\tau_p p_p}{\zeta_{mk}^s}}\beta_{\mathbf{v}_{mk}}\right)^2+\sum_{m=1}^{M_0^s}\beta_{\mathbf{v}_{mk}}. \quad (8)
$$

We must note that, by selecting the appropriate set of secondary APs for transmission, we are specifically maximizing the allowable transmit power of the selected secondary APs.

D. Downlink achievable Rate

In the following, we derive the achievable rate of the secondary distributed massive MIMO network in the presence of the primary co-located massive MIMO system. The kth SU in (2) effectively sees a deterministic channel with some uncorrelated noise (the first term and the remaining terms in (2), respectively). By using the worst-case Gaussian technique [6], the achievable rate for the kth SU can then be written as

$$
\mathcal{R}_k^s = \phi\log_2(1+\gamma_k^s), \quad (9)
$$

where $\phi = (\tau_c - \tau_p)/\tau_c$ is the pre-log factor which captures the effective portion of coherence interval for data transmission. Let γ_k^s be the effective signal-to-interference-plus-noise ratio (SINR) at the kth SU. To compute it, we consider the first term in (2) to be the desired signal and the remaining terms be an effective noise. Therefore, γ_k^s can be derived as (10), as shown at the top of the next page. By evaluating the expectation terms in (10), the effective SINR γ_k^s can be derived (Appendix C) as (11), as shown at the top of the next page, where $\zeta_k^p \triangleq 1 + \tau p_p(\beta_{\mathbf{u}_k} + \beta_{\mathbf{h}_k})$.

IV. Simulation Results

Herein, we provide simulation results to evaluate the performance of our secondary system. We consider a circular area of radius 500 m. The primary macro BS is located at the center and the K_p PUs are distributed uniformly at random in the cell. Moreover, the M^s secondary APs and K_s SUs are uniformly distributed in the given area. Uniform PA is considered for the primary network ($\lambda_k^p = 1/K_p$). The large-scale coefficients $\{\beta_{\mathbf{a}}\}$ are an uncorrelated shadow fading process with standard deviation $\sigma_{\text{sh}} = 8$ dB [6]. In all simulations, we assume that $L_p = 100$, $M^s = 100$, $L_s = 2$, $p_t^p = p_{t,\max}^s = 200$ mW, $p_p = 100$ mW, $K_p = 8$, $K_s = 12$. We also assume that $\tau_c = 56$ and the pilot sequence length is set to $\tau_p = \max(K_p, K_s) = 12$.

To specify how many APs are selected by the CPU according to the criteria given in (4), we first plot the number of selected secondary APs versus the threshold δ in Fig. 2a. As expected, by increasing δ, the number of secondary APs which are allowed to serve the SUs M_0^s decreases. In particular, for $\delta = 1.5$, on average 57 secondary APs are selected from the total number of 100 secondary APs to serve the SUs.

Fig. 2b demonstrates the secondary transmit power constraint (6) as a function of the interference temperature I_p. As a benchmark, we also plot the secondary transmit power constraint when all the secondary APs serve the SUs with uniform PA method. Both the uniform PA method and the proposed one in (5) are simple and can be performed at each AP independently rather than at the CPU, and hence does not require a fronthaul link (an important feature of practical communication systems).

Here, we consider $\delta = 1.5$; hence, with AP selection process, there are $M_s^0 = 57$ active secondary APs which is much lower than the case without AP selection ($M_s = 100$). Monte-Carlo simulations are also superimposed on each curve in order to validate the derived analytical expression. As expected, the secondary transmit power constraint increases with I_p and the transmit power of each secondary AP is limited by the primary network. Fig. 2b clearly reveals that, by deactivating the secondary APs which may impose considerable interference on the primary network, the remaining secondary APs are allowed to transmit with higher power while preserving the primary network's performance. More precisely, for $I_p = -5$ dB, with AP selection procedure and the PA strategy given in (5), the active secondary APs are allowed to transmit with the power $p_t^s = 14.42$ mW which is about 14 times more than the case when all the secondary APs serve the SUs using the same PA method ($p_t^s = 1.05$ mW). Furthermore, by allocating power to the SUs based on (5), each SU is dominantly served by the neighbouring active secondary APs with good channel strength and hence less interference is imposed on the primary network. Therefore, the secondary transmit power constraint increases compared to the case when uniform PA is employed.

Fig. 2c depicts the achievable sum rate (9) of the secondary network as a function of I_p. We observe that, the achieved sum rate by the SUs increases with I_p, since they are allowed to transmit with higher powers. Besides, the joint use of the proposed AP selection process and PA method significantly improve the sum rate of the secondary network. We also see that, although, the allowable transmit power of the secondary APs is higher with AP selection procedure and uniform PA, compared to the case without AP selection procedure and PA method (5), its achieved sum rate is less. This is due to the fact that, employing the proposed PA method, each secondary AP effectively allocates its available power to the SUs such that the SUs with better channel strengths receive a larger amounts of power than the SUs with weak channel condition which results in higher rates. Therefore, the proposed PA strategy is practically useful for the cognitive distributed massive MIMO networks.

V. Conclusions

In this paper, in a real scenario, we analysed the achieved sum rate of a distributed massive MIMO system underlaid a licensed primary massive MIMO network. We proposed an AP selection process and PA method for the secondary network. We showed that, the joint use of the proposed methods significantly improve the performance of the SUs while protecting the performance of the PUs.

Appendix A
Derivation of $\mathcal{M}(m)$

Since, $\mathbf{g}_{mk} = \hat{\mathbf{g}}_{mk} + \tilde{\mathbf{g}}_{mk}$, we have $\mathbb{E}\{|\mathbf{g}_{mk}^{\mathrm{H}}\mathbf{w}_{mk}^s|^2\} = L_s\theta_{\hat{\mathbf{g}}_{mk}} + \beta_{\mathbf{g}_{mk}}$. Besides, to calculate the denominator of (3), for the kth PU, if $1 \leq k \leq Q$

$$\sum_{j=1}^{K_s} \lambda_{mj}^s \mathbb{E}\left\{|\mathbf{v}_{mk}^{\mathrm{H}}\mathbf{w}_{mj}^s|^2\right\} = \lambda_{mk}^s \frac{L_s\tau_p p_p \beta_{\mathbf{v}_{mk}}^2}{\zeta_{mk}^s} + \sum_{j=1}^{K_s}\lambda_{mj}^s\beta_{\mathbf{v}_{mk}}, \tag{12}$$

Otherwise, i.e., $k > Q$, $\sum_{j=1}^{K_s} \lambda_{mj}^s \mathbb{E}\left\{|\mathbf{v}_{mk}^{\mathrm{H}}\mathbf{w}_{mj}^s|^2\right\} = \sum_{j=1}^{K_s}\lambda_{mj}^s\beta_{\mathbf{v}_{mk}}$. (13)

By substituting (12) and (13) in (3), $\mathcal{M}(m)$ is given as (4).

Appendix B
Derivation of $P(z_n)$

To calculate $P(z_k)$ (7), for the kth PU, if $1 \leq k \leq Q$

$$P(z_k) = \sum_{\substack{j=1\\ j\neq k}}^{K_s}\sum_{m=1}^{M_0^s} \lambda_{mj}^s \mathbb{E}\left\{\left|\mathbf{v}_{mk}^{\mathrm{H}}\mathbf{w}_{mj}^s\right|^2\right\} + \sum_{m=1}^{M_0^s}\lambda_{mk}^s\mathbb{E}\left\{\left|\mathbf{v}_{mk}^{\mathrm{H}}\mathbf{w}_{mk}^s\right|^2\right\}$$
$$+ \sum_{m=1}^{M_0^s}\sum_{\substack{m'=1\\ m'\neq m}}^{M_0^s}\sqrt{\lambda_{mk}^s}\sqrt{\lambda_{m'k}^s}\,\mathbb{E}\left\{\left(\mathbf{v}_{mk}^{\mathrm{H}}\mathbf{w}_{mk}^s\right)\left(\mathbf{v}_{m'k}^{\mathrm{H}}\mathbf{w}_{m'k}^s\right)^{\mathrm{H}}\right\}, \tag{14}$$

$$\gamma_k^s = \frac{p_t^s \left| \mathbb{E}\left\{ \sum_{m=1}^{M_0^s} \sqrt{\lambda_{mk}^s} \mathbf{g}_{mk}^{\rm H} \mathbf{w}_{mk}^s \right\} \right|^2}{p_t^s \mathbb{V}ar \left\{ \sum_{m=1}^{M_0^s} \sqrt{\lambda_{mk}^s} \mathbf{g}_{mk}^{\rm H} \mathbf{w}_{mk}^s \right\} + p_t^s \sum_{j=1, j\neq k}^{K_s} \mathbb{E}\left\{ \left| \sum_{m=1}^{M_0^s} \sqrt{\lambda_{mj}^s} \mathbf{g}_{mk}^{\rm H} \mathbf{w}_{mj}^s \right|^2 \right\} + P_t^p \sum_{j=1}^{K_p} \lambda_j^p \mathbb{E}\left\{ \left| \mathbf{u}_k^{\rm H} \mathbf{w}_j^p \right|^2 \right\} + 1}. \quad (10)$$

$$\gamma_k^s = \frac{L_s p_t^s \left(\sum_{m=1}^{M_0^s} \sqrt{\lambda_{mk}^s \theta_{\hat{\mathbf{g}}_{mk}}} \right)^2}{p_t^s \sum_{j=1}^{K_s} \sum_{m=1}^{M_0^s} \lambda_{mj}^s \beta_{\mathbf{g}_{mk}} + \mathcal{I}_k P_t^p \left((L_p - K_p) \lambda_k^p \frac{\tau_p p_p \beta_{\mathbf{u}_k}^2}{\zeta_k^p} + \sum_{j=1}^{K_p} \lambda_j^p \left(\beta_{\mathbf{u}_k} - \frac{\tau_p p_p \beta_{\mathbf{u}_k}^2}{\zeta_k^p} \right) \right) + (1 - \mathcal{I}_k) P_t^p \sum_{j=1}^{K_p} \lambda_j^p \beta_{\mathbf{u}_k}}. \quad (11)$$

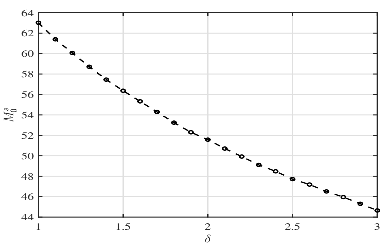

Fig. 2(a): The number of active secondary APs M_0^sversus δ. The curve is generated analytically, not via simulations.

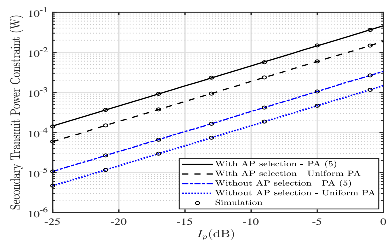

Fig. 2(b): Transmit power constraint versus I_p. The curves are generated using analysis and simulation.

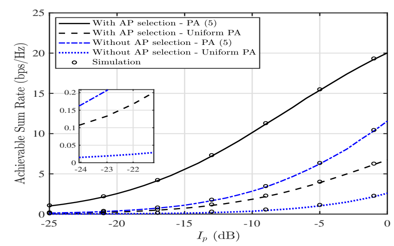

Fig. 2(c): Achievable sum rate versus interference temperature I_p. The curves are generated using analysis and simulation.

where the first and the second terms are calculated in (12) and

$$\mathbb{E}\left\{ \left(\mathbf{v}_{mk}^{\rm H} \mathbf{w}_{mk}^s \right) \left(\mathbf{v}_{m'k}^{\rm H} \mathbf{w}_{m'k}^s \right)^{\rm H} \right\} = \frac{L_s \tau_p p_p}{\sqrt{\zeta_{mk}^s \zeta_{m'k}^s}} \beta_{\mathbf{v}_{mk}} \beta_{\mathbf{v}_{m'k}}. \quad (15)$$

Otherwise, i.e., $k > Q$, $P(z_k) = \sum_{j=1}^{K_s} \sum_{m=1}^{M_0^s} \lambda_{mj}^s \beta_{\mathbf{v}_{mk}}$. (16)

Finally, substituting (14), (15) and (16) in (7), $P(z_n)$ is given as (8).

Appendix C
Derivation of γ_k^s(10)

To find γ_{nk}^s, similar to Appendix A, the following terms can be derived as:

$$\mathbb{E}\left\{ \sum_{m=1}^{M_0^s} \sqrt{\lambda_{mk}^s} \mathbf{g}_{mk}^{\rm H} \mathbf{w}_{mk}^s \right\} = \sum_{m=1}^{M_0^s} \sqrt{\lambda_{mk}^s L_s \theta_{\hat{\mathbf{g}}_{mk}}}, \quad (17)$$

$$\mathbb{V}ar \left\{ \sum_{m=1}^{M_0^s} \sqrt{\lambda_{mk}^s} \mathbf{g}_{mk}^{\rm H} \mathbf{w}_{mk}^s \right\} = \sum_{m=1}^{M_0^s} \lambda_{mk}^s \beta_{\mathbf{g}_{mk}}, \quad (18)$$

$$\sum_{\substack{j=1 \\ j\neq k}}^{K_s} \mathbb{E}\left\{ \left| \sum_{m=1}^{M_0^s} \sqrt{\lambda_{mj}^s} \mathbf{g}_{mk}^{\rm H} \mathbf{w}_{mj}^s \right|^2 \right\} = \sum_{\substack{j=1 \\ j'\neq k}}^{K_s} \sum_{m=1}^{M_0^s} \lambda_{mj}^s \beta_{\mathbf{g}_{mk}}. \quad (19)$$

Moreover, to calculate $\sum_{j=1}^{K_p} \lambda_j^p \mathbb{E}\{|\mathbf{u}_k^{\rm H} \mathbf{w}_j^p|^2\}$, we first obtain the normalization term in the primary beamforming vector $\mathbf{w}_k^p$. Employing Lemma 2.10 of [15], we obtain $\mathbb{E}\{\| \hat{\mathbf{H}}(\hat{\mathbf{H}}^{\rm H}\hat{\mathbf{H}})^{-1}\mathbf{e}_k \|^2\} = [(L_p - K_p)\theta_{\hat{\mathbf{h}}_k}]^{-1}$. Then, for the kth SU, if $1 \leq k \leq Q$

$$\sum_{j=1}^{K_p} \lambda_j^p \mathbb{E}\left\{ \left| \mathbf{u}_k^{\rm H} \mathbf{w}_j^p \right|^2 \right\} = (L_p - K_p) \lambda_k^p \frac{\tau_p p_p \beta_{\mathbf{u}_k}^2}{\zeta_k^p} + \sum_{j=1}^{K_p} \lambda_j^p \left(\beta_{\mathbf{u}_k} - \frac{\tau_p p_p \beta_{\mathbf{u}_k}^2}{\zeta_k^p} \right). \quad (20)$$

Otherwise, i.e., $k > Q$, $\sum_{j=1}^{K_p} \lambda_j^p \mathbb{E}\left\{ \left| \mathbf{u}_k^{\rm H} \mathbf{w}_j^p \right|^2 \right\} = \sum_{j=1}^{K_p} \lambda_j^p \beta_{\mathbf{u}_k}$. (21)

Finally, substituting (17)-(21) in (10), γ_{nk}^s is given as (11).

References

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

[2] F. Rezaei and A. Tadaion, "Multi-layer beamforming in uplink/downlink massive MIMO systems with multi-antenna users," *Signal Processing*, vol. 164, pp. 58–66, 2019.

[3] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proc. IEEE*, vol. 97, pp. 894–914, May 2009.

[4] 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, pp. 7897–7912, Sep. 2017.

[5] H. Al-Hraishawi, G. A. Aruma Baduge, H. Q. Ngo, and E. G. Larsson, "Multi-cell massive MIMO uplink with underlay spectrum sharing," *IEEE Trans. Cognitive Commun. and Networking*, vol. 5, pp. 119–137, March 2019.

[6] 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, pp. 1834–1850, Mar. 2017.

[7] F. Rezaei, C. Tellambura, A. Tadaion, and A. R. Heidarpour, "Rate analysis of cell-free massive MIMO-NOMA with three linear precoders," *IEEE Trans. Commun.*, pp. 1–1, 2020.

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

[9] Sprint unveils six 5G-ready cities; significant milestone toward launching first 5G mobile network in the U.S. [Online]. Available: https://newsroom.sprint.com/sprint-5g-overview-1-2.htm.

[10] F. Rezaei, A. R. Heidarpour, C. Tellambura, and A. Tadaion, "Underlaid spectrum sharing for cell-free massive MIMO-NOMA," *IEEE Commun. Lett.*, pp. 1–1, 2020.

[11] T. Parfait, Y. Kuang, and K. Jerry, "Performance analysis and comparison of ZF and MRT based downlink massive MIMO systems," in *Proc. 6th Int. Conf. Ubiquitous Future Netw.*, pp. 383–388, July 2014.

[12] N. Akbar, E. Bjoernson, E. G. Larsson, and N. Yang, "Downlink power control in massive MIMO networks with distributed antenna arrays," in *Proc.IEEE ICC*, pp. 1–6, May 2018.

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

[14] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2005.

[15] A. M. Tulino and S. Verdu, *Random Matrix Theory and Wireless Communications*, vol. 1. Jan. 2004.