Modified Vector Quantization for Small-Cell Access Point Placement with Inter-Cell Interference

Govind R. Gopal, *Graduate Student Member, IEEE,* Elina Nayebi,
Gabriel Porto Villardi, *Senior Member, IEEE,* and Bhaskar D. Rao, *Fellow, IEEE*

***Abstract*—In this paper, we explore the small-cell uplink access point (AP) placement problem in the context of throughput optimality and provide solutions while taking into consideration inter-cell interference (ICI). First, we briefly review the vector quantization (VQ) approach and related single user throughput-optimal formulations for AP placement. Then, we investigate the small-cell case with multiple users and expose the limitations of mean squared error based VQ for solving this problem. While the Lloyd algorithm from the VQ approach is found not to strictly solve the small-cell case, based on the tractability and quality of the resulting AP placement, we deem it suitable as a simple and appropriate framework to solve more complicated problems. Accordingly, to minimize ICI and consequently enhance achievable throughput, we design two Lloyd-type algorithms, namely the Interference Lloyd algorithm and the Inter-AP Lloyd algorithm, both of which incorporate ICI in their distortion functions. Simulation results show that both of the proposed algorithms provide superior 95%-likely rate over the traditional Lloyd algorithm and the Inter-AP Lloyd algorithm yields a significant increase of up to 36.34% in achievable rate over the Lloyd algorithm.**

***Index Terms*—Base station placement, Beyond 5G, Lloyd algorithm, SINR minimization, throughput optimization, user cell association.**

I. Introduction

THE past decade has witnessed the surge of wireless communication technologies that have significantly raised the bits-per-second-per-hertz figure of merit and throughput of wireless networks in order to cope with the ongoing widespread adoption of mobile broadband by society. One of these key technologies, massive multiple-input-multiple-output (massive MIMO) [2]–[6], utilizes hundreds of antenna elements and is particularly appealing since intra-cell interference and small-scale fading are naturally canceled out due to favorable propagation and channel hardening [7], leading to low-cost hardware implementations [8]. Distributed antenna systems (DASs), especially in the form of distributed MIMO [9]–[15], give rise to even higher average rates over co-located MIMO systems [10], [16], [17]. In general, distributed massive MIMO can either be cooperative or non-cooperative, with cell-free massive MIMO [18], [19] being an example of the former. Although cooperation between the distributed antenna elements in the cell-free approach assists in mitigating interference between users and further increases spectral efficiency over non-cooperative massive MIMO systems, the required user-related information exchange occupies a significant portion of the usually limited back-haul capacity of wireless systems [9], [20], [21]. As a result, despite the benefits of cell-free massive MIMO systems, near future deployments of 5G (3GPP Rel-15 and Rel-16 [22]) and WiFi 6 (IEEE 802.11ax [23]) will still be based on the concept of small-cells, possibly leaving cell-free approaches to posterior deployments of the technology. Hence, controlling inter-cell interference (ICI) continues to be a major system design problem, which will actually assume much greater significance with the expected network densification of Beyond 5G systems. Currently, ICI is dealt with in the standards by advanced scheduling techniques such as basic service set (BSS) coloring [23] and dynamic time division duplex (D-TDD) [24]. Also, controlling ICI is of great importance to public protection and disaster relief (PPDR) wireless networks occupying the 700 MHz (and below) frequency bands due to their desirable propagation characteristics and higher signal penetration capabilities, which can cause severe service outages to adjacent emergency networks even in not-so-dense deployments [25]–[28].

In this work, we exploit another degree of freedom in system design, namely access point (AP) placement, in order to tackle ICI in small-cell wireless systems with non-uniform user distributions. To contextualize the discussion, consider a large gathering such as a sporting event, where sections in the stadium see a different number and arrangement of spectators depending on the crowd on the day of the event. To avoid service interruption, more APs should be placed where the number of spectators is larger, and vice versa, leading to the concept of smart stadiums. Additionally, flexible AP deployment is of utmost importance in the infrequent emergency and disaster relief situations, where deployments should be tailored to the time-specific coverage and service requirements, therefore following the dynamics of the emergency event [25]. Thus, the question of interest is: *How do we optimally place the APs given the distribution of users?* In recent times, the

Govind R. Gopal and Bhaskar D. Rao are with the Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA 92093, USA (e-mail: ggopal@ucsd.edu; brao@ucsd.edu). The work of Govind R. Gopal and Bhaskar D. Rao was supported in part by the Center for Wireless Communications (CWC), University of California San Diego, in part by Qualcomm Inc. through the Faculty-Mentor-Advisor program, and in part by the National Science Foundation (NSF) under Grant CCF-2124929.

Elina Nayebi is with Apple Inc., Cupertino, CA 95014, USA (e-mail: e.nayebi@gmail.com). The work of Elina Nayebi was carried out when she was a PhD student at the University of California San Diego.

Gabriel Porto Villardi is with the Wireless Network Research Center of the National Institute of Information and Communications Technology (NICT), Yokosuka 239-0847, Japan. The work of Gabriel Porto Villardi was partly carried out when he was a visiting scholar with the Qualcomm Institute of Calit2, University of California San Diego, La Jolla, CA 92093, USA (e-mail: gpvillardi@nict.go.jp).

This article was presented in part at the 2021 55th Asilomar Conference on Signals, Systems, and Computers [1].

AP (or antenna) placement problem has attracted a great deal of attention [13]–[15], [29], however, optimizing the AP or antenna locations by maximizing a signal-to-noise ratio (SNR) objective function alone has traditionally been the standard approach. The authors of [13] consider a DAS and optimize the cell averaged ergodic capacity based only on SNR and neglect ICI. Using the square distance criterion, they notice similarities with codebook design in vector quantization (VQ), which enables the utilization of the well-known (for ease of implementation) Lloyd algorithm to solve the antenna placement problem. In [15], the average achievable per-user rate of uniformly distributed users is optimized in order to find the radius of a circular antenna array; however, due to the adoption of a single-cell model, no ICI is considered. Circular antenna array deployments based on average rate optimization are also considered in [14] based on one-cell and two-cell models, with the latter model accounting for leakage interference alone. Additionally, the authors of [29] simulated an indoor wireless environment where they generated a 10-fold improvement in the distributed system capacity over the co-located one. Further, placing APs in accordance with the user densities generated a significant increase (40% over uniform AP placement) in system capacity. The authors of [30] and the subsequent works by their group [31]–[34] have discussed heterogeneous wireless sensor network deployment as a source coding problem. In these works, the optimal deployment is solved while studying limited communication range and optimal total power consumption to place both APs and fusion centers, but without addressing ICI. Recently, unmanned aerial vehicles (UAVs) equipped with base stations have also been considered in the context of AP placement [35]–[42].

In the abovementioned works, the suitability of the Lloyd algorithm from VQ to throughput-optimal AP placement has not been investigated. VQ considers only a single user, and the objective function is averaged over the position of this user. This approach, however, does not conform with the small-cell scenario where there are multiple users, one from each cell, communicating with the serving AP in its cell. Further, ICI has been neglected, therefore leading to AP placements that yield sub-optimal throughput. Hence, in this work, we devise a non-cooperative small-cell system based on the Lloyd algorithm, which we show can solve for near-optimal AP locations, in terms of the fundamental performance measure of throughput, while considering ICI.

Contributions

To the best of our knowledge, solutions to the AP placement problem based on the Lloyd algorithm and that are derived from a detailed analysis of throughput optimality, while incorporating ICI, have not been provided in the literature. Hence, in this work on small-cell AP placement, our contributions are as follows.

- We first formulate various single user AP placement problems for throughput optimality in terms of rate, SNR, and a higher exponent for the user-AP distance (as opposed to squared distance). We explore the relationship of the Lloyd algorithm from VQ to these problems. We then study the multiple user case and address the small-cell AP placement problem. Although our analysis determines that the application of the Lloyd algorithm to small-cell AP placement is not ideal, we find that the Lloyd algorithm, apart from being easy to implement, is quite effective in solving the placement problem as a baseline algorithm and yields near-optimal AP locations.
- We present two methods to incorporate ICI into the optimization function of the Lloyd algorithm. Consequently, the distortion function of the Lloyd algorithm is modified and two Lloyd-type algorithms for AP placement that are aware of ICI and, as a result, maximize achievable per-user SINR, are proposed, namely the Interference Lloyd algorithm and Inter-AP Lloyd algorithm.

The remainder of this paper is organized as follows. Section II outlines the small-cell model used throughout the paper. VQ and the application of the Lloyd algorithm to the AP placement problem is described in Section III. Mathematical formulations of throughput optimality for single user and multiple user cases are provided in IV. The ensuing section V presents the formulations and solutions for including ICI in the VQ approach. Cell association strategies for each of the proposed algorithms are elucidated in Section VI. The simulation methodology and results are stated in Section VII. Finally, we provide concluding remarks in Section VIII.

II. System Model

We use the small-cell model detailed in [43], [44], and [45, Ch. 4], which is reproduced here for completeness. Also, throughout this paper, we use bold symbols to denote vectors, $\mathbb{E}\{\cdot\}$ is the expectation operator, $||\cdot||$ represents the ℓ_2-norm of a vector, and all logarithms are to the base 2. Now, consider a geographical area where K single-antenna users are distributed, according to some probability density function (pdf) $f_{\mathbf{P}}(\mathbf{p})$, where $\mathbf{p} \in \mathbb{R}^2$ is the random vector denoting the position of a user. There are M single-antenna APs that serve the users in this area. The location of an AP is denoted by $\mathbf{q} \in \mathbb{R}^2$. All APs are connected via error-free backhaul links to the network controller[1] (NC), so that it knows the positions of the APs and their respective users. For simplicity, a narrowband flat-fading channel is considered. With $m = 1, 2, \ldots, M$ and $k = 1, 2, \ldots, K$, the channel coefficient between the m^{th} AP and k^{th} user is

$$g_{mk} = \sqrt{\beta_{mk}} h_{mk}, \tag{1}$$

where β_{mk} and $h_{mk} \sim \mathcal{CN}(0,1)$ are the large-scale and small-scale fading coefficients, respectively. h_{mk} is assumed to remain constant during a coherent interval and change independently in the next, and is independent of β_{mk}. The large-scale fading coefficients are modeled as

$$\beta_{mk} = \begin{cases} c_0, & ||\mathbf{p}_k - \mathbf{q}_m|| \leq r_0, \\ \frac{c_1 z_{mk}}{||\mathbf{p}_k - \mathbf{q}_m||^{\gamma}}, & ||\mathbf{p}_k - \mathbf{q}_m|| > r_0, \end{cases} \tag{2}$$

where $\mathbf{p}_k$ and $\mathbf{q}_m$ represent the locations of the k^{th} user and m^{th} AP, respectively. Here, γ is the pathloss exponent, z_{mk}

[1]The NC is where the proposed placement algorithms to be described in detail in the remainder of this manuscript will be loaded and executed.

is the log-normal shadow fading coefficient, and c_0, c_1, and r_0 are constants. These coefficients can also be estimated by either ray-tracing [46] or data-driven [47] approaches.

The uplink transmission model used in this work schedules users in a round robin fashion with their serving APs using time-division multiple access (TDMA). Thus, each AP serves only one user in a time slot. In the small-cell set-up, each of the M cells corresponds to each of the M APs, and pursuant with the uplink model, the user in each cell communicating with its associated AP causes interference to all other APs. Now, letting k_m denote a user in the cell associated with AP m, the received signal y_m at this AP is

$$y_m = \sum_{m'=1}^{M} \sqrt{\rho_r} g_{mk_{m'}} s_{k_{m'}} + w_m, \tag{3}$$

where ρ_r is the uplink transmit power, s_{k_m} is the data symbol with $\mathbb{E}\{|s_{k_m}|^2\} = 1$ (unit power), and $w_m \sim \mathcal{CN}(0,1)$ is the additive noise. A matched filter (MF) employed at the AP m estimates the data symbol s_{k_m} of user k_m as

$$\begin{aligned}\hat{s}_{k_m} &= \frac{g^*_{mk_m}}{|g_{mk_m}|} y_m, \\ &= \underbrace{\sqrt{\rho_r}|g_{mk_m}| s_{k_m}}_{T_{\text{des}}\text{: desired term}} + \underbrace{\sum_{\substack{m'=1 \\ m' \neq m}}^{M} \sqrt{\rho_r} \frac{g^*_{mk_m}}{|g_{mk_m}|} g_{mk_{m'}} s_{k_{m'}}}_{T_{\text{int}}\text{: interference term}} + v_m,\end{aligned} \tag{4}$$

where $v_m \sim \mathcal{CN}(0,1)$. Considering T_{int} in (4) as noise, the signal-to-interference-plus-noise ratio (SINR) achieved by user k_m at AP m is derived to be

$$\phi_{k_m} = \frac{\rho_r \beta_{mk_m} |h_{mk_m}|^2}{1 + \rho_r \sum\limits_{\substack{m'=1 \\ m' \neq m}}^{M} \beta_{mk_{m'}} |h_{mk_{m'}}|^2}. \tag{5}$$

III. Vector Quantization and AP Placement

In this section, we provide an overview of VQ and how the Lloyd algorithm is currently used in its basic form, to solve for AP placement. Note that Section IV will investigate the suitability of the Lloyd algorithm to obtain AP locations.

A. Overview of Vector Quantization

In VQ, the random vector to be quantized is $\mathbf{x} \in \mathbb{R}^p$, where p is the dimension, and the two main steps to be designed are the encoding and decoding steps. The encoder $\mathcal{E}$ splits the domain under consideration into N regions (called Voronoi regions, each corresponding to a bit sequence of length $\log_2 N$) and assigns a region $\mathcal{R}$ to the input vector $\mathbf{x}$. The encoder performs the following mapping

$$\mathcal{E} : \mathbb{R}^p \to \{\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_N\}. \tag{6}$$

The decoder $\mathcal{D}$ then assigns to each region $\mathcal{R}_n$, where $n = 1, 2, \ldots, N$, a codepoint $\hat{\mathbf{x}}_n$, and performs the mapping

$$\mathcal{D} : \{\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_N\} \to \{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_N\}. \tag{7}$$

The set of codepoints $\{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_N\}$ is collectively the codebook. Thus, the quantizer $\mathcal{Q}$ assigns for every input $\mathbf{x}$, one of N codepoints, and is given as

$$\mathcal{Q}(\mathbf{x}) = \mathcal{D}(\mathcal{E}(\mathbf{x})) = \hat{\mathbf{x}}_{\mathcal{E}(\mathbf{x})}, \tag{8}$$

where $\hat{\mathbf{x}}_{\mathcal{E}(\mathbf{x})}$ specifies that the output codepoint is a function of the input vector and for simplicity in notation, we assume that $\mathcal{E}(\mathbf{x})$ denotes the index of the region that it specifies. The encoder $\mathcal{E}$ assigns to the input $\mathbf{x}$, the region that is closest to it, defined in terms of a distortion function d between the input vector and a codepoint. The codepoint corresponding to the region can formally be written as

$$\hat{\mathbf{x}}_{\mathcal{E}(\mathbf{x})} = \arg\min_{\hat{\mathbf{x}}_n} d(\mathbf{x}, \hat{\mathbf{x}}_n). \tag{9}$$

Taking the average of the distortion function over the distribution of the input vector, the VQ optimization problem is

$$\arg\min_{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \ldots, \hat{\mathbf{x}}_N} \mathbb{E}_{\mathbf{x}} \left\{ d(\mathbf{x}, \hat{\mathbf{x}}_{\mathcal{E}(\mathbf{x})}) \right\}. \tag{10}$$

To solve the optimization of (10), the goal is to find the optimal encoder and decoder jointly, which is difficult. Hence, it is split into two tasks, which are to find a optimal encoder given a fixed decoder and a optimal decoder given a fixed encoder, and form the two necessary conditions for quantizer optimality. The main methodology then is to alternate between these two tasks in order to converge to a reasonable solution. Accordingly, finding the best encoder given the decoder involves determining the best regions given fixed codepoints. This leads to the *Nearest Neighbor Condition (NNC)*

$$\mathcal{R}_n = \{\mathbf{x} : d(\mathbf{x}, \hat{\mathbf{x}}_n) \leq d(\mathbf{x}, \hat{\mathbf{x}}_l), \forall l \neq n\}. \tag{11}$$

Next, finding the best decoder given the encoder involves determining the best codepoints given the regions. This is the *Centroid Condition (CC)*, given by

$$\hat{\mathbf{x}}_n = \text{Cent}\{\mathbf{x} | \mathbf{x} \in \mathcal{R}_n\}, \tag{12}$$

where the centroid Cent[2] of region R_n gives the codepoint $\hat{\mathbf{x}}_n$ for the region. The Lloyd algorithm alternates between the NNC and CC steps until convergence and yields the optimal codepoints.

B. The Lloyd Algorithm for AP Placement

If the VQ approach were to be used to solve for small-cell AP placement, then the random vector to be quantized is the 2-D position $\mathbf{p}$ of a *single* user. The Voronoi regions are the cells $\mathcal{C}_m$ and the codepoints are the AP locations $\mathbf{q}_m$, where $m = 1, 2, \ldots, M$. The optimization problem in (10) can be written by using similar notations and taking the average over the user positions, as follows

$$\arg\min_{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_M} \mathbb{E}_{\mathbf{p}} \left\{ d(\mathbf{p}, \mathbf{q}_{\mathcal{E}(\mathbf{p})}) \right\}. \tag{13}$$

[2]The centroid is defined [48] as

$$\text{Cent}\{\mathbf{x} | \mathbf{x} \in \mathcal{R}_n\} = \arg\min_{\hat{\mathbf{x}}_n} \mathbb{E}\{d(\mathbf{x}, \hat{\mathbf{x}}_n) | \mathbf{x} \in \mathcal{R}_n\}.$$

It is worth reiterating that $\mathcal{E}(\mathbf{p})$ indexes the nearest AP that the user at $\mathbf{p}$ associates to. The objective function in (13) can be written as

$$\begin{aligned} J_{\text{VQ}} &= \mathbb{E}_{\mathbf{p}}\left\{d(\mathbf{p},\mathbf{q}_{\mathcal{E}(\mathbf{p})})\right\}, \\ &= \int\limits_{\mathbf{p}\in\mathbb{R}^2} d(\mathbf{p},\mathbf{q}_{\mathcal{E}(\mathbf{p})}) f_{\mathbf{P}}(\mathbf{p})\mathrm{d}\mathbf{p}, \\ &= \sum_{m=1}^{M}\left[\int\limits_{\mathbf{p}\in\mathcal{C}_m} d(\mathbf{p},\mathbf{q}_m) f_{\mathbf{P}}(\mathbf{p}|\mathbf{p}\in\mathcal{C}_m)\mathrm{d}\mathbf{p}\right]\Pr(\mathbf{p}\in\mathcal{C}_m), \\ &= \sum_{m=1}^{M} S_m \Pr(\mathbf{p}\in\mathcal{C}_m), \end{aligned} \tag{14}$$

where the penultimate step arises by splitting the integral in the previous step into the cells (Voronoi regions) with their respective codepoints and the quantity S_m is defined as

$$S_m = \int\limits_{\mathbf{p}\in\mathcal{C}_m} d(\mathbf{p},\mathbf{q}_m) f_{\mathbf{P}}(\mathbf{p}|\mathbf{p}\in\mathcal{C}_m)\mathrm{d}\mathbf{p}. \tag{15}$$

To solve for the optimal AP locations, the most often used distortion function is the squared Euclidean distance

$$d_{\text{SE}}(\mathbf{p},\mathbf{q}_{\mathcal{E}(\mathbf{p})}) = \left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^2, \tag{16}$$

and the objective function in (14) then becomes the mean squared error (MSE). In this paper, we retain the name 'Lloyd algorithm' for the algorithm that solves (14) using d_{SE}, the steps of which are provided in Algorithm 1. For algorithms that use all other distortion functions, we will use the name 'Lloyd-type algorithm'. Note that when the Lloyd algorithm

Algorithm 1 Lloyd Algorithm With Squared Error Distortion

1: Initialize random AP locations $\mathbf{q}_1^{(0)},\mathbf{q}_2^{(0)},\ldots,\mathbf{q}_M^{(0)}$.
2: Use the NNC to determine the cells $\mathcal{C}_1^{(i+1)},\mathcal{C}_2^{(i+1)},\ldots,\mathcal{C}_M^{(i+1)}$ such that

$$\mathcal{C}_m^{(i+1)} = \left\{\mathbf{p}_k : d_{\text{SE}}\left(\mathbf{p}_k,\mathbf{q}_m^{(i)}\right) \le d_{\text{SE}}\left(\mathbf{p}_k,\mathbf{q}_l^{(i)}\right), \forall l\neq m\right\}.$$

3: Use the CC to determine the AP locations $\mathbf{q}_1^{(i+1)},\mathbf{q}_2^{(i+1)},\ldots,\mathbf{q}_M^{(i+1)}$ such that

$$\mathbf{q}_m^{(i+1)} = \frac{1}{\left|\mathcal{C}_m^{(i+1)}\right|}\sum_{\mathbf{p}_k\in\mathcal{C}_m^{(i+1)}}\mathbf{p}_k.$$

4: Repeat from step 2 until convergence (MSE falls below a threshold).

is implemented, we use the K realization of users at positions $\mathbf{p}_k$, $k=1,2,\ldots,K$, as described in Section II. We will use this notation for all the Lloyd-type algorithms that follow. Also, observe that in the CC step in 1, as a result of d_{SE}, the centroid is replaced by the expectation which is evaluated by using the sample average over the user positions $\mathbf{p}_k$ present in cell $\mathcal{C}_m$.

An interesting observation here is that the VQ framework presented above considers only the positions of the users and APs, and hence is independent of both the small-scale fading and shadow fading components of the wireless system since these quantities are not dependent on the user and AP positions. These random quantities thus do not play a role in AP placement using VQ. It is also very important to note here that VQ considers only a single user to be quantized and the average over the distribution of that user is taken. However, this does not conform to our small-cell system model, where M users are each communicating with its serving AP at the same time. Hence, the VQ approach does not strictly solve the small-cell AP placement problem.

IV. Throughput Formulations and Solutions Without Inter-Cell Interference

In this section, we will describe throughput optimization via various formulations, such as average rate and SNR, and provide solutions to obtain optimal AP locations (preliminarily discussed in [1]). We start by considering the single user scenario inherent to VQ and expand to a more realistic one in which multiple users are present. We also illustrate, by formulation only, the case where ICI is present. In summary, we argue how the Lloyd algorithm, despite its simplicity, is suitable for small-cell AP placement.

A. Single User Case

1) Rate: The single user case is the simplest case wherein a user at location $\mathbf{p}$ alone is considered. Recall that $\mathbf{p}$ is a random vector with pdf $f_{\mathbf{P}}(\mathbf{p})$. We start our analysis with per-user rate, which is the common measure of interest, achieved by a user at $\mathbf{p}$ with its nearest AP at $\mathbf{q}_{\mathcal{E}(\mathbf{p})}$, as per the VQ principles discussed above. We also approximate the large-scale fading coefficients, given in (2), by

$$\beta_{\mathcal{E}(\mathbf{p})} \approx \frac{c_1 z_{\mathcal{E}(\mathbf{p})}}{\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}}, \tag{17}$$

since r_0 is much smaller than the dimensions of the area under consideration. Note that the second subscript has been dropped for the ensuing analyses, since we consider a single user. Let us define the average rate, utilizing the per-user SNR $\psi_{k_{\mathcal{E}(\mathbf{p})}}$ (obtained from (5) by neglecting ICI and replacing m with $\mathcal{E}(\mathbf{p})$) as follows

$$\overline{r}(\underline{\mathbf{q}}) = \mathbb{E}_{\mathcal{A},\mathbf{p}}\left\{\log\left(1+\psi_{k_{\mathcal{E}(\mathbf{p})}}\right)\right\}, \tag{18}$$

where we average over the user position $\mathbf{p}$, the random quantities $h_{\mathcal{E}(\mathbf{p})}$ and $z_{\mathcal{E}(\mathbf{p})}$, $\mathcal{A}=\{h_{\mathcal{E}(\mathbf{p})},z_{\mathcal{E}(\mathbf{p})}\}$ for brevity, and we use the notation $\underline{\mathbf{q}}=\{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_M\}$ to show that the average rate is a function of the M AP locations alone. Similar to VQ in the previous section, we average out the small-scale and shadow fading components defined in $\mathcal{A}$ since they are position independent and do not contribute to the optimal placement of APs. Assuming high SNR ($\psi_{k_{\mathcal{E}(\mathbf{p})}} \gg 1$), we can write (18) as

$$\overline{r}(\underline{\mathbf{q}}) = \mathbb{E}_{\mathcal{A},\mathbf{p}}\left\{\log\left(\frac{\rho_r c_1 |h_{\mathcal{E}(\mathbf{p})}|^2 z_{\mathcal{E}(\mathbf{p})}}{\left(\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^2\right)^{\frac{\gamma}{2}}}\right)\right\}, \tag{19}$$

and we wish to perform the optimization

$$\underset{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_M}{\arg\max}\ \overline{r}(\underline{\mathbf{q}}). \tag{20}$$

After averaging and removing the terms that are not involved in the optimization in (19), we obtain

$$\underset{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_M}{\arg\min}\ \mathbb{E}_{\mathbf{p}}\left\{\log\left(\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^2+\epsilon\right)\right\}, \tag{21}$$

where we have added a constant $\epsilon > 0$ (typically very small) to prevent the logarithm from approaching negative infinity if the user position $\mathbf{p}$ were to overlap with the position of the nearest AP $\mathbf{q}_{\mathcal{E}(\mathbf{p})}$. Note that ϵ could correspond to the pathloss at a reference distance or even height of the AP. The objective function to be optimized above is concave as a result of which the Majorization-Minimization (MM) technique [49] can be used to acquire a solution to the centroid computation (CC) step. Although the distortion function in (21) is the logarithm of the squared Euclidean distance, the NNC step here remains the same as in the Lloyd algorithm since both $\log(\cdot)$ and ϵ can be ignored when comparing two distortion functions. The MM technique upperbounds the objective function by a surrogate function and minimizes the surrogate through an iterative method. Solving the objective function in (21) using the MM method results in an iterative solution with the following two update equations

$$\begin{aligned}\mathbf{q}_m^{(j+1)} &= \frac{\sum\limits_{\mathbf{p}_k\in\mathcal{C}_m} w_k^{(j)}\mathbf{p}_k}{\sum\limits_{\mathbf{p}_k\in\mathcal{C}_m} w_k^{(j)}},\\ w_k^{(j+1)} &= \frac{1}{||\mathbf{q}_m^{(j+1)}-\mathbf{p}_k||^2+\epsilon},\ \forall \mathbf{p}_k\in\mathcal{C}_m,\end{aligned} \tag{22}$$

where j denotes the MM iteration index.

In summary, to solve for the AP locations, we can now formulate a Lloyd-type algorithm with the NNC step remaining the same as that in the Lloyd algorithm, i.e., with d_{SE}, and the CC step replaced by the above iterative solution of (22). We call this Lloyd-type algorithm as the *MM-Lloyd algorithm*. The proof of (22) is left to Appendix A and the algorithm is provided in Algorithm 2.

Algorithm 2 MM-Lloyd Algorithm

1: Initialize random AP locations $\mathbf{q}_1^{(0)},\mathbf{q}_2^{(0)},\ldots,\mathbf{q}_M^{(0)}$.

2: Use the NNC to determine the cells $\mathcal{C}_1^{(i+1)},\mathcal{C}_2^{(i+1)},\ldots,\mathcal{C}_M^{(i+1)}$ such that

$$\mathcal{C}_m^{(i+1)}=\left\{\mathbf{p}_k: d_{\text{SE}}\left(\mathbf{p}_k,\mathbf{q}_m^{(i)}\right)\le d_{\text{SE}}\left(\mathbf{p}_k,\mathbf{q}_l^{(i)}\right),\forall l\neq m\right\}.$$

3: Use MM iterations to determine the AP locations $\mathbf{q}_1^{(i+1)},\mathbf{q}_2^{(i+1)},\ldots,\mathbf{q}_M^{(i+1)}$ with the update equations

$$\begin{aligned}\mathbf{q}_m^{(j+1)} &= \frac{\sum\limits_{\mathbf{p}_k\in\mathcal{C}_m^{(i+1)}} w_k^{(j)}\mathbf{p}_k}{\sum\limits_{\mathbf{p}_k\in\mathcal{C}_m^{(i+1)}} w_k^{(j)}},\\ w_k^{(j+1)} &= \frac{1}{||\mathbf{q}_m^{(j+1)}-\mathbf{p}_k||^2+\epsilon},\ \forall \mathbf{p}_k\in\mathcal{C}_m^{(i+1)},\end{aligned}$$

where $\mathbf{q}_m^{(i+1)}=\mathbf{q}_m^{(j+1)}$ after convergence.

4: Repeat from step 2 until convergence.

2) SNR: If throughput is measured solely by the SNR averaged over the user location $\mathbf{p}$, then we can show that the simple case of SNR maximization is equivalent to the VQ optimization problem given in (14). Let us write the average achievable SNR as

$$\overline{\psi}(\underline{\mathbf{q}})=\mathbb{E}_{\mathcal{A},\mathbf{p}}\left\{\frac{\rho_r c_1|h_{\mathcal{E}(\mathbf{p})}|^2 z_{\mathcal{E}(\mathbf{p})}}{\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}}\right\}, \tag{23}$$

which is lower bounded by applying Jensen's inequality as

$$\overline{\psi}(\underline{\mathbf{q}})\ge\mathbb{E}_{\mathcal{A}}\left\{\frac{\rho_r c_1|h_{\mathcal{E}(\mathbf{p})}|^2 z_{\mathcal{E}(\mathbf{p})}}{\left(\mathbb{E}_{\mathbf{p}}\left\{\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^2\right\}\right)^{\frac{\gamma}{2}}}\right\}, \tag{24}$$

with $\mathcal{A}$ defined as before. Maximizing $\overline{\psi}(\underline{\mathbf{q}})$ to obtain the AP locations is the same as minimizing the term in the denominator, leading to the same objective function (14) in VQ. The optimization problem is

$$\underset{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_M}{\arg\min}\ \mathbb{E}_{\mathbf{p}}\left\{\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^2\right\}. \tag{25}$$

As before, this is solved using the Lloyd algorithm with d_{SE} (Algorithm 1). For consistency in future discussions, we introduce the notation $d(\mathbf{p},\underline{\mathbf{q}})$ as a general form of distortion measure with $\underline{\mathbf{q}}=\{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_M\}$. Hence, the squared error distortion function in (16) is written in the general form as

$$d_{\text{SE}}\left(\mathbf{p},\underline{\mathbf{q}}\right)=\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^2. \tag{26}$$

3) Higher Exponent for User-AP Distance: The objective function in the Lloyd algorithm is proportional to the square of the user-AP distance while that in the MM-Lloyd algorithm is proportional to the logarithm of the squared distance. This means that the MM-Lloyd algorithm disproportionately considers the contribution of users, as the logarithm suppresses the larger distances inherent to users at the cell borders, in comparison to the Lloyd algorithm. To overcome this effect, we can design another optimization function that exponentially scales up large distances relative to the Lloyd algorithm by raising the distance to a higher power. The topic of optimal quantizer design for higher powers of distance has been studied in [31]. This higher exponent $\chi > 2$ also characterizes higher frequency (e.g., mmWave) communications. The optimization problem can then be represented as

$$\underset{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_M}{\arg\min}\ \mathbb{E}_{\mathbf{p}}\left\{\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\chi}\right\}, \tag{27}$$

where χ is the power. This optimization problem can be solved by using a Lloyd-type algorithm that uses the distortion function

$$d_{\chi}\left(\mathbf{p},\underline{\mathbf{q}}\right)=\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\chi}. \tag{28}$$

While the NNC step uses d_{χ}, the CC step utilizes the steepest descent method, with the update equation

$$\mathbf{q}_m^{(j+1)}=\mathbf{q}_m^{(j)}-\delta\frac{\partial}{\partial\mathbf{q}_m^{(j)}}\left\{\int\limits_{\mathbf{p}\in\mathcal{C}_m} d_{\chi}\left(\mathbf{p},\mathbf{q}_m^{(j)}\right) f_{\mathbf{P}}(\mathbf{p})\mathrm{d}\mathbf{p}\right\}, \tag{29}$$

for all m, where j is the iteration index, δ is the step size, and the gradient expression is given by

$$\frac{\partial}{\partial \mathbf{q}_m}\left\{\int\limits_{\mathbf{p}\in\mathcal{C}_m} d_\chi\left(\mathbf{p},\mathbf{q}_m\right) f_{\mathbf{P}}(\mathbf{p})\mathrm{d}\mathbf{p}\right\} = \frac{\chi}{|\mathcal{C}_m|}\sum_{\mathbf{p}_k\in\mathcal{C}_m}\left(\mathbf{q}_m-\mathbf{p}_k\right)||\mathbf{p}_k-\mathbf{q}_m||^{\chi-2}. \quad (30)$$

This Lloyd-type algorithm is called the *Lloyd-χ algorithm* and the proof of the above result for gradient can be found in Appendix B. The algorithm is given in Algorithm 3.

Algorithm 3 Lloyd-χ Algorithm

1: Initialize random AP locations $\mathbf{q}_1^{(0)},\mathbf{q}_2^{(0)},\ldots,\mathbf{q}_M^{(0)}$.
2: Use the NNC to determine the cells $\mathcal{C}_1^{(i+1)},\mathcal{C}_2^{(i+1)},\ldots,\mathcal{C}_M^{(i+1)}$ such that

$$\mathcal{C}_m^{(i+1)}=\left\{\mathbf{p}_k : d_\chi\left(\mathbf{p}_k,\mathbf{q}_m^{(i)}\right)\le d_\chi\left(\mathbf{p}_k,\mathbf{q}_l^{(i)}\right),\forall l\neq m\right\}.$$

3: Use the steepest descent method to determine the AP locations $\mathbf{q}_1^{(i+1)},\mathbf{q}_2^{(i+1)},\ldots,\mathbf{q}_M^{(i+1)}$ with the update equation

$$\mathbf{q}_m^{(j+1)}=\mathbf{q}_m^{(j)}-\frac{\delta\chi}{\left|\mathcal{C}_m^{(i+1)}\right|}\sum_{\mathbf{p}_k\in\mathcal{C}_m^{(i+1)}}\left(\mathbf{q}_m^{(j)}-\mathbf{p}_k\right)\left|\left|\mathbf{p}_k-\mathbf{q}_m^{(j)}\right|\right|^{\chi-2},$$

where $\mathbf{q}_m^{(i+1)}=\mathbf{q}_m^{(j+1)}$ after convergence.
4: Repeat from step 2 until convergence.

The above formulations were developed by assuming a single user located at $\mathbf{p}$. However, in practice and according to the system model, M APs serve M users at the same time. Hence, we now consider the case where M users are picked from the distribution.

B. Multiple User Case

1) Random User Selection: If M users are selected independently from the overall distribution $f_{\mathbf{P}}(\mathbf{p})$, then the distribution of these users are i.i.d. Let $\underline{\mathbf{p}}\triangleq\{\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_M\}$ be the set of locations of the M users. If we assume that the users do not interact with each other[3], then the objective function can be the sum of distortions incurred by each user with its closest AP, i.e., the optimization is of the form

$$\underset{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_M}{\arg\min}\ \mathbb{E}_{\underline{\mathbf{p}}}\left\{\sum_{m=1}^{M} d(\mathbf{p}_m,\mathbf{q}_{\mathcal{E}(\mathbf{p}_m)})\right\} = \underset{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_M}{\arg\min}\ M\cdot\mathbb{E}_{\mathbf{p}}\left\{d(\mathbf{p},\mathbf{q}_{\mathcal{E}(\mathbf{p})})\right\}, \quad (31)$$

where $\mathbf{q}_{\mathcal{E}(\mathbf{p})}$ defined as before and the simplification arises from the fact that each user is i.i.d. The final objective function thus is essentially the same as the single user case.

The above model is applicable in the following scenario. First, since the selection does not limit one user per small cell, the cells must be capable of dealing with more than one user with no multiple access interference. Secondly, since there is no ICI considered, each small cell must be assigned orthogonal resources. This leads to an interesting resource allocation problem which we do not pursue in this work.

2) Random Selection of One User Per Cell without ICI: The formulation described above considers M users at a time, but fails to follow the system model as each user is not necessarily picked from the Voronoi region or cell in which its serving AP is present. Under this model, assuming again that the users at $\underline{\mathbf{p}}=\{\mathbf{p}_1,\mathbf{p}_2,\ldots,\mathbf{p}_M\}$ do not interact with one another, the objective function to minimize would be the sum of the average distortion in each cell, i.e., the optimization is

$$\underset{\mathbf{q}_1,\mathbf{q}_2,\ldots,\mathbf{q}_M}{\arg\min}\ \mathbb{E}_{\underline{\mathbf{p}}}\left\{\sum_{m=1}^{M} d(\mathbf{p}_m,\mathbf{q}_m)\right\}, \quad (32)$$

with the joint distribution of the user positions as

$$f_{\underline{\mathbf{P}}}(\underline{\mathbf{p}})=\prod_{m=1}^{M} f_{\mathbf{P}_m}(\mathbf{p}_m|\mathbf{p}_m\in\mathcal{C}_m). \quad (33)$$

The above objective function can be simplified as

$$\sum_{m=1}^{M}\mathbb{E}_{\underline{\mathbf{p}}}\left\{d(\mathbf{p}_m,\mathbf{q}_m)\right\}=\sum_{m=1}^{M}\int\limits_{\mathbf{p}\in\mathcal{C}_m} d(\mathbf{p},\mathbf{q}_m) f_{\mathbf{P}}(\mathbf{p}|\mathbf{p}\in\mathcal{C}_m)\mathrm{d}\mathbf{p}, = \sum_{m=1}^{M} S_m, \quad (34)$$

where S_m is from (15). It is worth noting here the difference between this objective function and that of the Lloyd algorithm in (14) where each term S_m is weighted by the probability that the user is present in the cell $\Pr(\mathbf{p}\in\mathcal{C}_m)$. The solution to the above objective function is then a Lloyd-type algorithm with the CC step unchanged, but with the NNC step using weighted distortion functions, with the weights being the inverse of the proportion of users present in the cell. More specifically, the squared error distortion $d_{\mathrm{SE}}(\mathbf{p},\mathbf{q}_m)$ is pre-multiplied with a weight $w_m=1/\Pr(\mathbf{p}\in\mathcal{C}_m)=K/N_m$, where N_m is the number of users in $\mathcal{C}_m$. The NNC step is

$$\mathcal{C}_m=\left\{\mathbf{p} : w_m d_{\mathrm{SE}}\left(\mathbf{p},\mathbf{q}_m\right)\le w_l d_{\mathrm{SE}}\left(\mathbf{p},\mathbf{q}_l\right),\forall l\neq m\right\}. \quad (35)$$

We call this algorithm as the *weighted MSE (WMSE) Lloyd algorithm*. Note that weighted distortion functions have been studied in [30], although the authors have considered weights that remain constant. The weights in the WMSE Lloyd algorithm, on the other hand, are learnt in every iteration. The proof of the above solution is provided in Appendix C and the algorithm is outlined in Algorithm 4.

3) Random Selection of One User Per Cell with ICI: In all the above formulations, we have considered only SNR and the fact that users do not interact with one another. However, under the effects of ICI, users do interact with one another in the form of providing interfering signals at the APs which are serving the other users. Thus, the distortion function between a user and its serving AP would be a function of all other

[3]This implies that a user does not influence the AP selection of any other user. In other words, the distortion function between a user at $\mathbf{p}_m$ and its closest AP $\mathbf{q}_{\mathcal{E}(\mathbf{p}_m)}$ is independent of the positions of all other users $\mathbf{p}_{m'}$, where $m'\neq m$. It is worth noting that multiple users can select the same AP as its closest one.

Algorithm 4 WMSE Lloyd Algorithm

1: Initialize random AP locations $\mathbf{q}_1^{(0)}, \mathbf{q}_2^{(0)}, \ldots, \mathbf{q}_M^{(0)}$.
2: Use the NNC to determine the cells $\mathcal{C}_1^{(i+1)}, \mathcal{C}_2^{(i+1)}, \ldots, \mathcal{C}_M^{(i+1)}$ such that

$$\mathcal{C}_m^{(i+1)} = \left\{ \mathbf{p}_k : w_m d_{\text{SE}}\left(\mathbf{p}_k, \mathbf{q}_m^{(i)}\right) \leq w_j d_{\text{SE}}\left(\mathbf{p}_k, \mathbf{q}_l^{(i)}\right), \forall l \neq m \right\}.$$

3: Use the CC to determine the AP locations $\mathbf{q}_1^{(i+1)}, \mathbf{q}_2^{(i+1)}, \ldots, \mathbf{q}_M^{(i+1)}$ such that

$$\mathbf{q}_m^{(i+1)} = \frac{1}{\left|\mathcal{C}_m^{(i+1)}\right|} \sum_{\mathbf{p}_k \in \mathcal{C}_m^{(i+1)}} \mathbf{p}_k.$$

4: Repeat from step 2 until convergence.

users as well and under a similar fashion as in (32), we can write the objective function as

$$\sum_{m=1}^{M} \mathbb{E}_{\underline{\mathbf{P}}} \left\{ d(\mathbf{p}_m, \mathbf{q}_m, \underline{\mathbf{p}}'_m) \right\} = \sum_{m=1}^{M} \int_{\mathbf{p}_1 \in \mathcal{C}_1} \cdots \int_{\mathbf{p}_M \in \mathcal{C}_M} d(\mathbf{p}_m, \mathbf{q}_m, \underline{\mathbf{p}}'_m) f_{\underline{\mathbf{P}}}(\underline{\mathbf{p}}) \mathrm{d}\underline{\mathbf{p}}, \quad (36)$$

where the (general) distortion function uses the term $\underline{\mathbf{p}}'_m$ which denotes the set of user positions other than the user at $\mathbf{p}_m$ and $f_{\underline{\mathbf{P}}}(\underline{\mathbf{p}})$ is as in (33). It is clear from the objective function in (36), due to the dependency of the distortion function on the interfering users, the joint distribution $f_{\underline{\mathbf{P}}}(\underline{\mathbf{p}})$ cannot be simplified to consider each cell $\mathcal{C}_m$ independently as in (34). This makes the said objective function difficult and intractable, and hence cannot be readily solved.

To deal with ICI in a tractable manner, we adopt a slightly different approach based on the following considerations. Based on results obtained so far, VQ provides a good framework to solve throughput optimization problems by Lloyd-type algorithms, although without ICI. We have also seen that the optimization in (36), which considers ICI, is difficult to solve and to derive an AP placement algorithm. Further, numerical simulations (shown in Experiment 1 of Section VII-C) show that the average achievable rate is very similar, whether the Lloyd or Lloyd-type algorithms described in this section are used. Motivated by these three facts, in the next section, we show how the Lloyd algorithm can be modified to account for ICI in AP placement. It is worth noting here that we could implement power control along with AP placement, i.e., optimizing uplink power with per-user power constraints jointly with the AP locations, in order to increment rate. Our focus in this work, however, is to solely investigate the appropriateness of the VQ approach to small-cell AP placement and the necessary modifications to add ICI to the VQ optimization framework such that a Lloyd-type algorithm can be used to solve the problem, both of which have not been performed in past literature. Thus, we will continue to assume the same uplink power for all users and address power control in future work.

V. Throughput Formulations Accounting for Inter-Cell Interference

To account for ICI in the VQ framework, we develop two distortion functions, namely the interference and inter-AP distortion functions.

A. Interference Distortion Measure

From (25), it is clear that the Lloyd algorithm maximizes only the desired signal component. In addition, we are now required to minimize the interference term. To construct a distortion function that considers both the desired and interference signals, we consider the achievable per-user rate, as considered in Section IV-A1, but using the SINR expression from (5). Formally, the rate maximization problem is

$$\underset{\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_M}{\arg\max} \; \mathbb{E}_{\mathcal{A}, \mathcal{B}, \underline{\mathbf{P}}} \left\{ \log\left(1 + \phi_{k_{\mathcal{E}(\mathbf{p})}}\right) \right\}, \quad (37)$$

where set $\mathcal{A} = \{h_{\mathcal{E}(\mathbf{p})}, z_{\mathcal{E}(\mathbf{p})}\}$ defined as before and set $\mathcal{B} = \{h_{m'}, z_{m'} : m' \neq \mathcal{E}(\mathbf{p})\}$ consists of the small-scale and shadow fading quantities for all interfering cells $\mathcal{C}_{m'}$, $m' \neq \mathcal{E}(\mathbf{p})$. For notational simplicity, the SINR $\phi_{k_{\mathcal{E}(\mathbf{p})}}$ above can be rewritten using T_{SNR} for the desired signal power in the numerator and T_{ICI} for the interference signal power in the denominator as follows

$$\phi_{k_{\mathcal{E}(\mathbf{p})}} = \frac{T_{\text{SNR}}}{1 + T_{\text{ICI}}}, \quad (38)$$

where $T_{\text{SNR}} = \rho_r \beta_{\mathcal{E}(\mathbf{p})} |h_{\mathcal{E}(\mathbf{p})}|^2$ and $T_{\text{ICI}} = \rho_r \sum_{m' \neq m} \beta_{m'} |h_{m'}|^2$. To recapitulate the notation, we use a single subscript for simplicity and while $h_{\mathcal{E}(\mathbf{p})}$ and $\beta_{\mathcal{E}(\mathbf{p})}$ are the small-scale and large-scale fading coefficients, respectively, for the user at $\mathbf{p}$ to the serving cell, $h_{m'}$ and $\beta_{m'}$ correspond to the same quantities for the same user, but to the non-serving AP m'. Approximating the rate with high SINR ($\phi_{k_{\mathcal{E}(\mathbf{p})}} \gg 1$) and $T_{\text{ICI}} \gg 1$, and simplifying, we get

$$\log \phi_{k_{\mathcal{E}(\mathbf{p})}} \approx \log T_{\text{SNR}} + \log \frac{1}{T_{\text{ICI}}}. \quad (39)$$

It is worth nothing here that the log-sum inequality could be applied to separate the second term above as the sum of inverses of the individual ICI terms. Further, considering the above sum of logarithm terms, it is clear that the MM technique can be applied. However, finding a surrogate function in this case is not as straightforward as in the solution to the MM-Lloyd algorithm discussed in Section IV-A1. We believe that the insight obtained from (39) is sufficient to generate a solution for AP placement. To simplify further, we negate the quantity in (39) and approximate using the relation $\log x < x$ which yields

$$-\log \phi_{k_{\mathcal{E}(\mathbf{p})}} < \frac{1}{T_{\text{SNR}}} + T_{\text{ICI}}. \quad (40)$$

We have now expressed the negative rate as the sum of the powers of the inverse of the desired and interference terms. Therefore, to maximize rate or equivalently, minimize the negative of the rate, we need to maximize SNR and minimize ICI, corresponding to the first and second terms in (40), respectively. The equation presented also reveals the structure of the distortion function that we will use.

TABLE I: Summary of Throughput Formulations and Solutions

Number of Users	Formulation	Solution Algorithm
Single user	Rate	MM-Lloyd
	SNR	Lloyd
	Higher exponent for user-AP distance	Lloyd-χ
Multiple users	Random user selection	Lloyd
	Random selection of one user per cell without ICI	WMSE Lloyd
	Random selection of one user per cell with ICI	Interference and Inter-AP Lloyd

Accordingly, in line with the objective function for a Lloyd-type algorithm in (14), we average (40) over the user positions and the random quantities defined in (37) above, to obtain

$$\mathbb{E}_{\mathcal{A},\mathcal{B},\underline{\mathbf{p}}}\left\{\frac{1}{T_{\text{SNR}}}+T_{\text{ICI}}\right\}=\mathbb{E}_{\mathbf{p}}\left\{\mathbb{E}_{\mathcal{A},\mathcal{B},\underline{\mathbf{p}}'}\left\{\frac{1}{T_{\text{SNR}}}+T_{\text{ICI}}\right\}\right\}, \tag{41}$$

where we denote $\underline{\mathbf{p}}'$ as the set of positions of the interfering users with respect to the user at $\mathbf{p}$. We have assumed in (41) that the served user position $\mathbf{p}$ is independent from the interfering user positions $\underline{\mathbf{p}}'$. We will also assume that as in (33) that the distribution of users in each interfering cell is independent and we can then write the joint distribution of users as

$$\begin{aligned} f_{\underline{\mathbf{P}}}(\underline{\mathbf{p}}) &= f_{\mathbf{P},\underline{\mathbf{P}}'}(\mathbf{p},\underline{\mathbf{p}}')=f_{\mathbf{P}}(\mathbf{p})f_{\underline{\mathbf{P}}'}(\underline{\mathbf{p}}'),\\ &= f_{\mathbf{P}}(\mathbf{p})\prod_{m'\neq\mathcal{E}(\mathbf{p})} f_{\mathbf{P}_{m'}}(\mathbf{p}_{m'}|\mathbf{p}_{m'}\in\mathcal{C}_{m'}), \end{aligned} \tag{42}$$

where $f_{\underline{\mathbf{P}}'}(\underline{\mathbf{p}}')$ is the joint distribution of the locations of all the interfering users and $f_{\mathbf{P}_{m'}}(\mathbf{p}_{m'}|\mathbf{p}_{m'}\in\mathcal{C}_{m'})$ is the distribution of the user in cell $\mathcal{C}_{m'}$. Consequently, by carrying out the expectations in (41) over $\mathcal{A}$, $\mathcal{B}$, and $\underline{\mathbf{p}}'$, we can write the distortion function as

$$\begin{aligned} d_{\text{IF}}\left(\mathbf{p},\underline{\mathbf{q}}\right) &= \kappa_1\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}\\ &+\kappa_2\sum_{m'\neq\mathcal{E}(\mathbf{p})}\int\cdots\int\frac{1}{\left|\left|\mathbf{p}_{m'}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}}f_{\underline{\mathbf{P}}'}(\underline{\mathbf{p}}')\mathrm{d}\underline{\mathbf{p}}', \end{aligned} \tag{43}$$

where $\kappa_1 = \mathbb{E}_{\mathcal{A}}\{1/\rho_r c_1 z_{\mathcal{E}(\mathbf{p})}|h_{\mathcal{E}(\mathbf{p})}|^2\}$, $\kappa_2 = \mathbb{E}_{\mathcal{B}}\{\rho_r c_1 z_{m'}|h_{m'}|^2\}$, and the integration limits have been omitted for notation simplicity. This is the *interference distortion function* denoted by d_{IF} and the corresponding Lloyd-type algorithm is called the *Interference Lloyd algorithm*. Further simplification using (42) leads to a simpler distortion measure

$$\begin{aligned} d_{\text{IF}}\left(\mathbf{p},\underline{\mathbf{q}}\right) &= \left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}\\ +\kappa\sum_{m'\neq\mathcal{E}(\mathbf{p})}\int_{\mathbf{p}_{m'}\in\mathcal{C}_{m'}}&\frac{1}{\left|\left|\mathbf{p}_{m'}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}}f_{\mathbf{P}_{m'}}(\mathbf{p}_{m'}|\mathbf{p}_{m'}\in\mathcal{C}_{m'})\mathrm{d}\mathbf{p}_{m'}, \end{aligned} \tag{44}$$

where $\kappa \triangleq \kappa_2/\kappa_1$. We call $\kappa \geq 0$ as the *trade-off factor* and it determines the trade-off between desired signal and ICI power. κ can be varied to determine the importance of ICI power over desired signal power.

To solve for the AP locations, the Interference Lloyd algorithm retains the NNC step and the steepest descent method is to be used for the CC step (update equation given in (29) above), both steps utilizing d_{IF}. For the sake of implementation, the integral in d_{IF} from (44) is numerically approximated using the sample average over a large number of realizations of the user locations, and is written as

$$\begin{aligned} d_{\text{IF}}\left(\mathbf{p},\mathbf{q}_m\right) &= ||\mathbf{p}-\mathbf{q}_m||^{\gamma}\\ &+\kappa\sum_{m'\neq m}\frac{1}{|\mathcal{C}_{m'}|}\sum_{\mathbf{p}_{k_{m'}}\in\mathcal{C}_{m'}}\frac{1}{\left|\left|\mathbf{p}_{k_{m'}}-\mathbf{q}_m\right|\right|^{\gamma}}, \end{aligned} \tag{45}$$

where $\mathbf{p}_{k_{m'}}$ represents the k^{th} realization of the user position in cell $\mathcal{C}_{m'}$. The gradient function in this update equation is

$$\begin{aligned} &\frac{\partial}{\partial\mathbf{q}_m}\left\{\int_{\mathbf{p}\in\mathcal{C}_m} d_{\text{IF}}\left(\mathbf{p},\mathbf{q}_m\right)f_{\mathbf{P}}(\mathbf{p})\mathrm{d}\mathbf{p}\right\}\\ &\quad=\frac{\gamma}{|\mathcal{C}_m|}\sum_{\mathbf{p}_k\in\mathcal{C}_m}\left(\mathbf{q}_m-\mathbf{p}_k\right)||\mathbf{p}_k-\mathbf{q}_m||^{\gamma-2}\\ &\quad+\kappa\sum_{m'\neq m}\frac{\gamma}{|\mathcal{C}_{m'}|}\sum_{\mathbf{p}_{k_{m'}}\in\mathcal{C}_{m'}}\frac{\left(\mathbf{p}_{k_{m'}}-\mathbf{q}_m\right)}{\left|\left|\mathbf{p}_{k_{m'}}-\mathbf{q}_m\right|\right|^{\gamma}}. \end{aligned} \tag{46}$$

The proof of this result is given in Appendix E and the steps for this Lloyd-type algorithm are provided in Algorithm 5.

Algorithm 5 Interference Lloyd Algorithm

1: Initialize random AP locations $\mathbf{q}_1^{(0)},\mathbf{q}_2^{(0)},\ldots,\mathbf{q}_M^{(0)}$.
2: Use the NNC to determine the cells $\mathcal{C}_1^{(i+1)},\mathcal{C}_2^{(i+1)},\ldots,\mathcal{C}_M^{(i+1)}$ such that

$$\mathcal{C}_m^{(i+1)}=\left\{\mathbf{p}_k: d_{\text{IF}}\left(\mathbf{p}_k,\mathbf{q}_m^{(i)}\right)\leq d_{\text{IF}}\left(\mathbf{p}_k,\mathbf{q}_l^{(i)}\right),\forall l\neq m\right\}.$$

3: Use the steepest descent method to determine the AP locations $\mathbf{q}_1^{(i+1)},\mathbf{q}_2^{(i+1)},\ldots,\mathbf{q}_M^{(i+1)}$ with the update equation

$$\begin{aligned} \mathbf{q}_m^{(j+1)} &= \mathbf{q}_m^{(j)}\\ &-\delta\left(\frac{\gamma}{\left|\mathcal{C}_m^{(i+1)}\right|}\sum_{\mathbf{p}_k\in\mathcal{C}_m^{(i+1)}}\left(\mathbf{q}_m^{(j)}-\mathbf{p}_k\right)\left|\left|\mathbf{p}_k-\mathbf{q}_m^{(j)}\right|\right|^{\gamma-2}\right.\\ &\left.+\kappa\sum_{m'\neq m}\frac{\gamma}{\left|\mathcal{C}_{m'}^{(i+1)}\right|}\sum_{\mathbf{p}_{k_{m'}}\in\mathcal{C}_{m'}^{(i+1)}}\frac{\left(\mathbf{p}_{k_{m'}}-\mathbf{q}_m^{(j)}\right)}{\left|\left|\mathbf{p}_{k_{m'}}-\mathbf{q}_m^{(j)}\right|\right|^{\gamma}}\right), \end{aligned}$$

which, after convergence, $\mathbf{q}_m^{(i+1)}=\mathbf{q}_m^{(j+1)}$.
4: Repeat from step 2 until convergence.

B. Inter-AP Distortion Measure

Here, we develop an alternate distortion function that also accounts for ICI. Consider the interference distortion function

d_{IF} in (44). Each of the ICI terms in the summation in d_{IF} can be approximated as follows

$$\mathbb{E}_{\mathbf{p}_{m'}}\left\{\frac{1}{\left|\left|\mathbf{p}_{m'}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}}\right\}\approx\frac{1}{\left|\left|\mathbf{q}_{m'}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}}, \tag{47}$$

the justification of which is provided in Appendix D. Substituting (47), we can simplify (44) as

$$d_{\text{IA}}\left(\mathbf{p},\underline{\mathbf{q}}\right)=\left|\left|\mathbf{p}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}+\kappa\sum_{m'\neq\mathcal{E}(\mathbf{p})}\frac{1}{\left|\left|\mathbf{q}_{m'}-\mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}}. \tag{48}$$

We call d_{IA} the *inter-AP distortion measure* as the ICI term now involves the distances between the interfering APs and AP indexed by $\mathcal{E}(\mathbf{p})$. The corresponding Lloyd-type algorithm is called the *Inter-AP Lloyd algorithm*.

The solution of the optimization problem using d_{IA} is similar to that of the Interference Lloyd algorithm. For the steepest descent method, the gradient corresponding to d_{IA} is given as

$$\begin{aligned}\frac{\partial}{\partial\mathbf{q}_m}&\left\{\int_{\mathbf{p}\in\mathcal{C}_m}d_{\text{IA}}\left(\mathbf{p},\mathbf{q}_m\right)f_{\mathbf{P}}(\mathbf{p})\mathrm{d}\mathbf{p}\right\}\\&=\frac{\gamma}{|\mathcal{C}_m|}\sum_{\mathbf{p}_k\in\mathcal{C}_m}\left(\mathbf{q}_m-\mathbf{p}_k\right)\left|\left|\mathbf{p}_k-\mathbf{q}_m\right|\right|^{\gamma-2}\\&\quad+\kappa\gamma\sum_{m'\neq m}\frac{\mathbf{q}_{m'}-\mathbf{q}_m}{\left|\left|\mathbf{q}_{m'}-\mathbf{q}_m\right|\right|^{\gamma+2}}.\end{aligned} \tag{49}$$

The proof of this result is omitted as it is similar in calculation to the gradient of the interference distortion function in (46) and the Inter-AP Lloyd algorithm is given in Algorithm 6. We also summarize the various formulations and solutions discussed in this section and the prior section in Table I.

Algorithm 6 Inter-AP Lloyd Algorithm

1: Initialize random AP locations $\mathbf{q}_1^{(0)},\mathbf{q}_2^{(0)},\ldots,\mathbf{q}_M^{(0)}$.

2: Use the NNC to determine the cells $\mathcal{C}_1^{(i+1)},\mathcal{C}_2^{(i+1)},\ldots,\mathcal{C}_M^{(i+1)}$ such that

$$\mathcal{C}_m^{(i+1)}=\left\{\mathbf{p}_k:d_{\text{IA}}\left(\mathbf{p}_k,\mathbf{q}_m^{(i)}\right)\leq d_{\text{IA}}\left(\mathbf{p}_k,\mathbf{q}_l^{(i)}\right),\forall l\neq m\right\}.$$

3: Use the steepest descent method to determine the AP locations $\mathbf{q}_1^{(i+1)},\mathbf{q}_2^{(i+1)},\ldots,\mathbf{q}_M^{(i+1)}$ with the update equation

$$\begin{aligned}\mathbf{q}_m^{(j+1)}&=\mathbf{q}_m^{(j)}\\&-\delta\left(\frac{\gamma}{\left|\mathcal{C}_m^{(i+1)}\right|}\sum_{\mathbf{p}_k\in\mathcal{C}_m^{(i+1)}}\left(\mathbf{q}_m^{(j)}-\mathbf{p}_k\right)\left|\left|\mathbf{p}_k-\mathbf{q}_m^{(j)}\right|\right|^{\gamma-2}\right.\\&\left.+\kappa\sum_{m'\neq m}\frac{\mathbf{q}_{m'}^{(j)}-\mathbf{q}_m^{(j)}}{\left|\left|\mathbf{q}_{m'}^{(j)}-\mathbf{q}_m^{(j)}\right|\right|^{\gamma+2}}\right),\end{aligned}$$

which, after convergence, $\mathbf{q}_m^{(i+1)}=\mathbf{q}_m^{(j+1)}$.

4: Repeat from step 2 until convergence.

Among the distortion functions discussed above, it is evident that the MSE distortion d_{SE} has the lowest complexity. On observing the expressions for the interference d_{IF} and inter-AP d_{IA} distortions, we find that in the former, the summation for each interfering cell is over all of the users in that cell while in the latter, the net summation is only over interfering cells. Hence, d_{IA} has lower implementation complexity than d_{IF}. We will also see in a later section that user association with d_{IA} is relatively much simpler.

VI. Cell Association Strategies

In the previous sections, we have addressed the problem of how to place APs based on the user locations. For completeness, we now aim at answering the following two questions on cell association: *When a new user enters the system, to which cell should it associate to? What metric should be used?* In this section, we elaborate on these two issues in the context of Lloyd and Lloyd-type algorithms. Accordingly, consider a user at location $\mathbf{p}_{\text{new}}$ that has entered the area after AP placement has already occurred and will associate to the AP at $\mathbf{q}_{m_{\text{new}}}$.

For the Lloyd and Lloyd-type algorithms developed in this paper, the user would associate to the AP that yields the lowest distortion value. This is a straightforward implementation of the NNC for each algorithm. Formally, if d represents any of the distortion functions, $\mathbf{q}_{m_{\text{new}}}$ is determined as

$$\mathbf{q}_{m_{\text{new}}}=\left\{\mathbf{q}_m:d\left(\mathbf{p}_{\text{new}},\mathbf{q}_m\right)\leq d\left(\mathbf{p}_{\text{new}},\mathbf{q}_l\right),\forall l\neq m\right\}. \tag{50}$$

It is worth pointing out that since the distortion function in the Interference Lloyd algorithm involves summing over all users in other (interfering) cells, the complexity of such a calculation cannot be overlooked. Instead, a cell association procedure (50) based on the simpler distortion measures of the Lloyd or the Inter-AP Lloyd algorithm can be undertaken as a low-complexity alternative. Note that the distortion function in the latter involves only the knowledge of the interfering APs positions. This is of greater practical value as opposed to knowing the positions of all interfering users in the Interference Lloyd algorithm. In summary, the Inter-AP Lloyd algorithm not only offers lower implementation complexity and thus a simpler cell association strategy, but is also of more practical value compared to the Interference Lloyd algorithm.

VII. Simulation Methodology and Results

A. Simulation Parameters

A geographical area of dimensions 2 km $\times$ 2 km is considered, consisting of $M=8$ APs and $K=2000$ users, and one randomly selected user in each cell communicates with its associated AP. The pathloss model in (2) is used with $\gamma=2$, shadow fading z_{mk} ignored as it is averaged out in Sections IV and V, $c_0=75.86$ and $c_1=7.59\times10^{-7}$ as in [45, eq. (4.36), eq. (4.37)] according to the COST 231 Hata propagation model, and $r_0=0.001$ km. Also, the value of the trade-off factor is chosen to be $\kappa=5\times10^8$ and the step-size for the gradient descent is $\delta=5\times10^{-5}$ for the Lloyd-χ algorithm and $\delta=0.5$ for the ICI-aware Lloyd-type algorithms. Moreover, the uplink transmit power is $\rho_r=200$

mW and the user distribution is a Gaussian mixture model (GMM) of the form

$$f_{\mathbf{P}}(\mathbf{p}) = \sum_{l=1}^{L} p_l \mathcal{N}\left(\mathbf{p}|\boldsymbol{\mu}_l, \sigma_l^2 \mathbf{I}\right), \tag{51}$$

where $\mathbf{I}$ is the identity matrix and L is the number of mixture components, called *groups* henceforth. For group l, p_l is the mixture component weight, $\boldsymbol{\mu}_l$ is the mean, and σ_l is the standard deviation. We set a user configuration with the parameters $L = 3$, $\boldsymbol{\mu}_1 = [0.5, -0.5]^T$, $\boldsymbol{\mu}_2 = [0, 0.5]^T$, $\boldsymbol{\mu}_3 = [-0.5, 0]^T$, $\sigma_1 = \sigma_2 = \sigma_3 = 100$, $p_1 = 0.6$, and $p_2 = p_3 = 0.2$.

B. Performance Measures

We use the per-user achievable rate of user k_m, which is calculated using SINR ϕ_{k_m} from (5). As given in [45, Ch. 4], we can also write the achievable rate as

$$R_{k_m} = \mathbb{E}\left\{\log_2\left(1 + \phi_{k_m}\right)\right\} = \frac{1}{\ln 2} e^{\mu_k} \mathrm{Ei}\left(\mu_k\right), \tag{52}$$

where $\mu_k = (1 + \rho_r \sum_{m' \neq m} \beta_{mk_{m'}})/\rho_r \beta_{mk_m}$ and $\mathrm{Ei}(x) = \int_x^\infty (e^{-t}/t)\mathrm{d}t$ is the exponential integral.

For each of the proposed algorithms and the benchmark Lloyd algorithm, the maximum iteration number is set at 50. Each of the above performance measures is calculated through Monte Carlo simulations with $10,000$ iterations, choosing a set of users randomly for transmission each time. Cumulative distribution function (CDF) plots are generated for each measure, though normalized by the largest value so as to focus on the relative performance of the considered algorithms. For comparison, we utilize the 95%-likely metric that represents the best rate of the worst 5% of the users (users closer to cell borders). We denote this by $R_{k_m}^{5\%}$. To quantify the improvement in relative performance of the proposed algorithms over the Lloyd algorithm, we use the following measure expressed as percentage

$$\text{Improvement Ratio} = \frac{R_{k_m}^{5\%,\text{Proposed}} - R_{k_m}^{5\%,\text{Lloyd}}}{R_{k_m}^{5\%,\text{Lloyd}}} \times 100. \tag{53}$$

All algorithms are initialized with the same initial AP locations for unbiased comparison.

C. Numerical Results

Experiment 1. We compare the throughput performances of the proposed Lloyd-type algorithms in Section IV with the baseline Lloyd algorithm. For the Lloyd-χ algorithm, we use $\chi = 4$ and we note that the rate calculations still use the exponent $\gamma = 2$. The AP locations resultant from the algorithms are shown in Fig. 1. Relative to the AP positions of the Lloyd algorithm which are shown as blue circles, the APs in both the MM-Lloyd and WMSE Lloyd algorithms are placed closer to the GMM centers. For the MM-Lloyd algorithm, this can be explained by the logarithm in its objective function which suppresses the effect of users which are at large distances (e.g., cell periphery users away from the GMM center) from the AP position during the placement process. This, in turn, causes the APs to position themselves closer to the GMM centers where the majority of the users at smaller distances are present. The WMSE Lloyd algorithm works in a different manner as the objective function in (34) is not weighted by the cell probabilities $\Pr(\mathbf{p} \in \mathcal{C}_m)$ as in (14) of the Lloyd algorithm. This allows cells in the WMSE Lloyd algorithm to have a larger number of users than the Lloyd algorithm. On the other hand, the objective function of the Lloyd-χ algorithm amplifies the contribution of the users at large distances and results in the APs moving away from the GMM center. The effects of these placements are observed in their achievable rate plots in Fig. 2. For both the MM-Lloyd and WMSE Lloyd algorithms, we observe that due to their AP positions, the lower rate suffers a reduction in comparison to the Lloyd algorithm. Nevertheless, note that there are more users achieving higher rates (right side of the CDF plot), particularly for the MM-Lloyd algorithm. The average rate values, however, are higher than that of the Lloyd algorithm, up to about 4%, as shown in Table II. On the other hand, the opposite of these effects are observed for the Lloyd-χ algorithm, with higher low rate values and lower average rate (only 0.65% lower) than the Lloyd algorithm. Although omitted here due to space constraints, these effects increase as the power χ increases.

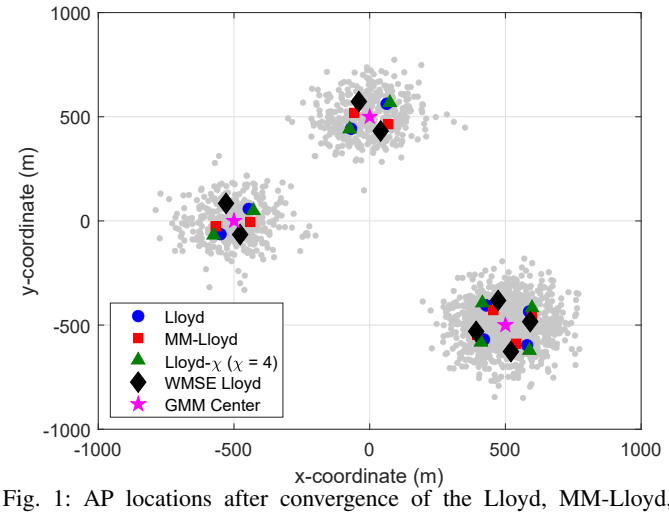

Fig. 1: AP locations after convergence of the Lloyd, MM-Lloyd, Lloyd-χ ($\chi = 4$), and WMSE Lloyd algorithms with $M = 8$.

TABLE II: Percentage Improvements in Average Achievable Rates for the Lloyd-Type Algorithms of Section IV

Algorithm	Average Achievable Rate
MM-Lloyd	4.26%
Lloyd-χ ($\chi = 4$)	−0.64%
WMSE Lloyd	1.14%

Experiment 2. Here, our simulations show throughput performances for the proposed ICI-aware Lloyd-type algorithms and the Lloyd algorithm, as well as their respective AP placements for comparison. The AP locations obtained after the algorithms converge are shown in Fig. 3. AP locations for the Lloyd algorithm are shown as circles around the GMM center, which in turn are shown by stars. Compared to these positions, we can observe that the AP locations for both

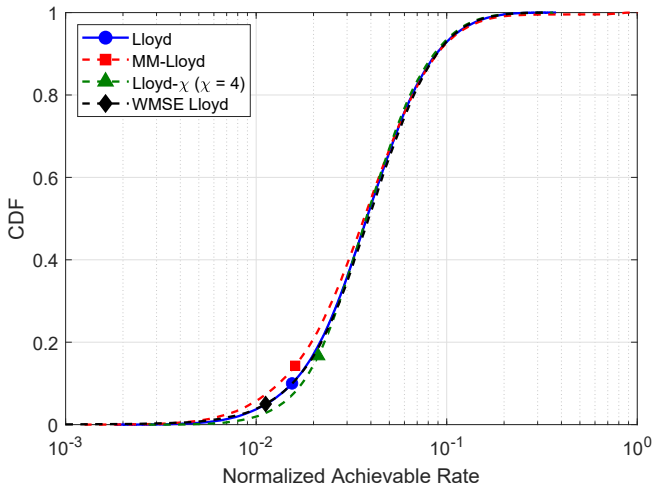

Fig. 2: CDF plots of per-user achievable rate for the Lloyd, MM-Lloyd, Lloyd-χ ($\chi = 4$), and WMSE Lloyd algorithms with $M = 8$.

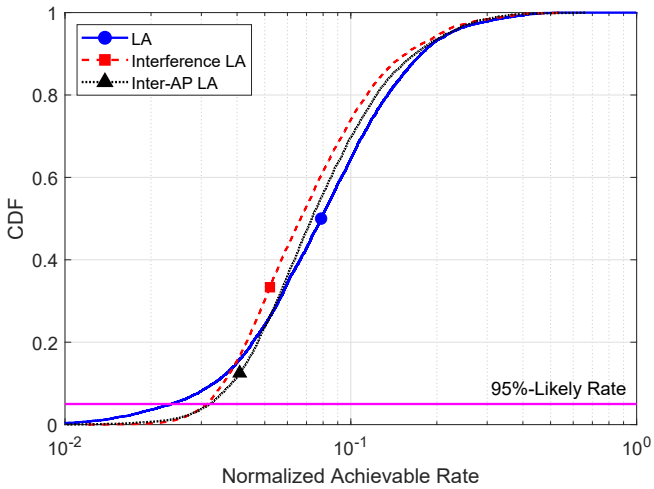

Fig. 4: CDF plots of per-user achievable rate for the Lloyd and ICI-aware Lloyd-type algorithms with $\kappa = 5 \times 10^8$ and $M = 8$.

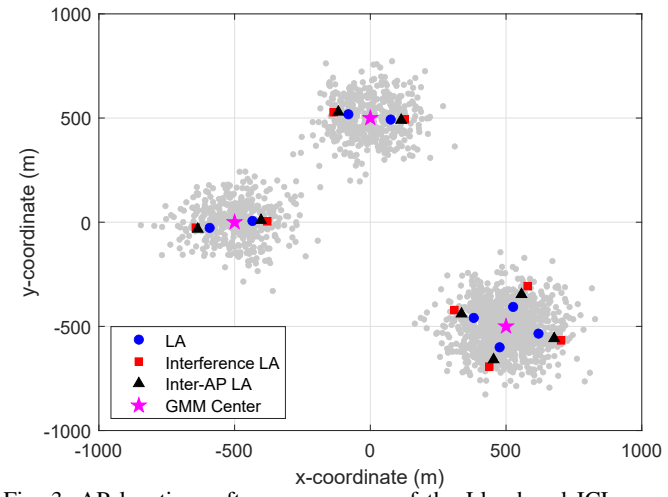

Fig. 3: AP locations after convergence of the Lloyd and ICI-aware Lloyd-type algorithms with $\kappa = 5 \times 10^8$ and $M = 8$.

TABLE III: Percentage Improvements in Average and 95%-Likely Achievable Rates for the ICI-Aware Lloyd-Type Algorithms

Algorithm	Average Rate	95%-Likely Rate
Interference Lloyd	-10.94%	33.37%
Inter-AP Lloyd	-4.35%	36.34%

the Lloyd-type algorithms are situated further away from the GMM centers. For the Interference Lloyd algorithm, the AP positions denoted by the squares are the farthest. This is due to the interference term in its distortion function that forces neighboring cells apart. This effect is different (smaller) for the Inter-AP Lloyd algorithm due to the inter-AP distances term in its distortion function in contrast to the interference term in the Interference Lloyd algorithm.

In Fig. 4, we show the CDFs of the achievable rate obtained per user for each of the considered algorithms. The horizontal line at the 5$^{\text{th}}$ percentile shows the 95%-likely rate and we compare the values where it intersects the throughput curves. It is clear that accounting for ICI during the AP placement procedure yields a superior performance to both Lloyd-type algorithms in comparison to the Lloyd algorithm in terms of the 95%-likely rate. It is to be noted that the average rate performances of both the proposed algorithms are lower than that of the Lloyd algorithm, however, the magnitude of increase in the 95%-likely rates overshadows the decrease in average rates. This occurrence is due to the fact that the original average rate maximization problem in (37) has been transformed into maximization of its lower bound in (40). In practice, Fig. 4 shows us that the worst 5% of the users, usually the ones located closer to the cell borders and thus more susceptible to the deleterious effects of ICI, will have an uplink performance boost when APs are placed according to the proposed algorithms. The percentage of improvements are quantified in Table III from where we can confirm a very significant rate enhancement of up to 36.34% in the 95%-likely achievable rate, in comparison to the Lloyd algorithm. Also, from the same table, we can quantify that the Inter-AP Lloyd algorithm, despite its significantly lower computational complexity, performs slightly to moderately better than the Interference Lloyd algorithm, giving an approximately 3% improvement in the 95%-likely achievable rate. It is worth pointing out that in our experiments, lower κ values resulted in less improvements as the Lloyd-type algorithms approached the results of the Lloyd algorithm. Higher κ values resulted in convergence issues during the AP placement process. Many iterations of the algorithms were performed with other GMM configurations and κ values. Similar performance trends were observed for various standard deviations of the GMMs. Thus, the choice of κ is an important part of the AP placement process and depends primarily on the area under consideration and the pathloss model. Finally, it is important to notice that although we have focused on the worst 5% of the users, the Inter-AP Lloyd algorithm actually boosts the performances of the worst (nearly) 25% of the users. The performance loss of the best users, as seen in the CDF plot, is justifiable due to the fact that users closer to the cell center tend to benefit from large SINR values that already suffice to provide them with more than their throughput requirements.

VIII. Conclusion

In this paper, we have addressed the access point (AP) placement problem in the small-cell uplink paradigm under the criteria of throughput, while considering inter-cell interference (ICI). After reviewing vector quantization (VQ), we explored related throughput formulations in the single user case and subsequently, the multiple user case corresponding to the considered small-cell model. Without ICI, we showed that the simple Lloyd algorithm performed similarly to the aforementioned formulations (only up to a 4% difference) and could be a baseline algorithm to solve more complex problems. Accordingly, we accounted for ICI in the optimization function of the Lloyd algorithm and mathematically arrived at two distinct distortion functions. Correspondingly, we proposed two Lloyd-type algorithms, namely the Interference Lloyd algorithm and the Inter-AP Lloyd algorithm. Both algorithms yield significant improvement to achievable rates, giving up to a marked 36.34% increase in the 95%-likely rate over the benchmark Lloyd algorithm. The Inter-AP Lloyd algorithm achieves throughput gains coupled with lower complexity and simpler user association over the Interference Lloyd algorithm. Finally, cell association strategies were outlined for all algorithms for completeness.

Appendix A
Proof of Solution for MM-Lloyd Algorithm

The expectation in the objective function in (21) can be replaced by the sample average using the user realizations at $\mathbf{p}_k$ and written as

$$J = \sum_{\mathbf{p}_k \in \mathcal{C}_m} \log\left(||\mathbf{q}_m - \mathbf{p}_k||^2 + \epsilon\right), \tag{54}$$

where $\mathbf{q}_{\mathcal{E}(\mathbf{p})}$ is replaced by $\mathbf{q}_m$ and the average is taken over all the users in cell $\mathcal{C}_m$ as the update steps correspond to the CC step of the Lloyd-type algorithm. Following the MM literature, a concave function can be upper bounded by its first-order Taylor expansion [50]

$$h(z) \leq h'(z_l)(z - z_l) + h(z_l), \tag{55}$$

where $h(\cdot)$ is concave on $\mathbb{R}^+$, z is the variable, z_l is the point around which the expansion is carried out, and $h'(\cdot)$ is the first derivative. In (54), we can take $h(z_k) = \log(z_k)$ and $z_k = ||\mathbf{q}_m - \mathbf{p}_k||^2 + \epsilon$, where we note that z_k is scalar. Thus, using the upper bound (55) in (54), the objective function is

$$J_1 = \sum_{\mathbf{p}_k \in \mathcal{C}_m} h(z_k) \leq \sum_{\mathbf{p}_k \in \mathcal{C}_m} [h'(z_{k,l})(z_k - z_{k,l}) + h(z_{k,l})]. \tag{56}$$

Removing the terms that are not involved in the optimization, we have

$$\arg\min_{\mathbf{q}_m} \sum_{\mathbf{p}_k \in \mathcal{C}_m} w_k z_k = \arg\min_{\mathbf{q}_m} \sum_{\mathbf{p}_k \in \mathcal{C}_m} w_k \left(||\mathbf{q}_m - \mathbf{p}_k||^2 + \epsilon\right), \tag{57}$$

where the weight w_k is defined as

$$\begin{aligned} w_k = h'(z_{k,l}) &= \left.\frac{\partial h(z_{k,l})}{\partial z_{k,l}}\right|_{z_{k,l} = ||\mathbf{q}_m - \mathbf{p}_k||^2 + \epsilon}, \\ &= \left.\frac{1}{z_{k,l}}\right|_{z_{k,l} = ||\mathbf{q}_m - \mathbf{p}_k||^2 + \epsilon} = \frac{1}{||\mathbf{q}_m - \mathbf{p}_k||^2 + \epsilon}, \end{aligned} \tag{58}$$

which gives the weight update equation. Now, given the weights, the objective function in (57) is

$$J_2 = \sum_{\mathbf{p}_k \in \mathcal{C}_m} w_k \left(||\mathbf{q}_m - \mathbf{p}_k||^2 + \epsilon\right). \tag{59}$$

Taking the derivative and equating it to 0, i.e., $\partial J_2 / \partial \mathbf{q}_m = 0$, gives the update equation for the AP position

$$\mathbf{q}_m = \frac{\sum\limits_{\mathbf{p}_k \in \mathcal{C}_m} w_k \mathbf{p}_k}{\sum\limits_{\mathbf{p}_k \in \mathcal{C}_m} w_k}. \tag{60}$$

Appendix B
Proof of Gradient for Lloyd-χ Algorithm

The gradient of the distortion function $d_\chi(\mathbf{p}, \mathbf{q}_m)$ is calculated as

$$\begin{aligned} &\frac{\partial}{\partial \mathbf{q}_m} \left\{ \int_{\mathbf{p} \in \mathcal{C}_m} d_\chi(\mathbf{p}, \mathbf{q}_m) f_{\mathbf{P}}(\mathbf{p}) \mathrm{d}\mathbf{p} \right\} \\ &\quad \overset{(a)}{=} \frac{\partial}{\partial \mathbf{q}_m} \left\{ \frac{1}{|\mathcal{C}_m|} \sum_{\mathbf{p}_k \in \mathcal{C}_m} d_\chi(\mathbf{p}_k, \mathbf{q}_m) \right\}, \\ &\quad = \frac{\partial}{\partial \mathbf{q}_m} \left\{ \frac{1}{|\mathcal{C}_m|} \sum_{\mathbf{p}_k \in \mathcal{C}_m} ||\mathbf{p}_k - \mathbf{q}_m||^\chi \right\}, \\ &\quad \overset{(b)}{=} \frac{\chi}{|\mathcal{C}_m|} \sum_{\mathbf{p}_k \in \mathcal{C}_m} (\mathbf{q}_m - \mathbf{p}_k) ||\mathbf{p}_k - \mathbf{q}_m||^{\chi - 2}. \end{aligned} \tag{61}$$

where (a) is obtained by replacing the expectation with the sample mean and the factor of 2 is assumed to be absorbed by the step-size δ in (b).

Appendix C
Proof of Solution for WMSE Lloyd Algorithm

Consider the simplified objective function in (34), which can be written as

$$\begin{aligned} \sum_{m=1}^{M} S_m &\overset{(a)}{=} \sum_{m=1}^{M} \frac{1}{N_m} \sum_{\mathbf{p}_k \in \mathcal{C}_m} d_{\text{SE}}(\mathbf{p}_k, \mathbf{q}_m), \\ &\overset{(b)}{=} \frac{1}{K} \sum_{m=1}^{M} \frac{1}{\Pr(\mathbf{p} \in \mathcal{C}_m)} \sum_{\mathbf{p}_k \in \mathcal{C}_m} d_{\text{SE}}(\mathbf{p}_k, \mathbf{q}_m), \end{aligned} \tag{62}$$

where in (a), we have replaced the integral of S_m (defined in (15)) with the sample average over the users present in the cell and N_m represents the number of users in cell $\mathcal{C}_m$, and in (b), we have used $\Pr(\mathbf{p} \in \mathcal{C}_m) = N_m / K$, with K as the total number of users. Comparing (62) with the objective function of the Lloyd algorithm J_{VQ} in (14), we have

$$\begin{aligned} \sum_{m=1}^{M} S_m \Pr(\mathbf{p} \in \mathcal{C}_m) &= \sum_{m=1}^{M} \frac{1}{N_m} \sum_{\mathbf{p}_k \in \mathcal{C}_m} d_{\text{SE}}(\mathbf{p}_k, \mathbf{q}_m) \times \frac{N_m}{K}, \\ &= \frac{1}{K} \sum_{m=1}^{M} \sum_{\mathbf{p}_k \in \mathcal{C}_m} d_{\text{SE}}(\mathbf{p}_k, \mathbf{q}_m), \end{aligned} \tag{63}$$

where we can observe that the objective function in (62) is a weighted MSE (WMSE) measure, with the weight related to AP m given by $w_m = 1/\Pr(\mathbf{p} \in \mathcal{C}_m)$. Thus, the NNC step is updated to use a weighted squared error distortion function, i.e., $w_m d_{\text{SE}}(\mathbf{p}_k, \mathbf{q}_m)$. The CC step however remains independent of the weights. This can be proven by taking the derivative of the objective function in (62) with respect to the AP location $\mathbf{q}_m$, which gives the AP location as

$$\mathbf{q}_m = \frac{1}{|\mathcal{C}_m|} \sum_{\mathbf{p}_k \in \mathcal{C}_m} \mathbf{p}_k. \tag{64}$$

APPENDIX D
JUSTIFICATION OF (47)

Consider the term in the denominator of (47)

$$||\mathbf{p}_{m'} - \mathbf{q}_{\mathcal{E}(\mathbf{p})}||^2 = ||\underbrace{\mathbf{p}_{m'} - \mathbf{q}_{m'}}_{\mathbf{y}} + \underbrace{\mathbf{q}_{m'} - \mathbf{q}_{\mathcal{E}(\mathbf{p})}}_{\mathbf{x}}||^2. \tag{65}$$

Let us assume that the distance between the interfering user and its serving AP, denoted by $\mathbf{y}$, is always smaller than the distance of that same AP from the nearest AP, denoted by $\mathbf{x}$, which means

$$||\mathbf{p}_{m'} - \mathbf{q}_{m'}|| \leq \left|\left|\mathbf{q}_{m'} - \mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right| \Rightarrow ||\mathbf{y}|| \leq ||\mathbf{x}||. \tag{66}$$

The above inequality always holds true when the interfering user is on the near side of the interfering AP with respect to the serving AP and does not always hold true when the interfering user is on the far side of the interfering AP with respect to the serving AP. We note that the importance of the second scenario is reduced when the interfering cells are farther away. Further, even among the interfering cells that are near (the neighboring cells), the proportion of users within such cells that does not satisfy the inequality above is small. Thus, we make the assumption that (66) is satisfied for all users. Given that m' indexes the interfering cells, we can classify these cells into cells that are the immediate neighbors of cell $\mathcal{C}_{\mathcal{E}(\mathbf{p})}$, denoted by $\mathcal{IN}(\mathcal{E}(\mathbf{p}))$ and those that are not, and are thus farther away. Hence, the two cases are

$$\begin{aligned} &||\mathbf{x}|| \geq ||\mathbf{y}||, \quad \forall m' \in \mathcal{IN}(\mathcal{E}(\mathbf{p})), \quad m' \neq \mathcal{E}(\mathbf{p}), \\ &||\mathbf{x}|| \gg ||\mathbf{y}||, \quad \forall m' \notin \mathcal{IN}(\mathcal{E}(\mathbf{p})), \quad m' \neq \mathcal{E}(\mathbf{p}). \end{aligned} \tag{67}$$

However, to simplify, we make the optimistic assumption that $||\mathbf{x}|| \gg ||\mathbf{y}||$ holds true for all $m' \neq \mathcal{E}(\mathbf{p})$. This gives

$$\begin{aligned} ||\mathbf{x} + \mathbf{y}||^2 &= ||\mathbf{x}||^2 + ||\mathbf{y}||^2 + 2||\mathbf{x}||||\mathbf{y}|| \cos\theta, \\ &= ||\mathbf{x}||^2 \left(1 + \frac{||\mathbf{y}||^2}{||\mathbf{x}||^2} + \frac{||\mathbf{y}|| \cos\theta}{||\mathbf{x}||}\right), \\ &\approx ||\mathbf{x}||^2. \end{aligned} \tag{68}$$

Note that this relation holds true even when γ assumes values other than $\gamma = 2$. Thus, from (65), we have

$$\mathbb{E}_{\mathbf{p}_{m'}} \left\{ \frac{1}{\left|\left|\mathbf{p}_{m'} - \mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}} \right\} \approx \frac{1}{\left|\left|\mathbf{q}_{m'} - \mathbf{q}_{\mathcal{E}(\mathbf{p})}\right|\right|^{\gamma}}. \tag{69}$$

APPENDIX E
PROOF OF GRADIENT FOR INTERFERENCE LLOYD ALGORITHM

The gradient is calculated using the distortion function as

$$\begin{aligned} &\frac{\partial}{\partial \mathbf{q}_m} \left\{ \int_{\mathbf{p} \in \mathcal{C}_m} d_{\text{IF}}(\mathbf{p}, \mathbf{q}_m) f_{\mathbf{P}}(\mathbf{p}) \mathrm{d}\mathbf{p} \right\} \\ &= \frac{\partial}{\partial \mathbf{q}_m} \left\{ \frac{1}{|\mathcal{C}_m|} \sum_{\mathbf{p}_k \in \mathcal{C}_m} d_{\text{IF}}(\mathbf{p}_k, \mathbf{q}_m) \right\}, \\ &= \frac{\partial}{\partial \mathbf{q}_m} \left\{ \frac{1}{|\mathcal{C}_m|} \sum_{\mathbf{p}_k \in \mathcal{C}_m} ||\mathbf{p}_k - \mathbf{q}_m||^{\gamma} \right. \\ &\quad \left. + \kappa \sum_{m' \neq m} \frac{1}{|\mathcal{C}_{m'}|} \sum_{\mathbf{p}_{k_{m'}} \in \mathcal{C}_{m'}} \frac{1}{\left|\left|\mathbf{p}_{k_{m'}} - \mathbf{q}_m\right|\right|^{\gamma}} \right\}, \\ &= \frac{\gamma}{|\mathcal{C}_m|} \sum_{\mathbf{p}_k \in \mathcal{C}_m} (\mathbf{q}_m - \mathbf{p}_k) ||\mathbf{p}_k - \mathbf{q}_m||^{\gamma - 2} \\ &\quad + \kappa \sum_{m' \neq m} \frac{\gamma}{|\mathcal{C}_{m'}|} \sum_{\mathbf{p}_{k_{m'}} \in \mathcal{C}_{m'}} \frac{\left(\mathbf{p}_{k_{m'}} - \mathbf{q}_m\right)}{\left|\left|\mathbf{p}_{k_{m'}} - \mathbf{q}_m\right|\right|^{\gamma}}, \end{aligned} \tag{70}$$

where the factor of 2 is assumed to be absorbed by the step-size δ as in Appendix B.

REFERENCES

[1] G. R. Gopal, G. P. Villardi, and B. D. Rao, "Is vector quantization good enough for access point placement?" in *Proc. 2021 55th Asilomar Conf. Signals, Syst., and Comput.*, Nov. 2021.

[2] 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.

[3] 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.

[4] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, June 2014.

[5] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: benefits and challenges," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[6] G. P. Villardi, K. Ishizu, and F. Kojima, "Reducing the codeword search complexity of FDD moderately large MIMO beamforming systems," *IEEE Trans. Commun.*, vol. 67, no. 1, pp. 273–287, Jan. 2019.

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

[8] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO systems with hardware-constrained base stations," in *Proc. 2014 IEEE Int. Conf. Acoust., Speech, and Signal Process. (ICASSP)*, May 2014, pp. 3142–3146.

[9] Y. Huang, G. Zheng, M. Bengtsson, K.-K. Wong, L. Yang, and B. Ottersten, "Distributed multicell beamforming with limited intercell coordination," *IEEE Trans. Signal Process.*, vol. 59, no. 2, pp. 728–738, Feb. 2011.

[10] K. T. Truong and R. W. Heath, "The viability of distributed antennas for massive MIMO systems," in *Proc. 2013 47th Asilomar Conf. Signals, Syst., and Comput.*, Nov. 2013, pp. 1318–1323.

[11] R. Rogalin, O. Y. Bursalioglu, H. Papadopoulos, G. Caire, A. F. Molisch, A. Michaloliakos, V. Balan, and K. Psounis, "Scalable synchronization and reciprocity calibration for distributed multiuser MIMO," *IEEE Trans. Wireless Commun.*, vol. 13, no. 4, pp. 1815–1831, Apr. 2014.

[12] 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.

[13] X. Wang, P. Zhu, and M. Chen, "Antenna location design for generalized distributed antenna systems," *IEEE Commun. Lett.*, vol. 13, no. 5, pp. 315–317, May 2009.

[14] E. Park, S. Lee, and I. Lee, "Antenna placement optimization for distributed antenna systems," *IEEE Trans. Wireless Commun.*, vol. 11, no. 7, pp. 2468–2477, July 2012.

[15] A. Yang, Y. Jing, C. Xing, Z. Fei, and J. Kuang, "Performance analysis and location optimization for massive MIMO systems with circularly distributed antennas," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5659–5671, Oct. 2015.

[16] W. Choi and J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multicell environment," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 69–73, Jan. 2007.

[17] J. Wang, H. Zhu, and N. J. Gomes, "Distributed antenna systems for mobile communications in high speed trains," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 675–683, May 2012.

[18] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and B. D. Rao, "Performance of cell-free massive MIMO systems with MMSE and LSFD receivers," in *Proc. 2016 50th Asilomar Conf. Signals, Syst., and Comput.*, Nov. 2016, pp. 203–207.

[19] 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, July 2017.

[20] D. Gesbert, S. Hanly, H. Huang, S. Shamai Shitz, O. Simeone, and W. Yu, "Multi-cell MIMO cooperative networks: A new look at interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[21] R. Irmer, H. Droste, P. Marsch, M. Grieger, G. Fettweis, S. Brueck, H. Mayer, L. Thiele, and V. Jungnickel, "Coordinated multipoint: concepts, performance, and field trial results," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, Feb. 2011.

[22] A. Ghosh, A. Maeder, M. Baker, and D. Chandramouli, "5G evolution: A view on 5G cellular technology beyond 3GPP release 15," *IEEE Access*, vol. 7, pp. 127 639–127 651, Sept. 2019.

[23] E. Khorov, A. Kiryanov, A. Lyakhov, and G. Bianchi, "A tutorial on IEEE 802.11ax high efficiency WLANs," *IEEE Commun. Surveys Tut.*, vol. 21, no. 1, pp. 197–216, Firstquarter 2019.

[24] T. Ding, M. Ding, G. Mao, Z. Lin, A. Y. Zomaya, and D. López-Pérez, "Performance analysis of dense small cell networks with dynamic TDD," *IEEE Trans. Veh. Technol.*, vol. 67, no. 10, pp. 9816–9830, Oct. 2018.

[25] G. P. Villardi, G. Thadeu Freitas de Abreu, and H. Harada, "TV white space technology: Interference in portable cognitive emergency network," *IEEE Veh. Technol. Mag.*, vol. 7, no. 2, pp. 47–53, June 2012.

[26] G. P. Villardi, C. Sum, C. Sun, Y. Alemseged, Z. Lan, and H. Harada, "Efficiency of dynamic frequency selection based coexistence mechanisms for TV white space enabled cognitive wireless access points," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 69–75, Dec. 2012.

[27] G. P. Villardi, Y. D. Alemseged, C. Sun, C. Sum, T. H. Nguyen, T. Baykas, and H. Harada, "Enabling coexistence of multiple cognitive networks in TV white space," *IEEE Wireless Commun.*, vol. 18, no. 4, pp. 32–40, Aug. 2011.

[28] X. Chen, D. Guo, and J. Grosspietsch, "The public safety broadband network: A novel architecture with mobile base stations," in *Proc. 2013 IEEE Int. Conf. Commun. (ICC)*, June 2013, pp. 3328–3332.

[29] S. T. Abraha, D. F. Castellana, X. Liang, A. Ng'oma, and A. Kobyakov, "Experimental study of distributed massive MIMO (DM-MIMO) in in-building fiber-wireless networks," in *Proc. Opt. Fiber Commun. Conf. Expo. (OFC)*, Mar. 2018, pp. 1–3.

[30] J. Guo and H. Jafarkhani, "Sensor deployment with limited communication range in homogeneous and heterogeneous wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 10, pp. 6771–6784, Oct. 2016.

[31] E. Koyuncu and H. Jafarkhani, "On the minimum average distortion of quantizers with index-dependent distortion measures," *IEEE Trans. on Signal Process.*, vol. 65, no. 17, pp. 4655–4669, Sep. 2017.

[32] J. Guo, E. Koyuncu, and H. Jafarkhani, "A source coding perspective on node deployment in two-tier networks," *IEEE Trans. Commun.*, vol. 66, no. 7, pp. 3035–3049, July 2018.

[33] S. Karimi-Bidhendi, J. Guo, and H. Jafarkhani, "Using quantization to deploy heterogeneous nodes in two-tier wireless sensor networks," in *Proc. 2019 IEEE Int. Symp. Inf. Theory (ISIT)*, July 2019, pp. 1502–1506.

[34] S. Karimi-Bidhendi, J. Guo, and H. Jafarkhani, "Energy-efficient node deployment in heterogeneous two-tier wireless sensor networks with limited communication range," *IEEE Trans. on Wireless Commun.*, vol. 20, no. 1, pp. 40–55, Jan. 2021.

[35] B. Galkin, J. Kibilda, and L. A. DaSilva, "Deployment of UAV-mounted access points according to spatial user locations in two-tier cellular networks," in *Proc. 2016 Wireless Days (WD)*, Mar. 2016, pp. 1–6.

[36] R. I. Bor-Yaliniz, A. El-Keyi, and H. Yanikomeroglu, "Efficient 3-D placement of an aerial base station in next generation cellular networks," in *Proc. 2016 IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–5.

[37] J. Lyu, Y. Zeng, R. Zhang, and T. J. Lim, "Placement optimization of UAV-mounted mobile base stations," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 604–607, Mar. 2017.

[38] L. Zhang, Q. Fan, and N. Ansari, "3-D drone-base-station placement with in-band full-duplex communications," *IEEE Commun. Lett.*, vol. 22, no. 9, pp. 1902–1905, Sept. 2018.

[39] C. Lai, C. Chen, and L. Wang, "On-demand density-aware UAV base station 3D placement for arbitrarily distributed users with guaranteed data rates," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 913–916, June 2019.

[40] L. Xie, J. Xu, and R. Zhang, "Throughput maximization for UAV-enabled wireless powered communication networks," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 1690–1703, Apr. 2019.

[41] J. Guo, P. Walk, and H. Jafarkhani, "Optimal deployments of UAVs with directional antennas for a power-efficient coverage," *IEEE Trans. Commun.*, vol. 68, no. 8, pp. 5159–5174, Aug. 2020.

[42] Y. Zeng, Q. Wu, and R. Zhang, "Accessing from the sky: A tutorial on UAV communications for 5G and beyond," *Proc. IEEE*, vol. 107, no. 12, pp. 2327–2375, Dec. 2019.

[43] E. Nayebi and B. D. Rao, "Access point location design in cell-free massive MIMO systems," in *Proc. 2018 52nd Asilomar Conf. Signals, Syst., and Comput.*, Oct. 2018, pp. 985–989.

[44] G. R. Gopal and B. D. Rao, "Throughput and delay driven access point placement," in *Proc. 2019 53rd Asilomar Conf. Signals, Syst., and Comput.*, Nov. 2019, pp. 1010–1014.

[45] E. Nayebi, "TDD massive MIMO systems: channel estimation, power optimization, and access point location design," Ph.D. dissertation, University of California, San Diego, 2018.

[46] Z. Yun and M. F. Iskander, "Ray tracing for radio propagation modeling: Principles and applications," *IEEE Access*, vol. 3, pp. 1089–1100, 2015.

[47] E. Ostlin, H. Zepernick, and H. Suzuki, "Macrocell path-loss prediction using artificial neural networks," *IEEE Trans. Veh. Technol.*, vol. 59, no. 6, pp. 2735–2747, July 2010.

[48] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression*. Norwell, MA, USA: Kluwer Acad. Publ., 1991.

[49] J. M. Ortega and W. C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*. New York, NY, USA: Academic, 1970.

[50] Y. Sun, P. Babu, and D. P. Palomar, "Majorization-minimization algorithms in signal processing, communications, and machine learning," *IEEE Trans. Signal Process.*, vol. 65, no. 3, pp. 794–816, Feb. 2017.

Govind R. Gopal (S'13–GS'21) received the B.Tech. degree in electronics and communication engineering from the National Institute of Technology Karnataka, Surathkal, India, in 2016, and the M.S. degree in electrical and computer engineering with an emphasis on communication theory and systems from the University of California San Diego, La Jolla, CA, USA, in 2018, where he is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, under the supervision of Prof. Bhaskar D. Rao. His previous experience is on routing in device-to-device communications. His current research interests include wireless communication theory, signal processing for wireless applications, multiple-input-multiple-output small-cell and cell-free systems, vector quantization, access point location design, beyond 5th generation wireless systems, and wireless standards.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2022.3148996, IEEE Transactions on Wireless Communications

Elina Nayebi received the B.S. degree in electrical engineering from the University of Tehran, Tehran, Iran, in 2012, and the M.S. and Ph.D. degrees from the Department of Electrical and Computer Engineering, University of California San Diego, La Jolla, CA, USA, in 2015 and 2018, respectively. She was an intern with Nokia Bell Labs, Murray Hill, NJ, USA, in 2014 and 2015, and with Qualcomm Technologies Inc., San Diego, CA, USA, in 2016. From 2018 to 2020, she was with the SOC-Cellular team, Samsung Semiconductor, Inc., San Diego, focusing on the R&D of RF and digital front-end systems. In 2020, she joined Apple as a PHY systems scientist/engineer. Her research interests include the areas of wireless communications and signal processing.

Gabriel Porto Villardi (S'07–M'09–SM'12) received the B.E. degree in electrical engineering with an emphasis on telecommunications from the Federal Center of Technological Education of Rio de Janeiro, Brazil, in 2002, and the M.E. and Ph.D. degrees in physics and electrical and computer engineering as a Japanese Government (Mombukagakusho) Scholar from Yokohama National University, Japan, in 2006 and 2009, respectively. In 2009, he joined the National Institute of Information and Communications Technology, where he is currently a Senior Researcher with the Wireless Systems Laboratory at the Yokosuka Research Park, Japan. His current research interests include wireless communications theory, statistical modeling, multiple-input multiple-output, cognitive radios, and beyond 5th generation wireless systems. He is member of the IEEE-SA, and has actively contributed to the IEEE 802.22, IEEE 802.11, IEEE 802.15, IEEE 802.19, and IEEE 1900.6 standardization working groups. From 2014 to 2015, he was the Secretary of the IEEE 802.22b Task Group on Enhancements for Broadband Services and Monitoring Applications in TVWS, until the standard completion, and the secretary of the IEEE 802.22 Working Group on Wireless Regional Area Networks in TVWS, from 2014 to 2016. From 1999 to 2000, he received the Coordination for the Improvement of Higher Education Personnel/Institute of International Education Scholarship to pursue his studies at Clemson University, SC, USA. From March 2019 to December 2020, he was a visiting scholar with the Qualcomm Institute of Calit2 at the University of California San Diego, La Jolla, CA, USA.

Bhaskar D. Rao (S'80–M'83–SM'91–F'00) is currently a Distinguished Professor in the Department of Electrical and Computer Engineering and the holder of the Ericsson Endowed Chair in wireless access networks at the University of California San Diego, La Jolla, CA, USA. He received the 2016 IEEE Signal Processing Society Technical Achievement Award. His research interests include digital signal processing, estimation theory, and optimization theory, with applications to digital communications, speech signal processing, and biomedical signal processing.