

Random Sequential Adsorption-based Pilot Assignment for Distributed Massive MIMO Systems

Priyabrata Parida and Harpreet S. Dhillon

Abstract—Inspired by the random sequential adsorption (RSA) process, we propose a new pilot assignment scheme for a distributed massive multiple-input-multiple-output (MIMO) system with desirable structural properties. In particular, this scheme ensures a minimum distance among co-pilot users to limit the effect of pilot contamination. Leveraging the rich literature of the RSA process, we analyze the statistics of the resulting co-pilot user point process when the system has only one pilot. Later, we extend the approach to obtain co-pilot user density for the scenario with multiple pilots. One of our key results is an accurate characterization of the probability of pilot assignment to a typical user in the network. Further, in order to provide a holistic overview of the performance of the proposed scheme, we compare the average user spectral efficiency (SE) of the proposed scheme with a baseline random pilot assignment scheme with no minimum distance guarantees. Our results demonstrate that the optimal distance threshold that maximizes the user SE for the proposed scheme decreases as the user density increases.

Index Terms—Random sequential adsorption, distributed massive MIMO, pilot assignment.

I. INTRODUCTION

The concept of distributed implementation of multiple-input-multiple-output (MIMO) technique has been actively explored for more than a decade. Recent push towards network densification has also made its practical implementation viable. In a distributed MIMO network, a number of geographically separated remote radio heads (RRHs) simultaneously serve users in the network and are connected to a central base band unit (BBU) that performs the complex task of signal processing. In a dense network, the access points (or RRHs) can be leveraged for distributed implementation resulting in higher spectral efficiency (SE) through macro-diversity and elimination of frequent handovers as users move from one access point to another. Apart from network densification, massive MIMO (mMIMO) is being considered as one of the major enablers of upcoming 5G wireless networks. Hence, it is natural to study the performance of a wireless network where a large number of RRHs with multiple antennas simultaneously serve multiple users. Not surprisingly, such studies have been conducted over the years under different pseudonyms such as network MIMO, coordinated multi-point [1], cloud radio access network (RAN) [2], fog-mMIMO [3], and cell-free mMIMO [4], [5]. Successful implementation of any MIMO technique requires accurate channel state information (CSI) at the transmitter (or receiver). Similar to cellular mMIMO networks, the CSI acquisition in distributed mMIMO systems needs to be done through uplink pilot transmission due to

its scalability. However, under the assumptions of independent Rayleigh fading and sub-optimal linear precoders, *pilot contamination* becomes the only capacity limiting factor of both cellular and cell-free mMIMO (a form of distributed mMIMO) networks [5]–[7]. Hence, judicious pilot assignment is essential to reduce the effect of pilot contamination.

From the perspective of distributed mMIMO, different approaches to allocate pilots can be categorized into centralized and distributed schemes. In [5], a distributed random pilot allocation and a centralized greedy pilot allocation scheme is presented. In [3] and [8], a distributed random access type pilot assignment scheme is proposed, where a user is not served if its CSI cannot be estimated reliably. In [9], a centralized pilot sequence design is proposed where the users in the neighborhood of an RRH use orthogonal pilot sequences. A centralized structured pilot allocation scheme leveraging k-means clustering algorithm is presented in [10]. In [11], a dynamic pilot allocation approach is presented where two users are assigned the same pilot sequence if the signal to interference and noise ratios (SINRs) of both the users are above a certain threshold. Most of these works rely on the common underlying principle that the same pilot can be assigned to the users that have sufficient geographical separation. This principle also motivates the pilot assignment scheme of this work. Our contributions are summarized next.

We present a random pilot assignment algorithm with a minimum distance constraint between co-pilot users to reduce the effect of pilot contamination. Our main contribution lies in the modeling and subsequent analysis of the resulting point process of co-pilot users by random sequential adsorption (RSA) process, which has been traditionally used across different scientific disciplines such as condensed matter physics, surface chemistry, and cellular biology, to name a few, to study the adsorption of large-particles such as colloids, proteins, and bacteria on a surface. To be specific, we accurately characterize the density of co-pilot users for a given total user density and a minimum distance threshold. This result is used to analyze the probability of pilot assignment to a typical user in the network. Further, we compare the performance of the proposed pilot assignment algorithm with respect to a baseline random pilot allocation scheme in terms of the average user SE. Our results highlight that for the proposed scheme, the optimal distance to maximize the user SE decreases as the user density increases.

II. SYSTEM MODEL

Network model: We limit our attention to the downlink (DL) of a distributed mMIMO system. The locations of the RRHs form a Poisson point process (PPP) Φ_r of density λ_r . Similarly, the user point process Ψ_u is also modeled as

Authors are with Wireless@VT, Dept. of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA (Email: {pparida, hdhillon}@vt.edu). The support of the US NSF (Grant ECCS-1731711) is gratefully acknowledged.

an independent PPP of density λ_u . Each RRH is equipped with N antennas and each user with a single antenna. The RRHs are connected to a BBU and collectively serve users in the network. The distance between a user at $\mathbf{u}_k \in \Psi_u$ and an RRH at $\mathbf{r}_m \in \Phi_r$ is denoted by d_{mk} . In line with the mMIMO literature, where the number of antennas is an order of magnitude more than the number of users, we consider that the antenna density $N\lambda_r \gg \lambda_u$. Further, invoking stationarity of this setup, we analyze the system performance for the typical user $\mathbf{u}_o$, which is located at the origin $\mathbf{o}$.

Channel estimation: Let $\mathbf{g}_{mk} = \sqrt{\beta_{mk}}\mathbf{h}_{mk}$ be the channel gain between the m -th RRH and the k -th user, where β_{mk} captures the large-scale channel gain and $\mathbf{h}_{mk} \sim \mathcal{CN}(0, \mathbf{I}_N)$ captures the small-scale channel fluctuations. We consider that the large-scale channel gain β_{mk} is only due to the distance dependent path loss, i.e. $\beta_{mk} = l(d_{mk})^{-1}$, where $l(\cdot)$ is a non-decreasing path loss function. While the mathematical analysis presented in this paper is agnostic to the choice of $l(\cdot)$, we will need to choose a specific $l(\cdot)$ for the numerical results, which is presented in Section V.

In order to obtain the channel estimates, each user uses a pilot from a set of P orthogonal pilot sequences $\mathcal{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_P]^T$, where $\mathbf{p}_i$ denotes the i -th sequence. The length of each pilot is τ_p symbol durations, which is less than the coherence interval. Further, the transmit signal-to-noise ratio (SNR) of each symbol in a pilot is ρ_p . Since we assume that the P sequences are orthogonal to each other, $P \leq \tau_p$ and $\mathbf{p}_i^H \mathbf{p}_j = \tau_p \rho_p \mathbf{1}(i = j)$, where $\mathbf{1}(\cdot)$ denotes the indicator function. Due to finite number of pilots, the pilot set needs to be reused across the network. Let the pilot used by the k -th user be $\mathbf{p}(k)$. During the pilot transmission phase, the received signal matrix at the m -th RRH is

$$\mathbf{Y}_m = \sum_{\mathbf{u}_k \in \Psi_u} \mathbf{g}_{mk} \mathbf{p}(k)^T + \mathbf{W}_m, \quad (\in \mathbb{C}^{N \times \tau_p})$$

where $\mathbf{W}_m$ is an additive white Gaussian noise matrix and its elements follow $\mathcal{CN}(0, 1)$. At the m -th RRH, the least-square estimate of the channel of the users using the l -th sequence is

$$\mathbf{y}_{ml} = \mathbf{Y}_m \mathbf{p}_l^* = \tau_p \rho_p \sum_{\mathbf{u}_k \in \Phi_{ul}} \mathbf{g}_{mk} + \mathbf{W}_m \mathbf{p}_l^*, \quad (\in \mathbb{C}^{N \times 1})$$

where Φ_{ul} is the set of locations of users that use the l -th sequence. Assuming $\mathbf{u}_o \in \Phi_{ul}$, the minimum-mean-squared-error (MMSE) estimate of the channel of the typical user at the m -th RRH is given as $\hat{\mathbf{g}}_{mo} = \mathbb{E}[\mathbf{y}_{ml} \mathbf{g}_{mo}^H] (\mathbb{E}[\mathbf{y}_{ml} \mathbf{y}_{ml}^H])^{-1} \mathbf{y}_{ml} =$

$$\frac{\beta_{mo}}{\sum_{\mathbf{u}_k \in \Phi_{ul}} \beta_{mk} + \frac{1}{\tau_p \rho_p}} \mathbf{y}_{ml} = \alpha_{mo} \mathbf{y}_{ml}. \quad (1)$$

In this case, the error vector $\tilde{\mathbf{g}}_{mk} = \mathbf{g}_{mk} - \hat{\mathbf{g}}_{mk}$ is uncorrelated to the estimated vector. Now the estimate and the error vectors are distributed as follows [4]:

$$\hat{\mathbf{g}}_{mo} \sim \mathcal{CN}(\mathbf{0}, \gamma_{mo} \mathbf{I}_N), \quad \tilde{\mathbf{g}}_{mo} \sim \mathcal{CN}(\mathbf{0}, (\beta_{mo} - \gamma_{mo}) \mathbf{I}_N),$$

where $\gamma_{mo} = \frac{\tau_p \rho_p \beta_{mo}^2}{1 + \sum_{\mathbf{u}_k \in \Phi_{ul}} \tau_p \rho_p \beta_{mk}}$. From the expression of γ_{mo} , it is clear that the quality of channel estimates depend on the locations of the co-pilot users in Φ_{ul} .

DL user SINR: Using the channel estimates, each RRH precodes the data for all the users in the network. In this work, we consider conjugate beamforming precoding scheme. Since the m -th RRH cannot distinguish among the channels of the users that use the l -th pilot, it uses the normalized direction of $\mathbf{y}_{ml}$ for beamforming, i.e. the precoding vector used to transmit data to the users that use l -th pilot is given as

$$\mathbf{w}_{ml} = \mathbf{y}_{ml} / \sqrt{\mathbb{E}[\|\mathbf{y}_{ml}\|^2]} = \hat{\mathbf{g}}_{mo} / \sqrt{\mathbb{E}[\|\hat{\mathbf{g}}_{mo}\|^2]}.$$

Now the data transmitted by the m -th RRH is given as

$$\mathbf{x}_m = \sqrt{\rho_d} \sum_{p=1}^P \mathbf{w}_{mp}^* \sum_{\mathbf{u}_k \in \Phi_{up}} \sqrt{\eta_{mk}} q_k,$$

where η_{mk} is the transmission power used by the m -th RRH for the k -th user. For each RRH, we assume the following power constraint: $\mathbb{E}[\|\mathbf{x}_m\|^2] \leq \rho_d$. The symbol received at $\mathbf{o}$ (that uses the l -th pilot) is given as

$$\begin{aligned} r_o &= \sum_{\mathbf{r}_m \in \Phi_r} \mathbf{g}_{mo}^T \mathbf{x}_m + n_o \\ &= n_o + \sqrt{\rho_d} \sum_{\mathbf{r}_m \in \Phi_r} (\hat{\mathbf{g}}_{mo}^T + \tilde{\mathbf{g}}_{mo}^T) \frac{\hat{\mathbf{g}}_{mo}^*}{\sqrt{N \gamma_{mo}}} \sqrt{\eta_{mo}} q_o \\ &\quad + \sqrt{\rho_d} \sum_{p=1, p \neq l}^P \sum_{\mathbf{u}_k \in \Phi_{up}} \sum_{\mathbf{r}_m \in \Phi_m} \mathbf{g}_{mp}^T \frac{\hat{\mathbf{g}}_{mp}^*}{\sqrt{N \gamma_{mp}}} \sqrt{\eta_{mk}} q_k \\ &\quad + \sqrt{\rho_d} \sum_{\mathbf{u}_{k'} \in \Phi'_{ul}} \sum_{\mathbf{r}_m \in \Phi_m} (\hat{\mathbf{g}}_{mo}^T + \tilde{\mathbf{g}}_{mo}^T) \frac{\hat{\mathbf{g}}_{mo}^*}{\sqrt{N \gamma_{mo}}} \sqrt{\eta_{mk'}} q_{k'}, \end{aligned}$$

where $\Phi'_{ul} = \Phi_{ul} \setminus \mathbf{u}_o$, the second term in the summation corresponds to the desired term, the third term corresponds to multi-user interference due to non-copilot users, and the fourth term is the source of interference due to pilot contamination.

III. SYSTEM PERFORMANCE ANALYSIS

A. DL power control and SINR of an arbitrary user

Since our objective is to propose a scheme to reduce pilot contamination, we focus on the operational regime where pilot contamination dominates rest of the interference terms. In the following lemma, we present the SINR expression of a typical user under the assumption that the RRHs are equipped with $N \rightarrow \infty$ antennas.

Lemma 1. *Conditioned on Φ_r and Φ_u , the asymptotic SINR of a typical user is given as*

$$\text{SINR}_{o,\infty} = \frac{(\sum_{\mathbf{r}_m \in \Phi_r} \sqrt{\eta_{mo} \gamma_{mo}})^2}{\sum_{\mathbf{u}_k \in \Phi'_{ul}} (\sum_{\mathbf{r}_m \in \Phi_r} \sqrt{\eta_{mk} \gamma_{mo}})^2}. \quad (2)$$

Proof: The estimated symbol at the typical user can be obtained as $\hat{q}_o = r_o / \sqrt{N}$. Now, using the law of large numbers, as $N \rightarrow \infty$, $\frac{\hat{\mathbf{g}}_{mo}^T \hat{\mathbf{g}}_{mo}^*}{N} \rightarrow \gamma_o$, $\frac{\hat{\mathbf{g}}_{mo}^T \tilde{\mathbf{g}}_{mo}^*}{N} \rightarrow 0$, and $\frac{\tilde{\mathbf{g}}_{mo}^T \hat{\mathbf{g}}_{mo}^*}{N} \rightarrow 0$. Hence, the limiting SINR coversges to (2). ■

In this work, we consider a distributed power control scheme [10] where the transmission power used by the m -th RRH for the typical user at $\mathbf{o}$ is given as

$$\eta_{mo} = \frac{\gamma_{mo}}{P \sum_{\mathbf{u}_k \in \Phi_{ul}} \gamma_{mk}},$$

such that $\sum_{\mathbf{u}_k \in \Phi_{ul}} \eta_{mk} = 1/P$. We assume that the RRHs allocate equal power $1/P$ to serve each set of co-pilot users. Note that since the proposed pilot assignment scheme (that we discuss in the following section) solely depends on the user locations, it can be used in conjunction with other power allocation schemes such as max-min power control [5]. Now, we define the following metrics for the performance analysis:

1. *Pilot assignment probability of a typical user*: Let $\{\mathcal{I}_o = 1\}$ be the event that the user at $\mathbf{o}$ is assigned a pilot. Then the aforementioned probability is given as

$$\mathcal{M}_o = \mathbb{P}[\mathcal{I}_o = 1] = \mathbb{E}[\mathbf{1}(\mathcal{I}_o = 1)], \quad (3)$$

where the expectation is taken over Φ_r and Φ_u . In Section IV, we present our proposed approach to characterize the above quantity. This quantity can be used to get an estimate of the number of pilots necessary to satisfy target pilot assignment probability for a given user density.

2. *Average user spectral efficiency*: It is defined as

$$\begin{aligned} \overline{\text{SE}}_o &= \mathbb{E}[\mathcal{I}_o \log_2(1 + \text{SINR}_o)] \\ &= \mathbb{E}[\log_2(1 + \text{SINR}_o) | \mathcal{I}_o = 1] \mathbb{P}[\mathcal{I}_o = 1], \end{aligned} \quad (4)$$

where the expectation is taken over Φ_r, Φ_u .

IV. ESTIMATION OF PILOT ASSIGNMENT PROBABILITY

A. The pilot assignment algorithm

Our goal is to select the sets of co-pilot users among all the users in the network such that a minimum distance R_{inh} is maintained between two co-pilot users. This can be achieved by dependent selection of the users from the original user point process Ψ_u as outlined in Algorithm 1. The algorithm assigns a random mark t_k , which is uniformly distributed in $[0, 1]$, to each point $\mathbf{u}_k \in \Psi_u$. Let $\mathcal{B}_{R_{\text{inh}}}(\mathbf{u}_k)$ be defined as the contention domain of the point at $\mathbf{u}_k$. For pilot assignment, the algorithm considers each user in increasing order of their marks, i.e. the lowest mark is considered first. From the available set of pilots, a pilot is randomly assigned to a user at $\mathbf{x}_k$, where the set of available pilots are those which have not been assigned to the users in $\mathcal{B}_{R_{\text{inh}}}(\mathbf{u}_k)$. Note that to implement this algorithm, the BBU requires only the location information of the users, which does not require any additional signaling overhead as this information is typically present at a centralized node in the network such as the BBU.

B. Analysis of the pilot assignment probability

Let Φ_u be the set of users that are assigned a pilot by Algorithm 1, i.e. $\Phi_u = \cup_{p=1}^P \Phi_{up}$. Let t_o be the mark associated with the typical user. Now, the user at $\mathbf{o}$ is assigned a pilot if $N_{\Phi_u}(\mathcal{B}_{R_{\text{inh}}}(\mathbf{o})) \leq P - 1$, where $N_{\Phi_u}(\mathcal{B}_{R_{\text{inh}}}(\mathbf{o}))$ denotes the number of points in the contention domain of $\mathbf{o}$. This is ensured by the following two events:

- 1) $\mathcal{E}_1$: there are at most $P - 1$ points in $\mathcal{B}_{R_{\text{inh}}}(\mathbf{o}) \cap \Psi_u$ that have marks less than t_o ,
- 2) $\mathcal{E}_2$: there are more than $P - 1$ points in $\mathcal{B}_{R_{\text{inh}}}(\mathbf{o}) \cap \Psi_u$ that have marks less than t_o . However, some of these points are not assigned a pilot as their contention domain

Input: User locations Ψ_u , the set of pilots $\mathcal{P}$, inhibition radius

R_{inh}

Result: Pilot assignment table $\mathcal{T}$

Initialization: $\mathcal{T} = \emptyset$, a random mark $t_i \sim U(0, 1)$ for each $\mathbf{u}_i \in \Psi_u$;

Let $\tilde{\Psi}_u$ be the set of users in the increasing order of marks;

for User $\mathbf{u} \in \tilde{\Psi}_u$ **do**

Set: $\mathcal{P}' = \mathcal{P}$;

while User $\mathbf{u}$ is not assigned a pilot **do**

if $\mathcal{P}' == \emptyset$ **then**

No pilot can be assigned: $\mathcal{T} = \mathcal{T} \cup \mathbf{o}$;

Break

else

Select a pilot sequence $\mathbf{p}_k$ randomly from the set $\mathcal{P}'$;

end

if No other users in $\mathcal{B}_{R_{\text{inh}}}(\mathbf{u})$ are using $\mathbf{p}_k$ **then**

Assign the pilot: $\mathcal{T} = \mathcal{T} \cup \mathbf{p}_k$;

Break;

else

Remove $\mathbf{p}_k$ from list of potential pilots:

$\mathcal{P}' = \mathcal{P}' \setminus \mathbf{p}_k$;

end

end

Algorithm 1: Proposed algorithm for pilot assignment in a distributed mMIMO systems.

has more than P points with marks smaller than their respective marks.

While obtaining the probability of $\mathcal{E}_1$ is straightforward, characterizing $\mathcal{E}_2$ is highly non-trivial even for $P = 1$. To characterize the probability of pilot assignment for a typical user, consider a finite observation window $\mathcal{B}_{R_s}(\mathbf{o}) \subset \mathbb{R}^2$, where $R_s \gg R_{\text{inh}}$. Let $N_u = \Psi_u(\mathcal{B}_{R_s}(\mathbf{o}))$ be the total number of users and $N_s = \Phi_u(\mathcal{B}_{R_s}(\mathbf{o}))$ be the number of users that are assigned a pilot sequence. Note that for a given N_u, N_s is a random variable as it depends on the realization of Ψ_u in $\mathcal{B}_{R_s}(\mathbf{o})$ as well as random marks associated with these points. As per the events described above, the probability of pilot assignment to a typical user is given as $\mathbb{P}[\mathcal{I}_o = 1] =$

$$\begin{aligned} &\mathbb{P}[N_u < P] + \mathbb{E}_{N_u}[\mathbb{E}[N_s | N_u] N_u^{-1} | N_u \geq P] \mathbb{P}[N_u \geq P] \\ &\approx \mathbb{P}[N_u < P] + \mathbb{E}[N_s] \mathbb{E}[N_u^{-1} | N_u \geq P] \mathbb{P}[N_u \geq P]. \end{aligned} \quad (5)$$

Note that in the above expression, we have provided an implicit way of characterizing $\mathcal{E}_2$, i.e. for a given N_u , $\mathbb{E}[N_s | N_u]$ is the average number of users that are assigned a pilot. Then the probability that any user out of N_u is assigned a pilot is $\mathbb{E}[N_s | N_u] / N_u$. The final expression requires averaging over N_u . The last step in (5) is an approximation as instead of $\mathbb{E}_{N_u}[\mathbb{E}[N_s | N_u] | N_u \geq P]$, we determine $\mathbb{E}[N_s]$ by replacing $N_u = \pi R_s^2 \lambda_u$, which is its expected value. Since N_u is Poisson distributed with mean $\lambda_u \pi R_s^2$,

$$\mathbb{E}[N_u^{-1} | N_u \geq P] = \frac{\lambda_u \pi R_s^2 - \sum_{n=0}^{P-1} \text{Poi}(\lambda_u \pi R_s^2, n)}{1 - \sum_{n=0}^{P-1} \text{Poi}(\lambda_u \pi R_s^2, n)}, \quad (6)$$

where $\text{Poi}(\lambda_u \pi R_s^2, n) = e^{-\lambda_u \pi R_s^2} \frac{(\lambda_u \pi R_s^2)^n}{n!}$. However, the estimation of $\mathbb{E}[N_s]$ for a given λ_u is not trivial even for $P = 1$. Note that for $P = 1$, the above formulation has been used to model CSMA-CA network. However, due to the intractability of $\mathcal{E}_2$, Matérn hardcore process of type-II

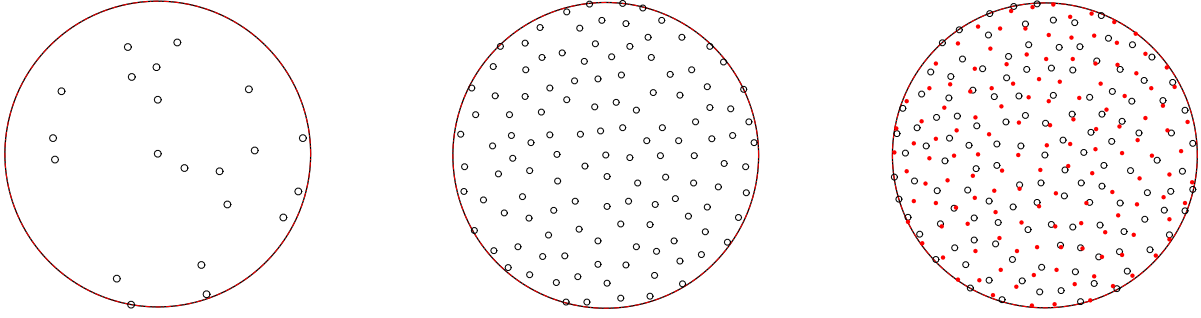


Fig. 1: Realizations of co-pilot user locations using Algorithm 1. Parameters: $R_{\text{inh}} = 200$, $\lambda_u = 2 \times 10^{-6}$ (left), $\lambda_u = 10^{-3}$ (center, right). Left and center figures represent realizations of co-pilot users for $P = 1$. Right figure represents a realization of co-pilot users for $P = 2$.

(MHPP-II) has been used for approximate characterization for Φ_u [12]. Hence, one may be inclined to extend the MHPP-II process for $P \geq 2$. However, one of the limitations of MHPP-II process is that it underestimates the number of points in Φ_u [13]. Hence, the extension of MHPP-II model for $P \geq 2$ will not result in an accurate estimation. Next, we discuss the exact point process that characterizes Φ_u and evaluate the corresponding $\mathbb{E}[N_s]$ for a given λ_u . We first present the analysis for the special case of $P = 1$. Later, we extend the approach to the general case of $P \geq 2$.

C. Pilot assignment probability for $P = 1$

The point process that overcomes the aforementioned limitation of MHPP-II is the simple sequential inhibition (SSI) process [13] or the RSA process [14]. Hence, for $P = 1$, Φ_u is an RSA point process. Traditionally, RSA process has been used across different disciplines such as condensed matter physics, surface chemistry, cellular biology to study the adsorption of different substances such as colloids, proteins, bacteria on a surface [14]. Next, we present a brief overview of RSA process before analyzing pilot assignment probability.

1) *Random sequential adsorption process*: An RSA process is defined as a stochastic space-time process, where n -dimensional hard spheres sequentially arrive at random locations in $\mathbb{R}^n$ such that any arriving sphere cannot overlap with already existing sphere. More formally, for 2D case, let Ψ be a homogeneous space-time point process on $\mathbb{R}^2 \times \mathbb{R}^+$. The circles with radii $R_{\text{inh}}/2$ are arriving at a rate of λ_Ψ per unit area per unit time. Let $\Psi(t)$ be the point process on $\mathbb{R}^2$ when Ψ is observed at an arbitrary time t . Observe that the density of $\Psi(t)$ is $\lambda_\Psi t$. At time t , an arriving point at $\mathbf{x} \in \mathbb{R}^2$ is retained if there are no other points within $\mathcal{B}_{R_{\text{inh}}}(\mathbf{x})$. Let $\varphi(t|\Psi)$ be the set of the retained points at time t . Clearly, $\varphi(t_1|\Psi) \subseteq \varphi(t_2|\Psi)$ for $t_1 \leq t_2$. Moreover, there exists a time $c \in \mathbb{R}^+$ such that $\varphi(t_i|\Psi) = \varphi(t_j|\Psi)$ for $t_i, t_j > c$, i.e. no more points can be added to the system. This is known as the *jamming limit*. Observe that the random marks assigned by Algorithm 1 can be thought of as the arrival times of the points in Ψ_u . The points that arrive early (have smaller marks) are more likely to get an assignment.

Let $\Phi(t)$ be the point process of the retained points at time t and $\rho(t)$ be the corresponding density. In order to obtain the density of retained point process for a given density of original point process, we need to observe the system at a

specific time. For example, if we want to obtain the density of retained points for an original point density of $2\lambda_\Psi$, then we need to observe the system at $t = 2$. Fig. 1 illustrates the realizations of co-pilot users for different λ_u and P . In the left figure, the system does not reach the jamming limit due to lower density of the original user point process Ψ_u . On the other hand, the center figure (almost) reaches the jamming limit and there cannot be more co-pilot users in the system. Notice the regular, almost grid-type, realization of points. The right figure illustrates the jamming state for $P = 2$. In the following lemma, we present the density of $\Phi(t)$.

Lemma 2. *The density $\rho(t)$ of the point process $\Phi(t)$ is obtained by solving the following differential equation [15] with the initial condition $\rho(0) = 0$:*

$$\int \frac{d\rho(t)}{\phi(\kappa\rho(t))} = \frac{\lambda_\Psi}{\kappa}t + C, \quad (7)$$

where $\kappa = \frac{\pi R_{\text{inh}}^2}{4}$ is the area covered by a circle, $\kappa\rho(t)$ is the fraction of the area that is covered by the retained circles at time t , $\phi(\kappa\rho(t))$ is the probability that a circle arriving at an arbitrary location in $\mathbb{R}^2$ is retained at time t , and C is the integration constant. The retention probability is given as [15, Eq. 19] $\phi(\kappa\rho(t)) =$

$$1 - 4\pi R_{\text{inh}}^2 \rho(t) + \frac{\rho(t)^2}{2} \int_0^{2R_{\text{inh}}} 4\pi r A_2(r) dr + \frac{\rho(t)^3}{3} \int_0^{2R_{\text{inh}}} 2\pi r A_2^2(r) dr - S_3^{\text{eq}} + O(\rho(t)^4), \quad (8)$$

where $S_3^{\text{eq}} = \frac{\rho(t)^3}{8} \pi (\sqrt{3}\pi - \frac{14}{3}) R_{\text{inh}}^2$, $A_2(r)$ is the area of intersection of two circles of radius R_{inh} whose centres are separated by distance r .

Proof: For the detailed proof of this lemma, please refer to [15]. Due to space limitations, we just present the proof sketch here. Note that $\kappa\rho(t)$ is the fraction of area covered by the retained circles at time t . Now, the rate of change of the fraction of the covered area with respect to time depends on the number of arrivals $\lambda_\Psi dt$ per unit area and the probability of an arrival being retained, which is given by $\phi(\kappa\rho(t))$. Hence,

$$\frac{d(\kappa\rho(t))}{dt} = \lambda_\Psi \phi(\kappa\rho(t)). \quad (9)$$

The expression for $\phi(\kappa\rho(t))$ is derived in [15]. Solution of the differential equation (9) gives us the density of $\Phi(t)$. ■

Since the function (8) is difficult to work with, a fitting function is analytically presented in [15] as $\phi_{\text{FIT}}(\rho(t)) =$

$$(1 + b_1 x(t) + b_2 x(t)^2 + b_3 x(t)^3)(1 - x(t)^3), \quad (10)$$

where $x(t) = \rho(t)/\rho(\infty)$ and $\rho(\infty)\kappa = 0.5474$ is the fraction of the area that is covered at the jamming limit as $t \rightarrow \infty$. The coefficients b_1, b_2 and b_3 are obtained by matching the order of $\rho(t)$ in equations (8) and (10). Now the expression for $\rho(t)$ is obtained by solving the differential equation (7). While the closed form solution of the equation is difficult, the problem can be efficiently solved using standard numerical softwares. Now, with the help of Lemma 2, we present pilot assignment probability to a user for $P = 1$.

Lemma 3. *When the system has one pilot, i.e. $P = 1$, the probability that typical user is assigned a pilot is given as $\mathcal{M}_o = \mathbb{P}[\mathcal{I}_o = 1] \approx$*

$$(1 - e^{-\pi R_s^2 \lambda_u})(\rho(1)\pi R_s^2)\mathbb{E}[N_u^{-1}|N_u \geq 1] - e^{-\pi R_s^2 \lambda_u},$$

where $\rho(1)$ is determined using Lemma 2 and $\mathbb{E}[N_u^{-1}|N_u \geq 1]$ using (6).

Proof: Since the density of user process Ψ_u is λ_u users per unit area, as per the RSA process definition, we can construct an equivalent space-time process where the arrivals occur at λ_u users per unit area per unit time. Now, to obtain the density of Φ_u , we observe this space-time system at time $t = 1$ time unit. Hence, the density of Φ_u is $\rho(1)$. Final expression is obtained by replacing $\mathbb{E}[N_s] = \pi R_s^2 \rho(1)$ in (5). ■

D. Pilot assignment probability for $P \geq 2$

For the general case of $P \geq 2$, consider that $\Phi_{u1}, \Phi_{u2}, \dots, \Phi_{uP}$ contain the locations of the users that are assigned the pilots $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_P$, respectively, by Algorithm 1. Since Algorithm 1 has no preference regarding the pilots, the densities of $\Phi_{u1}, \Phi_{u2}, \dots, \Phi_{uP}$ are the same. Let $\lambda_{\Phi_{uo}}$ be this density. In order to determine $\lambda_{\Phi_{uo}}$, modifications in Lemma 2 are necessary. To be specific, for (8), the knowledge of virial coefficients for a mixture of non-interacting hard spheres, and subsequently derivation of S_3^{eq} is necessary [15]. Since the above steps appear extremely difficult for this case, we provide an approximate yet accurate way to estimate the pilot assignment probability for $P \geq 2$.

Input: User locations Ψ_u , the set of pilots $\mathcal{P}$, inhibition radius R_{inh} ;

Result: Pilot assignment table $\mathcal{T}$;

Initialization: $\Psi'_u = \Psi_u, \mathcal{T} = \emptyset$;

```

for Each pilot  $\mathbf{p}_k \in \mathcal{P}$  do
  for Each user  $\mathbf{u} \in \Psi'_u$  do
    if No other users in  $\mathcal{B}_{R_{\text{inh}}}(\mathbf{u})$  are using  $\mathbf{p}_k$  then
      Assign the pilot:  $\mathcal{T} = \mathcal{T} \cup \mathbf{p}_k$ ;
      Remove  $\mathbf{u}$  from list of users:  $\Psi'_u = \Psi'_u \setminus \mathbf{u}$ ;
    end
  end
end

```

Algorithm 2: The regenerative algorithm for pilot assignment.

First, we present the regenerative pilot assignment algorithm (Algorithm 2) that is essential for our approximate analysis.

Different from Algorithm 1, in Algorithm 2, the pilots are assigned to users sequentially, i.e. the second pilot sequence is considered after the jamming limit of the first pilot has been reached, the third pilot sequence is considered after jamming limit of the second pilot sequence is reached, and so on. Recall that at the jamming limit, no additional co-pilot user can be added to the system. In order to proceed with our analysis, we make the following assumption:

Assumption 1. *The total number of pilot reuses required in $\mathcal{B}_{R_s}(\mathbf{o})$ to obtain a target pilot assignment probability is the same for both Algorithms 1 and 2. In other words, the density of users that are assigned a pilot is the same for both the algorithms.*

Let $\tilde{\Phi}_{u1}, \tilde{\Phi}_{u2}, \dots, \tilde{\Phi}_{uP}$ contain the locations of the users that are assigned pilots $\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_P$, respectively, by Algorithm 2. Let $\lambda_{\tilde{\Phi}_{u1}}, \lambda_{\tilde{\Phi}_{u2}}, \dots, \lambda_{\tilde{\Phi}_{uP}}$ be the densities of $\tilde{\Phi}_{u1}, \tilde{\Phi}_{u2}, \dots, \tilde{\Phi}_{uP}$, respectively. We approximate these densities for Algorithm 2 sequentially using Lemma 2. First, the density $\lambda_{\tilde{\Phi}_{u1}}$ of the users that are assigned the pilot $\mathbf{p}_1$ is directly obtained from Lemma 2 where the initial density of the process is λ_u . Now, to obtain the density $\lambda_{\tilde{\Phi}_{u2}}$ of the users that are assigned the pilot $\mathbf{p}_2$, we approximate the initial density of users as $\lambda_u - \lambda_{\tilde{\Phi}_{u1}}$. Also note that the points in $\Psi_u \setminus \tilde{\Phi}_{u1}$ do not form a PPP. However, for simplicity we approximate $\Psi_u \setminus \tilde{\Phi}_{u1}$ as a PPP. Similarly, to obtain $\lambda_{\tilde{\Phi}_{u3}}$, we approximate $\Psi_u \setminus \{\tilde{\Phi}_{u1} \cup \tilde{\Phi}_{u2}\}$ as a PPP of density $\lambda_u - \lambda_{\tilde{\Phi}_{u1}} - \lambda_{\tilde{\Phi}_{u2}}$ and use Lemma 2. The same approximation is made to get the rest of the densities. In the next section, we will demonstrate that these approximations do not compromise the accuracy of our results. As per the assumption, the average density of users that are assigned a pilot for both the algorithms are approximately the same. Hence, we write

$$\sum_{l=1}^P \lambda_{\tilde{\Phi}_{ul}} \simeq P \lambda_{\Phi_{uo}}. \quad (11)$$

In the next lemma, we present the pilot assignment probability for the general case of $P \geq 1$.

Lemma 4. *For a system with $P \geq 1$ pilots, the pilot assignment probability for a typical user is given as $\mathbb{P}[\mathcal{I}_o = 1] \approx$*

$$\mathbb{P}[N_u \geq P](P \lambda_{\Phi_{uo}} \pi R_s^2) \mathbb{E}[N_u^{-1}|N_u \geq P] - \mathbb{P}[N_u < P],$$

where $\lambda_{\Phi_{uo}}$ is determined from (11) and Lemma 2, N_u is Poisson distributed with mean $\lambda_u \pi R_s^2$, and $\mathbb{E}[N_u^{-1}|N_u \geq P]$ using (6).

Proof: The proof follows on the similar lines as that of Lemma 3. ■

V. RESULTS

In this section, through Monte Carlo simulations, we validate the theoretical analysis on pilot assignment probability and assess the performance benefits of RSA-inspired pilot allocation compared to random pilot allocation. For the simulations, we consider a network of radius 4000 m. In order to avoid edge effects, points within 3000 m are considered. We use the following non-line-of-sight path loss function:

$$l(d) = 161.04 - 7.1 \log_{10}(W) + 7.5 \log_{10}(h)$$

$$\begin{aligned}
& - [24.37 - 3.7(h/h_{\text{AP}})^2] \log_{10}(h_{\text{AP}}) \\
& + [43.42 - 3.1 \log_{10}(h_{\text{AP}})] [\log_{10}(d) - 3] \\
& + 20 \log_{10}(f_c) - (3.2 [\log_{10}(11.75 h_{\text{AT}})^2] - 4.97),
\end{aligned}$$

where $W = 20$, $h_{\text{AP}} = 40$, $h_{\text{AT}} = 1.5$, $h = 5$, $f_c = 0.45$ GHz.

In Fig. 2, the co-pilot user density as a function of number of pilots is presented. As expected, the co-pilot user density decreases with increasing number of pilots. In Fig. 3, we present the pilot assignment probability as a function of the number of pilots. This results is useful in determining the number of pilots that is required to achieve a certain assignment probability. Finally, in Fig. 4, we present the average user SE as a function of R_{inh} . To generate this result, we set the uplink pilot SNR $\rho_p = 80$ dB, length of pilot sequence $\tau_p = P = 16$. We observe that with increasing λ_u , the optimal R_{inh} that maximizes user SE becomes smaller. Further, there exists a range of R_{inh} that provides higher user SE compared to random pilot assignment scheme. However, this range shrinks as λ_u increases.

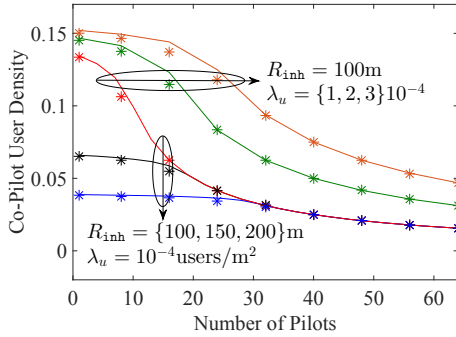


Fig. 2: The co-pilot user density as a function of P . Markers and solid lines represent simulations and theoretical results, respectively.

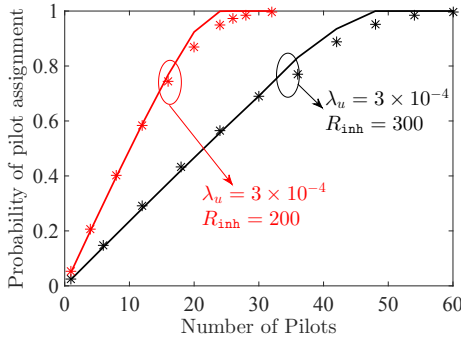


Fig. 3: Probability of pilot assignment as a function P . Markers and solid lines represent simulations and theoretical results, respectively.

VI. CONCLUSION

In this work, we propose a pilot assignment algorithm to mitigate the effect of pilot contamination for a distributed mMIMO system. Our algorithm is inspired by RSA process, which has been used to study the adsorptions of hard particles on a surface across different scientific fields. Using the well developed analytical tools for the RSA process, we present an accurate analytical expression for average pilot assignment probability for a typical user in the network. Further, we also

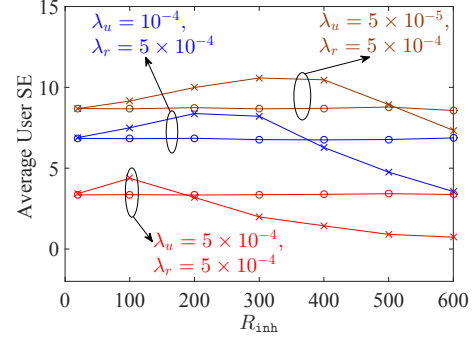


Fig. 4: Average user SE as a function of R_{inh} .

analyze the average user SE of the proposed scheme and compare it with a baseline random pilot assignment scheme. One promising extension of this work is the SINR analysis using tools from stochastic geometry.

REFERENCES

- [1] 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. Magazine*, vol. 49, no. 2, pp. 102–111, February 2011.
- [2] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for mobile networks—a technology overview," *IEEE Commun. Surveys and Tutorials*, vol. 17, no. 1, pp. 405–426, Firstquarter 2015.
- [3] O. Y. Bursalioglu, G. Caire, R. K. Mungara, H. C. Papadopoulos, and C. Wang, "Fog massive MIMO: A user-centric seamless hot-spot architecture," *IEEE Trans. on Wireless Commun.*, vol. 18, no. 1, pp. 559–574, Jan 2019.
- [4] E. Nayebe, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Trans. on Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, July 2017.
- [5] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. on Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, March 2017.
- [6] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. on Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, November 2010.
- [7] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO has unlimited capacity," *IEEE Trans. on Wireless Commun.*, vol. 17, no. 1, pp. 574–590, Jan 2018.
- [8] O. Y. Bursalioglu, C. Wang, H. Papadopoulos, and G. Caire, "RRH based massive MIMO with "on the fly" pilot contamination control," in *Proc., IEEE Intl. Conf. on Commun. (ICC)*, May 2016, pp. 1–7.
- [9] J. Zhang, X. Yuan, and Y. J. Zhang, "Locally orthogonal training design for cloud-RANs based on graph coloring," *IEEE Trans. on Wireless Commun.*, vol. 16, no. 10, pp. 6426–6437, Oct 2017.
- [10] M. Attarifar, A. Abbasfar, and A. Lozano, "Random vs structured pilot assignment in cell-free massive MIMO wireless networks," in *Proc., IEEE Intl. Conf. on Commun. Workshops (ICC Workshops)*, May 2018, pp. 1–6.
- [11] R. Sabbagh, H. Zhu, and J. Wang, "Pilot allocation and sum-rate analysis in distributed massive MIMO systems," in *Proc., IEEE Veh. Technology Conf. (VTC)*, Sep. 2017, pp. 1–5.
- [12] H. Q. Nguyen, F. Baccelli, and D. Kofman, "A stochastic geometry analysis of dense IEEE 802.11 networks," in *Proc., IEEE Int. Conf. on Comput. Commun.*, May 2007, pp. 1199–1207.
- [13] A. Busson and G. Chelius, "Capacity and interference modeling of CSMA/CA networks using SSI point processes," *Telecommun. Syst.*, vol. 57, no. 1, pp. 25–39, Sep. 2014.
- [14] J. Talbot, G. Tarjus, P. V. Tassel, and P. Viot, "From car parking to protein adsorption: an overview of sequential adsorption processes," *Colloids and Surfaces A: Physicochemical and Engineering Aspects*, vol. 165, no. 1, pp. 287 – 324, 2000.
- [15] P. Schaaf and J. Talbot, "Surface exclusion effects in adsorption processes," *J. Chem. Phys.*, vol. 91, no. 7, pp. 4401–4409, 1989.