

Channel Hardening and Favorable Propagation in Cell-Free Massive MIMO with Stochastic Geometry

Zheng Chen and Emil Björnson

Abstract—Cell-Free (CF) Massive MIMO is an alternative topology for Massive MIMO networks, where a large number of single-antenna access points (APs) are distributed over the coverage area. There are no cells but all users are jointly served by the APs using network MIMO methods. In prior work, it has been claimed that CF Massive MIMO inherits the basic properties of cellular Massive MIMO, namely channel hardening and favorable propagation. In this paper, we evaluate if one can rely on these properties when having a realistic stochastic AP deployment. Our results show that the level of channel hardening depends strongly on the propagation environment and there is generally little hardening, except when the pathloss exponent is small. We further show that using 5–10 antennas per AP, instead of one, we can greatly improve the hardening. The level of favorable propagation is affected by the propagation environment and by the distance between the users, where only spatially well separated users exhibit favorable propagation. The conclusion is that we cannot rely on channel hardening and favorable propagation when analyzing and designing CF Massive MIMO networks, but we need to use achievable rate expressions and resource allocation schemes that work well in the absence of these properties. Some options are described in this paper.

Index terms— Cell-Free Massive MIMO, channel hardening, favorable propagation, achievable rates, stochastic geometry.

I. INTRODUCTION

The throughput of conventional cellular networks is limited by the uncoordinated inter-cell interference. To mitigate this interference, Shamai and Zaidel introduced the co-processing concept in 2001 [2], which is more commonly known as network multiple-input multiple-output (MIMO) [3]. The key idea is to let all access points (APs) in the network jointly serve all users, in downlink as well as uplink, thereby turning interference into useful signals [4]. Despite the great theoretical potential, the 3GPP LTE standardization of this technology failed to provide any remarkable gains [5].

Two practical issues with network MIMO are to achieve scalable channel acquisition and sharing of data between APs. The former can be solved by utilizing only local channel state information (CSI) at each AP [6], which refers to knowledge of the channels between the AP and the users. These channels can be estimated by exploiting uplink pilot transmission and channel reciprocity in time-division duplex (TDD) systems,

thus making TDD a key enabler for network MIMO. User-centric clustering, where all APs reasonably close to a user transmit signals to it, is key to reduce the data sharing overhead [7]. These concepts are not easily incorporated into LTE, which relies on codebook-based channel acquisition and network-centric AP clustering.

The network MIMO concept has recently reappeared under the name Cell-Free (CF) Massive MIMO [8], [9], which refers to a network with a massive number of geographically distributed single-antenna APs that jointly serve a smaller number of users. In particular, it has been presented as a better option for providing coverage than using uncoordinated small cells [10]. The CF concept is fundamentally the same as in the network MIMO paper [6], where the APs perform joint transmission with access to data to every user but only local CSI. The main novelty introduced by CF Massive MIMO is the capacity analysis that takes practical pilot allocation and imperfect CSI into account [10], [11], using similar methodology as in the cellular Massive MIMO literature [12].

Conventional cellular Massive MIMO systems consist of non-cooperating APs equipped with a massive number of co-located antennas. Such systems deliver high spectral efficiency by utilizing the *channel hardening* and *favorable propagation* phenomena [12]. Channel hardening means that the beamforming turns the fading multi-antenna channel into an nearly deterministic scalar channel [13]. Favorable propagation means that the users' channel vectors are almost orthogonal [14]. These are both consequences of the law of large numbers.

CF Massive MIMO is essentially a single-cell Massive MIMO system with antennas distributed over a wide geographical area, which makes the joint channel from the APs to a user strongly spatially correlated—some APs are closer to the user than others. The capacity of Massive MIMO with spatial correlation and imperfect CSI has been analyzed in [15]–[17], among others, but with channel models that provide channel hardening and favorable propagation. It is claimed in [10] that the outstanding aspect of CF Massive MIMO is that it can also utilize these phenomena, but this has not been fully demonstrated so far. Hence, it is not clear if the known Massive MIMO capacity lower bounds are useful in the CF context or if they underestimate the achievable performance. For example, [18] derived a new downlink capacity lower bound for the case when the precoded channels are estimated using downlink pilots. That bound provides larger values than the common bound that estimates the precoded channels by relying on channel hardening, but it is unclear whether this indicates the need for downlink pilots and/or the lack of channel hardening in the CF Massive MIMO setup.

Z. Chen and E. Björnson are with the Department of Electrical Engineering (ISY), Linköping University, Linköping, Sweden (email: zheng.chen@liu.se, emil.bjornson@liu.se). A part of this work will appear in Globecom Workshops 2017 [1].

This work was supported in part by ELLIIT, CENIIT, and the Swedish Foundation for Strategic Research.

This paper aims at answering the following open questions:

- Can we observe channel hardening and favorable propagation in CF Massive MIMO with single-antenna APs?
- Is it more beneficial to deploy more antennas on few APs or more APs with few antennas, in order to achieve a reasonable degree of channel hardening and favorable propagation?
- Are there any other important factors that affect the conditions of these two properties?
- Which capacity bounds in conventional cellular Massive MIMO are appropriate to use in CF Massive MIMO?

In order to answer these questions, we model the AP distribution by a homogeneous Poisson Point Process (PPP), where each AP is equipped with $N \geq 1$ of antennas. Unlike the conventional regular grid model for the base station deployment, the stochastic point process model considered in this work can capture the irregular and semi-random AP deployment in real networks [19], [20]. *First, conditioning on a specific network realization with APs located at fixed locations and a reference user point at the origin*, we define the channel hardening and favorable propagation criteria as functions of the AP-user distances. Then, we examine the spatially averaged percentage/probability of randomly located users that satisfy these criteria. The separation of the randomness caused by small-scale fading and the spatial locations of APs allows us to study the statistical performance of large-scale CF Massive MIMO network with time-scale separation. It is similar to the concept of *meta distribution* proposed in [21], where the difference mainly lies in the definition of the studied performance metrics.

Our analysis is carried out by considering different number of antennas per AP and different non-singular pathloss models: the single-slope model with different pathloss exponents [22], [23] and the multi-slope model [10]. Compared to the conference paper [1], which focuses on the channel hardening aspect of CF Massive MIMO, in this paper, we provide thorough investigation for both channel hardening and favorable propagation, based on which we give insights into the selection of achievable rate expressions in CF Massive MIMO.

The remainder of this paper is organized as follows. In Section II we describe the CF Massive MIMO network model, including the AP distribution and the channel models. Next, Section III analyzes the channel hardening and Section IV analyzes the favorable propagation in CF Massive MIMO. Section V considers different capacity lower bounds from cellular Massive MIMO and demonstrates which ones are useful in CF systems. Section VI concludes this paper.

II. SYSTEM MODEL

We consider a CF Massive MIMO system in a finite-sized network region \mathcal{A} . The APs are distributed on the two-dimensional Euclidean plane according to a homogeneous PPP Φ_A with intensity λ_A , measured by per m^2 . [24]. Each AP is equipped with $N \geq 1$ antennas, which is a generalization of the $N = 1$ CF Massive MIMO considered in prior works [8]–[11]. All the APs are connected to a central processing unit (CPU) through backhaul, and the CPU codes and decodes the data

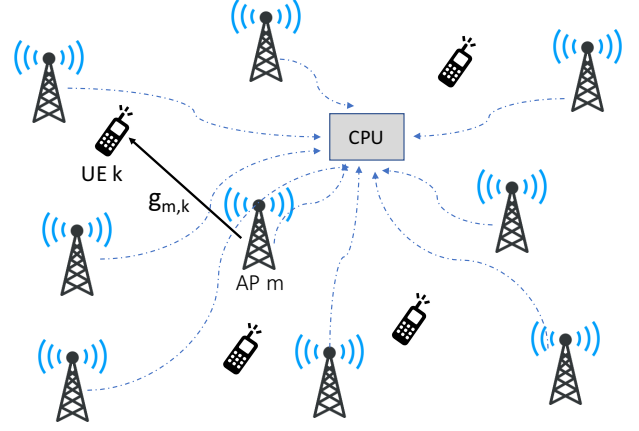


Fig. 1. Cell-Free Massive System. Here, each AP can be equipped with either single or multiple antennas.

signals; see Fig. 1 for an illustration. Different to a small-cell network, all the APs are coordinated to serve all users simultaneously using the same frequency-time resources. The number of users and their locations are generated by another independent point process. Denote by L the number of APs in a specific realization of the PPP Φ_A , we have that L is a Poisson random variable (RV) with mean value

$$\mathbb{E}[L] = \lambda_A S(\mathcal{A}), \quad (1)$$

where $S(\mathcal{A})$ denotes the area of the network region \mathcal{A} . Let M denote the total number of antennas existing in \mathcal{A} , then we have $M = LN$ and $\mathbb{E}[M] = N\lambda_A S(\mathcal{A})$.

Given the user distribution, we assume that there are K users in a specific network realization, where $K \ll M$. When the AP density is much larger than the user density, the boundary effect caused by the finite-size network region is weak, i.e., users located at the network boundary are still likely to have nearby dominant APs that makes their received signal distribution similar to network-center users. Consider a typical user at the origin, the spatially averaged network statistics seen at this typical user can represent the average network performance seen by randomly located users in the network. Denote by \mathbf{g}_k the $M \times 1$ channel vector between all the antennas and the typical user (labeled as user k), the m -th element $g_{m,k}$ is modeled by

$$g_{m,k} = \sqrt{l(d_{m,k})} h_{m,k}, \quad (2)$$

where $h_{m,k}$ represents the small-scale fading and $l(d_{m,k})$ represents the distance-dependent pathloss and it is a function of the distance $d_{m,k}$ between the m -th antenna and the user k . Since every N antennas are co-located at the same AP, we have $d_{(i-1) \cdot N + 1, k} = d_{(i-1) \cdot N + 2, k} = \dots = d_{i \cdot N, k}$, for $i = 1, \dots, L$.

We assume independent Rayleigh fading from each antenna to the typical user, which means that $\{h_{m,k}\}$ are independently and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ RVs. In the first part of this paper, we consider a non-singular pathloss model $l(r) = \min(1, r^{-\alpha})$, where r is the antenna-user distance and

$\alpha > 1$ is the pathloss exponent.¹ A three-slope pathloss model will also be studied in Section III-D.

Note that we do not include shadow fading in our analysis. With the commonly used log-normal shadowing model, the random shadowing coefficients from randomly located APs do not have any fundamental impact on the channel gain distribution. Therefore, the inclusion of shadowing coefficients would not change the general trends observed in this paper.

A. Main Advantage of CF Massive MIMO

Similar to other distributed antenna systems, the main advantage of CF Massive MIMO is the macro-diversity; that is, reduced distance between a user and its nearest APs. This can be demonstrated by analyzing the distribution of the squared norm of the channel vector,

$$\|\mathbf{g}_k\|^2 = \sum_{m=1}^M |h_{m,k}|^2 l(d_{m,k}), \quad (3)$$

which we refer to as the *channel gain*. Here, M depends on each realization of the PPP Φ_A .

Each AP is equipped with N antennas and therefore the sum of the small-scale fading coefficients of its N co-located antennas is a RV following a $\text{Gamma}(N, 1)$ distribution, which has mean N and variance N . We define the distance vector $\mathbf{r} = [r_1, \dots, r_L]^T$, where each element r_i denotes the distance from the i -th AP to the typical user at the origin. Thus, the squared norm in (3) can be written as

$$\|\mathbf{g}_k\|^2 = \sum_{i \in \Phi_A} H_i l(r_i), \quad (4)$$

where $H_i = \sum_{m=(i-1) \cdot N+1}^{i \cdot N} |h_{m,k}|^2 \sim \text{Gamma}(N, 1)$ and $r_i = d_{(i-1) \cdot N+1,k} = \dots = d_{i \cdot N,k}$ for $i = 1, \dots, L$.

Note that there are two sources of randomness in (4): $\{H_i\}$ and Φ_A . When studying the channel distribution for a randomly located user, it is natural to consider the distribution of $\|\mathbf{g}_k\|^2$ with respect to both sources of randomness. From prior studies on the application of stochastic geometry in wireless networks, it is well known that the sum of the received power from randomly distributed nodes is described by a shot noise process. The mean and variance of $\|\mathbf{g}_k\|^2$ averaging over the spatial distribution of the antennas are known for the unbounded and bounded pathloss models [23]. For our considered pathloss model $l(r) = \min(1, r^{-\alpha})$, in a finite network region with radius ρ centered around the typical user, we have

$$\mathbb{E}[\|\mathbf{g}_k\|^2] = \begin{cases} N\lambda_A \pi \left(1 + \frac{2(1-\rho^{2-\alpha})}{\alpha-2}\right) & \text{if } \alpha \neq 2 \\ N\lambda_A \pi (1 + 2 \ln(\rho)) & \text{if } \alpha = 2 \end{cases} \quad (5)$$

$$\text{Var}[\|\mathbf{g}_k\|^2] = (N^2 + N)\lambda_A \pi \left(1 + \frac{1 - \rho^{2-2\alpha}}{\alpha - 1}\right). \quad (6)$$

Proof: See Appendix A. \square

¹Note that the unbounded pathloss model $l(r) = r^{-\alpha}$ is not appropriate when analyzing CF Massive MIMO with stochastic geometry, because the antennas can then be arbitrarily close to the user, which might result in unrealistically high power gain when using the unbounded pathloss model.

When $\alpha > 2$, $\|\mathbf{g}_k\|^2$ is guaranteed to have finite mean even with infinite network size, i.e., $\rho \rightarrow \infty$. When $1 < \alpha < 2$, $\|\mathbf{g}_k\|^2$ increases unboundedly when the network size grows.

It is particularly interesting to study the case when the antenna density $\mu = N\lambda_A$ is fixed. We then observe that the mean channel gain in (5) is the same, irrespective of whether there is a high density of single-antenna APs or a smaller density of multi-antenna APs. The variance is, however, proportional to $(N^2 + N)\lambda_A = (N + 1)\mu$ and thus grows with N .

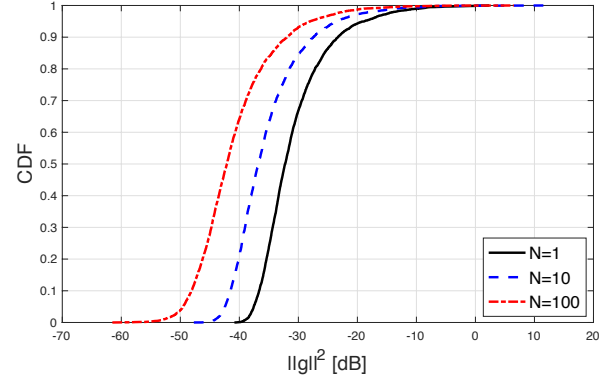


Fig. 2. The CDF of the squared norm of the channel vector $\|\mathbf{g}_k\|^2$ to user k with respect to small-scale fading and PPP realizations. The number of antennas per AP is $N = \{1, 10, 100\}$. The antenna density is fixed $\mu = N\lambda_A = 0.001/\text{m}^2$ ($10^3/\text{km}^2$).

Fig. 2 shows the cumulative distribution function (CDF) of $\|\mathbf{g}_k\|^2$, with respect to random spatial locations and small-scale fading realizations. We consider $\mu = 10^3/\text{km}^2$ and different numbers of antennas per AP: $N \in \{1, 10, 100\}$. Note that the horizontal axis is shown in decibel, thus a few users are very close to an AP and have large values of $\|\mathbf{g}_k\|^2$ while the majority have substantially smaller values. The long-tailed exponential distribution of the small-scale fading $|h_{m,k}|^2$ has a strong impact on the CDF, but also the AP density makes a difference. The larger N is, the longer tail the distribution has; at the 95%-likely point, $N = 1$ achieves a 12 dB higher value than $N = 100$. The reason behind the increasing tail with N is that $\text{Var}[\|\mathbf{g}_k\|^2]$ is proportional to $(N + 1)$, as described above. A practical interpretation is that having higher AP density reduces the average distance between the APs and this macro-diversity reduces the risk that a randomly located user has large distances to all of its closest APs. This observation evinces a key motivation of CF Massive MIMO with $N = 1$ and high AP density, in terms of providing more uniform coverage to users at random locations than a conventional cellular Massive MIMO deployment with $N = 100$ and a low AP density.

From [23] and [25], it is known that a Gamma distribution provides a good approximation of the interference distribution in a Poisson random field with non-singular pathloss. Here, the expression of $\|\mathbf{g}_k\|^2$ coincides with the definition of interference power in [23] and [25]. Thus, the Gamma distribution can be used to approximate the distribution of $\|\mathbf{g}_k\|^2$. The details are omitted since it is outside the scope of this work.

B. Conditional Channel Distribution at Fixed Location

The previous analysis characterized the channel gain distribution that a user will observe when moving around in a large network. Once the APs are deployed, for a user at a fixed location (e.g., located in a room), the small-scale fading varies over time but the large-scale fading from the APs to the user remains the same. Conditioning on a specific network realization of Φ_A , assuming that there are L APs in the network, the distances between the APs and the typical user are basically fixed. The conditional distribution of the channel statistics with respect to the small-scale fading distribution is essential for performance evaluation (e.g., computing the ergodic capacity) of CF Massive MIMO networks with a fixed topology and users at fixed but random locations.

With single-antenna APs, i.e., $N = 1$, the total number of antennas M is equal to the number of APs L . As a result of the exponentially distributed small-scale fading coefficient $|h_{m,k}|^2$, the conditional distribution of the channel gain $\|\mathbf{g}_k\|^2$ in (3) follows a Hypoexponential distribution, denoted by $\text{Hypo}(l(r_1)^{-1}, \dots, l(r_L)^{-1})$, which is usually a long-tailed distribution when the coefficients $l(r_i)$ are distinct [26].

With $N > 1$ antennas per AP, the channel gain $\|\mathbf{g}_k\|^2$ is given by (4) where $H_i \sim \text{Gamma}(N, 1)$. As a result, the conditional distribution of the channel gain from the i -th AP is $H_i l(r_i) \sim \text{Gamma}(N, l(r_i))$ for $i = 1, \dots, L$. Due to the sum of independent Gamma RVs with different scale parameters, the mean and variance of $\|\mathbf{g}_k\|^2$ conditioning on the distance vector \mathbf{r} are

$$\mathbb{E} \left[\|\mathbf{g}_k\|^2 \mid \mathbf{r} \right] = N \sum_{i=1}^L l(r_i) \quad (7)$$

$$\text{Var} \left[\|\mathbf{g}_k\|^2 \mid \mathbf{r} \right] = N \sum_{i=1}^L l^2(r_i). \quad (8)$$

The exact conditional probability density function (PDF) of $\|\mathbf{g}_k\|^2$ can be computed using the approach in [27] and the exact expression is available in [28, Eq. (6)]. By looking at (7) and (8), it is unclear how the channel gain behaves; for example, if it is the mean value or the channel variations that grow faster.

In cellular Massive MIMO with M co-located antennas, conditioning on a specific location of the user, denote by $\beta = \mathbb{E}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}] / M$ the pathloss from the M co-located antennas to the user. The squared norm of the channel gain $\|\mathbf{g}_k\|^2$ then follows a $\text{Gamma}(M, \beta)$ -distribution, with mean value βM and standard deviation $\beta \sqrt{M}$. When M increases, the distribution approaches a normal distribution and is (relatively speaking) concentrated around the mean since it grows faster than the standard deviation. The different channel gain distributions in CF and cellular Massive MIMO highlight the fundamental difference between the channel statistics of these two types of networks. In the remainder of this paper, we will proceed to investigate if in a CF Massive MIMO network, we could observe the classical Massive MIMO phenomena, namely channel hardening and favorable propagation.

III. MEASURE OF CHANNEL HARDENING

In cellular Massive MIMO, when the number of antennas grows, the channel between the AP and the user behaves as almost deterministic. This property of is called *channel hardening*. Conditioning on a specific network realization with distance vector $\mathbf{r} = [r_1, \dots, r_L]^T$, channel hardening appears in CF Massive MIMO when the following condition holds:

$$\frac{\|\mathbf{g}_k\|^2}{\mathbb{E}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}]} \rightarrow 1 \text{ as } M \rightarrow \infty, \quad (9)$$

where $\|\mathbf{g}_k\|^2 = \sum_{i=1}^L H_i l(r_i)$ is the channel gain from the L APs to the typical user. One way to prove channel hardening (with convergence² in (9) in mean square sense) is to show that the channel gain variation

$$\text{Var} \left[\frac{\|\mathbf{g}_k\|^2}{\mathbb{E}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}]} \mid \mathbf{r} \right] = \frac{\text{Var}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}]}{(\mathbb{E}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}])^2} \rightarrow 0 \text{ as } M \rightarrow \infty. \quad (10)$$

For a large wireless network, studying the channel statistics at a specific location is of limited interest and the results cannot be generalized to users at other arbitrary locations. To quantify the channel gain variation for users at arbitrary locations, we define the following channel hardening measure:

$$p_\theta = \mathbb{P} \left[\frac{\text{Var}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}]}{(\mathbb{E}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}])^2} \leq \theta \right]. \quad (11)$$

This is the CDF of $\frac{\text{Var}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}]}{(\mathbb{E}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}])^2}$ given a certain threshold θ . Here, the probability is obtained over different network realizations that generate different distance vector \mathbf{r} . As mentioned in Section II, the spatially averaged probability p_θ provides the percentage of randomly located users that experience $\frac{\text{Var}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}]}{(\mathbb{E}[\|\mathbf{g}_k\|^2 \mid \mathbf{r}])^2}$ smaller or equal to θ . Notice that $p_\theta = 1$ implies that all users have channel gain variations that are smaller than θ . The ideal case is $p_0 = 1$ where the variance is zero for all users. When the threshold θ is small enough, the larger p_θ is, with higher possibility we observe channel hardening for users at arbitrary locations.

A. Necessary Conditions for Channel Hardening

With $N \geq 1$ antennas per AP, from (7) and (8), the channel hardening measure in (11) can be written as

$$p_\theta = \mathbb{P} \left[\frac{N \sum_{i=1}^L l^2(r_i)}{\left(N \sum_{i=1}^L l(r_i) \right)^2} \leq \theta \right] = \mathbb{P} \left[\frac{\sum_{i=1}^L l^2(r_i)}{N \left(\sum_{i=1}^L l(r_i) \right)^2} \leq \theta \right]. \quad (12)$$

Since N appears in the denominator, for a given θ , p_θ always increases with N . This implies that regardless of the AP density, having more antennas per AP always helps the channel to harden. In the following, we fix the number of antennas N per AP and study the impact that the AP density λ_A has on the channel hardening criterion.

²Note that convergence in mean square implies convergence in probability.

For a given network realization with L APs, by defining $Y_1 = \sum_{i=1}^L l(r_i)$, $Y_2 = \sum_{i=1}^L l^2(r_i)$, and

$$X_{\text{ch}} = \frac{Y_2}{NY_1^2} = \frac{\sum_{i=1}^L l^2(r_i)}{N \left(\sum_{i=1}^L l(r_i) \right)^2}, \quad (13)$$

we can write the channel hardening measure as

$$p_\theta = \mathbb{P}[X_{\text{ch}} \leq \theta]. \quad (14)$$

The exact of distribution of X_{ch} is difficult to analyze even with the joint PDF of r_i , $i = 1, \dots, L$. One objective of this work is to provide intuitive insights into the relation between channel hardening and the AP density. Specifically, if p_θ should approach 1 when the AP density λ_A increases, we need $X_{\text{ch}} \rightarrow 0$ when $\lambda_A \rightarrow \infty$. Since $Y_1^2 = Y_2 + \sum_{i=1}^L l(r_i) \sum_{j=1, j \neq i}^L l(r_j)$, it follows that Y_2 and Y_1^2 are highly correlated. Though the distributions of Y_2 and Y_1^2 are not trivial to obtain, their mean and variance can be obtained by Campbell's theorem as in Section II-A. For the non-singular pathloss model $l(r) = \min(1, r^{-\alpha})$, in a network region with radius ρ , we have

$$\mathbb{E}[Y_1] = \begin{cases} \lambda_A \pi \left(1 + \frac{2(1-\rho^{2-\alpha})}{\alpha-2} \right) & \text{if } \alpha \neq 2 \\ \lambda_A \pi (1 + 2 \ln(\rho)) & \text{if } \alpha = 2 \end{cases} \quad (15)$$

$$\text{Var}[Y_1] = \lambda_A \pi \left(1 + \frac{1 - \rho^{2-2\alpha}}{\alpha-1} \right). \quad (16)$$

Then, using $\mathbb{E}[Y_1^2] = \text{Var}[Y_1] + (\mathbb{E}[Y_1])^2$, we obtain

$$\mathbb{E}[Y_1^2] = \begin{cases} \lambda_A \pi \left(\frac{\alpha - \rho^{2-2\alpha}}{\alpha-1} + \lambda_A \pi \left(\frac{\alpha-2\rho^{2-\alpha}}{\alpha-2} \right)^2 \right) & \text{if } \alpha \neq 2 \\ \lambda_A \pi \left(\frac{\alpha - \rho^{2-2\alpha}}{\alpha-1} + \lambda_A \pi (1 + 2 \ln(\rho))^2 \right) & \text{if } \alpha = 2. \end{cases} \quad (17)$$

For $Y_2 = \sum_{i=1}^L l^2(r_i)$, using again Campbell's theorem, we have

$$\mathbb{E}[Y_2] = \lambda_A \pi \frac{\alpha - \rho^{2-2\alpha}}{\alpha-1} = \text{Var}[Y_1]. \quad (18)$$

From the above results, we make the following observations:

- Y_2 scales proportionally to λ_A ;
- The higher order element of Y_1^2 scales proportionally to λ_A^2 ;
- When the pathloss is bounded, both Y_2 and Y_1^2 have finite mean, which increase with λ_A .

Given these observations, one intuitive conclusion is that when λ_A increases, $\frac{Y_2}{NY_1^2} \rightarrow 0$, which implies $\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]} \rightarrow 0$. In other words, if $\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]}$ does not converge to zero when λ_A increases, adding more APs will not help the channel to harden. We continue to investigate this necessary condition for channel hardening below.

When the network region grows infinity large, i.e., $\rho \rightarrow \infty$, depending on the pathloss exponent, we have the following cases:

1) $\alpha > 2$: As the network radius $\rho \rightarrow \infty$, we have $\rho^{2-2\alpha} \rightarrow 0$ and $\rho^{2-\alpha} \rightarrow 0$, which implies

$$\mathbb{E}[Y_1^2] \rightarrow \lambda_A \pi \left(\frac{\alpha}{\alpha-1} + \lambda_A \pi \left(\frac{\alpha}{\alpha-2} \right)^2 \right), \quad (19)$$

$$\mathbb{E}[Y_2] \rightarrow \lambda_A \pi \frac{\alpha}{\alpha-1}, \quad (20)$$

$$\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]} \rightarrow \frac{1/N}{1 + \lambda_A \pi \frac{\alpha(\alpha-1)}{(\alpha-2)^2}}. \quad (21)$$

With small N , in order to have $\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]}$ approaching 0, the AP density should satisfy $\lambda_A \frac{\alpha\pi(\alpha-1)}{(\alpha-2)^2} \gg 1$. Since the AP density is measured in APs per m^2 , the condition for channel hardening is only satisfied if $\lambda_A \sim 1$ AP/ m^2 , which is a rather unrealistic condition in practice.

2) $\alpha = 2$: This case behaves as in a free-space propagation environment. As $\rho \rightarrow \infty$, we have $\ln(\rho) \rightarrow \infty$ and $\rho^{2-2\alpha} \rightarrow 0$, which implies

$$\mathbb{E}[Y_2] \rightarrow \lambda_A \pi \frac{\alpha}{\alpha-1}, \quad (22)$$

$$\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]} \asymp \frac{1/N}{1 + \lambda_A \pi (1 + 2 \ln(\rho))^2 \frac{\alpha-1}{\alpha}} \rightarrow 0, \quad (23)$$

where the operator \asymp means that the difference between the expressions vanishes asymptotically. From (23), we observe that channel hardening is achieved as the network radius increases.

3) $1 < \alpha < 2$: One example of this case is the indoor near field propagation. With $\rho \rightarrow \infty$, we have $\rho^{2-2\alpha} \rightarrow \infty$ and $\rho^{2-\alpha} \rightarrow 0$, which implies

$$\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]} \asymp \frac{1/N}{1 + \lambda_A \pi \frac{4\rho^{4-2\alpha}}{(2-\alpha)^2} \frac{\alpha-1}{\alpha}} \rightarrow 0. \quad (24)$$

From the above equations, we see that $\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]}$ decreases rapidly with λ_A and ρ when $\alpha \leq 2$. When the network region grows infinitely large, $\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]}$ will eventually approach 0. This suggests that with smaller pathloss exponents, e.g., free-space propagation and indoor near field propagation, it is more likely to observe channel hardening in CF Massive MIMO. With the two-ray ground-reflection pathloss model and $\alpha = 4$ [29], the convergence to channel hardening only happens with impractically high antenna density.

In order to validate our analytical predictions, we present in Fig. 3 the simulated p_θ (i.e., the CDF of X_{ch}) for different AP densities, obtained with pathloss exponents $\alpha \in \{3.76, 2\}$. Note that in this figure we only consider $N = 1$. We have chosen large values of λ_A in order to see the behavior of p_θ when $\lambda_A \rightarrow \infty$. Fig. 3 shows that with $\alpha = 3.76$, for a given threshold θ , the channel hardening measure p_θ does not change much with the AP intensity, unless we reach $\lambda_A = 10^5/\text{km}^2$ ($0.1/\text{m}^2$). However, having $\lambda_A > 10^3/\text{km}^2$ is probably practically unreasonable. With $\alpha = 2$, the convergence of the channel hardening measure p_θ to one becomes more obvious when the AP density grows, which indicates that the probability to observe channel hardening at random locations is fairly large.

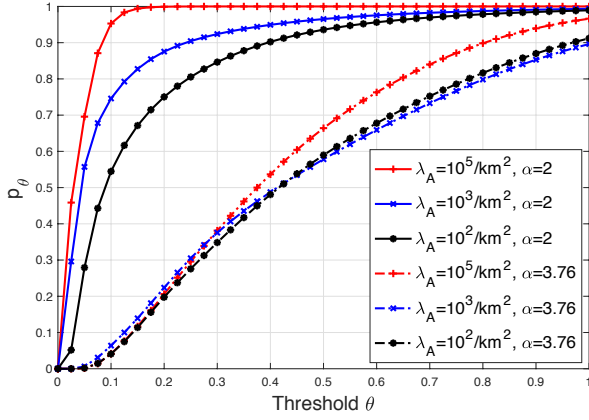


Fig. 3. The CDF of X_{ch} , with pathloss exponent $\alpha \in \{3.76, 2\}$. The network radius is $\rho = 0.5$ km and $N = 1$. The AP density is $\lambda_A \in \{10^2, 10^3, 10^5\}/\text{km}^2$, which is equivalent to $\{10^{-4}, 10^{-3}, 0.1\}/\text{m}^2$.

Property 1. Increasing the number of antennas per AP in CF Massive MIMO always helps the channel to harden. With one antenna per AP, increasing the AP density does not lead to channel hardening when using typical pathloss exponents and AP densities. In a propagation environment with a very small pathloss exponent, $\alpha \leq 2$, the channel hardening criterion has higher chance to be satisfied as the AP density increases.

B. More Antennas on Few APs or More APs with Few Antennas?

When the antenna density $\mu = N\lambda_A$ is fixed, whether to choose larger N with smaller AP density λ_A or vice versa to achieve a high level of channel hardening can be inferred from (21) for $\alpha > 2$. We can rewrite (21) as

$$\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]} = \frac{1}{N + N\lambda_A\pi\frac{\alpha(\alpha-1)}{(\alpha-2)^2}} = \frac{1}{N + \mu\pi\frac{\alpha(\alpha-1)}{(\alpha-2)^2}}. \quad (25)$$

Since the denominator contains N plus a constant term for fixed μ , we will clearly obtain more channel hardening by having more antennas on fewer APs if we have a limited number of antennas to deploy.³ For $\alpha \leq 2$, we can get the same observations from (23) and (24). Note that the stronger hardening comes at the price of less macro diversity.

In Fig. 4, we present p_θ (i.e., the CDF of X_{ch}) for different λ_A and N while keeping the overall antenna density fixed at $\mu = N\lambda_A = 10^3/\text{km}^2$ ($10^{-3}/\text{m}^2$). This figure confirms our prediction from (25) that having multiple antennas per AP will substantially help the channel to harden, and the level of channel hardening clearly increases with N . The curve $N = 50$ can be interpreted as a cellular Massive MIMO system, due to the massive number of antennas per AP. The largest gains occur when going from $N = 1$ to $N = 5$ (or to $N = 10$), thus we can achieve reasonable strong channel hardening within the scope of CF Massive MIMO if each AP is equipped with

³This result was obtained with uncorrelated fading between the user and the antennas on an AP. If there instead is spatially correlated fading, due to insufficient scattering around the AP, this will slightly reduce the hardening, but more antennas will still be beneficial.

an array of 5-10 antennas. With a smaller pathloss exponent, the required number of antennas per AP to achieve reasonably strong channel hardening is also smaller.

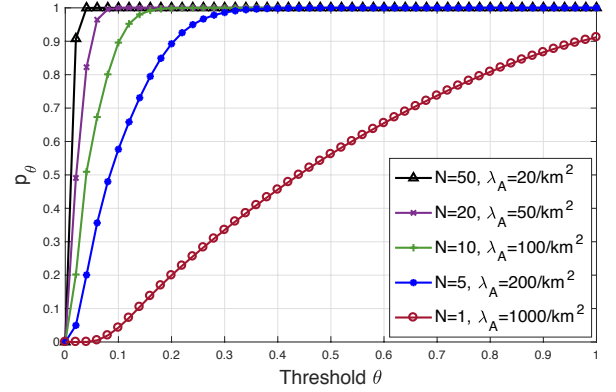


Fig. 4. The CDF of X_{ch} with pathloss exponent $\alpha = 3.76$ and network radius $\rho = 0.5$ km. The antenna density is $\mu = N\lambda_A = 1000/\text{km}^2$ ($10^{-3}/\text{m}^2$).

C. Distance vs. Number of Antennas

With higher AP density, the average number of antennas in a certain area is larger, thus the distances from the nearest APs to the user are generally smaller—this is the macro diversity effect. The effect of larger AP density on the channel hardening criteria can be related to the reduced distance and/or the increased antenna number. To understand whether the distance or the number of antennas plays the dominant role, we now examine the impact of the AP density while assuming that the typical user is only served by the M_c nearest antennas, where M_c is usually much smaller than the total antenna number M in the network region. Here, we focus only on the case with a single antenna per AP. Thus, the M_c nearest antennas represent also the M_c nearest APs.

From existing results on the distance distribution in Poisson networks, the joint PDF of the distances $\mathbf{r} = [r_1, r_2, \dots, r_{M_c}]^T$ from the M_c nearest antennas to the typical user is $f_{\mathbf{r}}(x_1, \dots, x_{M_c}) = e^{-\pi\lambda_A x_{M_c}^2} (2\pi\lambda_A)^{M_c}$, for $0 < x_1 < \dots < x_{M_c} < \infty$ [30]. To avoid the dependence between the distribution of the M_c distance variables, we consider an approximately equivalent case where a fixed number M_c antennas are uniformly and independently distributed in a disk $\mathcal{B}(0, R)$ centered at the typical user at the origin. Here, R is an average radius determined by the equivalent antenna density, i.e., $R = \sqrt{\frac{M_c}{\pi\lambda_A}}$.

As a result of the i.i.d. uniform distribution, the joint PDF of the M_c points within $\mathcal{B}(0, R)$ is

$$f_{\mathbf{r}}(x_1, \dots, x_{M_c}) = \prod_{i=1}^{M_c} \frac{2x_i}{R^2} \quad (26)$$

for $x_1, \dots, x_{M_c} \in [0, R]$. Then we have

$$\mathbb{E}[X_{\text{ch}}] \approx \int_0^R \dots \int_0^R \frac{\sum_{m=1}^{M_c} l^2(r_m)}{\left(\sum_{m=1}^{M_c} l(r_m)\right)^2} \prod_{i=1}^{M_c} \frac{2x_i}{R^2} dx_1 \dots dx_{M_c}. \quad (27)$$

Here, the approximation comes from two parts:

- $M_c \ll M$, i.e., we consider that only the M_c nearest APs contribute to the majority of the channel gain and M_c is usually much smaller than M ;
- Our assumption on having M_c antennas uniformly distributed on $\mathcal{B}(0, \sqrt{\frac{M_c}{\pi\lambda_A}})$ is an approximation of the real joint distance distribution of the M_c nearest antennas.

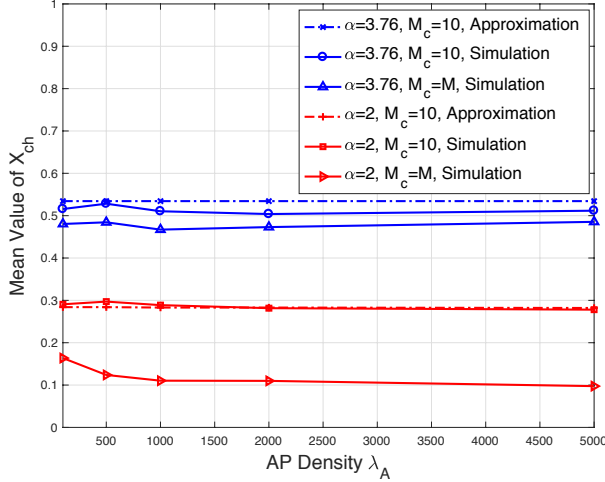


Fig. 5. The mean value of X_{ch} . The approximation results are obtained from (27) with $M_c = 10$. The simulation results are obtained for the case with $M_c = 10$ nearest APs transmitting to the user and the case with all the M antennas inside the entire network area of radius $\rho = 0.5$ km transmitting to the user. λ_A is measured in AP/km².

In Fig. 5, we present the approximation and simulation results of $\mathbb{E}[X_{ch}]$ for different antenna densities. We compare the approximation obtained from (27) with $M_c = 10$, and the mean value of $\mathbb{E}[X_{ch}]$ obtained in simulations when assuming only the nearest $M_c = 10$ antennas are transmitting to the user. For comparison, we present also simulation results of $\mathbb{E}[X_{ch}]$ with all the M antennas inside the entire network area of radius $\rho = 0.5$ km transmitting to the user at the origin. From this figure we make the following observations:

- Both analytical and simulation results show that the mean value of X_{ch} is barely affected by the AP density λ_A when $\lambda_A > 10^3/\text{km}^2$, but it depends on the pathloss exponent α . The smaller α is, the smaller mean of X_{ch} we have, the more likely we can observe channel hardening, which evinces our previous observations;
- The approximation in (27) provides an accurate measure of $\mathbb{E}[X_{ch}]$ when only a few nearest APs affect the channel gain variation, i.e., when the pathloss exponent α is large;
- With smaller pathloss exponent, e.g., $\alpha = 2$, the mean value of X_{ch} when the typical user is served by $M \gg M_c$ antennas is much lower than by $M_c = 10$ nearest antennas. This implies that the number of APs that affect the channel gain variation is much higher than 10. For example, when $\lambda_A = 10^3/\text{km}^2$, based on Monte Carlo simulations, approximately the nearest 200 APs can affect the distribution of X_{ch} . The pathloss exponent determines the speed of signal attenuation, therefore, with smaller α ,

more APs will affect the network statistics observed from the user.

Property 2. When the AP density grows, the shortened distances from nearby APs to the user have little impact on the channel hardening criterion. Increasing the total number of antennas in an environment with small pathloss exponent have stronger influence on the channel hardening than in a system with a larger pathloss exponent.

D. Multi-Slope Pathloss Model

In this section, we extend our analysis to the scenario with a multi-slope pathloss model, which models the fact that the pathloss exponent generally increases with the propagation distance. Similar to [10], we consider the three-slope pathloss model

$$l(r) = \begin{cases} Cr^{-3.5} & \text{if } r > d_1 \\ Cr^{-2}d_1^{-1.5} & \text{if } d_0 \leq r \leq d_1 \\ Cd_0^{-2}d_1^{-1.5} & \text{if } r < d_0, \end{cases} \quad (28)$$

where d_0 and d_1 are fixed distances at which the slope starts to change, C is a constant that depends on the carrier frequency and antenna height. Since the constant factor C does not affect the channel hardening measure, for simplicity, we consider $C = 1$ in the remainder of this section.

As previously in this paper, we consider $\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]} \rightarrow 0$ as the necessary condition of channel hardening. When $\rho \rightarrow \infty$,

$$\mathbb{E}[Y_1] = \lambda_A 2\pi \int_{\mathbb{R}} l(r) r dr = 2\lambda_A \pi d_1^{-1.5} \left(\ln d_1 - \ln d_0 + \frac{7}{6} \right); \quad (29)$$

$$\text{Var}[Y_1] = \lambda_A 2\pi \int_{\mathbb{R}} l^2(r) r dr = 2\lambda_A \pi d_1^{-3} \left(d_0^{-2} - \frac{3}{10} d_1^{-5} \right). \quad (30)$$

$$\begin{aligned} \mathbb{E}[Y_1^2] &= \text{Var}[Y_1] + (\mathbb{E}[Y_1])^2 \\ &= 2\lambda_A \pi d_1^{-3} \left[2\pi \lambda_A \left(\ln d_1 - \ln d_0 + \frac{7}{6} \right)^2 + d_0^{-2} - \frac{3}{10} d_1^{-5} \right]. \end{aligned} \quad (31)$$

Since $\mathbb{E}[Y_2] = \text{Var}[Y_1]$, we have

$$\begin{aligned} \frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]} &= \frac{1/N}{1 + 2\lambda_A \pi \frac{(\ln d_1 - \ln d_0 + \frac{7}{6})^2}{d_0^{-2} - \frac{3}{10} d_1^{-5}}} \\ &\approx \frac{1/N}{1 + 2\lambda_A \pi d_0^2 (\ln d_1 - \ln d_0 + \frac{7}{6})^2}, \end{aligned} \quad (32)$$

where the approximation holds when $d_1 \gg 1$ m. Based on this, using larger d_0 and d_1 will make $\frac{\mathbb{E}[Y_2]}{N\mathbb{E}[Y_1^2]}$ approach 0 with higher speed when λ_A increases. This is intuitive, given the previous observation that a smaller pathloss exponent improves the hardening, because as d_0 and d_1 increase, the number of APs with small pathloss exponents increases. Adding more antennas to the APs will also improve the channel hardening, both for a fixed λ_A and when the total antenna density $\mu = N\lambda_A$ is fixed.

Fig. 6 shows p_θ (i.e., the CDF of X_{ch}) obtained with the three-slope pathloss model, with $d_0 = 10$ m and $d_1 = 50$ m. In this figure, we compare p_θ obtained with different values

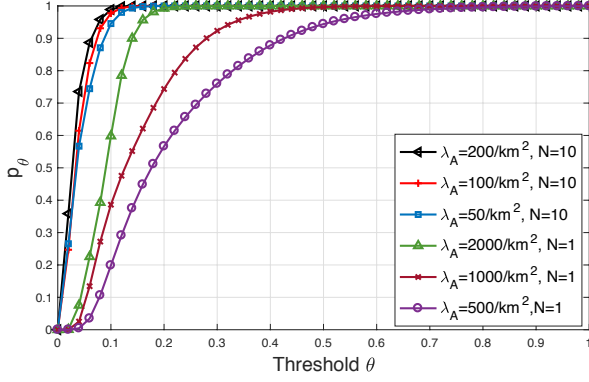


Fig. 6. The CDF of X_{ch} with the three-slope pathloss model. $d_0 = 10$ m and $d_1 = 50$ m. The total antenna density is $\mu = N\lambda_A = \{500, 1000, 2000\}/\text{km}^2$.

of the AP density λ_A for $N = 1$ and $N = 10$ antennas per AP. First, with $N = 1$, when the antenna density increases, the increase of p_θ is substantial, and more influential than the results in Fig. 3 and Fig. 4 with the one-slope model. Second, comparing the results obtained with $N = 1$ and $N = 10$, we see that the number of antennas per AP plays a more important role than increasing the AP density in helping the channel to harden, in terms of achieving a small X_{ch} with practically reasonable AP density values.

Property 3. *With the three-slope pathloss model, due to the small pathloss exponent of the propagation environment nearby the user, the channel gain variance declines rather fast with the AP density compared to the mean value. Furthermore, having large number of antennas per AP can guarantee small channel variation, which makes the channel hardening easier to achieve.*

IV. FAVORABLE PROPAGATION

In this section, we define and analyze the favorable propagation conditions in CF Massive MIMO networks. Similar to the previous section, we will consider both conventional CF networks with single-antenna APs and a generalization with multiple antennas per AP.

Recall that the channel vector from the M antennas to the user k is $\mathbf{g}_k = [g_{1,k}, \dots, g_{M,k}]^T$, where the m -th element is $g_{m,k} = \sqrt{l(d_{m,k})}h_{m,k}$. To have favorable propagation, the channel vectors between the BS and the user terminals should be orthogonal, which means

$$\mathbf{g}_k^H \mathbf{g}_j = \begin{cases} 0 & \text{if } k \neq j \\ \|\mathbf{g}_k\|^2 \neq 0 & \text{if } k = j. \end{cases} \quad (33)$$

When this condition is satisfied, each user can get the same communication performance as if it is alone in the network [31]. In practice, this condition is not fully satisfied, but can be approximately achieved when the number of antennas grows to infinity, in which case the channels are said to provide *asymptotically favorable propagation*. To be more specific, in

CF Massive MIMO, the asymptotically favorable propagation condition can be defined as follows:

$$\frac{\mathbf{g}_k^H \mathbf{g}_j}{\sqrt{\mathbb{E}[\|\mathbf{g}_k\|^2|\mathbf{d}_k] \mathbb{E}[\|\mathbf{g}_j\|^2|\mathbf{d}_j]}} \rightarrow 0, \quad \text{when } M \rightarrow \infty, k \neq j. \quad (34)$$

Here, we have conditioned on a specific network realization with distance vectors $\mathbf{d}_k = [d_{1,k}, \dots, d_{M,k}]^T$ and $\mathbf{d}_j = [d_{1,j}, \dots, d_{M,j}]^T$, and each element $d_{m,k}$ represents the distance from the m -th antenna to the user k . Recall that every N antennas are co-located at the same AP, we have $d_{(i-1)N+1,k} = \dots = d_{iN,k}$ for $i = 1, \dots, L$.

Different from cellular Massive MIMO with co-located antennas, the large-scale fading coefficients from each antenna in CF Massive MIMO to a user are different, which can be viewed as a type of spatial channel correlation. Thus, we have

$$\mathbf{g}_k^H \mathbf{g}_j = \sum_{m=1}^M \sqrt{l(d_{m,k})l(d_{m,j})} h_{m,k}^* h_{m,j} \quad (35)$$

and

$$\sqrt{\mathbb{E}[\|\mathbf{g}_k\|^2|\mathbf{d}_k] \mathbb{E}[\|\mathbf{g}_j\|^2|\mathbf{d}_j]} = \sqrt{\sum_{m=1}^M l(d_{m,k}) \sum_{m=1}^M l(d_{m,j})}. \quad (36)$$

Since $\{h_{m,k}\}$ are i.i.d. $\mathcal{CN}(0, 1)$ RVs, we have

$$\mathbb{E}[h_{m,k}^* h_{m,j}] = \begin{cases} 0 & \text{if } k \neq j \\ 1 & \text{if } k = j \end{cases} \quad (37)$$

and it follows that

$$\mathbb{E}\left[\frac{\mathbf{g}_k^H \mathbf{g}_j}{\sqrt{\mathbb{E}[\|\mathbf{g}_k\|^2|\mathbf{d}_k] \mathbb{E}[\|\mathbf{g}_j\|^2|\mathbf{d}_j]}} \middle| \mathbf{d}_k, \mathbf{d}_j\right] = \begin{cases} 0 & \text{if } k \neq j \\ 1 & \text{if } k = j. \end{cases} \quad (38)$$

With this mean value, the convergence in (34) holds (in mean square sense and in probability) if the variance of the left-hand side goes asymptotically to zero. Using (35) and (36), we have

$$\text{Var}\left[\frac{\mathbf{g}_k^H \mathbf{g}_j}{\sqrt{\mathbb{E}[\|\mathbf{g}_k\|^2|\mathbf{d}_k] \mathbb{E}[\|\mathbf{g}_j\|^2|\mathbf{d}_j]}} \middle| \mathbf{d}_k, \mathbf{d}_j\right] \quad (39)$$

$$= \frac{N \sum_{i=1}^L l(d_{i \cdot N, k}) l(d_{i \cdot N, j})}{N^2 \sum_{i=1}^L l(d_{i \cdot N, k}) \sum_{i=1}^L l(d_{i \cdot N, j})} \quad (40)$$

$$\leq \frac{L}{NL^2 \left(\frac{1}{L} \sum_{i=1}^L l(d_{i \cdot N, k})\right) \left(\frac{1}{L} \sum_{i=1}^L l(d_{i \cdot N, j})\right)}. \quad (41)$$

As the AP density λ_A grows, L increases. $\frac{1}{L} \sum_{i=1}^L l(d_{i \cdot N, k})$ will approach $\mathbb{E}[l(d_{i \cdot N, k})] > 0$, which is a positive value that only depends on the network size and the pathloss model. Thus, (40) is upper-bounded by a positive value that decreases as $1/L$ when L increases. When $L \rightarrow \infty$, the variance of the channel orthogonality will approach 0. Combined with (38), we can prove that the asymptotically favorable propagation condition defined in (34) holds for CF Massive MIMO.

For finite L , we use (40) to define the channel orthogonality metric

$$X_{\text{fp}} = \frac{\sum_{i=1}^L l(d_{i \cdot N, k}) l(d_{i \cdot N, j})}{N \sum_{i=1}^L l(d_{i \cdot N, k}) \sum_{i=1}^L l(d_{i \cdot N, j})} \quad (42)$$

and consider the probability that two users at random locations have X_{fp} no larger than a threshold γ :

$$p_\gamma = \mathbb{P}[X_{\text{fp}} \leq \gamma]. \quad (43)$$

Clearly, in order to have asymptotically favorable propagation, when the antenna density grows, p_γ should approach one for any $\gamma \geq 0$. For practical purposes, it is desirable that p_γ is large for values of the threshold γ that are close to zero. In the following, we will analyze how the antenna density, the inter-user distance, and the pathloss exponent affects the channel orthogonality.

A. Impact of Antenna Density on Channel Orthogonality

The impact of the antenna density $\mu = N\lambda_A$ will be analyzed in two cases: fixed λ_A with different N or fixed N with different λ_A . As mentioned above, X_{fp} is inversely proportional to N when L is fixed, so increasing N always helps the channels to become more orthogonal. In the other case, when L increases, the denominator of X_{fp} grows almost as L^2 , while the numerator increases almost linearly with L . Knowing that L is a Poisson RV with mean value proportional to λ_A , consequently, X_{fp} should scale roughly inversely proportional to λ_A , which evinces that for a given γ , p_γ will grow with the AP density λ_A . Combining the two cases, we can see that both larger N and larger AP density λ_A can help the channel to offer more favorable propagation. Fig. 7 presents the CDF of the channel orthogonality metric X_{fp} with different λ_A and N . By comparing the results obtained with $\lambda_A = \{500, 100\}/\text{km}^2$ and $N = \{1, 5\}$ (marked with circle, left triangle and plus sign), we validate that both increasing λ_A and increasing N can improve the channel orthogonality.

B. More Antennas on Few APs or More APs with Few Antennas?

From (42), we see that the value of X_{fp} is always upper-bounded by $\frac{1}{N}$. When the antenna density $\mu = N\lambda_A$ is fixed, increasing N means smaller λ_A . As the result, the average number of APs within close distance to the user will be less. Thus, it is hard to predict whether it is more beneficial to have more antennas on few APs or more APs with few antennas.

In Fig. 7, we present the CDF of X_{fp} when fixing the total antenna density $N\lambda_A = 500/\text{km}^2$. We see that increasing N does not necessarily lead to higher or lower p_γ for a given value of γ . We also observe that when choosing sufficiently large N , e.g., $N \geq 20$, the channel orthogonality metric X_{fp} becomes very small. In other words, sufficiently large N will help the channels to different users to be asymptotically orthogonal.

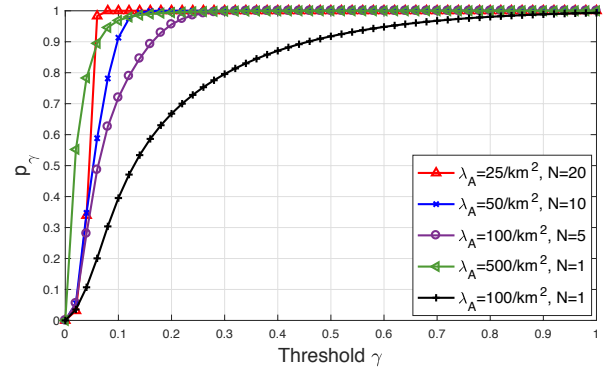


Fig. 7. The CDF of X_{fp} . Pathloss exponent $\alpha = 3.76$. Network radius $\rho = 0.5$ km. For the first four curves (marked with upward triangle, cross, circle and left triangle), the antenna density is $\mu = N\lambda_A = 500/\text{km}^2$ ($5 \times 10^{-4}/\text{m}^2$). The distance between user j and user k is 70 m.

C. Impact of Inter-User Distance on Channel Orthogonality

From (34), it is obvious that the distance between two users affects the variance of each term $\sqrt{l(d_{m,k})l(d_{m,j})}h_{m,k}^*h_{m,j}$. When two users are far apart, their channel vectors are more likely to be orthogonal with smaller variance. This result comes from the fact that $l(d_{m,k})l(d_{m,j})$ for all $m = 1, \dots, M$ will become much smaller when the distance $l_{k,j}$ between user k and user j is large. In addition, $\sum_{m=1}^M l(d_{m,k}) \sum_{m=1}^M l(d_{m,j})$ will not vary much with the inter-user distance $l_{k,j}$ when M is fixed. Therefore, X_{fp} becomes smaller when the distance $l_{k,j}$ increases.

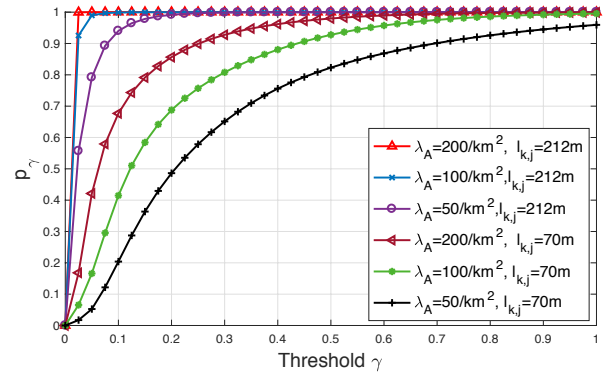


Fig. 8. The CDF of X_{fp} with different λ_A and different inter-user distances $l_{k,j}$, $N = 1$, and pathloss exponent $\alpha = 3.76$.

In Fig. 8, we present p_γ for different AP densities $\lambda_A \in \{50, 100, 200\}$ APs/ km^2 , $N = 1$, and inter-user distances $l_{k,j} \in \{70, 212\}$ m. First, it is shown that with larger λ_A , the variance of the orthogonality metric X_{fp} is smaller. Second, when the distance between two users is larger, they are more likely to have nearly orthogonal channels. This observation showcases the importance of serving spatially separated users in order to ensure near channel orthogonality.

D. Impact of Pathloss Exponent on Channel Orthogonality

When the number of antennas M and their locations are fixed, we consider two extreme cases: user k and user j are

very close or extremely far from each other. Since X_{ch} coincides with X_{fp} in the special case when $d_{i \cdot N, k} \approx d_{i \cdot N, j}$ for all $m = 1, \dots, M$, we infer from Section III that smaller pathloss exponent will also lead to smaller X_{fp} . In the other extreme case, when the two users are far apart, in the denominator of X_{fp} , $\sum_{i=1}^L l(d_{i \cdot N, j})$ is almost independent of $\sum_{i=1}^L l(d_{i \cdot N, k})$, and both terms increase much faster than the numerator, especially when α is small. Combining these two extreme cases, we expect that smaller pathloss exponent would help the channels to become asymptotically orthogonal when M is sufficiently large.

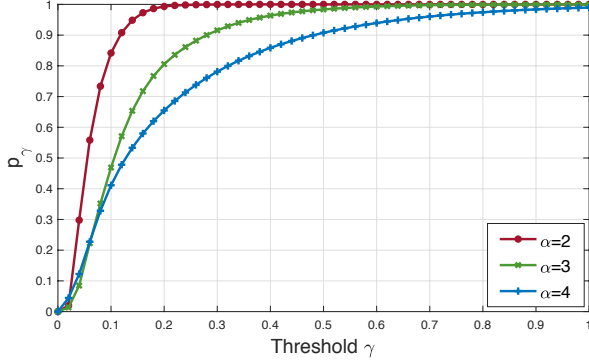


Fig. 9. The CDF of X_{fp} with different pathloss exponent $\alpha \in \{2, 3, 4\}$, $N = 1$, and the inter-user distance $l_{k,j} = 70$ m.

Fig. 9 shows p_γ obtained with pathloss exponents $\alpha \in \{2, 3, 4\}$, $N = 1$, and an inter-user distance of $l_{k,j} = 70$ m. The figure shows that with smaller α the channels become more orthogonal, which is line with our prediction from above. If we would instead use the three-slope pathloss model in (28), APs close to the user (distance smaller than d_1) will have pathloss exponent $\alpha \leq 2$, and users at larger distances have pathloss exponent 3.5. Therefore, compared to the single-slope non-singular pathloss model $l(r) = \min(r^{-3.76}, 1)$, the channels between two users will in the average be more orthogonal.

Summarizing the above analysis and observations, we have the following conclusions.

Property 4. *Increasing the antenna density by increasing either the AP density or the number of antennas per AP can both help the user channels to offer favorable propagation. Smaller pathloss also helps the channels to become asymptotically orthogonal. The larger the distance between two users, the more likely their channels will be nearly orthogonal.*

V. CAPACITY BOUNDS FOR CELL-FREE MASSIVE MIMO

A key conclusion from the previous sections is that CF Massive MIMO systems exhibit little channel hardening, as compared to cellular Massive MIMO. Hence, although CF Massive MIMO is equivalent to a single-cell Massive MIMO system with strong spatial channel correlation, we must be careful when reusing results from the Massive MIMO literature. In particular, capacity bounds that were derived by relying on channel hardening can potentially be very loose when applied to CF systems. In this section, we explain which capacity lower bounds are suitable for CF systems.

Consider a CF Massive MIMO system with M single-antenna APs and K users, which are assigned mutually orthogonal pilot sequences. The transmission is divided into coherence intervals of τ_c samples, whereof τ_p are used for uplink pilot signaling and $K \leq \tau_p \leq \tau_c$. The channels are modeled as in previous sections: $\mathbf{g}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{B}_k)$, where $\mathbf{B}_k = \text{diag}(\beta_{1,k}, \dots, \beta_{M,k})$ and $\beta_{m,k} = l(r_{m,k})$. Each user transmits its orthogonal pilot sequence. User k uses the transmit powers ρ_k and p_k for pilot and data, respectively. Since the elements in \mathbf{g}_k are independent, it is optimal to estimate them separately at the receiving antenna. The MMSE estimate of $g_{m,k}$ is

$$\hat{g}_{m,k} = \frac{\sqrt{\tau_p \rho_k \beta_{m,k}}}{\tau_p \rho_k \beta_{m,k} + 1} (\sqrt{\tau_p \rho_k} g_{m,k} + w_{m,k}), \quad (44)$$

where $w_{m,k} \sim \mathcal{CN}(0, 1)$ is i.i.d. additive noise. If we denote by $\gamma_{m,k} \triangleq \mathbb{E}[|\hat{g}_{m,k}|^2]$ the mean square of the MMSE estimate of $g_{m,k}$, then it follows from (44) that

$$\gamma_{m,k} = \frac{\tau_p \rho_k \beta_{m,k}^2}{\tau_p \rho_k \beta_{m,k} + 1}. \quad (45)$$

We will now compare different achievable rate expressions for uplink and downlink when using maximum ratio (MR) processing, which is commonly assumed in CF Massive MIMO since it can be implemented distributively. The expressions are lower bounds on the capacity, thus we should use the one that gives the largest value the accurately predict the achievable performance.

A. Uplink Achievable Rate

During the uplink data transmission, all K users simultaneously transmit to the M APs. When using MR, an achievable rate (i.e., a lower bound on the capacity) of user k is

$$R_{u,k}^{\text{UatF}} = \log_2 \left(1 + \frac{p_k \left(\sum_{m=1}^M \gamma_{m,k} \right)^2}{\sum_{j=1}^K p_j \sum_{m=1}^M \gamma_{m,k} \beta_{m,j} + \sum_{m=1}^M \gamma_{m,k}} \right), \quad (46)$$

which was used for CF Massive MIMO in [10]. This bound is derived based on the use and then forget (UatF) principle [12], where the channel estimates are used for MR but then “forgotten” and channel hardening is utilized to obtain a simple closed-form expression. Note that $R_{u,k}^{\text{UatF}}$ is a special case of the general expression in [32] for correlated single-cell Massive MIMO systems that apply MR processing. Alternatively, the achievable rate expression in [16] for spatially correlated channels can be used:

$$R_{u,k} = \mathbb{E} \left[\log_2 \left(\frac{p_k |\mathbf{a}_k^H \hat{\mathbf{g}}_k|^2}{\sum_{j \neq k}^K p_j |\mathbf{a}_k^H \hat{\mathbf{g}}_j|^2 + \mathbf{a}_k^H \left(\sum_{j=1}^K p_j (\mathbf{B}_j - \mathbf{\Gamma}_j) + \mathbf{I}_M \right) \mathbf{a}_k} \right) \right], \quad (47)$$

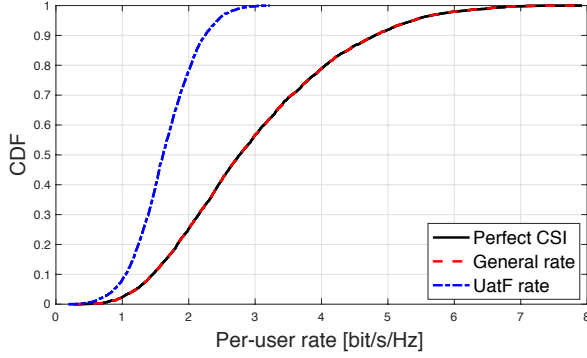


Fig. 10. The CDF of the uplink achievable rates obtained with the UatF bound, the general bound, and the rate with perfect CSI. $M = 100$, $N = 1$, $K = 20$, $\tau_p = 20$.

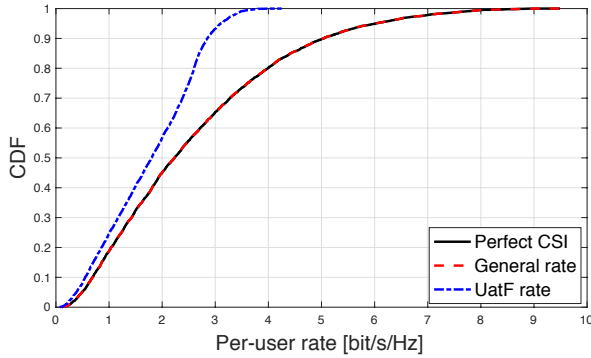


Fig. 11. Same as Fig. 10. $M = 100$, $N = 5$, $K = 20$, $\tau_p = 20$.

TABLE I
SIMULATION SETUP

Parameters	Values
M, K, τ_p, τ_c	100, 20, 20, 500
d_0, d_1	10 m, 50 m
Carrier frequency f	1.9 GHz
Antenna height h_{AP}, h_u	1.5 m, 1.65 m
Uplink pilot power ρ_k , data power p_k	100 mW, 100 mW
Downlink power per UE q	100 mW

where $\hat{\mathbf{g}}_k = [\hat{g}_{1,k}, \dots, \hat{g}_{M,k}]^T$, $\mathbf{\Gamma}_j = \text{diag}(\gamma_{1,j}, \dots, \gamma_{M,j})$, and \mathbf{a}_k is the combining vector, which is $\mathbf{a}_k = \hat{\mathbf{g}}_k$ for MR. This general bound does not rely on channel hardening.

In Fig. 10 and Fig. 11, we compare the uplink achievable rates obtained with (46) and (47), and also the rate with perfect CSI (obtained from (47) by letting $\rho_k \rightarrow \infty$). The simulation is performed in a network area of 1 km \times 1 km. $K = 20$ users are randomly and uniformly distributed in the network region, and we have $\tau_p = K$. The total number of antennas is $M = 100$. The results in Fig. 10 are obtained with $L = 100$ single-antenna APs and those in Fig. 11 are obtained with $L = 20$ APs with $N = 5$ antennas per AP. All the APs are independently and uniformly distributed in the network region. The length of the coherence block is $\tau_c = 500$. The large-scale fading coefficients between the antennas and the users are generated from 300 different PPP realizations. The three-slope pathloss model in (28) is used with $d_0 = 10$ m and $d_1 = 50$ m. The

constant factor C (dB) is given by

$$C = 105 + 46.3 + 33.9 \log_{10}(f) - 13.82 \log_{10}(h_{AP}) - (1.1 \log_{10}(f) - 0.7)h_u + (1.56 \log_{10}(f) - 0.8), \quad (48)$$

where f is the carrier frequency, h_{AP} and h_u are the AP and user antenna height, respectively [10]. The simulation parameters are summarized in Table I. Fig. 10 and Fig. 11 show the CDFs of the user rates for random distances between the L APs and the K users. The general rate expression in (47) provides almost identical rates to the perfect CSI case, which indicates that the estimation errors to the closest APs are negligible. In contrast, the UatF rate in (46) is a much looser bound on the capacity, particularly for the users that support the highest data rates. Some users get almost twice the rate when using the general expression. This is due to the lack of channel hardening for users that are close to only a few APs. Hence, a general guideline is to only use (47) for achievable rates in CF Massive MIMO with single-antenna APs. An important observation from Fig. 10 and Fig. 11 is that, with multiple antennas per AP, e.g., $N = 5$, the differences between the UatF bound and the perfect CSI case becomes smaller, particularly for the users that support the smallest rates. This comes from the fact that having multiple antennas per AP gives better channel hardening, as we have seen in Section III. However, the average rate reduces when going from many single-antenna APs to fewer multi-antenna APs, so it is not obvious what is the better choice in practice.

B. Downlink Achievable Rate

In the downlink of Massive MIMO, it is popular to use rate expressions where the only CSI available at the user is the channel statistics. When there is channel hardening, one can achieve rates similar to the uplink. If MR precoding is used, then one can use the rate expression for spatially correlated fading from [16] and obtain the achievable rate

$$R_{d,k}^{\text{UatF}} = \log_2 \left(1 + \frac{\left(\sum_{m=1}^M \sqrt{p_{m,k} \gamma_{m,k}} \right)^2}{\sum_{k'=1}^K \sum_{m=1}^M p_{m,k'} \beta_{m,k} + 1} \right) \quad (49)$$

for user k , where $p_{m,k}$ is the average transmit power allocated to user k by AP m . This type of expression was used for CF Massive MIMO in [10], using a slightly different notation. We call this the UatF rate since we use the received signals for detection, but “forget” to use them for blind estimation of the instantaneous channel realizations.

Suppose \mathbf{a}_k is the precoding vector (including power allocation) assigned to user k . To avoid relying on channel hardening, user k can estimate its instantaneous channel after precoding, $\mathbf{a}_k^H \mathbf{g}_k$, from the collection of τ_d received downlink

signals in the current coherence block. By following a similar approach as in [33, Lemma 3], we obtain the achievable rate

$$R_{d,k} = \mathbb{E} \left[\log_2 \left(\frac{|\mathbf{a}_k^H \mathbf{g}_k|^2}{\sum_{j \neq k}^K |\mathbf{a}_j^H \mathbf{g}_k|^2 + 1} \right) \right] - \frac{1}{\tau_d} \sum_{j=1}^K \log_2 \left(1 + \tau_d \text{Var}[\mathbf{a}_j^H \mathbf{g}_k] \right). \quad (50)$$

The first term in (50) is the rate with perfect CSI and the second term is a penalty term from imperfect channel estimation at the user. Note that the latter term vanishes as $\tau_d \rightarrow \infty$, thus this general rate expression is a good lower bound when the channels change slowly. The rate with MR precoding is obtained by setting $\mathbf{a}_k = [\hat{g}_{1,k} \sqrt{\frac{p_{1,k}}{\gamma_{1,k}}} \dots \hat{g}_{M,k} \sqrt{\frac{p_{M,k}}{\gamma_{M,k}}}]^T$.

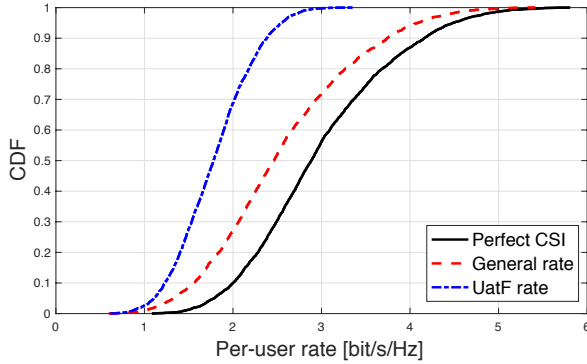


Fig. 12. The CDF of the downlink achievable rates obtained with the UatF bound, the general bound, and the rate with perfect CSI. $M = 100$, $N = 1$, $K = 20$, $\tau_p = 20$.

In Fig. 12, we compare the downlink achievable rates obtained with (49) and (50), and also the rate with perfect CSI at user (obtained from (50) by letting $\tau_d \rightarrow \infty$). The simulation parameters are summarized in Table I, which are the same as in Section V-A for the uplink. The downlink power $p_{m,k}$ is given by

$$p_{m,k} = \frac{q \cdot \gamma_{m,k}}{\mathbb{E}[\|\hat{\mathbf{g}}_k\|^2]} = q \frac{\gamma_{m,k}}{\sum_{m'=1}^M \gamma_{m',k}}, \quad (51)$$

where q is the downlink power allocated to each user k , which is chosen as 100 mW in the simulations. By doing so, we have $\mathbb{E}[\|\mathbf{a}_k\|^2] = q$, which is the same for all users. Fig. 12 shows the CDFs of the user rates for random antenna-user distances. As in the uplink, there is a substantial gap between the rates achieved by perfect CSI curve and the UatF rate. The curve for the general rate in (50) is in the middle, which implies that the users should estimate their downlink channels and not rely on channel hardening in CF Massive MIMO. It is unclear whether the gap from the perfect CSI curve is due to limited downlink estimation quality or is an artifact from the capacity bounding technique that lead to (50). In any case, there is a need to further study the achievable downlink rates in CF Massive MIMO.

VI. CONCLUSION

In this work, we provided thorough investigation of the channel hardening and favorable propagation phenomena in CF Massive MIMO systems from a stochastic geometry perspective. By studying the channel distribution from stochastically distributed APs with either single or multiple antennas per AP, we characterized the channel gain distribution, based on which we examined the conditions of channel hardening and favorable propagation. Our results showed that whether the channel hardens with increasing number of antennas strongly depends on the propagation environment and pathloss model, but one should generally not expect much hardening. However, one can improve the hardening by deploying multiple antennas per AP. There are several factors that can help the channel to provide favorable propagation, such as smaller pathloss exponent, higher antenna density, and spatially separated users with larger distances. Well separated users will generally exhibit favorable propagation since they are mainly communicating with different subsets of the APs.

A main implication of this work is that one should not rely on channel hardening and favorable propagation when computing the achievable rates in CF Massive MIMO, because this could lead to a great underestimation of the achievable performance. There is a good uplink rate expression for spatially correlated Massive MIMO systems that can be used. Further development of downlink rate expressions is needed to fully understand the achievable downlink performance in CF Massive MIMO.

APPENDIX

A. Proof of (5) and (6)

Recall Campbell's theorem as follows [23]. If $f(x) : \mathbb{R}^d \rightarrow [0, +\infty]$ is a measurable function and Φ is a stationary/homogeneous PPP with density λ , then

$$\mathbb{E} \left[\sum_{x \in \Phi} f(x) \right] = \lambda \int_{\mathbb{R}^d} f(x) dx. \quad (52)$$

Since Φ_A is a two-dimensional homogeneous PPP, for a finite network region with radius ρ , we have

$$\begin{aligned} \mathbb{E} [\|\mathbf{g}_k\|^2] &= \mathbb{E} \left[\sum_{i \in \Phi_A} H_i l(r_i) \right] \\ &= \mathbb{E} [H_i] \cdot \lambda_A \int_{\mathcal{B}(0, \rho)} l(\|x\|) dx. \\ &= \lambda_A \cdot \mathbb{E} [H_i] 2\pi \int_0^\rho l(r) r dr. \end{aligned} \quad (53)$$

Note that $l(r) = \min(1, r^{-\alpha})$ and $\mathbb{E} [H_i] = N$ as a result of the Gamma distribution, thus

$$\mathbb{E} [\|\mathbf{g}_k\|^2] = N \lambda_A 2\pi \left(\int_0^1 r dr + \int_1^\rho r^{1-\alpha} dr \right). \quad (54)$$

Depending on the value of α , we have

$$\mathbb{E} [\|\mathbf{g}_k\|^2] = \begin{cases} N \lambda_A \pi \left(1 + \frac{2(1-\rho^{2-\alpha})}{\alpha-2} \right) & \text{if } \alpha \neq 2 \\ N \lambda_A \pi (1 + 2 \ln \rho) & \text{if } \alpha = 2 \end{cases} \quad (55)$$

From [23], we have the expression for the variance

$$\text{Var} \left[\sum_{x \in \Phi} f(x) \right] = \lambda \int_{\mathbb{R}^d} f(x)^2 dx. \quad (56)$$

Since $H_i \sim \text{Gamma}(N, 1)$, we have $\mathbb{E}[H_i^2] = (\mathbb{E}[H_i])^2 + \text{Var}[H_i^2] = N^2 + N$, thus

$$\begin{aligned} \text{Var}[\|\mathbf{g}_k\|^2] &= \lambda_A \cdot \mathbb{E}[H_i^2] 2\pi \int_0^\rho l^2(r) r dr \\ &= (N^2 + N) \lambda_A 2\pi \left(\int_0^1 r dr + \int_1^\infty r^{1-2\alpha} dr \right) \\ &= (N^2 + N) \lambda_A \pi \left(1 + \frac{1}{\alpha - 1} (1 - \rho^{2-2\alpha}) \right). \end{aligned} \quad (57)$$

Note that (57) will have a different form when $\alpha = 1$, which is unlikely to happen in a real propagation environment. Thus, this case is not discussed here.

REFERENCES

- [1] Z. Chen and E. Björnson, "Can we rely on channel hardening in cell-free Massive MIMO?" *accepted in IEEE GLOBECOM Workshops*, 2017.
- [2] S. Shamai and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proc. IEEE VTC-Spring*, vol. 3, 2001, pp. 1745–1749.
- [3] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems," in *Proc. IEEE ACSSC*, 2007, pp. 83–87.
- [4] D. Gesbert, S. Hanly, H. Huang, S. Shamai, 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, 2010.
- [5] A. Osseiran, J. F. Monserrat, and P. Marsch, *5G Mobile and Wireless Communications Technology*. Cambridge University Press, 2016, ch. 9, Coordinated Multi-Point Transmission in 5G.
- [6] E. Björnson, R. Zakhour, D. Gesbert, and B. Ottersten, "Cooperative multicell precoding: Rate region characterization and distributed strategies with instantaneous and statistical CSI," *IEEE Trans. Signal Process.*, vol. 58, no. 8, pp. 4298–4310, 2010.
- [7] E. Björnson, N. Jaldén, M. Bengtsson, and B. Ottersten, "Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6086–6101, 2011.
- [8] H. Q. Ngo, A. E. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free Massive MIMO: Uniformly great service for everyone," in *Proc. IEEE SPAWC*, 2015.
- [9] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and H. Yang, "Cell-free Massive MIMO systems," in *Proc. Asilomar*, 2015.
- [10] 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, 2017.
- [11] 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, 2017.
- [12] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
- [13] H. Q. Ngo and E. Larsson, "No downlink pilots are needed in TDD Massive MIMO," *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2921–2935, May 2017.
- [14] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Aspects of favorable propagation in Massive MIMO," in *European Signal Processing Conference (EUSIPCO)*, September 2014, pp. 76–80.
- [15] H. Huh, G. Caire, H. Papadopoulos, and S. Ramprasad, "Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, 2012.
- [16] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, 2013.
- [17] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, 2013.
- [18] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "How much do downlink pilots improve cell-free Massive MIMO?" in *Proc. IEEE GLOBECOM*, 2016.
- [19] W. Lu and M. Di Renzo, "Stochastic geometry modeling of cellular networks: Analysis, simulation and experimental validation," in *ACM International Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, November 2015, pp. 179–188.
- [20] J. G. Andrews, F. Baccelli, and R. K. Ganti, "A tractable approach to coverage and rate in cellular networks," *IEEE Transactions on Communications*, vol. 59, no. 11, pp. 3122–3134, November 2011.
- [21] M. Haenggi, "The meta distribution of the SIR in Poisson bipolar and cellular networks," *IEEE Transactions on Wireless Communications*, vol. 15, no. 4, pp. 2577–2589, April 2016.
- [22] R. K. Ganti and M. Haenggi, "Interference in ad hoc networks with general motion-invariant node distributions," in *IEEE International Symposium on Information Theory*, July 2008, pp. 1–5.
- [23] M. Haenggi, R. K. Ganti *et al.*, "Interference in large wireless networks," *Foundations and Trends® in Networking*, vol. 3, no. 2, pp. 127–248, 2009.
- [24] M. Haenggi, *Stochastic geometry for wireless networks*. Cambridge University Press, 2012.
- [25] M. Kountouris and N. Pappas, "Approximating the interference distribution in large wireless networks," in *International Symposium on Wireless Communications Systems (ISWCS)*, August 2014, pp. 80–84.
- [26] K. Smaili, T. Kadri, and S. Kadry, "Hypoexponential distribution with different parameters," *Applied mathematics*, vol. 4, no. 04, p. 624, 2013.
- [27] S. V. Amari and R. B. Misra, "Closed-form expressions for distribution of sum of exponential random variables," *IEEE Trans. Rel.*, vol. 46, no. 4, pp. 519–522, 1997.
- [28] E. Björnson, D. Hammarwall, and B. Ottersten, "Exploiting quantized channel norm feedback through conditional statistics in arbitrarily correlated MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4027–4041, 2009.
- [29] A. Goldsmith, *Wireless Communications*. Cambridge University Press, 2005.
- [30] D. Moltchanov, "Distance distributions in random networks," *Ad Hoc Networks*, vol. 10, no. 6, pp. 1146–1166, 2012.
- [31] F. Rusek, D. Persson, B. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, 2013.
- [32] E. Björnson, L. Sanguinetti, and M. Debbah, "Massive MIMO with imperfect channel covariance information," in *Proc. ASILOMAR*, 2016.
- [33] G. Caire, "On the ergodic rate lower bounds with applications to Massive MIMO," *CoRR*, vol. abs/1705.03577, 2017.