

Uplink Spectral Efficiency of Large, Distributed Antenna Systems With MMSE Processing

QINGMU DENG¹ AND SIDDHARTAN GOVINDASAMY^{1,2}, (Member, IEEE)

¹Cerebras Systems Inc., Sunnyvale, CA 94085, USA

²Department of Engineering, Boston College, Chestnut Hill, MA 02467, USA


Corresponding author: Siddhartan Govindasamy (siddhartan.govindasamy@bc.edu)

ABSTRACT The spectral efficiency of a representative uplink in a wireless network with spatially distributed antennas and linear minimum-mean-square-error processing is found to approach an asymptote with a simple form as the number of antennas increases. These results generalize prior work which was applicable to systems with small numbers of base stations with large numbers of antennas each, to systems with large numbers of spatially distributed access points with small numbers of antennas each, and systems between these extremes. Additionally, these results are applied to systems where mobiles have multiple transmit antennas, and are used to characterize systems with both disjoint and user-centric clustering of cooperating base stations. Among other conclusions, these results indicate that with the same density and number of antennas cooperating to serve a mobile user, the uplink spectral efficiency with all antennas randomly distributed is several-fold higher than the case when the antennas are concentrated at one base station. These findings help improve our understanding of the tradeoffs involved in distributed antenna systems which have the potential to significantly increase data rates, but at a higher cost.

INDEX TERMS Cooperative communications, cell free, minimum-mean-square error (MMSE), distributed antennas, user-centric cooperation.

I. INTRODUCTION

Mobile communications systems where antennas serving a given mobile are not co-located at a particular base station (BS) have received significant attention in the literature recently including in Cloud-Radio Access Networks (CRAN), Networked Multiple-Input Multiple-Output (MIMO), and cell-free massive MIMO networks [1]–[13]. With some exceptions such as [5] and [6], most works in the literature do not analytically model the distribution of mobiles and antennas in space, which can provide additional insight into system performance such as the impact of antenna density, and the benefits of distributing antennas in space compared to clustering them at base stations. Of the works directly related to this work, [4] considers maximal-ratio-combining and zero-forcing receivers and [14] considers a networked-MIMO system with zero-forcing receivers and employed a number of approximations on the spatial user model as well as interference distribution to handle the complexity of analyzing networked MIMO systems. On the other hand, linear minimum-mean-square error (MMSE)

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Feng .

processing can provide significant performance increases over zero-forcing and maximal-ratio combining, as it is the optimal linear processor to maximize Signal-to-Interference Ratio (SIR), as observed by [15] in the context of massive MIMO, which considers the spatial distribution of antennas and users via simulations. Additionally, [12] considers cell-free massive MIMO systems using stochastic geometry and analyzes the channel hardening effect when matched-filters are used. [16] and [17] considered the uplink of cooperative base station systems with MMSE processing, and provided a simple analytical expression for the uplink spectral efficiency when a *small* number of cooperative base stations with large numbers of antennas each are used, for Poisson distributed mobiles in [16] and general mobile distributions in [17]. The general framework developed in [16] was applied to two models for base station architectures, namely a user-centric clustering approach, and a hexagonal-cell architecture where the spectral efficiency of a cell-edge mobile was considered, and asymptotic expressions for the spectral efficiency provided for up to six co-operating base stations at hexagonal lattice sites.

In this work, we consider a system similar to [16] and [17]. However, the receiver antennas in our system are arbitrarily



FIGURE 1. Illustration of cooperative base stations and cooperative access points.

distributed within a bounded region and do not have to be clustered at a small number of base stations. In other words, [16] and [17] considered systems depicted on the left of Figure 1. In contrast, this work considers systems with large numbers of antennas, regardless of whether they are located at a small number of base stations (depicted on the left of Figure 1), a small number of antennas per access point (AP) at a large number of APs (depicted on the right of Figure 1), or anywhere in between the two extremes. This difference in assumptions results in a different structure of the random channel vectors and matrices that determine link SIRs in this work as compared to [16], and requires us to use different random matrix techniques to derive the main results. We provide further details on regarding this difference in Section II.

In addition, two further generalizations in this work compared to [16] and [17] are that we allow multiple transmit antennas at each mobile, and we use a slightly more general fading model compared to Rayleigh fading used in [16] and [17]. We also applied the framework developed here to analyze the performance of user-centric clustering, where each mobile is served by its nearest K base stations (sometimes referred to as cell-free systems), antennas at hexagonal lattice sites as well as a disjoint-clustering system, where space is divided into cells and each cell contains K antennas at random positions. In contrast, [16] and [17] consider user-centric clustering and base stations at hexagonal lattice sites, both for small numbers of BSs.

Note that like [14], we assume that the channel state information (CSI) required for the beamformer design is known at the processors. While channel estimation errors do play a significant role in large MIMO systems, our focus here is to characterize the spectral efficiency that can be achieved if the CSI is known. Hence, these results can be used as a bound for systems where CSI errors are significant, or with an appropriate reduction in the spectral efficiencies calculated to account for training time and inaccuracies.

In addition, we make the simplifying assumption that the distributed antennas in this system are sufficiently well synchronized in time as is done in [14], [15], [18] and oth-

ers. While practical systems will require synchronization algorithms such as those described in Section 3.3. of [8], we believe that analyzing systems assuming distributed timing synchronization is helpful to understand the performance that could be expected with a well synchronized system.

Further, modern communication systems for human-to-human communication use efficient error-correcting codes, including capacity-achieving codes such as turbo-codes. As noted in [19] and others, the spectral efficiency (computed using the Shannon formula in this paper) is good approximation for rates achievable in practical modern communications systems. Therefore, we believe that analysis of the spectral efficiency is a useful approximation to the rates achievable in practical systems.

As such, the results in this work apply to distributed antenna systems ranging from systems with a large number of access points each with a small number of antennas (or just one antenna), to systems with a small number of base stations, each with large number of antennas and to systems between these extremes. Thus, the results presented in this work enable us to quantify the performance gains of completely distributing antennas, versus clustering antennas at a small number of base stations. Such a characterization is very helpful for system designers given the significant costs and complexities associated with such systems, and the potentially different costs associated with distributing antennas over a large number of access points versus at a small number of base stations.

In summary, the main contributions of this work are as follows:

- The introduction of an approach for analyzing the spectral efficiency of a system with distributed antennas and optimal linear processing when the number of antennas is large, which is a generalization of prior work which is only applicable to systems with a small number of cooperating base stations, with large numbers of antennas each [17]. In particular, this work enables us to characterize the spectral efficiencies achievable when antennas are completely distributed in space, versus concentrated at a small number of base stations.

- A method to characterize the spectral efficiency of a system with distributed antennas and optimal linear processing when mobiles have multiple transmit antennas.
- An application of these approaches to both disjoint and user-centered clustering of BS/APs which indicates that under these assumptions, for the same total number of antennas, several fold increases in spectral efficiency are possible if antennas are distributed completely randomly versus concentrated at one base station.

II. SYSTEM MODEL

Assume that there are K antennas at arbitrary positions in a planar network, which are connected to a central processor which jointly processes the signals received at the antennas. Assume that the distance of these antennas to the origin is bounded from above by D . At the origin of the network is a mobile transmitter, called the test transmitter with M antennas. The K antennas jointly detect the transmitted signal from this test transmitter. Other interfering mobiles are also located on the plane such that there is no mobile co-incident with an antenna, and with probability 1 (w.p.1), $\forall x \in \mathbb{R}^2$.

$$\lim_{d \rightarrow \infty} \frac{\#B(x, d)}{\pi d^2} = \rho, \tag{1}$$

where $\#B(x, d)$ is the number of mobiles in a radius d disk centered at x . Thus, ρ is the area density of the mobiles.

Note that the results of this work do not depend on the specific model for how the mobiles are distributed in space as long as (1) holds. The reason for this is that when the number of antennas used to detect the signal from the test mobile is large, the MMSE algorithm places deep nulls on the signals from the mobiles that are close to the antennas. Hence, the residual interference at the output of the MMSE receiver is dominated by faraway mobiles. Since signal power decays rapidly with distance according to the inverse power-law model, the specific locations of mobiles far away do not influence the spectral efficiency significantly. Note that this result was already shown for systems with a small number of base stations with large numbers of antennas each in [17]. Here, we prove it for the more general model used here, and verify it by simulations detailed in Section VI.

For ease of exposition, in the subsequent analysis, we shall consider the case where $M = 1$ in all sections except Section IV where we develop the extension to the $M > 1$ case, and Section VI-D where simulation results for $M > 1$ are provided.

The channel vector from the i th mobile to the K Rx antennas is $\mathbf{h}_i \in \mathbb{C}^{K \times 1}$:

$$\mathbf{h}_i = [r_{i,1}^{-\frac{\alpha}{2}} g_{i,1} \quad r_{i,2}^{-\frac{\alpha}{2}} g_{i,2} \quad \dots \quad r_{i,K}^{-\frac{\alpha}{2}} g_{i,K}]^T. \tag{2}$$

Here, $\alpha > 2$ is the path-loss exponent, $g_{i,k}$ are independent and identically distributed (i.i.d.) zero-mean, unit-variance random variables and $r_{i,k}$ is the distance from the i th mobile to the k th Rx antenna. Hence $r_{i,j}^{-\frac{\alpha}{2}} g_{i,j}$ is the flat fading coefficient between the i -th mobile and the j -th antenna.

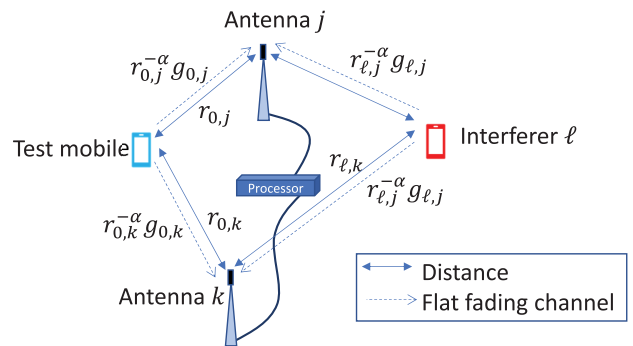


FIGURE 2. System diagram showing the test mobile, an interferer and two antennas. The solid lines represent the distances and the dashed lines represent flat fading. The central processor jointly combines the signals received at the different antennas.

Note that as in [12] and [14], we do not consider shadow fading in this work. Suppose that the distance of the test mobile to the K antennas is bounded from below, i.e. $r_{0,k} > D_{min} > 0$. A system diagram showing the test mobile, with just one interferer and just two antennas is shown in Figure 2, where we have suppressed showing additional interferers and antennas in the interest of clarity.

As in [16] and [17] we assume an interference-limited system and neglect noise. Note that if we do incorporate a constant, non-zero noise level in the analysis, in the asymptotic regime as $K \rightarrow \infty$, the system will become noise limited at a finite K . As such, the effect of interference will be negligible, resulting in a system that is asymptotically identical to an interference-free network. Since this happens at only an extremely large K , the results will not be useful for practical systems where interference is significant. Therefore, we do not incorporate noise into the analysis of the system, since our focus is on systems with significant interference. However, we do incorporate noise in numerical simulations provided in Section VI.

We write the signal received at the K antennas $\mathbf{y} \in \mathbb{C}^K$ as

$$\mathbf{y} = \mathbf{h}_0 x_0 + \sum_{i=1}^{\infty} \mathbf{h}_i x_i, \tag{3}$$

where the test mobile has index 0, and x_i is the zero-mean, unit variance symbol transmitted by the i -th mobile.

Define the channel matrix $\mathbf{H} \in \mathbb{C}^{K \times n}$ between the n mobiles closest to the origin and the antennas serving the test mobile as:

$$\mathbf{H}[n] = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 & \dots & \mathbf{h}_n \end{bmatrix}$$

The signal from the test mobile is estimated using a linear MMSE estimator

$$\mathbf{w}[n] = \left(\mathbf{H}[n] \mathbf{H}^\dagger[n] \right)^{-1} \mathbf{h}_0. \tag{4}$$

The SIR at the output of the MMSE estimator is

$$\text{SIR}[n] = \mathbf{h}_0^\dagger \left(\mathbf{H}[n] \mathbf{H}^\dagger[n] \right)^{-1} \mathbf{h}_0 \tag{5}$$

We assume that $n > K$ so $\mathbf{H}[n]\mathbf{H}^\dagger[n]$ is invertible w.p.1. The limits as $n \rightarrow \infty$ of $\mathbf{w}[n]$ and $\text{SIR}[n]$ are \mathbf{w} and SIR , respectively. Assuming Gaussian codebooks used by all transmitters, the spectral efficiency is defined using the Shannon formula as:

$$\eta = \log_2(1 + \text{SIR}) \quad b/s/Hz \quad (6)$$

Note that the channel vector between mobile i and the K antennas serving the test mobile is given in (2). For systems with a fixed number of base stations and large numbers of antennas which were analyzed in [16], there is a fixed and finite number of distinct $r_{i,k}$ terms. With this restriction, the asymptotic SIR was analyzed in [16] using random matrix techniques developed for Code-Division-Multiple-Access (CDMA) systems with a finite and constant number of antennas in [20]. However, in the case analyzed in this paper, the $r_{i,k}$ terms can all be different and the approach developed in [20] can no longer be used as it relied fundamentally on there being a constant, finite number of distinct $r_{i,k}$ terms. As such, in this work, we used different random matrix approaches including results from [21] and [22] to characterize the SIR in the asymptotic regime assumed here.

For convenience, we have provided a table with the list of the key symbols and notation used in this paper in Tables 1 and 2, respectively.

III. MAIN RESULTS

The SIR of the test link grows with the number of antennas K . We find that a normalized version of the SIR converges in the limit as $K \rightarrow \infty$ as described below.

Theorem 1: The following holds in probability (i.p.).

$$\lim_{K \rightarrow \infty} K^{-\frac{\alpha}{2}} \text{SIR} = \bar{P} \left[\frac{\alpha}{2\pi^2 \rho} \sin\left(\frac{2\pi}{\alpha}\right) \right]^{\frac{\alpha}{2}}. \quad (7)$$

where $\bar{P} = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{i=1}^K r_{0,i}^{-\alpha}$.

Proof: This result is proved using results characterizing the eigenvalue distributions of random matrices whose size go to infinity, for which most techniques involve matrices that go to infinity in both dimensions. On the other hand, since our system involves an infinite number of mobiles, the channel matrices constructed for our system would need to start with an infinite number of columns, making it challenging to directly apply techniques from infinite random matrix theory. To utilize these techniques, we first consider a network with only the n mobiles closest to the origin, which results in the channel matrix $\mathbf{H}[n]$ having K rows and n columns. We then take n and K together to ∞ such that $n/K = c > 2$, which enables us to use approaches from infinite random matrix theory. We will then show that the limit for the normalized SIR when n and K are taken to infinity together, equals the limit of the SIR for an infinite network of mobiles (i.e. $n = \infty$ to start), in the limit as $K \rightarrow \infty$.

To this end, note that the SIR of the test link for the system model with only the first n mobiles is given in (5). A normalized version of $\text{SIR}[n]$ converges according to the

TABLE 1. Table of key symbols and parameters.

Symbol	Definition/Notes
ρ	area density of mobiles
ρ_b	area density of BS/APs
α	path-loss exponent
K	number of antennas serving the test mobile
L	number of antennas per AP/base station serving the test mobile
N	number of BS/APs serving the test mobile
x_i	transmit symbol of i -th mobile
$r_{i,j}$	distance between the i -th mobile and the j -th of the K antennas serving the test mobile
D_{min}	lower bound on $r_{0,j}$
$g_{i,j}$	fading coefficient (excluding path loss) between the i -th mobile and the j -th of the K antennas serving the test mobile
\bar{P}	limit of average path loss between the test mobile and the K antennas serving it
$F_P(\cdot)$	CDF of P
\mathbf{y}	received signal vector
\mathbf{h}_i	channel vector between i -th mobile and K antennas serving the test mobile
$\mathbf{H}[n]$	channel matrix of mobiles $1, 2, \dots, n$ and the K antennas serving the test mobile
$\mathbf{w}[n]$	MMSE weight vector with mobiles $1, 2, \dots, n$
\mathbf{w}	MMSE weight vector with all mobiles
SIR	Signal-to-Interference Ratio of the test mobile uplink
$\text{SIR}[n]$	Signal-to-Interference Ratio with only mobiles $1, 2, \dots, n$
η	spectral efficiency of test link with 1 Tx. antenna
M	number of transmit antennas per mobile
η_M	spectral efficiency of test link with M Tx. antennas
\mathbf{x}_0	vector of transmitted symbols from test mobile
\mathbf{H}_0	channel matrix between test mobile and the K antennas serving it with multiple Tx. antennas
\mathbf{H}_I	channel matrix between interfering mobiles and the K antennas serving the test mobile with multiple Tx. antennas
\mathbf{K}	limit of the interference covariance matrix at the K antennas serving the test mobile with multiple Tx. antennas
R	cluster radius for disjoint clustering of BS/APs
$A_{i,j}$	coefficients used for $F_P(\cdot)$ with user-centered clustering

following lemma which is at the heart of the proof of the main result.

Lemma 1: Set $n = cK$, with $c > 2$.

$$\lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} K^{-\frac{\alpha}{2}} \text{SIR}[n] = \bar{P} \left[\frac{\alpha}{2\pi^2 \rho} \sin\left(\frac{2\pi}{\alpha}\right) \right]^{\frac{\alpha}{2}}. \quad (8)$$

Proof: Please see Appendix A.

The next lemma shows that (8) continues to hold i.p. even if we start with an infinite network (i.e. $n \rightarrow \infty$), and then take $K \rightarrow \infty$. This lemma is very close to Lemma 1 of [16], but has been modified to account for differences in the assumptions in this work.

Lemma 2: The following holds i.p.:

$$\lim_{K \rightarrow \infty} K^{-\frac{\alpha}{2}} \text{SIR} = \lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} K^{-\frac{\alpha}{2}} \text{SIR}[n] \quad (9)$$

TABLE 2. Summary of notation.

Symbol	Definition/Notes
$\#B(x, d)$	number of mobiles in a radius d disk centered at x
$\mathbb{C}^{K \times N}$	set of $K \times N$ matrices with complex entries
\mathbf{a}	a bold, lowercase letter denotes a vector
\mathbf{A}	a bold, uppercase letter denotes a matrix
\mathbf{I}	identity matrix with context-dependent size
\mathbf{A}^\dagger	Hermetian transpose of \mathbf{A}
$I(\cdot; \cdot)$	mutual information
$\lambda_{min}(\mathbf{A})$	smallest eigenvalue of \mathbf{A}
$\lambda_{max}(\mathbf{A})$	largest eigenvalue of \mathbf{A}
$\Gamma(\cdot)$	gamma function.
$\Gamma(\cdot, \cdot)$	upper incomplete gamma function.
$\text{Tr}(\cdot)$	matrix trace
$\text{Rank}(\cdot)$	matrix rank
$\ \cdot\ _F$	Frobenius norm of a matrix
$\text{var}\{\cdot\}$	variance
$E[\cdot]$	expectation operator
i.p.	in probability
w.p.1.	with probability 1
$\phi(i) = O_1(i^k)$	means $\lim_{i \rightarrow \infty} i^{-k} \phi(i) = A$ w.p. 1.

Proof: Please see Appendix B.

Combining Lemmas 1 and 2 completes the proof. ■

Note that \bar{P} , defined in Theorem 1, is the limit of the average path loss from the test mobile to the K antennas used to detect its signal. When compared to prior results from [16] and [17], we see that the limit of the normalized SIR for the distributed antenna system here is equal to the limit of the normalized SIR for a system with a single base station and K antennas, with path loss equal to \bar{P} .

Applying steps used to prove Theorem 2 of [16] to (7), we have the following in probability as $K \rightarrow \infty$

$$\left| \eta - \log_2 \left(1 + \bar{P} \left[\frac{\alpha}{2\pi^2 \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} K^{\frac{\alpha}{2}} \right) \right| \rightarrow 0. \quad (10)$$

Note that in addition to providing the scaling behaviour of the spectral efficiency with K , (10) can be used to approximate the spectral efficiency when the number of antennas is fixed but large. Further, while (10) holds for a fixed \bar{P} , if we treat \bar{P} as a random variable, when K is large, we can approximate the cumulative distribution function (CDF) of the spectral efficiency using the CDF of $P = \frac{1}{K} \sum_{i=1}^K r_{0,i}^{-\alpha}$, which we denote as F_P . Using this approach, by direct substitution, we get the following approximation for the CDF of the spectral efficiency with randomly distributed antennas

$$P(\eta \leq \tau) \approx F_P \left(\frac{(2^\tau - 1) K^{-\frac{\alpha}{2}} \rho^{\frac{\alpha}{2}}}{\left[\frac{\alpha}{2\pi^2} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}}} \right) \quad (11)$$

Hence, (11) can be used to characterize the spectral efficiency of the test link under different system parameters, including the statistical properties of P , which is highly dependent on how the antennas serving the test mobile are distributed in space. We apply this expression with different

choices of F_P in Section V which are then supported by simulations in Section VI.

IV. MULTIPLE TRANSMIT ANTENNAS

Now, assume that each mobile has $M \geq 1$ transmit antennas, and that each mobile transmits an independent data stream from each of its transmit antennas. Suppose that x_0, x_1, \dots, x_{M-1} are the M symbols transmitted by the test mobile at a given sampling time. Similarly x_M, \dots, x_{2M-1} are the symbols transmitted by Transmitter 2, and so on. We shall treat the signals from the interfering mobiles as if they were coming from separate mobiles in the single-transmit-antenna model above. (1) is still satisfied but with the effective density of mobiles of $M\rho$. We can write the following expression for the received signal \mathbf{y} :

$$\mathbf{y} = \mathbf{H}_0 \mathbf{x}_0 + \sum_{i=M}^{\infty} \mathbf{h}_i x_i \quad (12)$$

where $\mathbf{x}_0^T = (x_0, x_1, \dots, x_{M-1})$ is the vector of transmitted symbols from the test mobile. The channel vector between antenna $i - \lfloor i/M \rfloor M$ of mobile $\lfloor i/M \rfloor$ and the K antennas cooperating to detect the signals from the test mobile is given by

$$\mathbf{h}_i = [r_{\lfloor i/M \rfloor, 1}^{-\frac{\alpha}{2}} g_{i,1} \quad r_{\lfloor i/M \rfloor, 2}^{-\frac{\alpha}{2}} g_{i,2} \quad \dots \quad r_{\lfloor i/M \rfloor, K}^{-\frac{\alpha}{2}} g_{i,K}]^T. \quad (13)$$

The matrix \mathbf{H}_0 given below is the channel matrix between the antennas of the test mobile and the K antennas cooperating to detect its signal and is given by

$$\mathbf{H}_0 = [\mathbf{h}_0 \quad \mathbf{h}_1 \quad \dots \quad \mathbf{h}_{M-1}]. \quad (14)$$

Additionally, let's define a channel matrix corresponding to all the interfering signals from the first n mobiles as follows

$$\mathbf{H}_I[n] = [\mathbf{h}_M \quad \mathbf{h}_{M+1} \quad \dots \quad \mathbf{h}_{nM-1}] \quad (15)$$

and the covariance matrix of all the interfering signals as

$$\mathbf{K} = \lim_{n \rightarrow \infty} \mathbf{H}_I[n] \mathbf{H}_I^\dagger[n] \quad (16)$$

We define the spectral efficiency η_M to equal the mutual information between the transmitted and received signals from the test mobile. Note that with Gaussian code-books used at all mobiles, the mutual information corresponds to an achievable rate with an arbitrarily low probability of error. Hence,

$$\eta_M = I(\mathbf{y}; \mathbf{x}_0) = \log_2 \left| \mathbf{I} + \mathbf{H}_0 \mathbf{H}_0^\dagger \mathbf{K}^{-1} \right| \quad (17)$$

With these definitions, we can write the following theorem on the spectral efficiency with multiple transmit antennas.

Theorem 2: With $M \geq$ transmit antennas per mobile, the following holds in probability

$$\lim_{K \rightarrow \infty} \left| \eta_M - M \log_2 \left(1 + \bar{P} \left[\frac{\alpha}{2\pi^2 M \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} K^{\frac{\alpha}{2}} \right) \right| = 0. \quad (18)$$

Proof: Please see Appendix J.

If the locations of the antennas serving the test mobile are random, when the number of antennas is large enough that the spectral efficiency is close to the asymptote given in Theorem 2, the CDF of the spectral efficiency can be approximated by

$$P(\eta \leq \tau) \approx F_P \left(\frac{(2^{\frac{\tau}{M}} - 1) K^{-\frac{\alpha}{2}} (M\rho)^{\frac{\alpha}{2}}}{\left[\frac{\alpha}{2\pi^2} \sin\left(\frac{2\pi}{\alpha}\right) \right]^{\frac{\alpha}{2}}} \right) \quad (19)$$

V. APPLICATIONS

A. COOPERATIVE BASE STATIONS OR DISTRIBUTED ACCESS POINTS

In the results described above, we have not assumed much about how the antennas serving the test mobile are distributed in space, as long as there is a large number of them.

In the subsequent discussion, we shall assume that there are N AP/BSs serving the test user, and each AP/BS has L antennas such that $K = NL$. We consider regimes where L is large and N is small (cooperative BSs with large numbers of antennas per BS), N is large and L is small (large number of APs) and the cases in between the extremes. Here, we assume that the path losses from the antennas of each AP/BS to the test mobile are the same. In other words

$$\begin{aligned} r_{0,1} &= r_{0,2} = \dots = r_{0,L} \\ r_{0,L+1} &= r_{0,L+2} = \dots = r_{0,2L} \\ &\vdots \\ r_{0,K-L+1} &= r_{0,K-L+2} = \dots = r_{0,K} \end{aligned}$$

Then, we have

$$P = \frac{1}{K} \sum_{i=1}^K r_{0,i}^{-\alpha} = \frac{1}{N} \sum_{i=1}^N r_{0,i}^{-\alpha} \quad (20)$$

Note from (10) that with fixed \bar{P} , the spectral efficiency of the test mobile primarily depends on K and not the specific values of N and L . On the other hand, if \bar{P} is not fixed, how the AP/BSs are distributed in space could have different effects on the spectral efficiency. Suppose that the AP/BS locations are deterministic and the test mobile is served by the N closest AP/BSs to it. P , which we use to approximate \bar{P} with large K , is therefore non-decreasing with N , since it is the average of the path losses from the AP/BSs to the test mobile, which can only decrease if we add farther away AP/BSs. Therefore, increasing the number of APs, while keeping L constant, results in diminishing returns.

However if P is a random variable, the CDF of P depends significantly on N , L , and how the AP/BSs serving the test mobiles are distributed in space. For instance for a fixed total number of antennas K and a fixed density of antennas, if the APs are distributed independently in space, using a large number of access points will be more favorable than using a small number of BSs, because the probability of the test mobile being far away from all the antennas serving it will be lower when a large number of APs are used. On the other

hand, the cost of connecting a large number of AP/BSs that are distributed in space can be prohibitive, making it helpful to understand how much performance improvement can be gained by using a large number of access points. To get further insight into these effects, we need to apply specific models of how the AP/BSs are distributed in space.

We adopt two models for how the BS/APs are distributed in space to characterize the spectral efficiency of the test mobile. We consider a disjoint clustering approach, where the AP/BSs are divided into spatially disjoint co-operating clusters as done in [14], and a user-centric cooperation strategy, where the test mobile is served by the N closest APs/BSs to it. The latter system is also known as a cell-free system. By varying the number of AP/BSs N , with a constant total number of antennas K , we can then characterize the performance implications of concentrating a large number of antennas in a small number of BSs and vice-versa under these two models. In the subsequent discussion, we shall assume that the number of antennas K is large enough that the asymptotic expressions hold.

B. DISJOINT CLUSTERING

In disjoint clustering, cooperating AP/BSs are distributed in fixed clusters. For such systems, we assume that the plane is divided into congruent hexagonal cells, and within each cell N AP/BSs are distributed uniformly randomly and independent of each other. Each of the N AP/BSs has L antennas such that the total number of antennas used to detect the test mobile's signal is $K = NL$. The N AP/BSs cooperate to detect the signal of any mobile that falls within their cell. Each cell has a processor which is connected to all the AP/BSs within that cell. This approach is illustrated in Figure 3a. We use two models for the location of the test mobile, a randomly located test mobile and a test mobile at the cell edge. In the randomly distributed mobile case, the test mobile is assumed to be distributed uniformly randomly in its cell, independent of the locations of the APs/BSs in its cell. Note that while technically we require a minimum distance between the test mobile and the antennas serving it, we do not impose this restriction in the derivations of the CDF of P in order to simplify the resulting expressions. Note that situations where the test mobile gets very close to an antenna will result in large spectral efficiencies and are not in the low-outage-probability regime which is the regime of interest. Hence, not enforcing a minimum distance in our numerical evaluations does not influence the spectral efficiencies in the regimes of interest.

To make the analysis tractable, following [14], we approximate the hexagonal cells using circles with a radius R , such that the area of the circle equals that of a hexagonal cell. The resulting CDF for the random variable P is given by the following Lemma

Lemma 3:

$$F_P(x) = \Pr(P \leq x) = \int_0^R \left[\frac{1}{2} - \frac{1}{\pi} \int_0^\infty \frac{1}{s} \right]$$

$$\times \int_0^\infty \left\{ e^{-jsNx} \left(\int_0^\infty e^{jsr-\alpha} f_{r|d}(r|d) dr \right)^N \right\} ds \left] \frac{2d}{R^2} dd, \tag{21}$$

with

$$f_{r|d}(r) = \begin{cases} \frac{2r}{\pi R^2} \arccos\left(\frac{d^2 + r^2 - R^2}{2dr}\right) & \text{if } d + r > R \text{ and} \\ & d - r < R, \\ \frac{2r}{R^2} & \text{if } d + r \leq R, \\ 0 & \text{otherwise.} \end{cases} \tag{22}$$

Proof: Please see Appendix C.

This expression is used with (11) to approximate the CDF of the spectral efficiency for the disjoint clustering system.

When the test mobile is at the cell edge, the CDF of P is given by the following, which follows directly from the proof of Lemma 3.

$$F_P(x) = \Pr(P \leq x) = \frac{1}{2} - \frac{1}{\pi} \times \int_0^\infty \frac{1}{s} \int_0^\infty \left\{ e^{-jsNx} \left(\int_0^\infty e^{jsr-\alpha} f_{r|d}(r|R) dr \right)^N \right\} ds \tag{23}$$

Numerical results, including simulations for this case are provided in Section VI.

C. USER-CENTRIC CLUSTERING

With user-centric clustering, a mobile is served by a set of BS/APs which have favourable channels to it (e.g. the closest by Euclidian distance). Such a system is depicted in Figure 3b, where the APs illustrated in bold serve the mobile. Note that in this setup, APs need to be connected to several processors (or all APs need to be connected). In the example depicted in Figure 3b, the dashed bold lines indicate links that would need to be added to the system (as compared to a disjoint-clustering system) in order to serve the mobile shown. As such, user-centric clustering typically results in larger infrastructure and overhead requirements. On the other hand, with user-centric clustering, there is no notion of a cell-edge, which is the worst-case location for a mobile. Note that if we model the APs in the user-centric scenario as being randomly distributed, the performance of a given link can be quite different depending on whether we use a large number of APs with small numbers of antennas vs. a small number of BSs with large numbers of antennas each. The reason for this is that the probability of being far away from any antenna is relatively smaller in the former case, compared to the latter case.

A commonly used model (e.g. [14]) is to assume that the access points/base stations form a homogeneous Poisson Point Process (HPPP) on the plane. Here we adopt the HPPP model with density ρ_b BS/unit area. The test mobile then connects to the N closest BS/APs to it, where $K = LN$. With

this model, the CDF of P is given by the following lemma proved in [16].

Lemma 4: For user-centric clustering with BS/APs distributed as a HPPP with density ρ_b , the CDF of P is

$$\begin{aligned} F_P(x) &= \Pr(P \leq x) \Pr\left(\frac{1}{N} \sum_{i=1}^N r_{0,i}^{-\alpha} \leq x\right) \\ &= e^{-\rho_b \pi \left(\frac{1}{x}\right)^{\frac{2}{\alpha}}} \sum_{\ell=0}^{\infty} \frac{\left(\rho_b \pi \left(\frac{1}{x}\right)^{\frac{2}{\alpha}}\right)^\ell}{\ell!} \\ &\quad + \frac{1}{\pi} \sum_{\ell=1}^N \binom{N}{\ell} \left(-\Gamma\left(1 - \frac{2}{\alpha}\right)\right)^\ell \sin\left(\frac{2\pi\ell}{\alpha}\right) \\ &\quad \times \sum_{m=0}^{\infty} A_{N-\ell,m} \Gamma\left(m + \frac{2}{\alpha}\right) \\ &\quad \times (xN)^{-m - \frac{2}{\alpha}\ell} \frac{1}{N!} (\rho_b \pi)^{\frac{1}{2}m\alpha + \ell} \\ &\quad \times \Gamma\left(N - \frac{m\alpha}{2} - \ell + 1, \left(\frac{1}{x}\right)^{\frac{2}{\alpha}} \rho_b \pi\right) \end{aligned} \tag{24}$$

The coefficients $A_{i,j}$ are defined recursively as follows

$$A_{i,j} = \frac{1}{j} \sum_{\ell=1}^j (\ell(i+1) - j) \left(\frac{2}{\ell!(2 - \ell\alpha)}\right) A_{i,j-\ell} \tag{25}$$

with $A_{i,0} = 1$ for $0 < i \leq N$ and $\Gamma(\cdot, \cdot)$ is the upper incomplete gamma function. $N = K/L$, is the number of access points or base stations.

Assuming that the asymptotic approximation holds, we can approximate the CDF of the spectral efficiency combining (24) and (11). Note that while (24) is complicated, it can be evaluated numerically efficiently, as described in the following section.

VI. NUMERICAL RESULTS

A. FIXED ANTENNAS

To verify the accuracy of the analytical results, we simulated a system with AP/BSs at fixed locations serving a test mobile at the origin. We considered a range of values of K and N , and 1 antenna per AP/BS. Note that low numbers of antennas per base station is the regime of interest as systems with small numbers of base stations, with large numbers of antennas each have already been analyzed in [16]. We used a relative density of antennas to users of 20.

Further, in order to verify that as long as (1) holds, the specific distribution model of the mobiles does not impact the spectral efficiencies when the numbers of antennas are large, we consider three different models of how the interfering mobiles are distributed. As a baseline, we consider the HPPP model where the locations of the mobiles are completely independent of one another. To simulate mobile distributions where mobiles tend to form clusters, we use the Matern Cluster Process (MCP) with 10 mobiles per cluster (note that the MCP is a special case of the Neyman-Scott Process [23]).

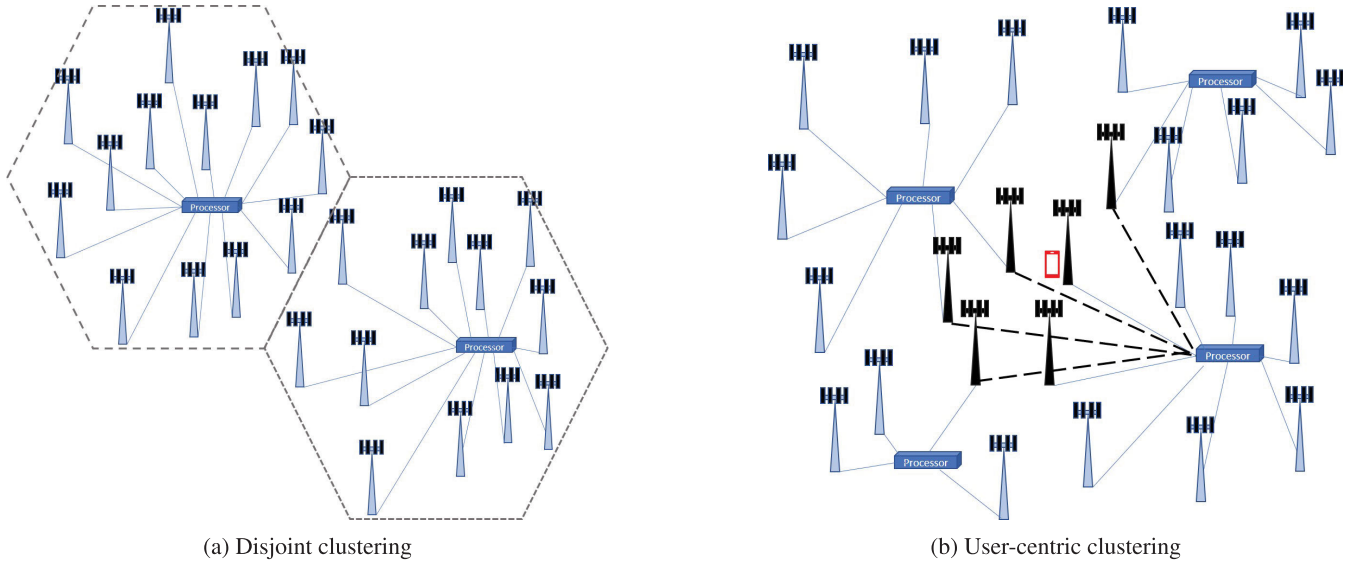


FIGURE 3. Illustration of disjoint clustering (left) vs. user-centric (right) architectures. In the disjoint clustering case, the access points within each cluster are all connected to a central processor. In the user-centric case, the mobile is served by base stations/access points close to it (with the bold illustration). Hence, some access points will have to be connected to multiple processors, or all access points have to be connected to a single central processor. The additional fronthaul links required to serve the mobile in the user-centric clustering case as compared to the disjoint clustering approach are illustrated using the bolded dashed lines in (b).

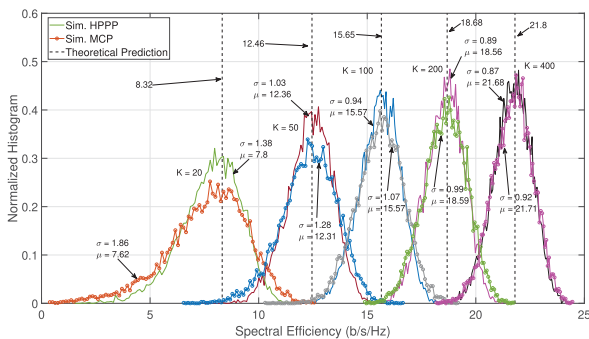


FIGURE 4. Normalized histogram of the spectral efficiency for APs in a hexagonal grid around the test mobile with one antenna per AP. Histograms for mobiles distributed as HPPPs and MCPs are shown.

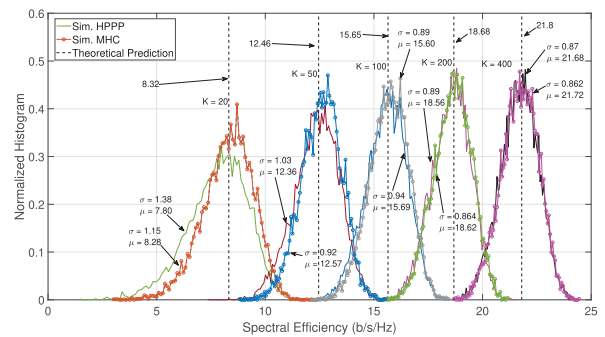


FIGURE 5. Normalized histogram of the spectral efficiency for APs in a hexagonal grid around the test mobile with one antenna per AP. Histograms for mobiles distributed as HPPPs and MHC processes are shown.

We also consider a repulsive model for mobile distribution, namely the Matern type-I hard-core (MHC) process with a repulsion distance of 10 units. In all cases, we set the effective density of mobiles to 10^{-3} mobiles per unit area. Note that the HPPP, MCP and MHC models are commonly used in the literature on stochastic geometry for wireless communications (see e.g. [24]). The specific definitions of these processes can be found in references such as [23].

Figure 4 shows a normalized histogram of the spectral efficiency for different numbers of antennas with one antenna per access point. The lines without the markers represent simulations using HPPP mobiles, and the lines with the markers represent the MCP model for the mobiles. Note that the reduction in the standard deviation of the spectral efficiency indicates that the spectral efficiency converges in mean-square, and hence in probability, which is the assertion

of the main results of this work. Additionally, for the HPPP case, we observe that while the convergence is not very fast, even with 50 antennas, the standard deviation of the spectral efficiency is below 10% of the asymptotic spectral efficiency, and with 200 antennas, the standard deviation of the spectral efficiency is below 5% of the asymptotic spectral efficiency. Note here that while this convergence is not fast in this case where P is deterministic, when P is random due to the random relative positions of the antennas to the test mobile, the approximation to the CDF of the spectral efficiency given in (11) is quite accurate for small numbers of antennas as shown in the remainder of this section. The reason for this observation is that for even moderately large numbers of antennas, the primary source of variation that impacts the spectral efficiency is the positions of the antennas relative to the test mobile. Figure 5 shows a normalized histogram of the

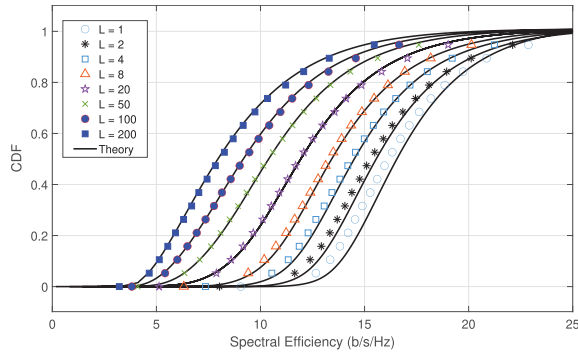


FIGURE 6. CDF of the spectral efficiency for $K = 200$ and different L s, for disjoint clustering with cells approximated as circles. The access points/base stations were distributed uniformly randomly within each circular cell. The test mobile is assumed to be distributed uniformly randomly in a circular cell.

spectral efficiency with lines without markers representing the HPPP model, and the lines with the markers representing the MHC model for the mobiles. In both Figures 4 and 5, the distribution of the spectral efficiencies of the MCP and MHC models approach that of the HPPP model, indicating that the specific model of mobile distribution is not significant when the number of antennas is large. Additionally, we expect, the MCP model (i.e. with clustering of mobiles) has a larger standard deviation than the HPPP model. Similarly, for the MHC model (where mobiles are more spaced out and hence more spatially regular), the standard deviation of the spectral efficiency is lower as compared to the HPPP model.

Figures 4 and 5 also provide some insight into how the accuracy of the asymptotic approximation changes with increasing K , with the only source of randomness (aside from fading) arising from the mobile distributions. The standard deviation for the HPPP case decreases from 1.38 to 0.87 as the number of antennas K increases from 20 to 400. Additionally, since the asymptotic spectral efficiencies increase with K , the ratio of the standard deviation to the asymptote plays a significant role as well. This quantity varies from 0.17 to 0.05 as the number of antennas increases from 20 to 400. For the MCP, these values range from 0.22 to 0.11, and for the MHC process, they range from 0.14 to 0.04. As such, the asymptotic approximation to the spectral efficiency is moderately accurate in predicting the spectral efficiency when the locations of the antennas serving the test mobile are fixed. On the other hand, as noted above and discussed subsequently, this result is still very useful in predicting the spectral efficiency when the antennas serving the test mobile are randomly located, since the ratios of the standard deviation to the asymptotic spectral efficiencies are small. The mean spectral efficiencies on the other hand are quite well predicted by the asymptotic spectral efficiency, with the deviation between the mean and asymptotic spectral efficiency being less than 1 b/s/Hz with 50 antennas and above, going down to less than 0.1 b/s/Hz at 400 antennas for all three models of mobile distributions.

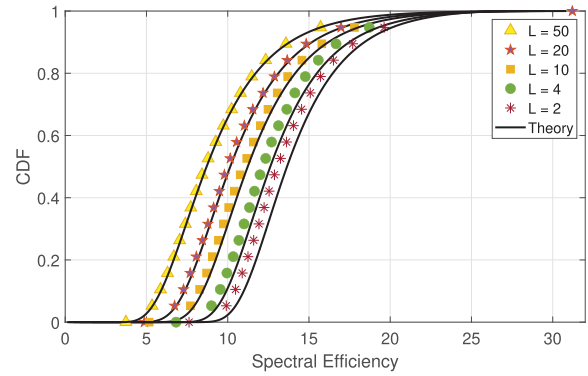


FIGURE 7. CDF of the spectral efficiency for $K = 200$ and different N s, for a cell-edge user, and disjoint clustering with cells approximated as circles. The access points/base stations were distributed uniformly randomly within each circular cell. Relative antenna to mobile density of 20:1 was used.

B. DISJOINT CLUSTERING

We simulated the disjoint clustering scenario described in Section V-B where we placed N AP/BS independently and uniformly randomly in a circular cell, serving a test mobile. We varied the number of AP/BS, N , and the number of antennas per BS, L , such that the total number of antennas serving the test mobile was constant at $K = NL = 200$. The test mobile was either distributed uniformly randomly in the cell or placed at the cell edge with the radius of the cell set so that the antenna density is 20 times that of the mobile density. 4000 interferers were then placed randomly in the plane with density 0.001 interferers per unit area, in order to simulate Poisson distributed interferers. A path-loss exponent of $\alpha = 4$. Single transmit antennas per mobile were used in these simulations.

Figure 6 shows the simulated spectral efficiencies (with markers) and the theoretical spectral efficiencies (with solid lines) obtained from evaluating (11) together with Lemma 3, where we used a noise power of -96dBm and transmit power of 100mW. From the graphs, it is clear that the theoretical predictions match the simulations well when the number of antennas per AP/BS is 4 or larger, but even when the number of antennas per BS $L = 2$ or $L = 1$, the graphs still predict the spectral efficiency moderately well. At an outage probability of approximately 0.1, the theoretical predictions match the simulations to within 0.4 b/s/Hz when $L = 2$ and within 0.6 b/s/Hz when $L = 1$.

For a fixed total number of antennas, the convergence of the spectral efficiency CDF to its asymptote is slow when the number of antennas per AP/BS is low. The reason for this difference is that the rate of convergence of the SIR (and hence, the spectral efficiency) is dependent on the convergence of the empirical distribution function of the path losses between the interfering mobiles and the antennas serving the test mobile. If the antennas are all co-located, the convergence of the empirical distribution function of the path-losses is faster with K as compared to the case when the antennas serving the test mobile are distributed randomly in space.

Hence, the convergence of the spectral efficiency when the number of antennas per AP/BS is low is slower with K . The portions of the derivation of Theorem 1 that directly impact this appear in Appendix E and with a note of explanation at the end of that section.

These results indicate that the asymptotic theoretical results are useful approximations to the spectral efficiency for reasonable (in the massive MIMO regime) system parameters. Further, observe that distributing all antennas randomly within the cell results in approximately 2.5 times the spectral efficiency compared to concentrating all 200 antennas at a single base station, at an outage probability of 0.1.

Figure 7 depicts the spectral efficiencies of users at a cell edge of a cell with radius such that the ratio of antenna to mobile density is 20. We used $K = 200$ antennas, while varying N and L such that $K = NL$, as indicated in the legend. The theoretical predictions match simulations between 0.1 b/s/Hz at $L = 50$ to 0.6 b/s/Hz at $L = 2$. The convergence is slower for the test mobile at the cell-edge than a randomly located mobile in the cell because at the cell edge, there is more variation in the distances (and hence path losses) between the test mobile and the antennas serving it as compared to a mobile randomly located in the cell.

C. USER-CENTRIC CLUSTERS

As described in Section V, in the user-centric scenario, the test mobile is served by the BS/APs closest to it. For this scenario, we conducted simulations where a test mobile was placed at the origin of the network and the K/L closest access points to it were used to detect the signal from the test mobile. The AP/BSs were randomly distributed according to a Poisson Point Process with density of antennas equal to 20 times that of the mobiles. Note that this value corresponds to an access point density of 0.1 times the mobile density, if the 200 antennas were located at a single access point. The path-loss exponent, $\alpha = 4$. Single transmit antennas per mobile were used in these simulations.

Figure 8 shows the simulated and theoretical CDF of the spectral efficiencies for $K = 100$ antennas total, and $L = 1, 4, 20, 100$ antennas per access point, noise power of -96dBm and transmit signal power of 100mW. The simulated CDFs (illustrated by the markers) were generated from 10000 trials of a circular network with 10000 mobiles. The theoretical prediction using (24) and (11) is plotted with the solid lines. Note that the asymptotic prediction matches the simulations well (within 0.3 b/s/Hz at an outage probability of 0.1), confirming the accuracy of the asymptotic analysis.

The results indicate that at an outage probability of 0.1, placing all antennas at a single base station has approximately 3.4 times lower spectral efficiency compared to completely randomly distributing the antennas. The decrease in spectral efficiency that results from concentrating all antennas at a single base station is due to the fact that the probability of the test mobile being far away from all antennas is higher when they are concentrated at a single base station ($L = 100$) as compared to distributing them in space ($L = 1$).

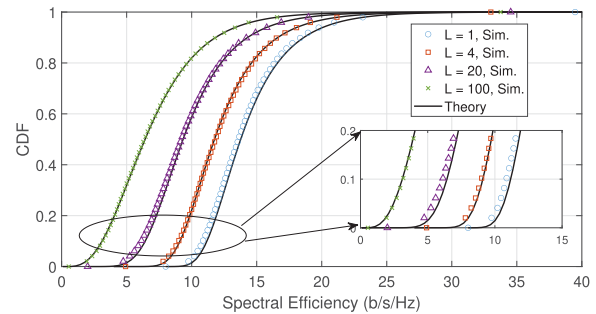


FIGURE 8. CDF of the spectral efficiency for $K = 100$ and different L s, and user-centric clustering using $N = K/L$ nearest access points to serve the test mobile.

Figure 9 shows the simulated and theoretical CDF of the spectral efficiencies for $K = 200$ antennas and different numbers of antennas per AP/BS. We used $L = 1, 4, 20, 100$ and 200 antennas per base station. We used a noise power of -96dBm and a transmit power of 100mW. The theoretical prediction using (24), and (11) is plotted with the solid lines, and the markers represent simulation results. From the graph, we see that the theoretical prediction matches the simulation well (within 0.1 b/s/Hz at an outage probability of 0.1) as illustrated in the figure, which confirms the accuracy of the asymptotic analysis. The results indicate that for $K = 200$ antennas at an outage probability of 0.1, placing all antennas at a single base station ($L = 200$) has approximately 2.8 times lower spectral efficiency compared to completely randomly distributing the antennas ($L = 1$).

Additionally, to verify that the approach introduced here works for systems where mobiles have different distributions, we plotted the CDF of the simulated spectral efficiencies and theoretical predictions for the system with mobiles distributed as MCPs and MHC processes in Figure 10. The parameters for the spatial distribution of antennas are the same as that described in the first paragraph of this section and the parameters of the MCP and MHC processes used to for the mobiles are as defined in Section VI-A. In both cases, we picked the mobile closest to the center of the network as the test mobile, and translated the origin to be located at the test mobile. At an outage probability of 0.1, the simulations for the MCP (depicted using the 'x' marker) and the simulations for the MHC (depicted using the 'o' marker) differ by at most 0.2 b/s/Hz from the theoretical predictions in the simulations we conducted. Additionally, we note that as expected, for low outage probabilities, the spectral efficiencies for the MCP mobiles are slightly lower than that of the MHC process because the MCP involves clustering of mobiles as compared to the MHC process where mobiles are more spaced apart which reduces the likelihood of other mobiles appearing close to the antennas serving the test mobile (which are close to the test mobile). These results verify the utility of the proposed framework for a variety of different modeling assumptions on the mobile distributions.

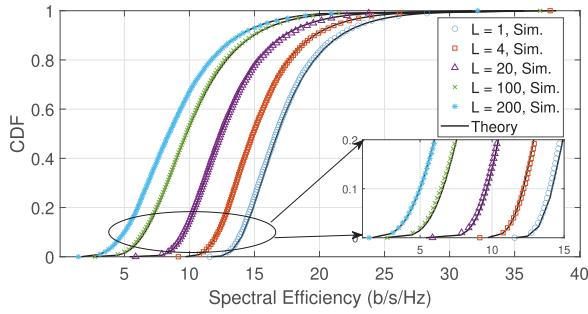


FIGURE 9. CDF of the spectral efficiency for $K = 200$ and different L s, and user-centric clustering using $N = K/L$ nearest access points to serve the test mobile.

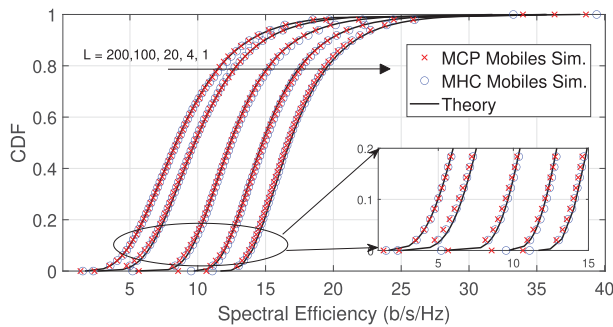


FIGURE 10. CDF of the spectral efficiency for $K = 200$ and different L s, and user-centric clustering using $N = K/L$ nearest access points to serve the test mobile. The mobiles were distributed either as a MCP (data plotted with an x) or a MHC process (data plotted with a circle). The graphs show results from systems with $L = 200, 100, 20, 4$ and 1 , from left to right.

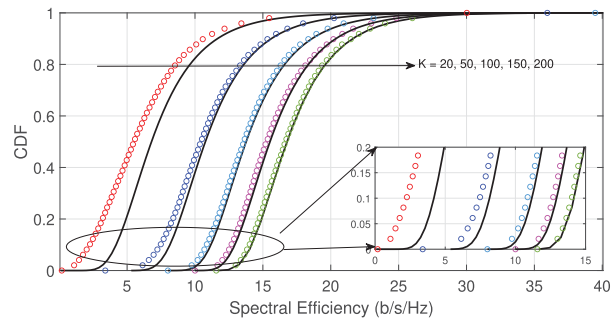


FIGURE 11. CDF of the spectral efficiency for different numbers of antennas K $L = 1$, and user-centric clustering using the $N = K/L$ nearest access points to serve the test mobile. The mobiles were distributed as a HPPP. The graphs show results from systems with $K = 20, 50, 100, 150$ and 200 , from left to right.

To illustrate how the accuracy of the asymptotic predictions changes as K increases, we plotted the simulated spectral efficiency of systems with user-centric clustering and different numbers of antennas K , using one antenna per AP/BS in all cases. We used noise level of -96 dBm and a transmit power of 100 mW for these simulations which are illustrated in Figure 11. At one extreme, with $K = 20$, the simulated spectral efficiency and its asymptotic prediction differ by 2 b/s/Hz at an outage probability of 0.1 . This discrepancy diminishes to 0.1 b/s/Hz for $K = 200$.

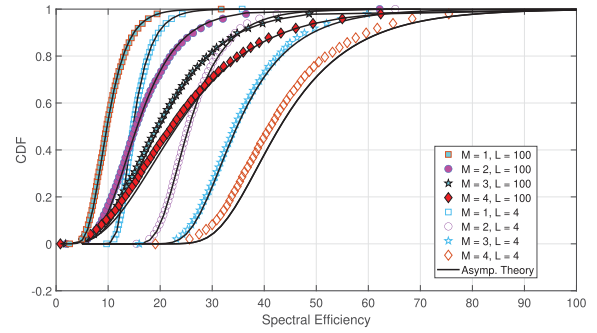


FIGURE 12. Simulated and theoretical CDF of the spectral efficiencies for $K = 200$ receive antennas and different numbers of transmit antennas per mobile, M s. Results for $L = 4$ and $L = 50$ antennas per access point/BS are shown. We used user-centric clustering with the $N = K/L$ closest AP/BSs serving a test mobile.

D. MULTIPLE TRANSMIT ANTENNAS

We additionally simulated systems where each mobile has multiple transmit antennas. We considered systems with user-centric clustering, and Poisson distributed AP/BSs as done for the single-transmit-antenna per mobile system described in Section VI-C. We consider $M = 1, 2, 3, 4$ transmit antennas per mobile. We use low numbers since it is unlikely that mobile devices will have large numbers of antennas. We illustrate two cases, when APs have 4 antennas and when BSs have 100 antennas, with the total number of antennas equaling 200 in all cases in Figure 12. The simulations agree well with the theoretical predictions from (19) for $M = 1, 2, 3$ transmit antennas per mobile, but less so for $M = 4$. Even for $L = 4$ and $M = 4$, the theoretical predictions match the simulations within 2 b/s/Hz at an outage probability of approximately 0.1 , indicating that the theoretical asymptotic predictions are moderately useful in predicting the spectral efficiencies when the mobiles have a small number of transmit antennas.

We further note the diminishing benefit of increasing the number of transmit antennas per mobile since the multiplexing gain obtained by using multiple transmit antennas at the test mobile is offset by the increased burden on suppressing interference since each multiantenna interferer appears as multiple single antenna interferers because we have assumed that independent transmissions are used per transmit antenna. For example, in the 4 antennas per access point case, going from one to 4 transmit antennas per mobile increases spectral efficiency by approximately a factor of 3.5 at an outage probability of 0.1 . However, when the antennas are concentrated at a few base stations, e.g. at 2 base stations with 100 antennas each, the spectral efficiency with 3 or 4 transmit antennas are both approximately 1.5 times that of just using a single antenna. The reason for the smaller increase in spectral efficiency in this case is because even with the same total number of antennas, when the antennas are concentrated at a few base stations, the probability of the test mobile being far away from all antennas is larger compared to the case where there are a large number of access points with a few antennas each,

distributed randomly in space. Consequently, the probability of the test mobile being in the low SIR regime is higher when antennas are concentrated at a few BSs. When the SIR is low, multiplexing gain with multiple transmit antennas is known to be reduced.

VII. SUMMARY AND CONCLUSION

We have presented an asymptotic technique to analyze the uplink spectral efficiency of a distributed antenna system with linear MMSE processing and a large number of antennas. The system model is general enough that it applies to systems with a large number of single-antenna access points, to systems with a small number of base stations, each with a large number of antennas, and systems between those extremes. Hence, the results of this work generalize the findings in [17] which only apply to the latter case. We applied the approach developed here to a user-centric distributed antenna system with Poisson distributed access points and mobiles, where a mobile is served by the 200 closest antennas to it. We found that with a fixed antenna density, a factor of about 2.8 improvement in spectral efficiency is achieved when the antennas are fully distributed (1 antenna/access point) as compared to 200 antennas all placed at a single base station. Similar increases were also found in the case of disjoint clustering of antennas, where space is divided into cells with a fixed number of antennas distributed independently and uniformly randomly in each cell.

Hence, the analysis presented here can be used by system designers to trade off between placing fewer number of antennas on a larger number of access points, vs. placing larger number of antennas in fewer access points. Given the potentially different cost implications associated with different allocations of antennas and access points, and the overall high cost associated with distributed antenna systems in general, we expect that the approach presented here to be helpful for system designers to trade off cost and complexity vs. spectral efficiency in distributed antenna systems.

APPENDIX A

PROOF OF LEMMA 1

Define $\mathbf{D} = \text{diag}(r_{0,1}^{-\alpha}, r_{0,2}^{-\alpha}, \dots, r_{0,K}^{-\alpha})$ and $\mathbf{g}_0 = [g_{0,1} \ g_{0,2} \ \dots \ g_{0,K}]^T$. We can then write

$$K^{-\frac{\alpha}{2}} \text{SIR}[n] = \frac{1}{K} \mathbf{g}_0^\dagger \bar{\mathbf{R}}^{-1} \mathbf{g}_0 \quad (26)$$

where $\bar{\mathbf{R}} = \mathbf{D}^{-\frac{1}{2}} \left(K^{-\frac{\alpha}{2}-1} \mathbf{H}[n] \mathbf{H}^\dagger[n] \right) \mathbf{D}^{-\frac{1}{2}}$. Let λ_k be the k -th largest eigenvalue of $\bar{\mathbf{R}}$.

The smallest of the λ_k s, is bounded from below for large K as follows.

Lemma 5: $\exists K_0$ such that $\forall K > K_0$, and some $\lambda_{\ell b} > 0$, w.p.1..

$$\min \lambda_k > \lambda_{\ell b}. \quad (27)$$

Proof: Please see Appendix D.

Next, consider the following

$$\begin{aligned} & \lim_{K \rightarrow \infty} \Pr \left\{ \left| \frac{1}{K} \mathbf{g}_0^\dagger \bar{\mathbf{R}}^{-1} \mathbf{g}_0 - \frac{1}{K} \text{Tr}(\bar{\mathbf{R}}^{-1}) \right| > \epsilon \right\} \\ &= \lim_{K \rightarrow \infty} E_{\bar{\mathbf{R}}} \left[\Pr \left\{ \left| \frac{1}{K} \mathbf{g}_0^\dagger \bar{\mathbf{R}}^{-1} \mathbf{g}_0 - \frac{1}{K} \text{Tr}(\bar{\mathbf{R}}^{-1}) \right| > \epsilon \mid \bar{\mathbf{R}} \right\} \right] \\ &= E_{\bar{\mathbf{R}}} \left[\lim_{K \rightarrow \infty} \Pr \left\{ \left| \frac{1}{K} \mathbf{g}_0^\dagger \bar{\mathbf{R}}^{-1} \mathbf{g}_0 - \frac{1}{K} \text{Tr}(\bar{\mathbf{R}}^{-1}) \right| > \epsilon \mid \bar{\mathbf{R}} \right\} \right] \\ &= 0 \end{aligned} \quad (28)$$

where the limit is taken into the expectation by the Dominated Convergence Theorem. The limit in (28) goes to zero from Theorem 3.4 of [21] and the fact that convergence w.p.1. implies convergence i.p..

Hence, from (26) and (28) the following holds i.p.

$$\begin{aligned} \lim_{K \rightarrow \infty} K^{-\frac{\alpha}{2}} \text{SIR}[n] &= \lim_{K \rightarrow \infty} \frac{1}{K} \text{Tr}(\bar{\mathbf{R}}^{-1}) \\ &= \lim_{K \rightarrow \infty} \int_{-\infty}^{\infty} \frac{dF_K(\tau)}{\tau}, \end{aligned} \quad (29)$$

where the empirical distribution function (e.d.f.) of the eigenvalues of $\bar{\mathbf{R}}$ is $F_K(\tau)$, so $dF_K(\tau)$ is a measure with a mass of $\frac{1}{K}$ at each eigenvalue of $\bar{\mathbf{R}}$.

The following lemma shows that (29) converges to the desired form as $K \rightarrow \infty$, completing the proof.

Lemma 6:

$$\begin{aligned} \lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} \text{Tr}(\bar{\mathbf{R}}^{-1}) &= \lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} \int_{-\infty}^{\infty} \frac{dF_K(\tau)}{\tau} \\ &= \bar{P} \left[\frac{\alpha}{2\pi^2 \rho} \sin\left(\frac{2\pi}{\alpha}\right) \right]^{\frac{\alpha}{2}}. \end{aligned} \quad (30)$$

Proof: Please see Appendix E.

Note that the specific form of the limit in (65) originates from the convergence of the e.d.f. of appropriately scaled path-losses from the interferers, which was originally derived for systems with small number of base stations in [17], as described in Appendix G.

APPENDIX B

PROOF OF LEMMA 2

We use a sandwiching argument here, with the general approach similar to the proof of Lemma 1 from [16], but with details differing due to the more general system assumptions in this work. Let $\underline{\text{SIR}}[n]$ be the SIR when the MMSE estimator constructed for a network with just the first n interferers is applied to the network with all interferers. Recall that $\text{SIR}[n]$ is the SIR of a system with just the first n interferers in it. Hence,

$$\underline{\text{SIR}}[n] \leq \text{SIR} \leq \text{SIR}[n] \quad (31)$$

We have

$$\underline{\text{SIR}}[n] = \frac{|\beta \mathbf{w}^\dagger[n] \mathbf{h}_0|^2}{\beta^2 \mathbf{w}^\dagger[n] \left(\mathbf{H}[n] \mathbf{H}^\dagger[n] + \sum_{m=n+1}^{\infty} \mathbf{h}_m \mathbf{h}_m^\dagger \right) \mathbf{w}[n]} \quad (32)$$

where β is a scale factor on the weight vector which does not impact the SIR as it scales the signal and the interference by the same value. From (4) and (5),

$$\begin{aligned} \underline{\text{SIR}}[n] &= \frac{\text{SIR}[n]}{1 + \frac{1}{\underline{\text{SIR}}[n]} \mathbf{w}^\dagger[n] \left(\sum_{m=n+1}^{\infty} \mathbf{h}_m \mathbf{h}_m^\dagger \right) \mathbf{w}[n]} \\ &= \frac{\text{SIR}[n]}{1 + \frac{1}{\underline{\text{SIR}}[n]} \sum_{m=n+1}^{\infty} |\mathbf{w}^\dagger[n] \mathbf{h}_m|^2} \end{aligned} \quad (33)$$

Consider the limits as $K \rightarrow \infty$, and $c \rightarrow \infty$ of the summation term in the denominator of the previous expression. We will subsequently show that this term goes to zero in probability.

$$\lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} \frac{1}{\underline{\text{SIR}}[n]} \sum_{m=n+1}^{\infty} |\mathbf{w}^\dagger[n] \mathbf{h}_m|^2. \quad (34)$$

The following lemma helps simplify the summation term in (34), in the limit.

Lemma 7: Consider an $L \times L$, non-negative definite, Hermitian matrix \mathbf{B}_L , whose minimum eigenvalue is bounded from below by $\bar{\lambda} > 0$ with probability 1. Suppose that $\mathbf{a} \in \mathbb{C}^{L \times 1}$ and $\mathbf{b} \in \mathbb{C}^{L \times 1}$ comprise zero-mean, uncorrelated random variables with variances $\sigma_{a_1}^2, \dots, \sigma_{a_L}^2$ and $\sigma_{b_1}^2, \dots, \sigma_{b_L}^2$ respectively which are bounded $\forall L$. Then, for some $L_0 > 0$, for $L > L_0$, with probability 1

$$E \left[\left| \frac{1}{L} \mathbf{a}^\dagger \mathbf{B}^{-1} \mathbf{b} \right|^2 \right] \leq \frac{1}{\bar{\lambda}^2 L} \max\{\sigma_{a_\ell}^2\} \max\{\sigma_{b_\ell}^2\} \quad (35)$$

where the expectation is with respect to \mathbf{a} and \mathbf{b} . Hence, the following holds in probability

$$\lim_{L \rightarrow \infty} \frac{1}{L} \mathbf{a}^\dagger \mathbf{B}^{-1} \mathbf{b} = 0. \quad (36)$$

Proof: Please see Appendix F.

Consider the expectation of the summation term in (34) scaled by $K^{-\frac{\alpha}{2}}$.

$$\begin{aligned} E \left[\left| K^{-\frac{\alpha}{2}} \sum_{m=n+1}^{\infty} |\mathbf{w}^\dagger[n] \mathbf{h}_m|^2 \right| \right] \\ = \frac{1}{K} \sum_{m=n+1}^{\infty} E \left[|\mathbf{g}_0^\dagger[n] \left(K^{\frac{\alpha}{2}-1} \bar{\mathbf{R}} \right)^{-1} \mathbf{q}_m|^2 \right] \end{aligned} \quad (37)$$

where the expectation goes into the sum due to positivity of the summands, and

$$\begin{aligned} \mathbf{q}_m &= \mathbf{D}^{-\frac{1}{2}} \mathbf{h}_m \\ &= [r_{0,1}^{\frac{\alpha}{2}} r_{m,1}^{-\frac{\alpha}{2}} g_{m,1} \quad r_{0,2}^{\frac{\alpha}{2}} r_{m,2}^{-\frac{\alpha}{2}} g_{m,2} \quad \dots \quad r_{0,K}^{\frac{\alpha}{2}} r_{m,K}^{-\frac{\alpha}{2}} g_{m,K}]^T. \end{aligned} \quad (38)$$

\mathbf{q}_m comprises independent random variables. Since $r_{i,j}$ are bounded from below, the variance of these variables is bounded from above. Hence for m sufficiently large that $r_m > D$,

$$\max\{r_{0,1}^\alpha, \dots, r_{0,K}^\alpha\} \max\{r_{m,1}^{-\alpha}, \dots, r_{m,K}^{-\alpha}\}$$

$$\leq \max\{r_{0,1}^\alpha, \dots, r_{0,K}^\alpha\} (r_m - D)^{-\alpha}, \quad (39)$$

with probability 1. Thus, for n sufficiently large that $r_{n+1} > D$, with probability 1, we can bound (37) as

$$\begin{aligned} E \left[\left| K^{-\frac{\alpha}{2}} \sum_{m=n+1}^{\infty} |\mathbf{w}^\dagger[n] \mathbf{h}_m|^2 \right| \right] \\ = \frac{1}{K} \sum_{m=n+1}^{\infty} E \left[|\mathbf{g}_0^\dagger[n] \left(K^{\frac{\alpha}{2}-1} \bar{\mathbf{R}} \right)^{-1} \mathbf{q}_m|^2 \right] \\ \leq \sum_{m=n+1}^{\infty} \max\{r_{0,1}^\alpha, \dots, r_{0,K}^\alpha\} (r_m - D)^{-\alpha} \frac{1}{\lambda_{\min}(\bar{\mathbf{R}}^{-1}) K} \\ = \max\{r_{0,1}^\alpha, \dots, r_{0,K}^\alpha\} \frac{1}{\lambda_{\min}(\bar{\mathbf{R}}^{-1}) K} \sum_{m=n+1}^{\infty} (r_m - D)^{-\alpha}, \end{aligned} \quad (40)$$

where the inequality in (40) follows from Lemma 7. Note that the expectation above is over \mathbf{g}_0 , and \mathbf{h}_m . For large enough n such that $r_{n+1} > D$, the summation of $(r_m - D)^{-\alpha}$ is known to be bounded with probability 1, and since $r_{0,1}, \dots, r_{0,K}$ is bounded, and with probability 1, $\lim_{K \rightarrow \infty} \lambda_{\min}(\bar{\mathbf{R}}^{-1})$ is bounded by a strictly positive number from Lemma 5,

$$\lim_{K \rightarrow \infty} K^{-\frac{\alpha}{2}} \sum_{m=n+1}^{\infty} |\mathbf{w}^\dagger[n] \mathbf{h}_m|^2 = 0 \quad (42)$$

in probability. From Lemma 1 with $n = cK$ and $c > 2$,

$$\lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} \text{SIR}[n] = \bar{P} \left[\frac{\alpha}{2\pi^2 \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} \quad (43)$$

Therefore (34) converges in probability to zero, and the following holds in probability

$$\begin{aligned} \lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} K^{-\frac{\alpha}{2}} \underline{\text{SIR}}[n] &= \lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} K^{-\frac{\alpha}{2}} \text{SIR}[n] \\ &= \bar{P} \left[\frac{\alpha}{2\pi^2 \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} \end{aligned} \quad (44)$$

Thus, the upper and lower bounds of SIR are equal in the limit, completing the proof.

APPENDIX C PROOF OF LEMMA 3

Note that

$$\begin{aligned} \Pr(\bar{P} \leq x) &= \Pr \left(\frac{1}{N} (r_{0,1}^{-\alpha} + r_{0,2}^{-\alpha} + \dots + r_{0,N}^{-\alpha}) \leq x \right) \\ &= \Pr \left(r_{0,1}^{-\alpha} + r_{0,2}^{-\alpha} + \dots + r_{0,N}^{-\alpha} \leq Nx \right). \end{aligned} \quad (45)$$

To find the CDF of $r_{0,1}^{-\alpha} + r_{0,2}^{-\alpha} + \dots + r_{0,N}^{-\alpha}$, we first condition on the test mobile being at distance d from the center of the circular cell. Since the AP/BSs are uniformly randomly distributed within the circular cell, the PDF of the distance from the mobile to one AP/BS is equal to the PDF of the distance between a random point in a disk of radius R to another point in the disk at distance d from the center of that

disk. Using geometric arguments one can find this PDF to equal (22). Note that this PDF can be found in different forms in references such as [25] and others.

Conditioned on d , the characteristic function of $r_{0,i}^{-\alpha}$ for $i = 1, 2, \dots, K$ is therefore

$$\int_0^\infty e^{jsr^{-\alpha}} f_{r|d}(r|d) dr \quad (46)$$

Since the AP/BSs are i.i.d., conditioned on d , the characteristic function of $r_{0,1}^{-\alpha} + r_{0,2}^{-\alpha} + \dots + r_{0,N}^{-\alpha}$ is

$$\left(\int_0^\infty e^{jsr^{-\alpha}} f_{r|d}(r|d) dr \right)^N \quad (47)$$

We then apply the Gil-Palaez Theorem, [26], to find the CDF of $r_{0,1}^{-\alpha} + r_{0,2}^{-\alpha} + \dots + r_{0,N}^{-\alpha}$ conditioned on d , evaluated at Nx , which yields:

$$\frac{1}{2} - \frac{1}{\pi} \int_0^1 \frac{1}{s} \Im \left\{ e^{-jsNx} \left(\int_0^\infty e^{jsr^{-\alpha}} f_{r|d}(r|d) dr \right)^N \right\} ds \quad (48)$$

The proof is completed by averaging the CDF over d whose PDF is $2d/R^2$ for $0 < d \leq R$.

**APPENDIX D
POOF OF LEMMA 5**

First let's define $\hat{K} = c_1 K$, where c_1 is a constant, with $1 < c_1 < c$. Since $n = cK > \hat{K}$, by the Weyl inequality (see e.g. [27]),

$$\lambda_{\min} \left(\hat{K}^{\alpha/2-1} \mathbf{H}[\hat{K}] \mathbf{H}^\dagger[\hat{K}] \right) \leq \lambda_{\min} \left(\hat{K}^{\alpha/2-1} \mathbf{H}[n] \mathbf{H}^\dagger[n] \right) \quad (49)$$

Additionally,

$$\begin{aligned} \mathbf{H}[\hat{K}] \mathbf{H}^\dagger[\hat{K}] &= \sum_{i=1}^{\hat{K}} \mathbf{h}_i \mathbf{h}_i^\dagger = \sum_{i=1}^{\hat{K}} (r_{\hat{K}+1} + D)^{-\alpha} \mathbf{g}_i \mathbf{g}_i^\dagger \\ &+ \sum_{i=1}^{\hat{K}} (\mathbf{h}_i \mathbf{h}_i^\dagger - (r_{\hat{K}+1} + D)^{-\alpha} \mathbf{g}_i \mathbf{g}_i^\dagger) \end{aligned} \quad (50)$$

Since $r_{i,j} < r_{\hat{K}+1} + D$ for $i = 1, 2, \dots, \hat{K}$ and $j = 1, 2, \dots, K$ with probability 1, the matrix resulting from rightmost sum of (50) is non-negative definite with probability 1. Hence, by the Weyl inequality (see e.g. [27])

$$\lambda_{\min} \left\{ \sum_{i=1}^{\hat{K}} (r_{\hat{K}+1} + D)^{-\alpha} \mathbf{g}_i \mathbf{g}_i^\dagger \right\} \leq \lambda_{\min} \left\{ \mathbf{H}[\hat{K}] \mathbf{H}^\dagger[\hat{K}] \right\} \quad (51)$$

Note that the following is known to hold with probability 1 ([28])

$$\lim_{\hat{K} \rightarrow \infty} \lambda_{\min} \left\{ \sum_{i=1}^{\hat{K}} \frac{1}{M} \mathbf{g}_i \mathbf{g}_i^\dagger \right\} = \left(1 - \sqrt{\frac{1}{c_1}} \right)^2 \quad (52)$$

By assumption, we have the following with probability 1

$$\lim_{\hat{K} \rightarrow \infty} \frac{\pi \rho r_{\hat{K}+1}^2}{\hat{K} + 1} = 1 \quad (53)$$

$$\lim_{\hat{K} \rightarrow \infty} \left(r_{\hat{K}+1} + D \right)^{-\alpha} \left(\frac{\hat{K}}{\pi \rho} \right)^{\alpha/2} = 1 \quad (54)$$

Combining (52) and (54), we have

$$\begin{aligned} \lim_{\hat{K} \rightarrow \infty} \hat{K}^{\alpha/2-1} \lambda_{\min} \left\{ \left(r_{\hat{K}+1} + D \right)^{-\alpha} \sum_{i=1}^{\hat{K}} \mathbf{g}_i \mathbf{g}_i^\dagger \right\} \\ = (\pi \rho)^{\alpha/2} \left(1 - \sqrt{\frac{1}{c_1}} \right)^2 \end{aligned} \quad (55)$$

with probability 1. Combining this expression with (51) and (49), and rearranging yields the following in probability

$$\begin{aligned} (\pi \rho)^{\alpha/2} \left(1 - \sqrt{\frac{1}{c_1}} \right)^2 c_1^{1-\frac{\alpha}{2}} \\ \leq \lim_{K \rightarrow \infty} \lambda_{\min} \left\{ K^{\alpha/2-1} \mathbf{H}[n] \mathbf{H}^\dagger[n] \right\} \end{aligned} \quad (56)$$

Since $\lambda_{\min}(\mathbf{D}^{-1}) \geq D_{\min}^\alpha$

$$\begin{aligned} \left(\pi \rho D_{\min}^2 \right)^{\alpha/2} \left(1 - \sqrt{\frac{1}{c_1}} \right)^2 c_1^{1-\frac{\alpha}{2}} \\ \leq \lim_{K \rightarrow \infty} \lambda_{\min} \left\{ \mathbf{D}^{-\frac{1}{2}} \left(K^{\alpha/2-1} \mathbf{H}[n] \mathbf{H}^\dagger[n] \right) \mathbf{D}^{-\frac{1}{2}} \right\} \end{aligned} \quad (57)$$

Additionally, it is known from [29] that for sufficiently large K , with probability 1, the matrix $\bar{\mathbf{R}} = \mathbf{D}^{-\frac{1}{2}} \left(K^{\alpha/2-1} \mathbf{H}[n] \mathbf{H}^\dagger[n] \right) \mathbf{D}^{-\frac{1}{2}}$ has no eigenvalues outside the support of its limiting eigenvalue distribution. Therefore for large enough K , with probability 1, no eigenvalue of $\bar{\mathbf{R}}$ will be less than or equal to the RHS of (57), which is bounded from below by the LHS.

**APPENDIX E
PROOF OF LEMMA 6**

Define the following vectors for $i = 1, 2, \dots$

$$\tilde{\mathbf{h}}_i = [r_i^{-\frac{\alpha}{2}} g_{i,1} \quad r_i^{-\frac{\alpha}{2}} g_{i,2} \quad \dots \quad r_i^{-\frac{\alpha}{2}} g_{i,K}]^T \quad (58)$$

$$\begin{aligned} \hat{\mathbf{h}}_i = \mathbf{h}_i - \tilde{\mathbf{h}}_i = [(r_{i,1}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}}) g_{i,1} \quad (r_{i,2}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}}) g_{i,2} \\ \dots (r_{i,K}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}}) g_{i,K}]^T, \end{aligned} \quad (59)$$

where r_i is the distance of the i -th mobile from the origin. Hence, the vector $\tilde{\mathbf{h}}_i$ is equivalent to the channel vector between the i -th mobile and the K receive antennas, if the antennas were all co-located at the origin instead of their actual locations.

Define the following matrices

$$\mathbf{H}_{bs}[n] = [\tilde{\mathbf{h}}_1 \quad \tilde{\mathbf{h}}_2 \quad \tilde{\mathbf{h}}_3 \quad \dots \quad \tilde{\mathbf{h}}_n] \quad (60)$$

$$\mathbf{H}_{\Delta}[n] = [\hat{\mathbf{h}}_1 \quad \hat{\mathbf{h}}_2 \quad \hat{\mathbf{h}}_3 \quad \dots \quad \hat{\mathbf{h}}_n] \quad (61)$$

Hence,

$$\mathbf{H}[n] = \mathbf{H}_{bs}[n] + \mathbf{H}_{\Delta}[n] \quad (62)$$

Using this expression, rewrite $\bar{\mathbf{R}}$ as

$$\begin{aligned} \bar{\mathbf{R}} = K^{-\frac{\alpha}{2}-1} \mathbf{D}^{-\frac{1}{2}} & \left(\mathbf{H}_{bs}[n] \mathbf{H}_{bs}^{\dagger}[n] + \mathbf{H}_{\Delta}[n] \mathbf{H}_{bs}^{\dagger}[n] \right. \\ & \left. + \mathbf{H}_{bs}[n] \mathbf{H}_{\Delta}^{\dagger}[n] + \mathbf{H}_{\Delta}[n] \mathbf{H}_{\Delta}^{\dagger}[n] \right) \mathbf{D}^{-\frac{1}{2}}. \end{aligned} \quad (63)$$

Let $F_K^{bs}(\tau)$ be the e.d.f. of the eigenvalues of the matrix

$$\bar{\mathbf{R}}_{bs} = \mathbf{D}^{-\frac{1}{2}} \left(K^{-\frac{\alpha}{2}-1} \mathbf{H}_{bs}[n] \mathbf{H}_{bs}^{\dagger}[n] \right) \mathbf{D}^{-\frac{1}{2}} \quad (64)$$

The limit of $F_K^{bs}(\tau)$ is given by the following Lemma
Lemma 8: The following hold w.p.1.

$$\lim_{K \rightarrow \infty} F_K^{bs}(\tau) = F^{bs}(\tau) \quad (65)$$

and

$$\lim_{\tau \rightarrow \infty} \int \frac{F^{bs}(\tau)}{\tau} = \bar{P} \left[\frac{\alpha}{2\pi^2 \rho} \sin\left(\frac{2\pi}{\alpha}\right) \right]^{\frac{\alpha}{2}}. \quad (66)$$

Proof: Please see Appendix G.

The following lemma shows that as $K \rightarrow \infty$, $F_K(\tau)$ does not depend on the three RHS terms in the parentheses of (63), and completes the proof when combined with (66).

Lemma 9: The following hold i.p.

$$\lim_{K \rightarrow \infty} F_K(\tau) = F(\tau) = F^{bs}(\tau) \quad \text{and} \quad (67)$$

$$\lim_{K \rightarrow \infty} \int \frac{F_K(\tau)}{\tau} d\tau = \int \frac{F^{bs}(\tau)}{\tau} d\tau \quad (68)$$

Proof: Please see Appendix H.

Note that if all antennas were co-located at a given location, all the $r_{i,j}$ terms in the vectors \mathbf{h}_i in (59) would be equal, resulting in faster convergence to zero of the terms involving $\mathbf{H}_{\Delta}[n]$ in equation (63).

**APPENDIX F
 PROOF OF LEMMA 7**

Denote the i -th entries of the vectors \mathbf{a} and \mathbf{b} by a_i and b_i respectively, and the ij -th entry of \mathbf{B}_L^{-1} by $B_{i,j}^{-1}$

$$\begin{aligned} E \left[\left| \frac{1}{L} \mathbf{a}^{\dagger} \mathbf{B}^{-1} \mathbf{b} \right|^2 \right] \\ = E \left[\frac{1}{L^2} \sum_{i_1=1}^L \sum_{i_2=1}^L \sum_{j_1=1}^L \sum_{j_2=1}^L a_{i_1} a_{i_2}^* b_{j_1} b_{j_2}^* B_{i_1,j_1}^{-1} B_{i_2,j_2}^{-1} \right] \end{aligned} \quad (69)$$

Terms in the summation with odd numbers of a_{i_1} and b_{j_1} terms will go to zero, due to the fact that a_i and b_i are zero mean, uncorrelated random variables. Therefore, for all the remaining terms, with $i_1 = i_2$ and $j_1 = j_2$, we have

$$\begin{aligned} E \left[\left| \frac{1}{L} \mathbf{a}^{\dagger} \mathbf{B}^{-1} \mathbf{b} \right|^2 \right] \\ = \frac{1}{L^2} \sum_{i_1=1}^L \sum_{j_1=1}^L E[|a_{i_1}|^2] E[|b_{j_1}|^2] |B_{i_1,j_1}^{-1}|^2 \end{aligned} \quad (70)$$

$$\begin{aligned} & \leq \frac{1}{L^2} \max_{\ell=1 \dots L} \{\sigma_{a_{\ell}}^2\} \max_{\ell=1 \dots L} \{\sigma_{b_{\ell}}^2\} \sum_{i_1=1}^L \sum_{j_1=1}^L |B_{i_1,j_1}^{-1}|^2 \\ & = \frac{1}{L^2} \max_{\ell=1 \dots L} \{\sigma_{a_{\ell}}^2\} \max_{\ell=1 \dots L} \{\sigma_{b_{\ell}}^2\} \text{Tr} \left\{ \mathbf{B}_L^{-2} \right\} \\ & \leq \frac{1}{L} \max_{\ell=1 \dots L} \{\sigma_{a_{\ell}}^2\} \max_{\ell=1 \dots L} \{\sigma_{b_{\ell}}^2\} \lambda_{\max}(\mathbf{B}_L^{-2}), \end{aligned} \quad (71)$$

Since $\lambda_{\max}(\mathbf{B}_L^{-2}) = \frac{1}{(\lambda_{\min}(\mathbf{B}_L))^2}$, for $L > L_0$, with probability 1, (35) is proved. The second part follows because (35) converges to zero as $L \rightarrow \infty$, and the fact that mean-square convergence and convergence with probability one both imply convergence in probability.

**APPENDIX G
 PROOF OF LEMMA 8**

From Theorem 6.1 of [21], as $n, K \rightarrow \infty$ with $n = cK$, and $c > 2$, w.p.1.,

$$\begin{aligned} F_K^{bs}(\tau) & \rightarrow F^{bs}(\tau), \\ \lim_{K \rightarrow \infty} \int \frac{F_K^{bs}(\tau)}{\tau - z} & = \int \frac{F^{bs}(\tau)}{\tau - z} = \lim_{K \rightarrow \infty} m_K^{bs}(z) \end{aligned} \quad (72)$$

where $m_K^{bs}(z)$ and $m^{bs}(z)$ are the Stieltjes transforms (see e.g. [22]) of $F_K^{bs}(z)$ and $F^{bs}(z)$, respectively. Thus, (65) is proved.

For the next part of the Lemma, using \uparrow to denote approach from below we have

$$\begin{aligned} \lim_{z \uparrow 0} \lim_{K \rightarrow \infty} \int \frac{F_K^{bs}(\tau)}{\tau - z} & = \lim_{z \uparrow 0} \int \frac{F^{bs}(\tau)}{\tau - z} = \lim_{z \uparrow 0} \lim_{K \rightarrow \infty} m_K^{bs}(z) \\ & = \lim_{z \uparrow 0} m^{bs}(z). \end{aligned} \quad (73)$$

Additionally from Theorem 6.1 of [21], as $n, K \rightarrow \infty$ with $n = cK$, and $c > 2$,

$$\lim_{K \rightarrow \infty} m_K^{bs}(z) = \lim_{K \rightarrow \infty} \frac{1}{K} \text{Tr} \left(\int \frac{\tau dH(\tau)}{1 + \frac{1}{c} \tau e(z)} \mathbf{D}^{-1} - z \mathbf{I}_K \right)^{-1} \quad (74)$$

where $e(z)$ is the unique solution to

$$e(z) = \lim_{K \rightarrow \infty} \frac{1}{K} \text{Tr} \left(\int \frac{\tau dH(\tau)}{1 + \frac{1}{c} \tau e(z)} \mathbf{I}_K - z \mathbf{D} \right)^{-1}, \quad (75)$$

where $H(\tau)$ is the limit of the e.d.f. of $K^{\frac{\alpha}{2}} r_1^{-\alpha}$, $K^{\frac{\alpha}{2}} r_2^{-\alpha}, \dots, K^{\frac{\alpha}{2}} r_K^{-\alpha}$. Additionally, this convergence is uniform in z for any $z < 0$. Therefore, from Theorem 7.11 of [30],

$$\begin{aligned} \lim_{z \uparrow 0} \lim_{K \rightarrow \infty} m_K^{bs}(z) & = \\ \lim_{z \uparrow 0} \lim_{K \rightarrow \infty} \frac{1}{K} \text{Tr} & \left(\int \frac{\tau dH(\tau)}{1 + \frac{1}{c} \tau e(z)} \mathbf{D}^{-1} - z \mathbf{I}_K \right)^{-1} \\ = \lim_{K \rightarrow \infty} \lim_{z \uparrow 0} \frac{1}{K} \text{Tr} & \left(\int \frac{\tau dH(\tau)}{1 + \frac{1}{c} \tau e(z)} \mathbf{D}^{-1} - z \mathbf{I}_K \right)^{-1} \end{aligned}$$

$$\begin{aligned}
 &= \lim_{K \rightarrow \infty} \lim_{z \uparrow 0} \frac{1}{K} \text{Tr} \left(\int \frac{\tau dH(\tau)}{1 + \frac{1}{c} \tau e(z)} \mathbf{D}^{-1} \right)^{-1} \\
 &= \lim_{K \rightarrow \infty} \lim_{z \uparrow 0} \frac{1}{K} \sum_{k=1}^K \frac{r_{0,k}^{-\alpha}}{\int \frac{\tau dH(\tau)}{1 + \frac{1}{c} \tau e(z)}} \\
 &= \lim_{K \rightarrow \infty} \lim_{z \uparrow 0} \frac{1}{\int \frac{\tau dH(\tau)}{1 + \frac{1}{c} \tau e(z)}} \frac{1}{K} \sum_{k=1}^K r_{0,k}^{-\alpha} \quad (76)
 \end{aligned}$$

Next, take the limit of (76) as $c \rightarrow \infty$ to get

$$\begin{aligned}
 &\lim_{c \rightarrow \infty} \lim_{z \uparrow 0} \lim_{K \rightarrow \infty} m_K^{bs}(z) \\
 &= \lim_{c \rightarrow \infty} \lim_{K \rightarrow \infty} \lim_{z \uparrow 0} \frac{1}{\int \frac{\tau dH(\tau)}{1 + \frac{1}{c} \tau e(z)}} \frac{1}{K} \sum_{k=1}^K r_{0,k}^{-\alpha} \quad (77)
 \end{aligned}$$

Suppose that $r_{0,1} = r_{0,2} = \dots = r_{0,K} = 1$, which corresponds to the case of a single base station at unit distance from the test mobile. Then, it is known from [31] for Poisson distributed mobiles, and [17] for the general mobile distribution assumed here, that

$$\lim_{c \rightarrow \infty} \lim_{z \uparrow 0} \lim_{K \rightarrow \infty} m_K^{bs}(z) = \left[\frac{\alpha}{2\pi^2 \rho} \sin \left(\frac{2}{\alpha} \right) \right]^{\frac{\alpha}{2}} \quad (78)$$

Comparing with (77), we conclude that when it is not necessarily the case that $r_{0,1} = r_{0,2} = \dots = r_{0,K} = 1$, we have

$$\lim_{c \rightarrow \infty} \lim_{z \uparrow 0} \lim_{K \rightarrow \infty} m_K^{bs}(z) = \left[\frac{\alpha}{2\pi^2 \rho} \sin \left(\frac{2}{\alpha} \right) \right]^{\frac{\alpha}{2}} \frac{1}{K} \sum_{k=1}^K r_{0,k}^{-\alpha} \quad (79)$$

The lemma is proved from the definition of \bar{P}

APPENDIX H PROOF OF LEMMA 9

From Lemma 8, it is known that $F_K^{bs}(\tau) \rightarrow F^{bs}(\tau)$ w.p.1.. We then use the results of Exercises 2.4.3 and 2.4.4 of [22] to show that all the terms involving \mathbf{H}_Δ in (63) do not play a role in the e.d.f. of the eigenvalues of $\bar{\mathbf{R}}$ in the limit. To this end, let's define

$$\mathbf{B} = \mathbf{H}_\Delta[n] \mathbf{H}_{bs}^\dagger[n] + \mathbf{H}_{bs}[n] \mathbf{H}_\Delta^\dagger[n] + \mathbf{H}_\Delta[n] \mathbf{H}_\Delta^\dagger[n] \quad (80)$$

and for some $0 < \delta < 1$, define the following matrices

$$\mathbf{B}_{a1} = \sum_{i=[n^{1-\delta}]+1}^n \tilde{\mathbf{h}}_i \tilde{\mathbf{h}}_i^\dagger \quad (81)$$

$$\mathbf{B}_{a2} = \sum_{i=[n^{1-\delta}]+1}^n \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^\dagger, \quad (82)$$

$$\mathbf{B}_{a3} = \sum_{i=[n^{1-\delta}]+1}^n \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^\dagger \quad (83)$$

$$\mathbf{B}_{b1} = \sum_{i=1}^{[n^{1-\delta}]} \tilde{\mathbf{h}}_i \tilde{\mathbf{h}}_i^\dagger \quad (84)$$

$$\mathbf{B}_{b2} = \sum_{i=1}^{[n^{1-\delta}]} \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^\dagger \quad (85)$$

$$\mathbf{B}_{b3} = \sum_{i=1}^{[n^{1-\delta}]} \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^\dagger \quad (86)$$

Note that $\mathbf{B} = \mathbf{B}_{a1} + \mathbf{B}_{b1} + \mathbf{B}_{a2} + \mathbf{B}_{b2} + \mathbf{B}_{a3} + \mathbf{B}_{b3}$. The following lemmas will be used to show that the matrices above have a diminishing effect on $F_K(\tau)$ as $K \rightarrow \infty$.

Lemma 10: The following holds i.p. for $i = 1, 2, 3$

$$\lim_{K \rightarrow \infty} \frac{1}{K} \|\mathbf{D}^{-\frac{1}{2}} (K^{-\frac{\alpha}{2}-1} \mathbf{B}_{ai}) \mathbf{D}^{-\frac{1}{2}}\|_F^2 = 0. \quad (87)$$

Proof: Please see Appendix I.

Lemma 11: The following holds i.p. for $i = 1, 2, 3$

$$\lim_{K \rightarrow \infty} \frac{1}{K} \text{Rank} \left(\mathbf{D}^{-\frac{1}{2}} (K^{-\frac{\alpha}{2}-1} \mathbf{B}_{bi}) \mathbf{D}^{-\frac{1}{2}} \right) = 0. \quad (88)$$

Proof: Since \mathbf{B}_{b1} is a sum of $[n^{1-\delta}]$ rank 1 matrices, its rank is $\leq n^{1-\delta}$. Hence,

$$\begin{aligned}
 \frac{1}{K} \text{Rank} \left(\mathbf{D}^{-\frac{1}{2}} (K^{-\frac{\alpha}{2}-1} \mathbf{B}_{bi}) \mathbf{D}^{-\frac{1}{2}} \right) &\leq \frac{1}{K} \text{Rank} (K^{-\frac{\alpha}{2}-1} \mathbf{B}_{b1}) \\
 &\leq \frac{1}{K} n^{1-\delta} \quad (89)
 \end{aligned}$$

Since $\delta > 0$ and $n = cK$, the upper bound above converges to zero as $K \rightarrow \infty$ completing the proof ■

Let $\bar{\mathbf{R}} = \bar{\mathbf{R}}_{bs} + \mathbf{R}_a + \mathbf{R}_b$, where

$$\begin{aligned}
 \mathbf{R}_a &= \mathbf{D}^{-\frac{1}{2}} \mathbf{B}_{a1} \mathbf{D}^{-\frac{1}{2}} + \mathbf{D}^{-\frac{1}{2}} \mathbf{B}_{a2} \mathbf{D}^{-\frac{1}{2}} + \mathbf{D}^{-\frac{1}{2}} \mathbf{B}_{a3} \mathbf{D}^{-\frac{1}{2}} \\
 \mathbf{R}_b &= \mathbf{D}^{-\frac{1}{2}} \mathbf{B}_{b1} \mathbf{D}^{-\frac{1}{2}} + \mathbf{D}^{-\frac{1}{2}} \mathbf{B}_{b2} \mathbf{D}^{-\frac{1}{2}} + \mathbf{D}^{-\frac{1}{2}} \mathbf{B}_{b3} \mathbf{D}^{-\frac{1}{2}}
 \end{aligned}$$

From Lemmas 10 and 11, $\lim_{K \rightarrow \infty} \frac{1}{K} \|\mathbf{R}_a\|_F^2 = 0$ and $\lim_{K \rightarrow \infty} \frac{1}{K} \text{Rank}(\mathbf{R}_b) = 0$ i.p., respectively. Hence, from Exercises 2.4.3 and 2.4.4 of [22], the e.d.f. of the eigenvalues of $\bar{\mathbf{R}}$ converges to the e.d.f. for the eigenvalues of $\bar{\mathbf{R}}_{bs}$ i.p., completing the proof of the first part. The second part follows from Theorem 25.8 of [32] and Lemma 5.

APPENDIX I PROOF OF LEMMA 10

1) \mathbf{B}_{a1}

Since $\frac{1}{K} \|\mathbf{D}^{-\frac{1}{2}} (K^{-\frac{\alpha}{2}-1} \mathbf{B}_{a1}) \mathbf{D}^{-\frac{1}{2}}\|_F^2$

$\leq \|\mathbf{D}^{-\frac{1}{2}}\|_F^4 \frac{1}{K} \|K^{-\frac{\alpha}{2}-1} \mathbf{B}_{a1}\|_F^2$, it suffices to show that as $K \rightarrow \infty$, $\frac{1}{K} \|K^{-\frac{\alpha}{2}-1} \mathbf{B}_{a1}\|_F^2 \rightarrow 0$ i.p..

Consider the following, where the expectation is w.r.t. the fading variables g_{ij}

$$\begin{aligned}
 E \left[\left| \frac{1}{K} \|K^{-\frac{\alpha}{2}-1} \mathbf{B}_{a1}\|_F^2 \right| \right] &= K^{\alpha-3} \sum_{\ell=1}^K \sum_{\substack{m=1, \\ m \neq \ell}}^K E \left[|[\mathbf{B}_{a1}]_{\ell m}|^2 \right] \\
 &\quad + K^{\alpha-3} \sum_{\ell=1}^K E \left[|[\mathbf{B}_{a1}]_{\ell \ell}|^2 \right] \quad (90)
 \end{aligned}$$

Subsequently, when we say $\phi(i) = O_1(i^\kappa)$, we mean $\lim_{i \rightarrow \infty} i^{-\kappa} \phi(i) = A$, for a constant (possibly zero) A , w.p.1. With this notation, we have

Lemma 12: With expectation w.r.t. the fading variables:

$$E \left[|[\mathbf{B}_{a1}]_{\ell m}|^2 \right] = \begin{cases} O_1(K^{-\alpha+\alpha\delta+\delta+1}), & \text{for } \ell = m \\ O_1(K^{-\alpha+\alpha\delta+\delta}), & \text{for } \ell \neq m \end{cases} \quad (91)$$

Proof:

$$[\mathbf{B}_{a1}]_{\ell m} = \sum_{i=[n^{1-\delta}]_j+1}^n (r_{i,1}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}})^2 g_{i,\ell} g_{i,m}^* \quad (92)$$

From (1), we have

$$r_i^{-\frac{\alpha}{2}} = O_1(i^{-\frac{\alpha}{4}}) \quad (93)$$

Additionally, since the antennas serving the test user are contained within a bounded area, for $\ell = 1, 2, \dots, K$:

$$|r_{i,\ell}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}}| = O_1(i^{-\frac{\alpha}{4}-\frac{1}{2}}) \quad (94)$$

Taking the variance of (92) w.r.t. $g_{i,\ell}$ and $g_{i,m}$, and substituting (93) and (94) yields the following w.p.1.,

$$\begin{aligned} \text{var} \{[\mathbf{B}_{a1}]_{\ell m}\} &= \sum_{i=[n^{1-\delta}]_j+1}^n r_i^{-\alpha} (r_{i,1}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}})^2 \text{var} \{g_{i,\ell} g_{i,m}^*\} \\ &= \sum_{i=[n^{1-\delta}]_j+1}^n r_i^{-\alpha} (r_{i,1}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}})^2 \quad (95) \\ &\leq \sum_{i=[n^{1-\delta}]_j+1}^n r_{[n^{1-\delta}]_j+1}^{-\alpha} (r_{[n^{1-\delta}]_j+1,1}^{-\frac{\alpha}{2}} - r_{[n^{1-\delta}]_j+1}^{-\frac{\alpha}{2}})^2 \\ &\leq n r_{[n^{1-\delta}]_j+1}^{-\alpha} (r_{[n^{1-\delta}]_j+1,1}^{-\frac{\alpha}{2}} - r_{[n^{1-\delta}]_j+1}^{-\frac{\alpha}{2}})^2 \quad (96) \\ &= O_1(n([n^{1-\delta}]_j+1)^{-\alpha-1}) = O_1(n^{\alpha\delta+\delta-\alpha}) \quad (97) \end{aligned}$$

For $\ell = m$, note from (92) that

$$\begin{aligned} E \left[|[\mathbf{B}_{a1}]_{\ell m}| \right] &= \sum_{i=[n^{1-\delta}]_j+1}^n r_i^{-\frac{\alpha}{2}} (r_{i,1}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}}) E[|g_{i,\ell}|^2] \\ &= \sum_{i=[n^{1-\delta}]_j+1}^n r_i^{-\frac{\alpha}{2}} (r_{i,1}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}}) \\ (E \left[|[\mathbf{B}_{a1}]_{\ell m}| \right])^2 &= \sum_{i=[n^{1-\delta}]_j+1}^n r_i^{-\alpha} (r_{i,1}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}})^2 \\ &+ \sum_{i=[n^{1-\delta}]_j+1}^n \sum_{\substack{j=[n^{1-\delta}]_j+1 \\ j \neq i}}^n r_i^{-\frac{\alpha}{2}} |r_{i,1}^{-\frac{\alpha}{2}} - r_i^{-\frac{\alpha}{2}} - r_j^{-\frac{\alpha}{2}}| |r_{j,1}^{-\frac{\alpha}{2}} - r_j^{-\frac{\alpha}{2}}| \quad (98) \end{aligned}$$

Using steps similar to the derivation from (95) to (97) yields the following w.p.1.

$$(E \left[|[\mathbf{B}_{a1}]_{\ell m}| \right])^2 = O_1(n^{\alpha\delta+\delta-\alpha}) + O_1(n^{\alpha\delta+\delta-\alpha+1})$$

$$= O_1(n^{\alpha\delta+\delta-\alpha+1}). \quad (99)$$

Since $E \left[|[\mathbf{B}_{a1}]_{\ell m}|^2 \right] = \text{var} \{[\mathbf{B}_{a1}]_{\ell m}\} + (E \left[|[\mathbf{B}_{a1}]_{\ell m}| \right])^2$, the proof for $\ell = m$ is completed by adding (97) and (99) and substituting $n = cK$. For $\ell \neq m$, $E \left[|[\mathbf{B}_{a1}]_{\ell m}| \right] = 0$. Hence, for $\ell \neq m$ we can use (97) with $n = cK$ to say that w.p.1.,

$$E \left[|[\mathbf{B}_{a1}]_{\ell m}|^2 \right] = O_1(K^{-\alpha+\alpha\delta+\delta}). \quad (100)$$

The proof is complete since $\alpha\delta + \delta < 1$ and $\alpha > 2$, and convergence in mean and w.p.1. imply convergence i.p.. ■

From substituting Lemma 12 into (101) and simplifying results in

$$E \left[\left| \frac{1}{K} \left\| K^{\frac{\alpha}{2}-1} \mathbf{B}_{a1} \right\|_F^2 \right| \right] = O_1(K^{\delta\alpha+\delta-1}) \quad (101)$$

If we pick δ such that $\delta\alpha + \delta < 1$, we have the following w.p.1.

$$\lim_{K \rightarrow \infty} E \left[\left| \frac{1}{K} \left\| K^{\frac{\alpha}{2}-1} \mathbf{B}_{a1} \right\|_F^2 \right| \right] = 0. \quad (102)$$

Since convergence in mean, and w.p.1. both imply convergence i.p., (87) is proved. The proof of $i = 2$ and 3 follow similar steps.

APPENDIX J PROOF OF EXTENSION TO MULTIPLE TRANSMIT ANTENNAS

Consider a normalized version of the mutual information between the transmitted signal from the test transmitter and the received signal vector as follows

$$\begin{aligned} I(\mathbf{y}; \mathbf{x}_0) + \log_2 \left(K^{-\frac{\alpha}{2}} \mathbf{I} \right) &= \log_2 \left| \mathbf{I} + \mathbf{H}_0 \mathbf{H}_0^\dagger \mathbf{K}^{-1} \right| + \log_2 \left(K^{-\frac{\alpha}{2}} \mathbf{I} \right) \\ &= \log_2 \left| K^{-\frac{\alpha}{2}} \mathbf{I} + K^{-\frac{\alpha}{2}} \mathbf{H}_0 \mathbf{H}_0^\dagger \mathbf{K}^{-1} \right| \\ &= \log_2 \left| K^{-\frac{\alpha}{2}} \mathbf{I} + K^{-\frac{\alpha}{2}} \mathbf{H}_0^\dagger \mathbf{K}^{-1} \mathbf{H}_0 \right| \\ &= \log_2 \left| K^{-\frac{\alpha}{2}} \mathbf{I} + \frac{1}{K} \mathbf{H}_0^\dagger \left(K^{\frac{\alpha}{2}-1} \mathbf{K} \right)^{-1} \mathbf{H}_0 \right| \quad (103) \end{aligned}$$

Let's take a closer look at the determinant of the matrix $\frac{1}{K} \mathbf{H}_0^\dagger \left(K^{\frac{\alpha}{2}-1} \mathbf{K} \right)^{-1} \mathbf{H}_0$, for large K . Define $\bar{\mathbf{K}} = K^{\alpha/2-1} \mathbf{K}$ for notational convenience. We then have

$$\begin{aligned} \lim_{K \rightarrow \infty} \frac{1}{K} \left| \mathbf{H}_0^\dagger \bar{\mathbf{K}}^{-1} \mathbf{H}_1 \right| &= \lim_{K \rightarrow \infty} \frac{1}{K} \\ &\times \begin{vmatrix} \mathbf{h}_0^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_0 & \mathbf{h}_0^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_1 & \dots & \mathbf{h}_0^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_{M-1} \\ \mathbf{h}_1^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_0 & \mathbf{h}_1^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_1 & \dots & \mathbf{h}_1^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_{M-1} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{h}_{M-1}^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_0 & \mathbf{h}_{M-1}^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_1 & \dots & \mathbf{h}_{M-1}^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_{M-1} \end{vmatrix} \quad (104) \end{aligned}$$

Note that the smallest eigenvalue of $\bar{\mathbf{K}}$ is bounded from below with probability 1. Hence, by Lemma 7, $\lim_{K \rightarrow \infty} \frac{1}{K} \mathbf{h}_i^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_j = 0$ for $i \neq j$. Since the determinant of

a matrix is a continuous function of its entries, by the continuous mapping theorem we have the following in probability

$$\begin{aligned}
& \lim_{K \rightarrow \infty} \frac{1}{K} \left| \mathbf{H}_0^\dagger \bar{\mathbf{K}}^{-1} \mathbf{H}_1 \right| = \lim_{K \rightarrow \infty} \frac{1}{K} \\
& \times \begin{vmatrix} \mathbf{h}_0^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_0 & 0 & \dots & 0 \\ 0 & \mathbf{h}_1^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \mathbf{h}_{M-1}^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_{M-1} \end{vmatrix} \\
& = \lim_{K \rightarrow \infty} \left| K^{-\frac{\alpha}{2}} \mathbf{I} + \frac{1}{K} \mathbf{H}_0^\dagger \bar{\mathbf{K}}^{-1} \mathbf{H}_1 \right| \\
& = \prod_{i=0}^{M-1} \lim_{K \rightarrow \infty} \frac{1}{K} \mathbf{h}_i^\dagger \bar{\mathbf{K}}^{-1} \mathbf{h}_i = \prod_{i=0}^{M-1} \lim_{K \rightarrow \infty} K^{-\frac{\alpha}{2}} \mathbf{h}_i^\dagger \mathbf{K}^{-1} \mathbf{h}_i \\
& = \left(\bar{P} \left[\frac{\alpha}{2\pi^2 \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} \right)^M \quad (105)
\end{aligned}$$

where the last step is from substituting Theorem 1.

Now consider the following, where the identity matrix in (107) is $M \times M$:

$$\begin{aligned}
& \left| \eta_M - M \log_2 \left(1 + \bar{P} \left[\frac{\alpha}{2\pi^2 M \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} K^{\frac{\alpha}{2}} \right) \right| \\
& = \left| \log_2 \left(\left| \mathbf{I} + \mathbf{H}_0^\dagger \mathbf{K}^{-1} \mathbf{H}_0 \right| \right) \right. \\
& \quad \left. - M \log_2 \left(1 + \bar{P} \left[\frac{\alpha}{2\pi^2 M \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} K^{\frac{\alpha}{2}} \right) \right| \\
& = \left| \log_2 \left(\left| \mathbf{I} + \mathbf{H}_0^\dagger \mathbf{K}^{-1} \mathbf{H}_0 \right| \right) - M \log_2 \left(K^{\frac{\alpha}{2}} \right) \right. \\
& \quad \left. - M \log_2 \left(1 + \bar{P} \left[\frac{\alpha}{2\pi^2 M \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} K^{\frac{\alpha}{2}} \right) \right. \\
& \quad \left. + M \log_2 \left(K^{\frac{\alpha}{2}} \right) \right| \\
& = \left| \log_2 \left(\left| \mathbf{I} + \mathbf{H}_0^\dagger \mathbf{K}^{-1} \mathbf{H}_0 \right| \right) - \log_2 \left| K^{\frac{\alpha}{2}} \mathbf{I} \right| \right. \\
& \quad \left. - M \log_2 \left(1 + \bar{P} \left[\frac{\alpha}{2\pi^2 M \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} K^{\frac{\alpha}{2}} \right) \right. \\
& \quad \left. + M \log_2 \left(K^{\frac{\alpha}{2}} \right) \right| \quad (106) \\
& = \left| \log_2 \left(\left| K^{-\frac{\alpha}{2}} \mathbf{I} + K^{-\frac{\alpha}{2}} \mathbf{H}_0^\dagger \mathbf{K}^{-1} \mathbf{H}_0 \right| \right) \right. \\
& \quad \left. - M \log_2 \left(K^{-\frac{\alpha}{2}} + \bar{P} \left[\frac{\alpha}{2\pi^2 M \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} \right) \right|
\end{aligned}$$

$$= \left| \log_2 \left(\frac{\left| K^{-\frac{\alpha}{2}} \mathbf{I} + K^{-\frac{\alpha}{2}} \mathbf{H}_0^\dagger \mathbf{K}^{-1} \mathbf{H}_0 \right|}{\left(K^{-\frac{\alpha}{2}} + \bar{P} \left[\frac{\alpha}{2\pi^2 M \rho} \sin \left(\frac{2\pi}{\alpha} \right) \right]^{\frac{\alpha}{2}} \right)^M} \right) \right| = 0. \quad (107)$$

The last step follows from (105).

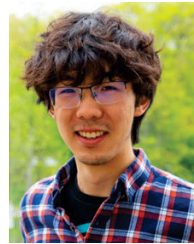
ACKNOWLEDGMENT

The authors thank the anonymous reviewers and associate editor for helpful comments which have greatly improved this paper.

REFERENCES

- [1] M. Peng, Y. Li, Z. Zhao, and C. Wang, "System architecture and key technologies for 5G heterogeneous cloud radio access networks," *IEEE Netw.*, vol. 29, no. 2, pp. 6–14, Mar./Apr. 2015.
- [2] T. Q. Quek, M. Peng, O. Simeone, and W. Yu, *Cloud Radio Access Networks: Principles, Technologies, and Applications*. Cambridge, U.K.: Cambridge Univ. Press, 2017.
- [3] M. Peng, Y. Sun, X. Li, Z. Mao, and C. Wang, "Recent advances in cloud radio access networks: System architectures, key techniques, and open issues," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 3, pp. 2282–2308, 3rd Quart., 2016.
- [4] M. Mohammadi, H. A. Suraweera, and C. Tellambura, "Full-duplex cloud-RAN with uplink/downlink remote radio head association," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2016, pp. 1–6.
- [5] F. Baccelli and A. Giovanidis, "A stochastic geometry framework for analyzing pairwise-cooperative cellular networks," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 794–808, Feb. 2015.
- [6] G. Nigam, P. Minero, and M. Haenggi, "Coordinated multipoint joint transmission in heterogeneous networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4134–4146, Nov. 2014.
- [7] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-free massive MIMO: A new next-generation paradigm," *IEEE Access*, vol. 7, pp. 99878–99888, 2019.
- [8] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO communications," *EURASIP J. Wireless Commun. Netw.*, vol. 2019, no. 1, pp. 1–13, Dec. 2019.
- [9] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.
- [10] X.-H. You, D.-M. Wang, B. Sheng, X.-Q. Gao, X.-S. Zhao, and M. Chen, "Cooperative distributed antenna systems for mobile communications [coordinated and distributed MIMO]," *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 35–43, Jun. 2010.
- [11] U. Madhoo, D. R. Brown, S. Dasgupta, and R. Mudumbai, "Distributed massive MIMO: Algorithms, architectures and concept systems," in *Proc. Inf. Theory Appl. Workshop (ITA)*, Feb. 2014, pp. 1–7.
- [12] Z. Chen and E. Björnson, "Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5205–5219, Nov. 2018.
- [13] S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: User-centric approach," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 706–709, Dec. 2017.
- [14] C. Zhu and W. Yu, "Stochastic modeling and analysis of user-centric network MIMO systems," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6176–6189, Dec. 2018.
- [15] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, Jan. 2020.
- [16] S. Govindasamy and I. Bergel, "Uplink performance of multi-antenna cellular networks with co-operative base stations and user-centric clustering," *IEEE Trans. Wireless Commun.*, vol. 17, no. 4, pp. 2703–2717, Apr. 2018.
- [17] S. Govindasamy and I. Bergel, "Asymptotic uplink rates in cooperative multiantenna cellular networks with mobiles at general locations," *IEEE Wireless Commun. Lett.*, vol. 9, no. 1, pp. 30–33, Jan. 2020.
- [18] J. Zhang, J. Zhang, E. Björnson, and B. Ai, "Local partial zero-forcing combining for cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 69, no. 12, pp. 8459–8472, Sep. 2021.

- [19] C. She, C. Yang, and T. Q. S. Quek, "Radio resource management for ultra-reliable and low-latency communications," *IEEE Commun. Mag.*, vol. 55, no. 6, pp. 72–78, Jun. 2017.
- [20] Z. Bai and J. W. Silverstein, "On the signal-to-interference-ratio of CDMA systems in wireless communications," *Ann. Appl. Probab.*, vol. 17, no. 1, pp. 81–101, 2007.
- [21] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. Cambridge, U.K.: Cambridge Univ. Press, 2011.
- [22] T. Tao, *Topics in Random Matrix Theory*, vol. 132. Providence, RI, USA: American Mathematical Society, 2012.
- [23] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and its Applications*. Hoboken, NJ, USA: Wiley, 1995.
- [24] M. Haenggi, J. G. Andrews, F. Baccelli, O. Dousse, and M. Franceschetti, "Stochastic geometry and random graphs for the analysis and design of wireless networks," *IEEE J. Sel. Areas Commun.*, vol. 27, no. 7, pp. 1029–1046, Sep. 2009.
- [25] J. Tang, G. Chen, J. P. Coon, and D. E. Simmons, "Distance distributions for Matérn cluster processes with application to network performance analysis," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2017, pp. 1–6.
- [26] J. Gil-Pelaez, "Note on the inversion theorem," *Biometrika*, vol. 38, pp. 481–482, Dec. 1951.
- [27] R. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1990.
- [28] Z. D. Bai and Y. Q. Yin, "Limit of the smallest eigenvalue of a large dimensional sample covariance matrix," *Ann. Probab.*, vol. 21, no. 3, pp. 1275–1294, 1993.
- [29] D. Paul and J. W. Silverstein, "No eigenvalues outside the support of the limiting empirical spectral distribution of a separable covariance matrix," *J. Multivariate Anal.*, vol. 100, no. 1, pp. 37–57, Jan. 2009.
- [30] W. Rudin, *Principles of Mathematical Analysis*. New York, NY, USA: McGraw-Hill, 1976.
- [31] O. B. S. Ali, C. Cardinal, and F. Gagnon, "Performance of optimum combining in a Poisson field of interferers and Rayleigh fading channels," *IEEE Trans. Wireless Commun.*, vol. 9, no. 8, pp. 2461–2467, Aug. 2010.
- [32] P. Billingsley, "Probability and measure," in *A Wiley-Interscience Publication*. Hoboken, NJ, USA: Wiley, 1995.



QINGMU DENG received the B.S. degree in electrical and computer engineering from the Olin College of Engineering, Needham, MA, USA, in 2021. Since June 2021, he has been with Cerebras Systems Inc., Sunnyvale, CA, USA, where he works on hardware design and analysis on and around Cerebras's wafer-scale processors.



SIDDHARTAN GOVINDASAMY (Member, IEEE) received the bachelor's, master's, and Ph.D. degrees from the Massachusetts Institute of Technology (MIT), in 1999, 2000, and 2008, respectively. At MIT, he worked on a master's thesis on speech enhancement algorithms in partnership with Qualcomm Inc. and a Ph.D. thesis on *ad-hoc* wireless networks with multi-antenna devices. From 2000 to 2003, he was a DSP Engineer and later a Senior DSP Engineer at Aware Inc. where

he worked on developing broadband modem technology. From 2008 to 2020, he was an Assistant Professor and later an Associate Professor of electrical and computer engineering at the Olin College of Engineering, Needham, MA, USA, where he conducted research on large multiple-input multiple-output (MIMO) systems and optical wireless communications. He joined the Department of Engineering, Boston College, as a Founding Faculty Member, in fall 2020, where he is a Professor of engineering. He is a coauthor of the textbook *Adaptive Wireless Communications: MIMO Channels and Networks* (Cambridge University Press, 2013). His research interests include in large MIMO systems and signal detection and processing.

• • •