Downlink Spectral Efficiency of Distributed Massive MIMO Systems with Linear Beamforming under Pilot Contamination

Jiamin Li, Dongming Wang, Pengcheng Zhu, *Member, IEEE,* Jiangzhou Wang, *Fellow, IEEE* and Xiaohu You, *Fellow, IEEE*

***Abstract*—In this paper, the downlink spectral efficiency of multi-cell multi-user distributed massive MIMO systems with linear beamforming is studied in the presence of pilot contamination. According to the levels of effective channel gain information at user side, we provide the lower bound and upper bound on user ergodic achievable downlink rate. Due to the different access distance from each user to different remote antenna units, the entries of user channel vectors are no longer identically distributed in distributed massive MIMO systems, which makes the spectral efficiency analysis challenging. Using the properties of Gamma distributions together with the approximate methods for non-isotropic vectors, we derive tractable but accurate closed-form expressions for the rate bounds with maximum ratio transmission (MRT) and zero-forcing (ZF) beamforming in distributed massive MIMO systems. Based on these expressions, user ultimate achievable rates are also given when the ratio of the total number of transmit antennas to the number of users goes to infinity. It is shown that MRT and ZF beamforming achieve the same ultimate rate no matter what levels of effective channel gain information at user side. Numerical results show that ZF achieves better performance gain and faster convergence speed than MRT. When the coherence interval is large, the downlink beamforming training scheme is more preferable for the distributed massive MIMO systems.**

***Index Terms*—Distributed massive MIMO; pilot contamination; spectral efficiency; Gamma distribution; non-isotropic vector**

I. Introduction

The use of massive antennas was first proposed for multi-cell multi-user cellular systems in [1] and since then, it has received much research interest [2]–[5]. It was shown that massive multi-input multi-output (MIMO) has very large performance gains compared with the conventional MIMO provided that a sufficiently large number of transmit antennas per active user are employed at each base station (BS). The very large signal vector dimension at a massive MIMO antenna array favors low complexity beamforming such as maximum ratio transmission (MRT) and zero-forcing (ZF), and as the number of transmit antennas becomes large, MRT and ZF become near optimal [4]. Therefore, in this paper, we mainly focus on the two linear beamforming schemes. MRT beamforming maximizes the gain of the desired signal and relies on that interfering signals are rejected automatically since the co-user channels are quasi-orthogonal when the number of transmit antennas is large. In contrast, ZF beamforming cancels the intra-cell interference by projecting the intended channel onto the null space of the subspace spanned by the channels of all other users inside the same cell. A large co-located antenna array (i.e., co-located massive MIMO) and a large geographically distributed antenna array (i.e., distributed massive MIMO) are two non-conflicting approaches to achieve the gains of massive MIMO and present the two extremes of massive MIMO paradigm [5], [6]. Compared with co-located massive MIMO, due to the increased macro-diversity gain and reduced access distance, distributed massive MIMO has the potential to improve spectral efficiency, system coverage, energy efficiency, and battery life of user terminals [7]–[13]. The main difference between co-located massive MIMO systems and distributed massive MIMO systems is that distributed massive MIMO systems suffer from different degrees of path loss caused by different access distances to distributed antenna arrays. Consequently, the channels of distributed massive MIMO systems are typically modeled as composite channels including uncorrelated large-scale fading and small-scale fading, which makes the performance analysis of distributed massive MIMO systems more challenging [14]–[16].

In order to realize the potential of massive MIMO, good enough channel state information (CSI) is required at the BSs and the users [3], [5]. Thanks to the effect of channel hardening, i.e., the effective channel gain seen by each user fluctuates only slightly around its mean when the number of antennas at the BS is very large, the users can reliably decode the transmitted signals from BSs based on only statistical CSI [17]. This is the reason that most previous studies on massive MIMO assumed that there is no need for users to estimate the CSI based on downlink training [18]–[23]. However, when the antennas of BSs are separately distributed in cells, each user may be effectively served by only a small number of remote antenna units (RAUs). As a result, the effect of channel hardening in distributed massive MIMO systems is less pronounced than that in co-located massive MIMO systems and the gain from estimating CSI at users become larger [24]. [25] investigated the problem of estimating CSI at

J. Li, D. Wang, P. Zhu and X. You are with National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China (e-mail:lijiamin, wangdm, p.zhu, xhyu@seu.edu.cn).

J. Wang is with the School of Engineering and Digital Arts, University of Kent, Canterbury CT2 7NT, U.K. (e-mail: j.z.wang@kent.ac.uk).

This work was supported in part by National Natural Science Foundation of China (NSFC) (Grant Nos. 61501113, 61271205, 61521061, 61372100, 61571120), and Jiangsu Provincial Natural Science Foundation (Grant Nos. BK20150630, BK20151415).

Manuscript received XXX, XX, 2016; revised XXX, XX, 2017.

users and proposed a downlink beamforming training scheme to acquire the estimates of the effective channel gain, defined as the inner product of channel vectors and beamforming vectors, instead of the CSI at each user. This scheme is feasible in practice since the channel estimation overhead is only proportional to the number of users. With this downlink beamforming training, there are basically three levels of effective channel gain information at user side: (i) "Statistical", without downlink pilots, users know only the statistical knowledge of effective channel gain; (ii) "Estimated", users are aware of the estimates of effective channel gain; (iii) "Perfect", users can estimate the effective channel gain perfectly with a genie receiver [26]. Note that (i) and (iii) are two extreme cases of (ii) which provide the lower bound and upper bound on user ergodic achievable downlink rate [24], respectively. Based on these rate bounds, we can analyze the spectral efficiency of distributed massive MIMO systems and evaluate the benefits of the downlink beamforming scheme. Consequently, in this paper, we focus on the cases (i) and (iii), and derive closed-form expressions for the rate bounds with MRT and ZF beamforming in multi-cell distributed massive MIMO systems.

In the case of (i), i.e., assuming that users detect the signals transmitted from BS based on only the statistical effective channel gain informantion, [18] proposed a lower bound on user ergodic achievable downlink rate which has been widely used to analyze the spectral efficiency of both co-located and distributed massive MIMO systems [19]–[23]. Practical per user power normalization was considered in [18] which can provide user fairness. However, [18] did not provide a general closed-form rate expression expect for a simple single-user scenario with MRT beamforming. In order to give analytical tractability, [19]–[23] assumed average transmit power normalization instead of per user power normalization. However, this assumption is only valid when the number of transmit antennas is very large [23]. Recently, under per user power normalization, [27] derived closed-form expressions for the lower bound in co-located massive MIMO systems. For co-located massive MIMO systems, channel vectors are isotropic, i.e., comprising independent and identically distributed (i.i.d.) entries. Available techniques for the spectral efficiency analysis of MIMO systems mostly assume channel vectors with i.i.d. entries, which simplifies the analysis significantly [28]–[30]. However, in distributed massive MIMO systems, each user suffers from different degrees of path losses caused by different access distances to different RAUs, and hence the entries of its aggregate channel vector to all RAUs are non-identically distributed, i.e., non-isotropic, in general. To the best of our knowledge, under per user power normalization, the closed-form expressions for the ergodic achievable downlink rate lower bound with both MRT and ZF beamforming in distributed massive MIMO systems are not available in the literature.

In the case of (iii), i.e., with the assumption of perfect effective channel gain information at users, the upper bound on user ergodic achievable downlink rate was obtained in [24]. Deriving the closed-form expressions for the upper bound is difficult since we need to characterize the distributions of the signal and interference powers which are related to the projection of non-isotropic channel vectors onto a beamforming subspace. [31]–[35] investigated this problem in some simplified scenarios, i.e., distriuted massive MIMO systems without pilot contamination and network MIMO systems with the assumption of perfect CSI at BSs. [36] extended the investigation to the case of imperfect CSI at BSs. However, the analytical expression in [36] has a prohibitively high complexity since Meijer's G-function defined by line integral in complex plane [37] is involved. Moreover, in the network MIMO systems considered in [32]–[36], data and CSI of all users are shared among BSs and BSs act as a single distributed multi-antenna transmitter to jointly serve the users in the coverage area. This means that the network MIMO system considered in [32]–[36] is just a single-cell case of the distributed massive MIMO systems. In the multi-cell case, there is pilot contamination, i.e., the correlated interference from other cells due to the reuse of the same pilot sequences, which has been a fundamental bottleneck of massive MIMO systems and makes the spectral efficiency analysis more challenging [1]. In addition, only ZF beamforming was considered in [31]–[36]. MRT beamforming, also called conjugate beamforming, is a particularly attractive beamforming scheme for massive MIMO systems because of its low computational complexity, robustness to channel impairments, and high asymptotic performance [4]. Thus, deriving the closed-form expression for the rate upper bound with MRT beamforming in multi-cell distributed massive MIMO systems is needed. In multi-cell distributed massive MIMO systems with pilot contamination, no simple closed-form expressions for the rate upper bound with both MRT and ZF beamforming have been given in the literature to the authors' best knowledge.

Herein, considering practical per user power normalization, we analyze the downlink spectral efficiency of multi-cell multi-user distributed massive MIMO systems with both MRT and ZF beamforming under pilot contamination. The main contributions of this paper are summarized as follows:

- Taking into consideration the effect of pilot contamination, we propose the method of characterizing the distributions of signal and interference powers in distributed massive MIMO systems with both MRT and ZF beamforming, which enables the spectral efficiency analysis of distributed massive MIMO systems in more practical scenarios.
- Considering practical per user power normalization, we derive tractable but accurate closed-form expressions for the rate bounds in distributed massive MIMO systems with both MRT and ZF beamforming, thereby enabling the spectral efficiency analysis of distributed massive MIMO systems and the efficient evaluation of the benefits of estimating the effective channel gain information at user side.
- Simplified closed-form expressions for the ergodic achievable downlink rate upper bound are derived based on some approximation techniques, which achieve nearly the same performance with much less complexity.
- Based on these derived expressions, we give the user ultimate achievable rate, from which we can analyze the

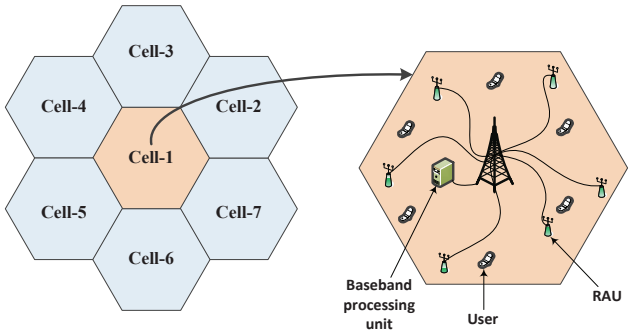


Fig. 1. System Configuration.

asymptotic performance of distributed massive MIMO systems with different beamforming and different levels of effective channel gain information at user side.

- We corroborate our analysis by performing simulations coinciding with analytical expressions, and draw insightful conclusions from the comparison between distributed and co-located massive MIMO systems and the analysis of how the coherence interval affects the per user spectral efficiency with both MRT and ZF beamforming.

The remainder of the paper is organized as follows. In Section II, we describe the system model including the system configuration, channel model, and channel estimation with pilot contamination. Section III contains the analytical work where closed-form expressions for the lower bound and upper bound on ergodic achievable downlink rate are derived. Representative numerical results are given in Section IV before we conclude the paper in Section V.

The following notations are used. All boldface letters stand for vectors (lower case) or matrices (upper case). $\mathbf{I}_N$ is the size-N identity matrix. Italic letters (e.g., X or x) denote scalars. The transpose, Hermitian transpose and trace operators are denoted by $(\cdot)^{\mathrm{T}}$, $(\cdot)^{\mathrm{H}}$ and $\mathrm{tr}\,(\cdot)$, respectively. $\mathbb{C}^{m\times n}$ denotes the set of $m \times n$ complex valued matrices. $|x|$ is the absolute value of a scalar x, $\|\mathbf{X}\|$ is the spectral norm of a matrix $\mathbf{X}$. $x \sim \mathcal{CN}(0,\sigma^2)$ means that x is a circularly symmetric complex Gaussian random variable with mean zero and variance σ^2. $\mathbb{E}\,[\cdot]$ denotes the expectation operator. A random variable X which follows a Gamma distribution with shape parameter k and scale parameter θ is denoted by $X \sim \Gamma(k,\theta)$. Nakagami(m,Ω) denotes Nakagami distribution with shape parameter m and controlling spread parameter Ω.

II. System Model

In the first part of this section, we describe the system configuration and present the mathematical description of the channel model. The uplink channel estimate with pilot contamination and the analysis of pilot contamination effect are given in the second part of this section.

A. System Configuration and Channel Model

Consider a distributed massive MIMO system with L adjacent cells, as illustrated in Fig. 1. Each cell consists of M RAUs equipped with N antennas and K single-antenna users which share the same bandwidth. The system configuration specified above is denoted by (M,N,K). As an example, a $(7,3,6)$ system is shown on the right side of Fig. 1, where $M=7$, $N=3$ and $K=6$. This system configuration is quite general, with traditional CAS [23] $(M=1)$, DAS with fully distributed antennas [38], [39] $(N=1)$ and network MIMO [36] $(L=1)$ as special cases. Consequently, the results obtained in this paper can also be applied to the above special cases. We consider transmissions over frequency-flat fading channels, and assume that the system operates in time-division duplex (TDD) mode, not frequency-division duplex in legacy mobile systems [40]–[42]. The channel vector from the k-th user in the l-th cell to all of the RAUs in the i-th cell is denoted as

$$\mathbf{g}_{i,l,k} = \left[\sqrt{\lambda_{i,1,l,k}}\mathbf{h}_{i,1,l,k}^{\mathrm{T}} \cdots \sqrt{\lambda_{i,M,l,k}}\mathbf{h}_{i,M,l,k}^{\mathrm{T}}\right]^{\mathrm{T}}, \quad (1)$$

where $\lambda_{i,m,l,k} \triangleq c d_{i,m,l,k}^{-\alpha} s_{i,m,l,k}$ represents the large-scale and shadow fading between the k-th user in the l-th cell and the m-th RAU in the i-th cell which change slowly and can be learned over long period of time, c is the median of the mean path gain at a reference distance $d_{i,m,l,k}=1$ km, α is the path loss exponent, typically between 3.0 and 5.0, $s_{i,m,l,k}$ is a log-normal shadow fading variable, $\mathbf{h}_{i,m,l,k}$ represents the small-scale fast fading, and it is a vector with size N which contains i.i.d. zero mean circularly symmetric complex Gaussian (ZMCSCG) random variables with unit variance.

B. Channel Estimation with Pilot Contamination

We focus on the case where pilot symbol aided transmission is employed to assist the BS in performing channel estimation. In multi-cell TDD massive MIMO systems, due to the limited channel coherence time, non-orthogonal pilot sequences must be reused to estimate the CSI in adjacent cells. This leads to channel estimation impairments known as pilot contamination [1], which has been a fundamental bottleneck of massive MIMO systems. Several techniques have been proposed to mitigate the pilot contamination effect such as time-shifted pilot protocol [43], pilot reuse design [44], [45], eigenvalue-decomposition-based method [46] and pilot contamination precoding [47]. It was shown that the effect of pilot contamination can be mitigated completely based on several critical but optimistic assumptions. However, in realistic cases, e.g., the number of BS antennas is large but still finite, the effect of pilot contamination should also be considered.

Considering minimum mean-square error (MMSE) channel estimator, the channel estimate of (1) with pilot contamination has been given in [22] by

$$\hat{\mathbf{g}}_{i,l,k} = \left[\sqrt{\beta_{i,1,l,k}}\hat{\mathbf{h}}_{i,k,1}^{\mathrm{T}} \cdots \sqrt{\beta_{i,M,l,k}}\hat{\mathbf{h}}_{i,k,M}^{\mathrm{T}}\right]^{\mathrm{T}}, \quad (2)$$

where

$$\beta_{i,m,l,k} \triangleq \lambda_{i,m,l,k}^2 \Big(\sum\nolimits_{j=1}^{L} \lambda_{i,m,j,k} + 1/\gamma_{\mathrm{P}}\Big)^{-1}, \quad (3)$$

γ_{P} is the training SNR, and $\hat{\mathbf{h}}_{i,k} \triangleq \left[\hat{\mathbf{h}}_{i,k,1}^{\mathrm{T}}, \cdots, \hat{\mathbf{h}}_{i,k,M}^{\mathrm{T}}\right]^{\mathrm{T}} \sim \mathcal{CN}\left(0, \mathbf{I}_{MN}\right)$ represents the equivalent Rayleigh fading part of the estimated channel. Herein we implicitly assume that

the realizations of $\lambda_{i,m,l,k}$ are perfectly known to the BSs as in [1], [18]. It can be seen that, due to the effect of pilot contamination, the equivalent Rayleigh fading part is not related to the second subscript of the estimated channel. Consequently, although the channel vectors $\mathbf{g}_{i,l,k}$ and $\mathbf{g}_{i,j,k}$ are independent for $j \neq l$, the estimated channel vectors $\hat{\mathbf{g}}_{i,l,k}$ and $\hat{\mathbf{g}}_{i,j,k}$ become correlated random vectors with the following correlation matrix

$$\mathrm{cov}\left(\hat{\mathbf{g}}_{i,l,k}, \hat{\mathbf{g}}_{i,j,k}\right) = \mathbf{\Lambda}_{i,l,k}\mathbf{Q}_{i,k}^{-1}\mathbf{\Lambda}_{i,j,k}, \tag{4}$$

where $\mathbf{\Lambda}_{i,l,k} = \mathrm{diag}\left(\left[\lambda_{i,1,l,k}\cdots\lambda_{i,M,l,k}\right]^{\mathrm{T}}\right) \otimes \mathbf{I}_N$ and $\mathbf{Q}_{i,k} \triangleq \sum_{l=1}^{L}\mathbf{\Lambda}_{i,l,k} + 1/\gamma_{\mathrm{P}}\mathbf{I}_{MN}$.

Moreover, from the orthogonality property of MMSE estimate, the channel $\mathbf{g}_{i,l,k}$ can be decomposed as

$$\mathbf{g}_{i,l,k} = \hat{\mathbf{g}}_{i,l,k} + \tilde{\mathbf{g}}_{i,l,k}, \tag{5}$$

where

$$\tilde{\mathbf{g}}_{i,l,k} \sim \mathcal{CN}\Big(0, \mathrm{diag}\big(\left[\eta_{i,1,l,k}, \cdots, \eta_{i,M,l,k}\right]^{\mathrm{T}}\big) \otimes \mathbf{I}_N\Big) \tag{6}$$

is the uncorrelated estimation error which is statistically independent of $\hat{\mathbf{g}}_{i,l,k}$ due to the joint Gaussianity of both vectors, and

$$\eta_{i,m,l,k} \triangleq \lambda_{i,m,l,k} - \beta_{i,m,l,k}. \tag{7}$$

Similarly, due to the effect of pilot contamination, $\tilde{\mathbf{g}}_{i,l,k}$ and $\tilde{\mathbf{g}}_{i,j,k}$ also become correlated random vectors given by

$$\mathrm{cov}\left(\tilde{\mathbf{g}}_{i,l,k}, \tilde{\mathbf{g}}_{i,j,k}\right) = \begin{cases} -\mathbf{\Lambda}_{i,l,k}\mathbf{Q}_{i,k}^{-1}\mathbf{\Lambda}_{i,j,k}, & l \neq j, \\ \mathbf{\Lambda}_{i,l,k} - \mathbf{\Lambda}_{i,l,k}\mathbf{Q}_{i,k}^{-1}\mathbf{\Lambda}_{i,j,k}, & l = j. \end{cases} \tag{8}$$

The effect of pilot contamination presented in (4) and (8) makes the spectral efficiency analysis in Section III more challenging.

III. User Ergodic Achievable Downlink Rate

In this section, we derive the closed-form expressions for the lower bound and upper bound on user ergodic achievable downlink rate with both MRT and ZF beamforming in multi-cell multi-user distributed massive MIMO systems in the presence of pilot contamination. The first part of this section describes the downlink signal model. In the second part, we derive isotropic approximation for the non-isotropic channel vector in distributed massive MIMO systems after presenting some related lemmas. In the third and fourth parts, according to the levels of effective channel gain information at user side, we first provide lower bound and upper bound on user ergodic achievable downlink rate, and then based on the properties of Gamma distributions and non-isotropic channel approximation techniques we derive closed-form expressions for these rate bounds with both MRT and ZF beamforming. Moreover, some approximation techniques are utilized to further simplify the closed-form expressions for the upper bound. Based on the derived expressions, user ultimate achievable rates are also given.

A. Downlink Signal Model

The received signal $y_{l,k} \in \mathbb{C}$ of the k-th user in the l-th cell can be written as

$$y_{l,k} = \underbrace{\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}s_{l,k}}_{\text{desired signal}} + \underbrace{\sum\nolimits_{(i,j)\neq(l,k)}\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}s_{i,j}}_{\text{intra-cell and inter-cell interference}} + z_{l,k}, \tag{9}$$

where $\mathbf{w}_{l,k} \in \mathbb{C}^{MN\times 1}$ is the beamforming vector assigned for the k-th user in the l-th cell, $s_{l,k} \sim \mathcal{CN}(0,1)$ is the associated data symbol, $z_{l,k} \sim \mathcal{CN}(0, 1/\gamma_{\mathrm{DL}})$ indicates the complex additive white Gaussian noise (AWGN), γ_{DL} represents the downlink SNR after normalizing the transmit power per user.

This paper restricts attention to two linear beamformers of practical interest, namely MRT and ZF, because they are relatively easy to implement and analyze. Mathematically, the MRT and ZF beamforming vectors are defined as

$$\mathbf{w}_{l,k} = \begin{cases} \dfrac{\hat{\mathbf{g}}_{l,l,k}}{\|\hat{\mathbf{g}}_{l,l,k}\|}, & \text{for MRT,} \\ \dfrac{\mathbf{a}_{l,l,k}}{\|\mathbf{a}_{l,l,k}\|}, & \text{for ZF.} \end{cases} \tag{10}$$

respectively, where $\mathbf{a}_{l,l,k}$ is the k-th column of $\hat{\mathbf{G}}_l\big(\hat{\mathbf{G}}_l^{\mathrm{H}}\hat{\mathbf{G}}_l\big)^{-1}$, and $\hat{\mathbf{G}}_l \triangleq \left[\hat{\mathbf{g}}_{l,l,1}, \cdots, \hat{\mathbf{g}}_{l,l,K}\right]$ is the estimated compound channel matrix.

B. Isotropic Approximation for Channel Vectors

In this subsection, we give the isotropic approximation for the non-isotropic channel vector in distributed massive MIMO systems which will be required to derive closed-form expressions for the rate bounds on ergodic achievable downlink rate.

We first consider the case when the channel vectors are isotropic, i.e., comprising i.i.d. entries. Let $\mathbf{x} \sim \mathbb{C}^{m\times 1}$ is isotropic with i.i.d. $\mathcal{CN}(0,\sigma^2)$ elements, then $\mathbf{x}^{\mathrm{H}}\mathbf{x}$ is the summation of m i.i.d. $\Gamma(1,\sigma^2)$ random variables. Thus, we have $\mathbf{x}^{\mathrm{H}}\mathbf{x} \sim \Gamma(m,\sigma^2)$ [22]. It can be seen that, each of the m spatial dimensions adds one to the shape parameter of the power distribution. Further, about the projection of isotropic vector, we have the following lemma.

Lemma 1 ([33], [48]): For an m-dimensional vector $\mathbf{x} \sim \mathbb{C}^{m\times 1}$ with i.i.d. $\mathcal{CN}(0,\sigma^2)$ elements, when projected onto an s-dimensional subspace, its power is distributed as $\Gamma\left(s,\sigma^2\right)$.

Remark 1: From the perspective of each user, an intended beam lies in a subspace of dimension $s = MN$ with MRT beamforming and $s = MN - K + 1$ with ZF beamforming, respectively, whereas any independent vectors lie in a one-dimensional subspace [35], [49].

From Lemma 1, we can find that each of the s spatial dimensions also contributes one to the shape parameter of the projection power distribution.

Lemma 1 is applied only when the channel vectors are isotropic. Thus, it cannot be utilized to characterize the distributions of signal and interference powers in distributed massive MIMO systems directly. In distributed massive MIMO systems, for the channel strength from the RAUs in the i-th

cell to the k-th user in the l-th cell, we have

$$\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{g}_{i,l,k} = \sum_{m=1}^{M} \lambda_{i,m,l,k}\mathbf{h}_{i,m,l,k}^{\mathrm{H}}\mathbf{h}_{i,m,l,k}, \tag{11}$$

which is a sum of M independent and non-identically distributed variables where the m-th variable is distributed as $\lambda_{i,m,l,k}\mathbf{h}_{i,m,l,k}^{\mathrm{H}}\mathbf{h}_{i,m,l,k} \sim \Gamma(N, \lambda_{i,m,l,k})$ [15]. The exact distribution of the sum of independent and non-identically distributed Gamma random variables can be found in [50], however, it will not yield a mathematically tractable expression. Therefore, we employ the second-order matching technique shown in the following lemma to obtain an approximation distribution.

Lemma 2 ([31]): If $\{x_i\}$ are independent Gamma distributed random variables with shape and scale parameters k_i and θ_i, i.e., $x_i \sim \Gamma(k_i, \theta_i)$, the sum $\sum_i x_i$ can be approximated as another Gamma distributed random variable which has the same first and second order moments, with the shape and scale parameters given by $k = \left(\sum_i k_i\theta_i\right)^2 \Big/ \sum_i k_i\theta_i^2$ and $\theta = \sum_i k_i\theta_i^2 \big/ \sum_i k_i\theta_i$.

As a consequence of Lemma 2, the distribution of $\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{g}_{i,l,k}$ can be approximated as

$$\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{g}_{i,l,k} \sim \Gamma(k_{i,l,k,\mathrm{a}}, \theta_{i,l,k,\mathrm{a}}), \tag{12}$$

where

$$k_{i,l,k,\mathrm{a}} = \frac{N(\sum_{m=1}^{M}\lambda_{i,m,l,k})^2}{\sum_{m=1}^{M}\lambda_{i,m,l,k}^2}, \tag{13}$$

$$\theta_{i,l,k,\mathrm{a}} = \frac{\sum_{m=1}^{M}\lambda_{i,m,l,k}^2}{\sum_{m=1}^{M}\lambda_{i,m,l,k}}, \tag{14}$$

and the letter "a" in the subscript means approximation. From (13), it can be seen that, $k_{i,l,k,\mathrm{a}} \leq MN$ with the equality if $\mathbf{g}_{i,l,k}$ is isotropic which means that the non-isotropic nature of $\mathbf{g}_{i,l,k}$ reduces the contribution of each spatial dimension to the shape parameter of the resulting power distribution.

Based on the analysis above and inspired by [31], [33], we characterize the distributions of the powers of the non-isotropic channel vectors $\mathbf{g}_{i,l,k}$ with dimension $m = MN$ projected onto an s-dimensional beamforming subspace as follows. First, given the approximate distribution of $\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{g}_{i,l,k}$ in (12), we approximate the non-isotropic channel vector $\mathbf{g}_{i,l,k}$ as an isotropic vector $\mathbf{g}_{i,l,k,\mathrm{a}}$ with i.i.d. $\mathcal{CN}(0, \theta_{i,l,k,\mathrm{a}})$ elements while the distribution of the projection of $\mathbf{g}_{i,l,k,\mathrm{a}}$ onto an s-dimensional subspace is distributed as $\Gamma(s, \theta_{i,l,k,\mathrm{a}})$ from Lemma 1. Then, considering that the contribution of each spatial dimension is reduced when $\mathbf{g}_{i,l,k}$ is non-isotropic, we approximate the distribution of the powers of non-isotropic vector $\mathbf{g}_{i,l,k}$ projected onto an s-dimensional beamforming subspace as $\Gamma(s\varphi, \theta_{i,l,k,a})$, where $\varphi \leq 1$ denotes the contribution of each spatial dimension when the vectors are non-isotropic. φ can be solved by matching the mean in the case with MRT beamforming [31], i.e., solve $\mathbb{E}[|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,k}^{\mathrm{MRT}}|^2] = mr\theta_{i,l,k,\mathrm{a}}$ for r, where $mr\theta_{i,l,k,\mathrm{a}}$ is the mean of the projection power since it distributed as $\Gamma(m\varphi, \theta_{i,l,k,a})$. In this situation, from (10) and (12), we obtain $|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,k}^{\mathrm{MRT}}|^2 = \mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{g}_{i,l,k} \sim \Gamma(k_{i,l,k,\mathrm{a}}, \theta_{i,l,k,\mathrm{a}})$. Thus, we have $m\varphi\theta_{i,l,k,\mathrm{a}} = k_{i,l,k,\mathrm{a}}\theta_{i,l,k,\mathrm{a}}$ and obtain $\varphi = k_{i,l,k,\mathrm{a}}/m$.

In light of the discussion above, we have the following lemma.

Lemma 3: For the m-dimensional non-isotropic channel vector $\mathbf{g}_{i,l,k}$, when projected onto an s-dimensional subspace, the distribution of the projection power can be approximated as $\Gamma(sk_{i,l,k,\mathrm{a}}/m, \theta_{i,l,k,\mathrm{a}})$ with $k_{i,l,k,\mathrm{a}}$ and $\theta_{i,l,k,\mathrm{a}}$ defined in (13) and (14).

Remark 2: Lemma 3 provides a good approximation when the path loss to each RAU is similar; otherwise it overpredicts the degrees of freedom. We find in simulation that the approximation is quite good for useful signal powers and a little worse for pilot contamination since the distances from each user to the RAUs in the interfering cells is relative large and vary drastically. However, as shown in the simulations, we can increase the approximation accuracy by adjusting the effective dimension of the projection subspace.

From Lemma 1 and Lemma 3, it can be seen that the distributions of the projections of channel vectors onto an s-dimensional subspace can be characterized in terms of Gamma random variables, regardless of whether the channel vectors are isotropic or non-isotropic. The main difference is that the shape parameter is changed from s to $sk_{i,l,k,\mathrm{a}}/m$ when the channel vectors are non-isotropic. In other words, the non-isotropic nature of the channels in distributed massive MIMO systems is captured by changing the contribution of each spatial dimension to the shape parameter of the resulting Gamma distribution from 1 to $k_{i,l,k,\mathrm{a}}/m$.

Different from [32]–[35] where perfect CSI was assumed to be available at transmitters, this paper focus on a practical case of imperfect CSI in the presence of pilot contamination. Thus, we need to characterize the distributions of the powers of the non-isotropic channel estimation vectors $\hat{\mathbf{g}}_{i,l,k}$ and the estimation error vectors $\tilde{\mathbf{g}}_{i,l,k}$ projected onto an s-dimensional beamforming subspace further. Using Lemma 2, the approximation distributions of $\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\hat{\mathbf{g}}_{i,l,k}$ and $\tilde{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\tilde{\mathbf{g}}_{i,l,k}$ can be given by

$$\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\hat{\mathbf{g}}_{i,l,k} \sim \Gamma(\hat{k}_{i,l,k,\mathrm{a}}, \hat{\theta}_{i,l,k,\mathrm{a}}), \tag{15}$$

$$\tilde{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\tilde{\mathbf{g}}_{i,l,k} \sim \Gamma(\tilde{k}_{i,l,k,\mathrm{a}}, \tilde{\theta}_{i,l,k,\mathrm{a}}), \tag{16}$$

where

$$\hat{k}_{i,l,k,\mathrm{a}} = \frac{N(\sum_{m=1}^{M}\beta_{i,m,l,k})^2}{\sum_{m=1}^{M}\beta_{i,m,l,k}^2}, \tag{17}$$

$$\hat{\theta}_{i,l,k,\mathrm{a}} = \frac{\sum_{m=1}^{M}\beta_{i,m,l,k}^2}{\sum_{m=1}^{M}\beta_{i,m,l,k}}, \tag{18}$$

$$\tilde{k}_{i,l,k,\mathrm{a}} = \frac{N(\sum_{m=1}^{M}\eta_{i,m,l,k})^2}{\sum_{m=1}^{M}\eta_{i,m,l,k}^2}, \tag{19}$$

$$\tilde{\theta}_{i,l,k,\mathrm{a}} = \frac{\sum_{m=1}^{M}\eta_{i,m,l,k}^2}{\sum_{m=1}^{M}\eta_{i,m,l,k}}, \tag{20}$$

Given these distributions, the distributions of projection powers can be obtained by applying Lemma 3.

Based on the analysis above, we are now ready to derive the closed-form expressions for the ergodic rate lower bound and upper bound in the following subsections.

C. Lower Bound on Ergodic Achievable Rate

It is assumed that users do not have any channel estimate and detect the transmitted signals from BSs with statistical

effective channel gain, i.e., the mean of the inner product of channel vectors and beamforming vectors $\mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}]$. Based on the techniques developed in [18, Theorem 1], we provide a lower bound on user ergodic achievable downlink rate. Considering the signal component received over the effective channel gain mean $\mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}]$ is the only desired signal, (9) can be rewritten as

$$y_{l,k} = \mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}]s_{l,k} + (\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k} - \mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}])s_{l,k} + \sum_{(i,j)\neq(l,k)} \mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}s_{i,j} + z_{l,k}. \quad (21)$$

By treating the interference plus noise $\sum_{(i,j)\neq l,k}\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}s_{i,j} + z_{l,k}$ and the remaining signal component $\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}s_{l,k} - \mathbb{E}\big[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}\big]s_{l,k}$ as worst-case Gaussian distributed noise [18], [23], the ergodic achievable rate lower bound of the k-th user in the l-th cell can be given by

$$R_{l,k}^{\mathrm{LB}} = \log_2\left(1+\gamma_{l,k}^{\mathrm{LB}}\right), \quad (22)$$

where

$$\gamma_{l,k}^{\mathrm{LB}} = \frac{|\mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}]|^2}{\mathrm{var}\big[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}\big] + \sum_{(i,j)\neq(l,k)}\mathbb{E}\big[|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}|^2\big] + \frac{1}{\gamma_{\mathrm{DL}}}}.$$

Note that [18] did not provide a general closed-form expression for (22), and in order to give analytical tractability, [19]–[23] all assumed average transmit power normalization instead of the per user power normalization considered in [18] which can provide user fairness. In the following theorems, considering per user power normalization, we derive the closed-form expressions for the lower bound (22) with MRT and ZF beamforming in distributed massive MIMO systems.

Theorem 1: Considering practical per user power normalization, the closed-form expression for the rate bound (22) of the k-th user in the l-th cell with MRT beamforming is given by

$$R_{l,k}^{\mathrm{LB,MRT}} = \log_2\left(1+\frac{\xi(\hat{k}_{l,l,k,\mathrm{a}})\hat{\theta}_{l,l,k,\mathrm{a}}}{\mathcal{I}_{l,k}^{\mathrm{LB,MRT}} + \sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}}\right), \quad (23)$$

where $\mathcal{I}_{l,k}^{\mathrm{LB,MRT}} \triangleq \hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}} + \frac{1}{MN}\sum_{i=1}^{L}\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}} + \frac{K-1}{MN}\big(k_{l,l,k,\mathrm{a}}\theta_{l,l,k,\mathrm{a}} + \sum_{i\neq l}(\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}} + \tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}})\big) - \xi(\hat{k}_{l,l,k,\mathrm{a}})\hat{\theta}_{l,l,k,\mathrm{a}} + 1/\gamma_{\mathrm{DL}}$, and

$$\xi(x) \triangleq \frac{\Gamma(x+1/2)}{\Gamma(x)}. \quad (24)$$

Proof: The proof is given in Appendix A. ∎

Theorem 2: Considering practical per user power normalization, the closed-form expression for the rate bound (22) of the k-th user in the l-th cell with ZF beamforming is given by

$$R_{l,k}^{\mathrm{LB,ZF}} = \log_2\left(1+\frac{\xi(\frac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}})\hat{\theta}_{l,l,k,\mathrm{a}}}{\mathcal{I}_{l,k}^{\mathrm{LB,ZF}} + \sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}}\right), \quad (25)$$

where $\mathcal{I}_{l,k}^{\mathrm{LB,ZF}} \triangleq \frac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}} - \xi(\frac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}})\hat{\theta}_{l,l,k,\mathrm{a}} + \frac{K}{MN}\sum_{i=1}^{L}\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}} + 1/\gamma_{\mathrm{DL}}$, and

$$\rho \triangleq MN - K + 1. \quad (26)$$

Proof: The proof is given in Appendix B. ∎

Remark 3: Practical per user power normalization was considered in this paper. Considering that $\|\hat{\mathbf{g}}_{l,l,k}\|^2$ and $\|\mathbf{a}_{l,l,k}\|^2$ fluctuate only slightly around their means $\mathbb{E}[\|\hat{\mathbf{g}}_{l,l,k}\|^2]$ and $\mathbb{E}[\|\mathbf{a}_{l,l,k}\|^2]$ as the number of transmit antennas goes to infinity, i.e.,

$$\frac{|\mathbb{E}[\|\mathbf{v}_{l,l,k}\|^2] - \|\mathbf{v}_{l,l,k}\|^2|}{\mathbb{E}[\|\mathbf{v}_{l,l,k}\|^2]} \xrightarrow{MN\to\infty} 0, \quad (27)$$

where $\mathbf{v}_{l,l,k} = \hat{\mathbf{g}}_{l,l,k}$ or $\mathbf{a}_{l,l,k}$, [19]–[23] assumed average power normalization instead of per user power normalization in (10) to give analytical tractability. Based on this assumption, the MRT and ZF beamforming vectors defined in (10) can be rewritten as

$$\mathbf{w}_{l,k} = \begin{cases} \dfrac{\hat{\mathbf{g}}_{l,l,k}}{\sqrt{\mathbb{E}[\|\hat{\mathbf{g}}_{l,l,k}\|]}}, & \text{for MRT,} \\ \dfrac{\mathbf{a}_{l,l,k}}{\sqrt{\mathbb{E}[\|\mathbf{a}_{l,l,k}\|]}}, & \text{for ZF.} \end{cases} \quad (28)$$

The difference in performance between average power normalization in (28) and per user power normalization in (10) is negligible in the context of co-located massive MIMO systems (the number of transmit antennas is very large) with MRT and ZF beamforming vectors. However, this assumption is only valid when the number of transmit antennas is very large [23]. Moreover, when the antennas of BSs are separately distributed in cells, each user may be effectively served by only a small number of RAUs. As a result, the difference in performance between average power normalization and per user power normalization will become larger [24]. From the numerical example (not shown here due to the space constraints), it can be seen that, in order to achieve 10% approximation error (defined in (27)), about 60 antennas are needed for co-located systems, while about 200 antennas are needed for distributed systems with $M = 10$ RAUs.

Next, we consider the case when the total number of transmit antennas is much larger than the number of users, i.e., $\frac{MN}{K} \to \infty$.

Corollary 1: Let $\frac{MN}{K} \to \infty$, the rate bounds (22) of the k-th user in the l-th cell with MRT and ZF beamforming approach the same ultimate achievable rate $R_{l,k}^{\mathrm{LB},\,\infty}$ as

$$R_{l,k}^{\mathrm{LB},\,\infty} = \log_2\left(1+\frac{\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}}}{\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}}\right). \quad (29)$$

Proof: Due to the similarity, we only provide the proof for MRT beamforming in the following. From the expression of $\hat{k}_{i,l,k,\mathrm{a}}$ in (17), it can be seen that $\hat{k}_{i,l,k,\mathrm{a}} \to \infty$ as $MN \to \infty$. Therefore, we have $\lim_{MN\to\infty}\left(\xi(\hat{k}_{l,l,k,\mathrm{a}}) - \hat{k}_{l,l,k,\mathrm{a}}\right) = 0$ which results from $\lim_{x\to\infty}\xi(x) = x$ [18, Theorem 4]. Then, replacing $\hat{k}_{i,l,k,\mathrm{a}}$ with $MN\frac{(\sum_{m=1}^{M}\beta_{i,m,l,k}/M)^2}{\sum_{m=1}^{M}\beta_{i,m,l,k}^2/M}$ in (23), and since the mean terms $\frac{(\sum_{m=1}^{M}\beta_{i,m,l,k}/M)^2}{\sum_{m=1}^{M}\beta_{i,m,l,k}^2/M}$ are finite values, it is straightforward to obtain the ultimate rate of the k-th user in the l-th cell in (29) by dividing the denominator and numerator of (23) by $\frac{MN}{K}$ and letting $\frac{MN}{K} \to \infty$. ∎

In the following subsection, we first derive closed-form expressions for the upper bound on ergodic achievable downlink rate with both MRT and ZF beamforming in distributed

$$k_z^{\text{ZF}} = \frac{\left(\frac{\rho}{MN}\sum_i \hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}} + \frac{K}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}} + \frac{(K-1)}{MN}\sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}}\right)^2}{\frac{\rho}{MN}\sum_i \hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}} + \frac{K}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}} + \frac{(K-1)}{MN}\sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}}}, \tag{33}$$

$$\theta_z^{\text{ZF}} = \gamma_{\text{DL}}\frac{\frac{\rho}{MN}\sum_i \hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}} + \frac{K}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}} + \frac{(K-1)}{MN}\sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}}}{\frac{\rho}{MN}\sum_i \hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}} + \frac{K}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}} + \frac{(K-1)}{MN}\sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}}}, \tag{34}$$

$$k_y^{\text{ZF}} = \frac{\left(\frac{K}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}} - \frac{1}{MN}\tilde{k}_{l,l,k,\text{a}}\tilde{\theta}_{l,l,k,\text{a}} + \sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}}\right)^2}{\frac{K}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}} - \frac{1}{MN}\tilde{k}_{l,l,k,\text{a}}\tilde{\theta}^2_{l,l,k,\text{a}} + \sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}}}, \tag{35}$$

$$\theta_y^{\text{ZF}} = \gamma_{\text{DL}}\frac{\frac{K}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}} - \frac{1}{MN}\tilde{k}_{l,l,k,\text{a}}\tilde{\theta}^2_{l,l,k,\text{a}} + \sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}}}{\frac{K}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}} - \frac{1}{MN}\tilde{k}_{l,l,k,\text{a}}\tilde{\theta}_{l,l,k,\text{a}} + \sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}}}, \tag{36}$$

$$k_z^{\text{MRT}} = \frac{\left(\frac{K-1}{MN}k_{l,l,k,\text{a}}\theta_{l,l,k,\text{a}} + \sum_i \hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}} + \frac{1}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}} + \frac{K-1}{MN}\sum_{i\neq l}(\hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}} + \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}})\right)^2}{\frac{K-1}{MN}k_{l,l,k,\text{a}}\theta^2_{l,l,k,\text{a}} + \sum_i \hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}} + \frac{1}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}} + \frac{K-1}{MN}\sum_{i\neq l}(\hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}} + \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}})}, \tag{38}$$

$$\theta_z^{\text{MRT}} = \gamma_{\text{DL}}\frac{\frac{K-1}{MN}k_{l,l,k,\text{a}}\theta^2_{l,l,k,\text{a}} + \sum_i \hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}} + \frac{1}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}} + \frac{K-1}{MN}\sum_{i\neq l}(\hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}} + \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}})}{\frac{K-1}{MN}k_{l,l,k,\text{a}}\theta_{l,l,k,\text{a}} + \sum_i \hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}} + \frac{1}{MN}\sum_i \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}} + \frac{K-1}{MN}\sum_{i\neq l}(\hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}} + \tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}})}, \tag{39}$$

$$k_y^{\text{MRT}} = \frac{\left(\frac{K-1}{MN}k_{l,l,k,\text{a}}\theta_{l,l,k,\text{a}} + \frac{MN+K-1}{MN}\sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}} + \frac{K}{MN}\sum_{i\neq l}\tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}}\right)^2}{\frac{K-1}{MN}k_{l,l,k,\text{a}}\theta^2_{l,l,k,\text{a}} + \frac{MN+K-1}{MN}\sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}} + \frac{K}{MN}\sum_{i\neq l}\tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}}}, \tag{40}$$

$$\theta_y^{\text{MRT}} = \gamma_{\text{DL}}\frac{\frac{K-1}{MN}k_{l,l,k,\text{a}}\theta^2_{l,l,k,\text{a}} + \frac{MN+K-1}{MN}\sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}^2_{i,l,k,\text{a}} + \frac{K}{MN}\sum_{i\neq l}\tilde{k}_{i,l,k,\text{a}}\tilde{\theta}^2_{i,l,k,\text{a}}}{\frac{K-1}{MN}\sum_i k_{l,l,k,\text{a}}\theta_{l,l,k,\text{a}} + \frac{MN+K-1}{MN}\sum_{i\neq l}\hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}} + \frac{K}{MN}\sum_{i\neq l}\tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}}}. \tag{41}$$

massive MIMO systems in the presence of pilot contamination, and then based on some approximation techniques we further simplify these expressions.

D. Upper Bound on Ergodic Achievable Rate

Assuming that users can estimate the channel gain perfectly using downlink pilots with a genie receiver [26], we can obtain the upper bound $\text{R}_{l,k}^{\text{UB}}$ on the ergodic achievable downlink rate of the k-th user in the l-th cell given by [24]

$$R_{l,k}^{\text{UB}} = \mathbb{E}\big[\log_2\left(1+\gamma_{l,k}^{\text{UB}}\right)\big], \tag{30}$$

where

$$\gamma_{l,k}^{\text{UB}} = \frac{|\mathbf{g}_{l,l,k}^{\text{H}}\mathbf{w}_{l,k}|^2}{\sum_{(i,j)\neq(l,k)}|\mathbf{g}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}|^2 + \frac{1}{\gamma_{\text{DL}}}}.$$

In order to derive an accurate analytical expression for (30), we need to know the distributions of the desired signal term, i.e., $|\mathbf{g}_{l,l,k}^{\text{H}}\mathbf{w}_{l,k}|^2$, and the residual interference term, i.e., $\sum_{(i,j)\neq(l,k)}|\mathbf{g}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}|^2$. In the following theorems, we first characterize these distributions and then derive the closed-form expressions for the rate bound (30) with MRT and ZF beamforming in the presence of pilot contamination.

Theorem 3: Under pilot contamination, the closed-form expression for the upper bound (30) of the k-th user in the l-th cell with ZF beamforming is given by

$$R_{l,k}^{\text{UB, ZF}} = f\left(k_z^{\text{ZF}}, \theta_z^{\text{ZF}}\right) - f\left(k_y^{\text{ZF}}, \theta_y^{\text{ZF}}\right), \tag{31}$$

where

$$f(k,\theta) = \frac{1}{\Gamma(k)\ln 2}G_{3,2}^{1,3}\left(\theta\left|{}^{1-k,1,1}_{1,0}\right.\right), \tag{32}$$

$G_{p,q}^{m,n}\left(x\left|{}^{a_1,\cdots,a_p}_{b_1,\cdots,b_q}\right.\right)$ denotes the Meijer's G-function [37], and k_z^{ZF}, θ_z^{ZF}, k_y^{ZF}, θ_y^{ZF} are defined in (33)-(36) at the top of the page.

Proof: The proof is given in Appendix C. ■

Theorem 4: Under pilot contamination, the closed-form expression for the upper bound (30) of the k-th user in the l-th cell with MRT beamforming is given by

$$R_{l,k}^{\text{UB, MRT}} = f\left(k_z^{\text{MRT}}, \theta_z^{\text{MRT}}\right) - f\left(k_y^{\text{MRT}}, \theta_y^{\text{MRT}}\right), \tag{37}$$

where $f(k,\theta)$ is defined in (32), and k_z^{MRT}, θ_z^{MRT}, k_y^{MRT}, θ_y^{MRT} are defined in (38)-(41) at the top of the page.

Proof: Using a derivation process similar to that in proof of Theorem 3 and applying the same approximation for the distributions of the projections of channel vectors onto MRT beamforming subspace as that in proof of Theorem 1, we can obtain the result. Consequently, we omit the detailed proof of this theorem here. ■

Remark 4: In Theorem 3, based on Lemmas 2 and 3, the sum of the desired signal power and the interference power with ZF beamforming is approximated as a Gamma random variable Z_{ZF}, and the interference power is approximated as a Gamma random variable Y_{ZF}. The k_z^{ZF}, θ_z^{ZF} defined in (33), (34) and k_y^{ZF}, θ_y^{ZF} defined in (35), (36) are the shape and scale parameters of Z_{ZF} and Y_{ZF}, respectively. In Theorem 4, the (38)-(41) are the corresponding scale and shape parameters with MRT beamforming. From (31)-(41), it can be seen that, the rate expressions with MRT and ZF beamforming are very similar, and the main difference is that the latter scheme cancels intra-cell interference, at the price of reducing the array gain from MN to $\rho = MN - K + 1$.

Assuming perfect channel gain information at user side, we can obtain the upper bound on ergodic achievable rate

$$R_{l,k}^{\mathrm{UB}} \approx \tilde{R}_{l,k}^{\mathrm{UB}} = \log_2\left(1+\frac{\mathbb{E}\left[|\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}|^2\right]}{\mathbb{E}\left[\sum_{j\neq k}|\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,j}|^2\right]+\mathbb{E}\left[\sum_{i\neq l}\sum_{j=1}^{K}|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}|^2\right]+1/\gamma_{\mathrm{DL}}}\right) \quad (44)$$

which provides the maximal performance benefit of estimating the channel gain information at user side. Thanks to the effect of channel hardening, the users can reliably decode the transmitted signals from BSs based on only statistical effective channel gain information, i.e., the benefit of channel estimation at user side is low in the context of massive MIMO systems [24], [25]. However, the effect of channel hardening is less pronounced and the benefit of estimating channel gain information at user side become larger in distributed massive MIMO systems, especially with ZF beamforming.

As known, Meijer's G function is defined by line integral in complex plane [37]. Consequently, the closed-form expressions (31) and (37) have a prohibitively high complexity since Meijers G-function is involved. In distributed massive MIMO systems, users have less access distance to BS antennas [31], [51], thus, high SNR approximation is reasonable. Next, we first present a lemma about the expectation of log function, and then provide a high SNR approximation for the ergodic achievable downlink rate upper bound (30).

Lemma 4: If x is a Gamma distributed random variable with parameters k and θ, i.e., $x \sim \Gamma(k,\theta)$, then

$$\begin{aligned}\mathbb{E}\left[\log_2 x\right] &\overset{(a)}{=} \log_2 e\psi(k)+\log_2(\theta)\\ &\overset{(b)}{\approx} \log_2(k\theta)-\log_2\left(1+\tfrac{1}{2k}+\tfrac{5}{24k^2}\right),\end{aligned} \quad (42)$$

where $\psi(k)$ is the digamma function, (a) can be obtained from [31, Lemma 2], (b) results from $\psi(x) \sim \ln(x) - \ln\left(1+\sum_{n=1}^{\infty} a_n/x^n\right)$ as $x \to \infty$ [52, Remark 2.1].

From the expressions (33), (35), (38) and (40), it can be seen that the shape parameters k approach infinity when $MN \to \infty$. Consequently, the approximation in Lemma 4 will be asymptotically exact in distributed massive MIMO systems. Simulation results in Section IV show that the second-order expansion of $\psi(k)$ in (42) is accurate enough. Based on the Lemma 4, we obtain simplified closed-form expressions for the upper bound (30) in the following theorem by performing high SNR approximation.

Theorem 5: The high SNR rate approximation for the rate upper bound (30) of the k-th user in the l-th cell is given by

$$R_{l,k}^{\mathrm{UB,\,HS}} = \log_2\left(\frac{\theta_z}{\theta_y}\frac{k_z^3}{24k_z^2+12k_z+5}\frac{24k_y^3+12k_y+5}{k_y^3}\right), \quad (43)$$

where k_z, θ_z, k_y, and θ_y are defined in (33), (34), (35) and (36) for ZF beamforming, and in (38), (39), (40) and (41) for MRT beamforming, respectively.

Proof: The proof is given in Appendix D. ∎

Remark 5: The accuracy of high SNR rate approximation is very high in distributed antenna systems [31, Section IIV.B].

Theorem 5 has provided simplified closed-form expressions for the rate upper bound (30) (avoids the complex Meijer's G function), which can achieve nearly the same performance as that of Theorems 3 and 4. However, the expression (43) obtained in Theorem 5 is still complex. Recently, a tight approximation for ergodic achievable rate was proposed in [53] for massive MIMO systems which provides a useful general tool for studying ergodic rate. In the following, we derive closed-form expressions for this approximation. Although there is a little performance penalty, we can analyze the effect of system parameters on ergodic achievable rate more intuitively.

From [53, Lemma 4], we obtain the approximation (44) at the top of the page. Based on this approximation, we can calculate the terms $\mathbb{E}\left[|\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}|^2\right]$, $\mathbb{E}\left[\sum_{j\neq k}|\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,j}|^2\right]$ and $\mathbb{E}\left[\sum_{i\neq l}\sum_{j=1}^{K}|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}|^2\right]$ directly instead of characterizing the distributions of signal and interference powers to obtain an analytical expression. In the following theorems, we derive the closed-form expressions for (44) in distributed massive MIMO systems with ZF and MRT beamforming, respectively.

Theorem 6: The closed-form expression of (44), as an approximation for the ergodic achievable downlink rate upper bound of the k-th user in the l-th cell, with ZF beamforming is given by

$$\tilde{R}_{l,k}^{\mathrm{UB,\,ZF}} = \log_2\left(1+\frac{\rho\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}}+\tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}}}{\mathcal{I}_{l,k}^{\mathrm{A,\,ZF}}+MN\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}}\right), \quad (45)$$

where $\mathcal{I}_{l,k}^{\mathrm{A,\,ZF}} \triangleq K\sum_i \tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}} - \tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}} + MN/\gamma_{\mathrm{DL}}$.

Proof: The proof is given in Appendix E. ∎

Theorem 7: The closed-form expression of (44) with MRT beamforming is given by

$$\tilde{R}_{l,k}^{\mathrm{UB,\,MRT}} = \log_2\left(1+\frac{\hat{k}_{l,l,k\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}}+\frac{1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}}}{\mathcal{I}_{l,k}^{\mathrm{A,\,MRT}}+\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}}\right), \quad (46)$$

where $\mathcal{I}_{l,k}^{\mathrm{A,\,MRT}} \triangleq \frac{K-1}{MN}(\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}}+\sum_{i\neq l}(\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}+\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}}))+1/\gamma_{\mathrm{DL}}+\frac{1}{MN}\sum_{i\neq l}\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}}$.

Proof: Similar to the analysis in proof of Theorem 4, we omit the detailed proof of this theorem here. ∎

Remark 6: MRT beamforming aims to maximize the SNR ratio but does not pay attention to the multi-user interference. Meanwhile, ZF beamforming sacrifices some of the array gain to mitigate the multi-user interference. Thus, MRT beamforming will be preferred in low SNR (noise-limited) region. Moreover, as shown in (25) and (45), the user rate of ZF at $MN = K$ will close to a very small value. Thus, the user rate performance of MRT will be better than that of ZF when K is large.

In the following corollary, we investigate the asymptotic case when the total number of transmit antennas is much larger than the number of users, i.e., $\frac{MN}{K} \to \infty$, and give the user ultimate achievable rate.

Corollary 2: Letting $\frac{MN}{K} \to \infty$, the upper bound on ergodic achievable downlink rate of the k-th user in the l-th cell with MRT and ZF beamforming approach the same

ultimate achievable rate $\tilde{R}_{l,k}^{\mathrm{UB},\infty}$, given by

$$\tilde{R}_{l,k}^{\mathrm{UB},\infty} = \log_2\left(1 + \frac{\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}}}{\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}}\right). \quad (47)$$

Proof: Using a method similar to that in proof of Corollary 1, it is straightforward to obtain (47) by dividing the denominators and numerators of (43) or (45) and (46) by $\frac{MN}{K}$ and letting $\frac{MN}{K} \to \infty$. ■

Remark 7: As seen from Corollaries 1 and 2, when the ratio $\frac{MN}{K}$ is very large, the gain of adding more antennas diminishes and the user achievable rate is limited by pilot contamination. Thus, pilot contamination mitigation techniques, e.g., pilot scheduling in [45], can significantly increase the user achievable rate.

From Corollaries 1 and 2, it can be seen that, as $\frac{MN}{K} \to \infty$ the effect of uncorrelated received noise is eliminated completely, and transmissions from the users within the same cell do not interfere. However, the correlated interference from other cells (pilot contamintion) due to the reuse of the same pilot sequences remains existent. Moreover, ZF beamforming and MRT beamforming achieve the same ultimate achievable rate as $\frac{MN}{K} \to \infty$ no matter with statistical or perfect effective channel gain information at users. This is because with very large $\frac{MN}{K}$, the effective channel gain $a_{i,j,l,k} \triangleq \mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}$ becomes nearly deterministic due to the channel hardening effect [17]. In this case, using the statistical information $\mathbb{E}[a_{i,j,l,k}]$ for signal detection at user side is good enough.

In summary of this section, closed-form expressions for the lower bound and upper bound on ergodic achievable downlink rate have been derived in distributed massive MIMO systems with MRT and ZF beamforming in the presence of pilot contamination. Based on the derived expressions, the user ultimate achievable rates are also given as $\frac{MN}{K} \to \infty$. In the following section, we validate the accuracy of the theoretical results for different scenarios.

IV. Numerical Results

In this section, the theoretical analysis presented in Section III is verified through a set of Monte Carlo simulations. A hexagonal system with L cells is considered. Unless mentioned otherwise, the locations of RAUs and users are assumed to be uniformly distributed in each cell. The cell radius and the distance between two adjacent cells are normalized to 1 and $\sqrt{3}$, respectively, and the minimum distance between users and RAUs is set to 0.01. To allow for reproducibility of the results, we consider a distance-based path loss model with path loss exponent $\alpha = 3.7$, without shadowing [20], [23]. We set the parameter c to be one and $\gamma_{\mathrm{DL}} = 10$ dB to be the reference SNR at the cell edge in the downlink. In all examples, $\gamma_{\mathrm{P}} = K\gamma_{\mathrm{DL}}$.

We begin with justifying the various different approximations employed in this paper. Fig. 2 shows the cumulative distribution function (CDF) of the powers of the non-isotropic channel vectors $\mathbf{g}_{i,l,k}$ obtained numerically and by employing Gamma second-order matching technique (Lemma 2). As seen from the figure, Gamma second-order matching technique provides an accurate approximation for the powers of $\mathbf{g}_{i,l,k}$.

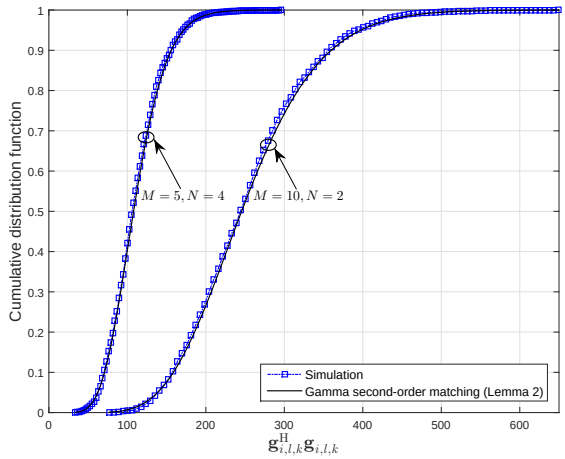


Fig. 2. Cumulative distribution function of the powers of non-isotropic channel vectors $\mathbf{g}_{i,l,k}$ with different M and N.

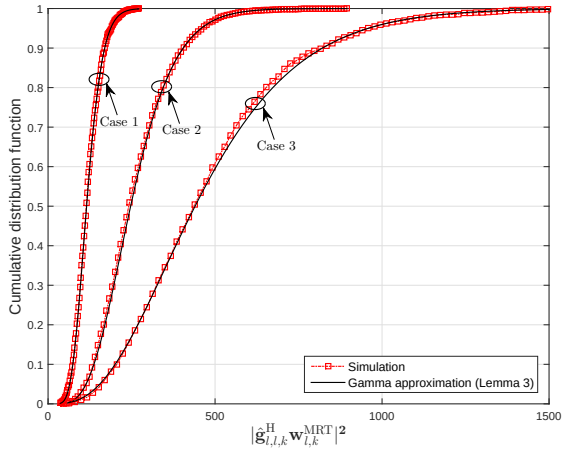


Fig. 3. Cumulative distribution function of the signal power with different M and N. Case 1, $M = 1$, $N = 10$; Case 2, $M = 6$, $N = 4$; Case 3, $M = 20$, $N = 1$.

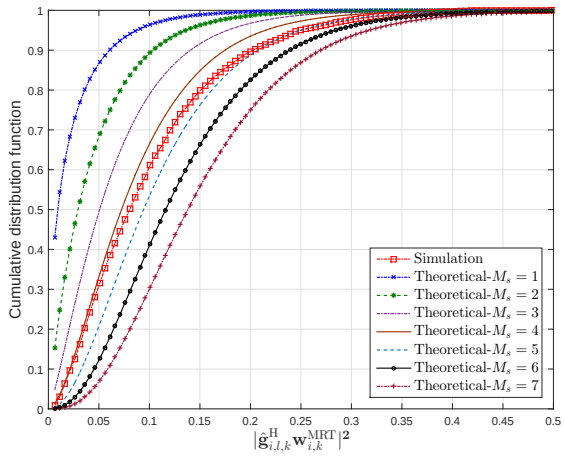


Fig. 4. Cumulative distribution function of pilot contamination power, $L = 7$, $M = 7$, $N = 1$.

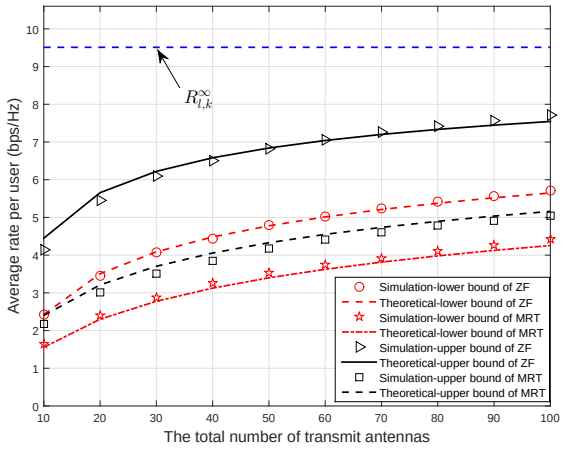


Fig. 5. Average rate per user against the total number of transmit antennas, $L = 7$, $K = 4$, and $M = 5$.

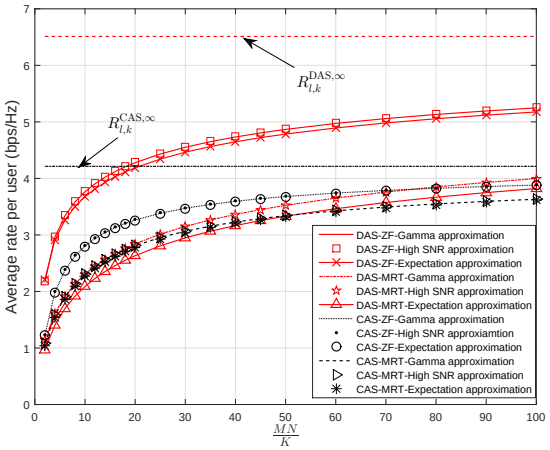


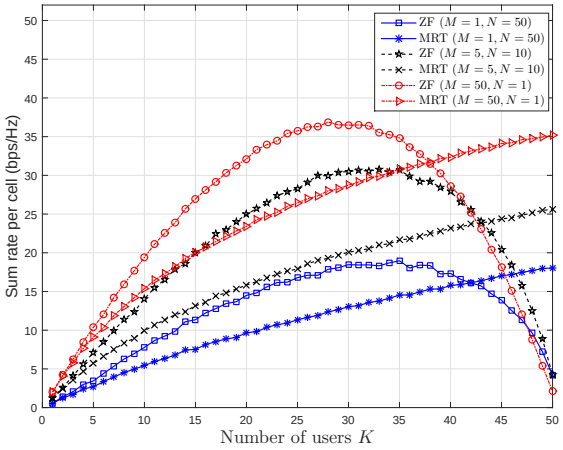
Fig. 6. Average rate per user at cell edge against the total number of transmit antennas, $L = 7$, $K = 4$, and $M = 5$.

This provides the incentives to use the moment matched approximation in distributed massive MIMO systems. Focus on the k-th user in the l-th cell, the CDFs of signal power $|\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}|^2$ and pilot contamination power $|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,k}^{\mathrm{MRT}}|^2$ from the transmission in i-cell for the k-th user obtained numerically and by employing Gamma approximation technique (Lemma 3) are presented in Fig. 3 and Fig. 4, respectively. As seen from Fig. 3, the CDF curves obtained by Gamma approximation technique perfectly match the numerical results for the signal power in all cases. From Fig. 4, it can be seen that, if we approximate the pilot contamination power by Lemma 3 directly (the dimension of projection subspace is MN), the approximation is slightly less accurate. It is because that the path losses from the k-th user in the l-th cell to the RAUs in the i-th cell are relative large and vary drastically, i.e., the user may be interfered mainly by a small part of the RAUs in the i-th cell. As seen from Fig. 4, we can increase the approximation accuracy by setting the dimension of the projection subspace equal to M_sN ($M_s = 4$ or 5 in the Fig. 4) instead of MN where M_s is the number of RAUs in the i-th cell with less and similar path loss which can be obtained easily based on the known large-scale fading. Note that the simulation results with ZF beamforming are omitted here due to space constraints but provide similar results.

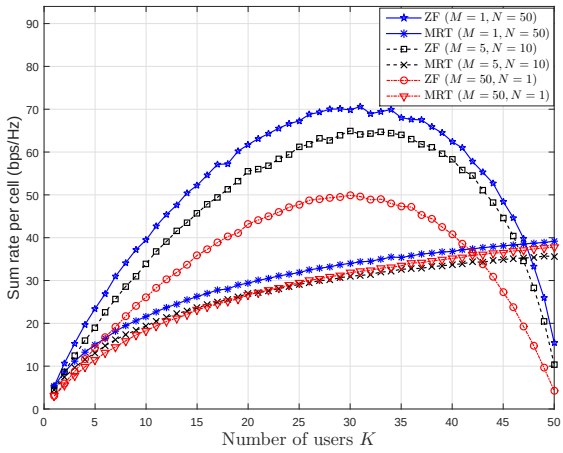
Then, we verify the accuracy of the closed-form expressions given in Theorems 1, 2, 3 and 4. Fig. 5 depicts the rate bounds on ergodic achievable downlink rate with MRT and ZF beamforming as a function of the total number of transmit antennas MN for $L = 7$ cells, and $K = 4$ users, $M = 5$ RAUs. We denote the ultimate achievable rate by $R_{l,k}^{\infty}$ since $R_{l,k}^{\mathrm{LB},\infty} = \tilde{R}_{l,k}^{\mathrm{UB},\infty}$. As seen from the figure, although there is a small mismatch between the closed-form expressions and simulation results due to the approximations applied for the non-isotropic channel vectors, they also match well with less than five percent error. Note that, the closed-form expressions are almost indistinguishable from simulation results in co-located massive MIMO systems with both MRT and ZF beamforming which is omitted due to space constraints. Consequently, in the following, we will use these closed-form expressions for all numerical simulations.

Next, we evaluate the performance of the closed-form expressions for the upper bound on user ergodic achievable rate obtained by applying high SNR approximation and expectation approximation in Theorems 5, 6 and 7 in a distributed massive MIMO system comprising $L = 7$ cells and $M = 6$ RAUs. For the comparison between distributed massive MIMO (DAS in the figure) and co-located massive MIMO systems (CAS in the figure), it is assumed that $K = 6$ users are uniformly distributed within the cell edge (defined as the region outside the circle of radius $r = 3/4$) in each cell. As seen from Fig. 6, the following findings have been obtained. First, in both distributed and co-located massive MIMO systems, the high SNR approximations in Theorem 5 achieve nearly the same performance as that of the closed-form expressions in Theorems 3 and 4 (denoted by "Gamma approximation" in the figure) with ZF and MRT beamforming. Second, the closed-form expressions in Theorems 6 and 7 (denoted by "Expectation approximation" in the figure) also achieve nearly the same performance in co-located massive MIMO systems, and have only a little performance penalty in distributed massive MIMO systems. Considering the lowest computational complexity, the closed-form expressions in Theorems 6 and 7 are preferable when the number of antennas is large. Third, ZF beamforming leads to a significant performance gain over MRT beamforming as it reduces multiuser interference and has a faster convergence speed than MRT beamforming. Here we should note that, although it can be seen from Fig. 6 that the average rate per user of the distributed massive MIMO systems is much larger than that of the co-located massive MIMO systems, in practice, the overheads for CSI estimate and user data sharing will reduce the spectral efficiency of distributed massive MIMO systems. How to establish a scalable signal processing framework for distributed massive MIMO systems is a key challenge and needs to be further addressed.

Fig. 7 shows the sum rate per cell calculated by Theorems 1 and 2 as a function of the number of users K at SNR = -25 dB and 10 dB, where purely co-located systems ($M = 1, N = 50$), partly distributed systems ($M = 5, N = 10$) and purely

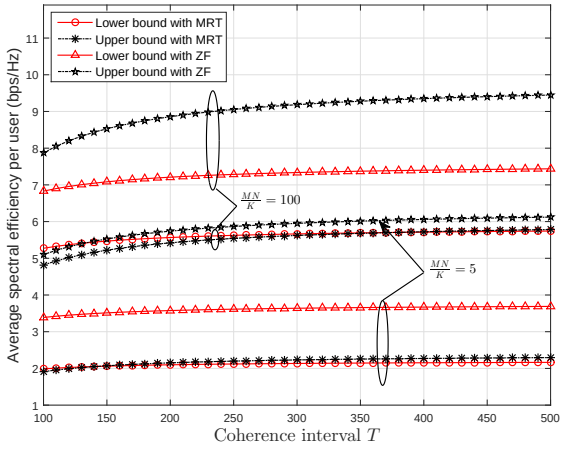


(a) SNR = -25 dB

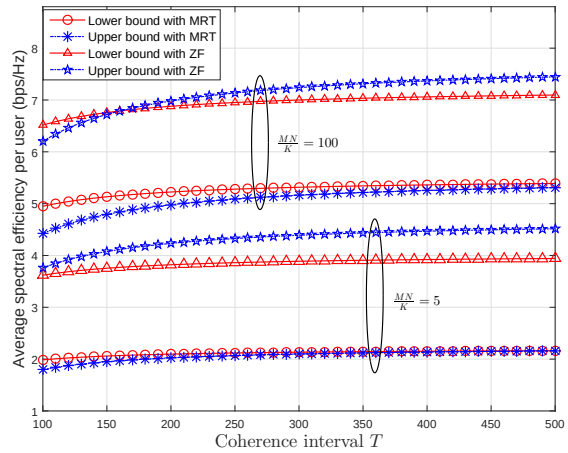


(b) SNR = 10 dB

Fig. 7. Sum rate per cell as a function of the number of users, $L = 7$, $MN = 50$, $K = [1, 50]$.



(a) Co-located massive MIMO



(b) Distributed massive MIMO

Fig. 8. Average spectral efficiency per user against the coherence interval T, $L = 7$, $K = 10$, $\tau_{\mathrm{u}} = \tau_{\mathrm{d}} = K$, MN/K = 5, 100.

distributed systems ($M = 50, N = 1$) are considered. As seen from the figure, we get the following findings. First, although ZF often provides better sum rate than MRT, it is interesting to note that MRT is competitive when K is large and the SNR is low, both in terms of sum rate and computational complexity since the complexity of ZF scales as $\mathcal{O}(MNK^2)$ while the complexity of MRT scales as $\mathcal{O}(MNK)$ [23]. Second, when the SNR is low (-25 dB), distributing the transmit antennas increases the desired signal power thanks to the reduction of minimum access distance. However, the interference (intra-cell and inter-cell interference if MRT beamforming is adopted and only inter-cell interference with ZF beamforming since the intra-cell interference is eliminated by joint precoding over users) power is also enhanced and the interference becomes more and more severe as the number of users K increases. This is the reason that the three sum rate per cell curves with ZF beamforming cross each other when K is large in Fig. 7(a). When the SNR is relative high (10 dB), purely co-located systems provide better sum rate performance which is because that the average access distance of distributed massive MIMO systems is larger than that of co-located massive MIMO systems when the BS is located at the center of the cell [54]. For a reason similar to that in the low SNR case, the sum rate per cell curves of partly distributed systems and purely distributed systems with MRT beamformign cross each other in Fig. 7(b). Moreover, considering the performance-complexity tradeoff, a partly distributed massive system is an appealing design choice.

Finally, based on the closed-form expressions for the rate bounds on ergodic achievable downlink rate derived in Theorems 1, 2, 6 and 7, the average spectral efficiency per user performances with statistical and perfect effective channel gain information against different coherence interval are compared. Taking into account the performance loss due to the uplink and downlink pilots, the average spectral efficiency per user is defined as

$$S = \frac{T - \tau_{\mathrm{u}} - \tau_{\mathrm{d}}}{T} \frac{\sum_{l=1}^{L} \sum_{k=1}^{K} R_{l,k}}{LK}, \tag{48}$$

where T is the coherence interval (in symbols), τ_{u} and τ_{d} are uplink and downlink pilot overhead, i.e., the number of symbols per coherence interval spent for training phases, respectively. Given the definition in (48), we analyze the average spectral efficiency per user performance in distributed and co-located massive MIMO systems comprising $L = 7$

cells and $K = 10$ users with different ratios $\frac{MN}{K}$. As seen from Fig. 8, we get the following findings. First, a longer coherence interval yields a larger average spectral efficiency per user since the additional pilot overhead ratio reduces; Second, both the co-located and distributed massive MIMO systems do not need downlink beamforming training scheme when the coherence interval T is small which will result in a large pilot overhead ratio. Moreover, when MRT beamforming is applied, the performance gain obtained by the downlink beamforming training scheme is relatively small; Third, the downlink beamforming training scheme is more preferable for the distributed massive MIMO systems, which is because that only a small number of RAUs may substantially contribute to serve a given user which results in less channel hardening in distributed massive MIMO systems .

V. Conclusions

In this paper, according to the levels of effective channel gain information at user side, we provided the lower bound and upper bound on user ergodic achievable downlink rate. Considering pilot contamination and practical per user power normalization, accurate and computationally efficient closed-form expressions for the rate bounds with both MRT and ZF beamforming were derived based on the properties of Gamma distributions and the non-isotropic channel approximation techniques. Based on these closed-form expressions, we studied the spectral efficiency of distributed and co-located massive MIMO systems. Our investigation showed that, distributed massive MIMO systems can provide better performance when the SNR is low and the ratio $\frac{MN}{K}$ is large, and although ZF often provides better performance, MRT is competitive when K is large and the SNR is low. Moreover, the benefits of estimating the effective channel gain at user side were analyzed. Numerical results showed that downlink beamforming scheme is more preferable for the distributed massive MIMO systems with ZF beamforming when the coherence interval is relatively large.

Appendix A
Proof of Theorem 1

We derive the closed-form expression for the rate bound (22) with MRT beamforming by calculating the following three terms in $\gamma_{l,k}^{\mathrm{LB}}$, $|\mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}]|^2$, $\mathrm{var}\big[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}\big]$, and $\sum_{(i,j)\neq(l,k)}\mathbb{E}\big[|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{MRT}}|^2\big]$.

For the term $|\mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}]|^2$, based on the independence of channel estimate $\hat{\mathbf{g}}_{l,l,k}$ and estimation error $\tilde{\mathbf{g}}_{l,l,k}$, we have

$$|\mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}]|^2 = |\mathbb{E}[\|\hat{\mathbf{g}}_{l,l,k}\|]|^2. \tag{A.1}$$

From Lemma 3 and (15), we obtain

$$|\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}|^2 = \|\hat{\mathbf{g}}_{l,l,k}\|^2 \sim \Gamma(\hat{k}_{l,l,k,\mathrm{a}}, \hat{\theta}_{l,l,k,\mathrm{a}}). \tag{A.2}$$

Based on the well-known relationship between Gamma and Nakagami distribution, we have

$$\|\hat{\mathbf{g}}_{l,l,k}\| \sim \text{Nakagami}(\hat{k}_{l,l,k,\mathrm{a}}, \hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}}). \tag{A.3}$$

Thus,

$$|\mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}]|^2 = |\mathbb{E}[\|\hat{\mathbf{g}}_{l,l,k}\|]|^2 = \xi(\hat{k}_{l,l,k,\mathrm{a}})\hat{\theta}_{l,l,k,\mathrm{a}}. \tag{A.4}$$

For the term $\mathrm{var}\big[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}\big]$, we have

$$\begin{aligned}
&\mathrm{var}\big[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}\big] \\
&\overset{(a)}{=} \mathbb{E}\big[|\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}|^2\big] + \mathbb{E}\big[|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}|^2\big] - \big|\mathbb{E}[\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}]\big|^2 \\
&\overset{(b)}{=} (\hat{k}_{l,l,k,\mathrm{a}} - \xi(\hat{k}_{l,l,k,\mathrm{a}}))\hat{\theta}_{l,l,k,\mathrm{a}} + \tfrac{1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}},
\end{aligned} \tag{A.5}$$

where (a) results from the independence of $\hat{\mathbf{g}}_{l,l,k}$ and $\tilde{\mathbf{g}}_{l,l,k}$, and (b) is obtained by applying Lemma 3 to approximate the distributions of $|\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}|^2$ and $|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}|^2$ with $\Gamma(\hat{k}_{l,l,k,\mathrm{a}}, \hat{\theta}_{l,l,k,\mathrm{a}})$ and $\Gamma(\frac{1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}, \tilde{\theta}_{l,l,k,\mathrm{a}})$, and $\big|\mathbb{E}[\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{MRT}}]\big|^2 = |\mathbb{E}[\|\hat{\mathbf{g}}_{l,l,k}\|]|^2$ has been given in (A.4).

Considering pilot contamination which makes $\mathbf{w}_{i,j}^{\mathrm{MRT}}$ and $\hat{\mathbf{g}}_{i,l,k}$ dependent for $j = k$, and based on the independence of $\hat{\mathbf{g}}_{i,l,k}$ and $\tilde{\mathbf{g}}_{i,l,k}$, we decompose the term $\sum_{(i,j)\neq(l,k)}\mathbb{E}\big[|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{MRT}}|^2\big]$ into four components as

$$\begin{aligned}
&\sum_{(i,j)\neq(l,k)}\mathbb{E}\big[|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{MRT}}|^2\big] \\
&= \sum_{j\neq k}\mathbb{E}[|\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,j}^{\mathrm{MRT}}|^2] + \sum_{i\neq l}\sum_{j\neq k}\mathbb{E}[|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{MRT}}|^2] \\
&\quad + \sum_{i\neq l}\mathbb{E}[|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,k}^{\mathrm{MRT}}|^2] + \sum_{i\neq l}\sum_{j=1}^{K}\mathbb{E}[|\tilde{\mathbf{g}}_{i,l,k,}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{MRT}}|^2].
\end{aligned} \tag{A.6}$$

Similarly, from Lemma 3, we can obtain the distributions of the four components in (A.6),

$$|\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,j}^{\mathrm{MRT}}|^2 \sim \Gamma(\tfrac{1}{MN}k_{l,l,k,\mathrm{a}}, \theta_{l,l,k,\mathrm{a}}), \tag{A.7}$$

$$|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,k}^{\mathrm{MRT}}|^2 \sim \Gamma(\hat{k}_{i,l,k,\mathrm{a}}, \hat{\theta}_{i,l,k,\mathrm{a}}), \tag{A.8}$$

$$|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{MRT}}|^2 \sim \Gamma(\tfrac{1}{MN}\hat{k}_{i,l,k,\mathrm{a}}, \hat{\theta}_{i,l,k,\mathrm{a}}), \tag{A.9}$$

$$|\tilde{\mathbf{g}}_{i,l,k,}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{MRT}}|^2 \sim \Gamma(\tfrac{1}{MN}\tilde{k}_{i,l,k,\mathrm{a}}, \tilde{\theta}_{i,l,k,\mathrm{a}}). \tag{A.10}$$

Substituting (A.7), (A.8), (A.9) and (A.10) into (A.6) yields

$$\begin{aligned}
&\textstyle\sum_{(i,j)\neq(l,k)}\mathbb{E}[|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{MRT}}|^2] \\
&= \tfrac{K-1}{MN}\big(k_{l,l,k,\mathrm{a}}\theta_{l,l,k,\mathrm{a}} + \textstyle\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}\big) \\
&\quad + \textstyle\sum_{i\neq l}\big(\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}} + \tfrac{K}{MN}\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}}\big),
\end{aligned} \tag{A.11}$$

Combing (A.4), (A.5) and (A.11) concludes the proof.

Appendix B
Proof of Theorem 2

The useful signal power term $|\mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}]|^2$ can be calculated by

$$\begin{aligned}
|\mathbb{E}[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}]|^2 &= |\mathbb{E}[(\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}} + \tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}})\mathbf{w}_{l,k}^{\mathrm{ZF}}]|^2 \\
&\overset{(a)}{=} |\mathbb{E}[1/\|\mathbf{a}_{l,l,k}\|]|^2 \\
&\overset{(b)}{=} \left|\mathbb{E}\Big[\big(\big[\big(\hat{\mathbf{G}}_l^{\mathrm{H}}\hat{\mathbf{G}}_l\big)^{-1}\big]_{k,k}\big)^{-1/2}\Big]\right|^2 \\
&\overset{(c)}{=} \xi(\tfrac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}})\hat{\theta}_{l,l,k,\mathrm{a}},
\end{aligned} \tag{B.1}$$

where (a) is obtained because of the independence of $\mathbf{w}_{l,k}^{\mathrm{ZF}}$ and the estimation error $\tilde{\mathbf{g}}_{l,l,k}$ and $\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}} = 1/\|\mathbf{a}_{l,l,k}\|$, (b)

results from $\|\mathbf{a}_{l,l,k}\|^2 = \left[\left(\hat{\mathbf{G}}_l^{\mathrm{H}}\hat{\mathbf{G}}_l\right)^{-1}\right]_{k,k}$, (c) is obtained because $1/\|\mathbf{a}_{l,l,k}\| \sim \text{Nakagami}\big(\frac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}}, \frac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}}\big)$ where we have applied Lemma 3 to approximate the distribution of $\left[\left(\hat{\mathbf{G}}_l^{\mathrm{H}}\hat{\mathbf{G}}_l\right)^{-1}\right]_{k,k}$ as $\Gamma(\frac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}}, \hat{\theta}_{l,l,k,\mathrm{a}})$ since $\left[\left(\hat{\mathbf{G}}_{l,\mathrm{a}}^{\mathrm{H}}\hat{\mathbf{G}}_{l,\mathrm{a}}\right)^{-1}\right]_{k,k} \sim \Gamma(\rho, \hat{\theta}_{l,l,k,\mathrm{a}})$ [27, lemma 10].

For the variance term $\text{var}\left[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}\right]$, we have

$$\begin{aligned}
&\text{var}\left[\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}\right] \\
&= \mathbb{E}\left[|\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2\right] + \mathbb{E}\left[|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2\right] - \left|\mathbb{E}[\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}]\right|^2 \\
&\overset{(a)}{=} \big(\tfrac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}} - \xi(\tfrac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}})\big)\hat{\theta}_{l,l,k,\mathrm{a}} + \tfrac{1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}},
\end{aligned} \tag{B.2}$$

where (a) is obtained by applying Lemma 3 to approximate the distributions of $|\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2$ and $|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2$ with $\Gamma(\frac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}}, \hat{\theta}_{l,l,k,\mathrm{a}})$ and $\Gamma(\frac{1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}, \tilde{\theta}_{l,l,k,\mathrm{a}})$, and $\left|\mathbb{E}[\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}]\right|^2 = |\mathbb{E}[1/\|\mathbf{a}_{l,l,k}\|]|^2$ has been given in (B.1).

Considering the term $\sum_{(i,j)\neq(l,k)}\mathbb{E}\left[|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2\right]$, we have

$$\begin{aligned}
&\sum\nolimits_{(i,j)\neq(l,k)}\mathbb{E}\left[|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2\right] \\
&\overset{(a)}{=} \sum\nolimits_{j\neq k}\mathbb{E}[|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,j}^{\mathrm{ZF}}|^2] + \sum\nolimits_{i\neq l}\sum\nolimits_{j\neq k}\mathbb{E}[|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2] \\
&\quad + \sum\nolimits_{i\neq l}\mathbb{E}[|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,k}^{\mathrm{ZF}}|^2] + \sum\nolimits_{i\neq l}\sum\nolimits_{j=1}^{K}\mathbb{E}[|\tilde{\mathbf{g}}_{i,l,k,}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2] \\
&\overset{(b)}{=} \tfrac{K-1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}} + \sum_{i\neq l}(\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}} + \tfrac{K}{MN}\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}}),
\end{aligned} \tag{B.3}$$

where (a) is obtained because $\hat{\mathbf{g}}_{l,l,k}\mathbf{w}_{l,j}^{\mathrm{ZF}} = 0$ for $j \neq k$, (b) is obtained by applying Lemma 3 to approximate the distributions of $|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,j}^{\mathrm{ZF}}|^2$, $|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,k}^{\mathrm{ZF}}|^2$, $|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2$ and $|\tilde{\mathbf{g}}_{i,l,k}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2$ with $\Gamma(\frac{\tilde{k}_{l,l,k,\mathrm{a}}}{MN}, \tilde{\theta}_{l,l,k,\mathrm{a}})$, $\Gamma(\frac{\rho}{MN}\hat{k}_{i,l,k,\mathrm{a}}, \hat{\theta}_{i,l,k,\mathrm{a}})$, $\Gamma(\frac{\tilde{k}_{i,l,k,\mathrm{a}}}{MN}, \hat{\theta}_{i,l,k,\mathrm{a}})$ and $\Gamma(\frac{\tilde{k}_{i,l,k,\mathrm{a}}}{MN}, \tilde{\theta}_{i,l,k,\mathrm{a}})$, respectively.

Substituting (B.1), (B.2) and (B.3) into (22) yields the closed-form expression (25). This completes the proof.

Appendix C
Proof of Theorem 3

The ergodic achievable downlink rate upper bound (30) of the k-th user in the l-th cell with ZF beamforming can be rewritten as

$$R_{l,k}^{\mathrm{UB}} = \mathbb{E}\left[\log_2\left(1 + X_{\mathrm{ZF}} + Y_{\mathrm{ZF}}\right)\right] - \mathbb{E}\left[\log_2(1 + Y_{\mathrm{ZF}})\right], \tag{C.1}$$

where

$$X_{\mathrm{ZF}} \triangleq \gamma_{\mathrm{DL}}|\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2,$$

$$Y_{\mathrm{ZF}} \triangleq \gamma_{\mathrm{DL}}\sum_{j\neq k}|\mathbf{g}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,j}^{\mathrm{ZF}}|^2 + \gamma_{\mathrm{DL}}\sum_{i\neq l}\sum_{j=1}^{K}|\mathbf{g}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2.$$

For the variable X_{ZF}, we have

$$\begin{aligned}
X_{\mathrm{ZF}} \quad &= \quad \gamma_{\mathrm{DL}}|(\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}} + \tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}})\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2 \\
&\xrightarrow[MN\to\infty]{(a)} \gamma_{\mathrm{DL}}(|\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2 + |\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2),
\end{aligned} \tag{C.2}$$

where (a) results from neglecting $\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}(\mathbf{w}_{l,k}^{\mathrm{ZF}})^{\mathrm{H}}\tilde{\mathbf{g}}_{l,l,k}$ and $\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}(\mathbf{w}_{l,k}^{\mathrm{ZF}})^{\mathrm{H}}\hat{\mathbf{g}}_{l,l,k}$ since it is insignificant compared with $|\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2$ and $|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2$ as the number of transmit antenna approach infinity [20].

Based on Lemma 3, we characterize the distributions of the two terms in (C.2) as

$$|\hat{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2 \sim \Gamma\big(\tfrac{\rho}{MN}\hat{k}_{l,l,k,\mathrm{a}}, \hat{\theta}_{l,l,k,\mathrm{a}}\big), \tag{C.3}$$

$$|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,k}^{\mathrm{ZF}}|^2 \sim \Gamma\big(\tfrac{1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}, \tilde{\theta}_{l,l,k,\mathrm{a}}\big). \tag{C.4}$$

Therefore, the variable X_{ZF} in (C.2) can be approximated as a sum of independent Gamma random variables with different shape and scale parameters. From Lemma 2, we have

$$X_{\mathrm{ZF}} \sim \Gamma\left(k_x^{\mathrm{ZF}}, \theta_x^{\mathrm{ZF}}\right), \tag{C.5}$$

where

$$k_x^{\mathrm{ZF}} = \frac{(\rho\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}} + \tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}})^2}{MN(\rho\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}}^2 + \tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}}^2)}, \tag{C.6}$$

$$\theta_x^{\mathrm{ZF}} = \gamma_{\mathrm{DL}}\frac{\rho\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}}^2 + \tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}}^2}{\rho\hat{k}_{l,l,k,\mathrm{a}}\hat{\theta}_{l,l,k,\mathrm{a}} + \tilde{k}_{l,l,k,\mathrm{a}}\tilde{\theta}_{l,l,k,\mathrm{a}}}. \tag{C.7}$$

The variable Y_{ZF} can be devided into four terms as

$$\begin{aligned}
\frac{1}{\gamma_{\mathrm{DL}}}Y_{\mathrm{ZF}} = &\sum\nolimits_{j\neq k}|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,j}^{\mathrm{ZF}}|^2 + \sum\nolimits_{i\neq l}|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,k}^{\mathrm{ZF}}|^2 \\
&+ \sum_{i\neq l}\sum_{j\neq k}|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2 + \sum_{i\neq l}\sum_{j=1}^{K}|\tilde{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2.
\end{aligned} \tag{C.8}$$

By using Lemmas 2 and 3 and some algebraic simplifications, we obtain the distributions of the four terms in (C.8) as

$$\sum\nolimits_{j\neq k}|\tilde{\mathbf{g}}_{l,l,k}^{\mathrm{H}}\mathbf{w}_{l,j}^{\mathrm{ZF}}|^2 \sim \Gamma\Big(\tfrac{(K-1)}{MN}\tilde{k}_{l,l,k,\mathrm{a}}, \tilde{\theta}_{l,l,k,\mathrm{a}}\Big), \tag{C.9}$$

$$\begin{aligned}
&\sum\nolimits_{i\neq l}|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,k}^{\mathrm{ZF}}|^2 \\
&\sim \Gamma\Big(\tfrac{\rho}{MN}\tfrac{\left(\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}\right)^2}{\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}^2}, \tfrac{\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}^2}{\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}}\Big),
\end{aligned} \tag{C.10}$$

$$\begin{aligned}
&\sum\nolimits_{i\neq l}\sum\nolimits_{j\neq k}|\hat{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2 \\
&\sim \Gamma\Big(\tfrac{(K-1)}{MN}\tfrac{\left(\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}\right)^2}{\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}^2}, \tfrac{\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}^2}{\sum_{i\neq l}\hat{k}_{i,l,k,\mathrm{a}}\hat{\theta}_{i,l,k,\mathrm{a}}}\Big),
\end{aligned} \tag{C.11}$$

$$\begin{aligned}
&\sum\nolimits_{i\neq l}\sum\nolimits_{j=1}^{K}|\tilde{\mathbf{g}}_{i,l,k}^{\mathrm{H}}\mathbf{w}_{i,j}^{\mathrm{ZF}}|^2 \\
&\sim \Gamma\Big(\tfrac{K}{MN}\tfrac{\left(\sum_{i\neq l}\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}}\right)^2}{\sum_{i\neq l}\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}}^2}, \tfrac{\sum_{i\neq l}\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}}^2}{\sum_{i\neq l}\tilde{k}_{i,l,k,\mathrm{a}}\tilde{\theta}_{i,l,k,\mathrm{a}}}\Big).
\end{aligned} \tag{C.12}$$

The proofs of (C.9)-(C.12) are similar to that of the terms in X_{ZF}, and hence we omit the details. From Lemma 2, we obtain

$$Y_{\mathrm{ZF}} \sim \Gamma\left(k_y^{\mathrm{ZF}}, \theta_y^{\mathrm{ZF}}\right), \tag{C.13}$$

where k_y^{ZF} and θ_y^{ZF} are given by (35) and (36), respectively. Then, we approximate $X_{\mathrm{ZF}}+Y_{\mathrm{ZF}}$ with another Gamma random variable Z_{ZF} by Lemma 2,

$$Z_{\mathrm{ZF}} = X_{\mathrm{ZF}} + Y_{\mathrm{ZF}} \sim \Gamma\left(k_z^{\mathrm{ZF}}, \theta_z^{\mathrm{ZF}}\right), \tag{C.14}$$

where k_z^{ZF} and θ_z^{ZF} are given in (33) and (34), respectively.

According to [55, Eq. (8.4.6.5)], the logarithmic term $\log_2(1+x)$ can be expressed with a Meijer's G-function as

$$\log_2(1+x) = \frac{1}{\ln 2}G_{1,2}^{2,2}\left(x\left|\begin{matrix}1,1\\1,0\end{matrix}\right.\right). \tag{C.15}$$

Then, for Gamma random variable x, $\mathbb{E}\left[\log_2(1+x)\right]$ is computed as

$$\begin{aligned}&\mathbb{E}\left[\log_2(1+x)\right]\\ &= \frac{1}{\theta^k\Gamma(k)\ln 2}\int_0^\infty G_{1,2}^{2,2}\left(x\left|\begin{smallmatrix}1,1\\1,0\end{smallmatrix}\right.\right)x^{k-1}e^{-x/\theta}dx\\ &\overset{(a)}{=} \frac{1}{\Gamma(k)\ln 2}G_{1,3}^{3,2}\left(\theta\left|\begin{smallmatrix}1-k,1,1\\1,0\end{smallmatrix}\right.\right),\end{aligned} \tag{C.16}$$

where (a) results from [37, Eq. (7.813.1)]

Combining (C.1), (C.13), (C.14) and (C.16) concludes the proof.

APPENDIX D
PROOF OF THEOREM 5

For the ergodic achievable downlink rate (30) of the k-th user in the l-th, with the definition of

$$X \triangleq \gamma_{\text{DL}}|\mathbf{g}_{l,l,k}^{\text{H}}\mathbf{w}_{l,k}|^2, \tag{D.1}$$

$$Y \triangleq \gamma_{\text{DL}}\sum_{j\neq k}|\mathbf{g}_{l,l,k}^{\text{H}}\mathbf{w}_{l,j}|^2 + \gamma_{\text{DL}}\sum_{i\neq l}\sum_{j=1}^{K}|\mathbf{g}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}|^2, \tag{D.2}$$

we have

$$\begin{aligned}R_{l,k} &= \mathbb{E}\left[\log_2\left(1+\tfrac{X}{Y+1}\right)\right]\\ &\overset{(a)}{=} \mathbb{E}\left[\log_2\left(1+Z\right)\right] - \mathbb{E}\left[\log_2\left(1+Y\right)\right]\\ &\overset{(b)}{\approx} \mathbb{E}\left[\log_2\left(Z\right)\right] - \mathbb{E}\left[\log_2\left(Y\right)\right]\\ &\overset{(c)}{\approx} \log_2(k_z\theta_z) - \log_2\left(1+\tfrac{1}{2k_z}+\tfrac{5}{24k_z^2}\right)\\ &\quad - \log_2(k_y\theta_y) + \log_2\left(1+\tfrac{1}{2k_y}+\tfrac{5}{24k_y^2}\right),\end{aligned} \tag{D.3}$$

where (a) results from $\mathbb{E}\left[\log_2(1+x/y)\right] = \mathbb{E}\left[\log_2(x+y)\right] - \mathbb{E}\left[\log_2(y)\right]$ and $Z \triangleq X+Y$, (b) is obtained by applying high SNR approximation, (c) follows from Lemma 4, and k_z, θ_z, k_y, θ_y are given by (33), (34), (35),(36) with ZF beamforming and (38), (39), (40), (41) with MRT beamforming, respectively.

This completes the proof.

APPENDIX E
PROOF OF THEOREM 6

For the term $\mathbb{E}[|\mathbf{g}_{l,l,k}^{\text{H}}\mathbf{w}_{l,k}^{\text{ZF}}|^2]$, we have

$$\begin{aligned}&\mathbb{E}[|\mathbf{g}_{l,l,k}^{\text{H}}\mathbf{w}_{l,k}^{\text{ZF}}|^2]\\ &\overset{(a)}{=} \mathbb{E}[|\hat{\mathbf{g}}_{l,l,k}^{\text{H}}\mathbf{w}_{l,k}^{\text{ZF}}|^2] + \mathbb{E}[|\tilde{\mathbf{g}}_{l,l,k}^{\text{H}}\mathbf{w}_{l,k}^{\text{ZF}}|^2]\\ &\overset{(b)}{=} \tfrac{\rho}{MN}\hat{k}_{l,l,k,\text{a}}\hat{\theta}_{l,l,k,\text{a}} + \tfrac{1}{MN}\tilde{k}_{l,l,k,\text{a}}\tilde{\theta}_{l,l,k,\text{a}},\end{aligned} \tag{E.1}$$

where (a) results from the independence of $\hat{\mathbf{g}}_{l,l,k}$ and $\tilde{\mathbf{g}}_{l,l,k}$, (b) is obtained because $|\hat{\mathbf{g}}_{l,l,k}^{\text{H}}\mathbf{w}_{l,k}^{\text{ZF}}|^2 \sim \Gamma\big(\rho\frac{\hat{k}_{l,l,k,\text{a}}}{MN}, \hat{\theta}_{l,l,k,\text{a}}\big)$ and $|\tilde{\mathbf{g}}_{l,l,k}^{\text{H}}\mathbf{w}_{l,k}^{\text{ZF}}|^2 \sim \Gamma\big(\frac{\tilde{k}_{l,l,k,\text{a}}}{MN}, \tilde{\theta}_{l,l,k,\text{a}}\big)$ which result from lemma 3 and the analysis in proof of Theorem 2.

The term $\sum_{j\neq k}\mathbb{E}[|\mathbf{g}_{l,l,k}^{\text{H}}\mathbf{w}_{l,j}^{\text{ZF}}|^2]$ can be computed as

$$\begin{aligned}\sum_{j\neq k}\mathbb{E}[|\mathbf{g}_{l,l,k}^{\text{H}}\mathbf{w}_{l,j}^{\text{ZF}}|^2] &\overset{(a)}{=} \sum_{j\neq k}\mathbb{E}[|\tilde{\mathbf{g}}_{l,l,k}^{\text{H}}\mathbf{w}_{l,j}^{\text{ZF}}|^2]\\ &\overset{(b)}{=} \tfrac{K-1}{MN}\tilde{k}_{l,l,k,\text{a}}\tilde{\theta}_{l,l,k,\text{a}},\end{aligned} \tag{E.2}$$

where (a) is obtained because $\hat{\mathbf{g}}_{l,l,k}^{\text{H}}\mathbf{w}_{l,j}^{\text{ZF}} = 0$ for $j \neq k$, and (b) results from Lemma 2 where we have applied Lemma 3 to approximate the distribution of $|\tilde{\mathbf{g}}_{l,l,k}^{\text{H}}\mathbf{w}_{l,j}^{\text{ZF}}|^2$ with $\Gamma\big(\frac{\tilde{k}_{l,l,k,\text{a}}}{MN}, \tilde{\theta}_{l,l,k,\text{a}}\big)$.

Considering the term $\sum_{i\neq l}\sum_{j=1}^{K}\mathbb{E}[|\mathbf{g}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}^{\text{ZF}}|^2]$, we have

$$\begin{aligned}&\sum_{i\neq l}\sum_{j=1}^{K}\mathbb{E}[|\mathbf{g}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}^{\text{ZF}}|^2]\\ &= \sum_{i\neq l}\sum_{j=1}^{K}\left(\mathbb{E}[|\hat{\mathbf{g}}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}^{\text{ZF}}|^2] + \mathbb{E}[|\tilde{\mathbf{g}}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}^{\text{ZF}}|^2]\right)\\ &= \sum_{i\neq l}\mathbb{E}[|\hat{\mathbf{g}}_{i,l,k}^{\text{H}}\mathbf{w}_{i,k}^{\text{ZF}}|^2] + \sum_{i\neq l}\sum_{j\neq k}\mathbb{E}[|\hat{\mathbf{g}}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}^{\text{ZF}}|^2]\\ &\quad + \sum_{i\neq l}\sum_{j=1}^{K}\mathbb{E}[|\tilde{\mathbf{g}}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}^{\text{ZF}}|^2]\\ &\overset{(a)}{=} \sum_{i\neq l}\left(\hat{k}_{i,l,k,\text{a}}\hat{\theta}_{i,l,k,\text{a}} + \tfrac{K}{MN}\tilde{k}_{i,l,k,\text{a}}\tilde{\theta}_{i,l,k,\text{a}}\right),\end{aligned} \tag{E.3}$$

where (a) results from Lemma 2 where we have applied Lemma 3 to approximate the distributions of $|\hat{\mathbf{g}}_{i,l,k}^{\text{H}}\mathbf{w}_{i,k}^{\text{ZF}}|^2$, $|\hat{\mathbf{g}}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}^{\text{ZF}}|^2$ and $|\tilde{\mathbf{g}}_{i,l,k}^{\text{H}}\mathbf{w}_{i,j}^{\text{ZF}}|^2$ with $\Gamma\big(\frac{\rho}{MN}\hat{k}_{i,l,k,\text{a}}, \hat{\theta}_{i,l,k,\text{a}}\big)$, $\Gamma\big(\frac{\tilde{k}_{i,l,k,\text{a}}}{MN}, \hat{\theta}_{i,l,k,\text{a}}\big)$ and $\Gamma\big(\frac{\tilde{k}_{i,l,k,\text{a}}}{MN}, \tilde{\theta}_{i,l,k,\text{a}}\big)$.

Combining all results yields the closed-form expression (45).

REFERENCES

[1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

[2] D. Wang, Y. Zhang, H. Wei, X. You, X. Gao, and J. Wang, "An overview of transmission theory and techniques of large-scale antenna systems for 5G wireless communications," *Sci. China Inf. Sci.*, vol. 59, no. 081301, pp. 1–18, Aug. 2016.

[3] L. Lu, G. Y. Li, A. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, Oct. 2014.

[4] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, Jan. 2013.

[5] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[6] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Transactions on Wireless Communications*, vol. 14, no. 8, pp. 4353–4368, Aug. 2015.

[7] 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 Communications*, vol. 17, no. 3, pp. 35–43, Jun. 2010.

[8] C. Wan and J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multicell environment," *IEEE Transactions on Wireless Communications*, vol. 6, no. 1, pp. 69–73, Jan. 2007.

[9] L. Dai, "A comparative study on uplink sum capacity with co-located and distributed antennas," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 6, pp. 1200–1213, June 2011.

[10] H. Zhu, "Performance comparison between distributed antenna and microcellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 6, pp. 1151–1163, June 2011.

[11] L. Dai, "An uplink capacity analysis of the distributed antenna system (DAS): From cellular DAS to DAS with virtual cells," *IEEE Transactions on Wireless Communications*, vol. 13, no. 5, pp. 2717–2731, May. 2014.

[12] T. Alade, H. Zhu, and J. Wang, "Uplink spectral efficiency analysis of in-building distributed antenna systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 4063–4074, July 2015.

[13] M. Matthaiou, C. Zhong, M. R. McKay, and T. Ratnarajah, "Sum rate analysis of ZF receivers in distributed MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 180–191, Feb. 2013.

[14] S.-R. Lee, S.-H. Moon, J.-S. Kim, and I. Lee, "Capacity analysis of distributed antenna systems in a composite fading channel," *IEEE Transactions on Wireless Communications*, vol. 11, no. 3, pp. 1076–1086, Mar. 2012.

[15] J. Li, D. Wang, P. Zhu, and X. You, "Spectral efficiency analysis of large-scale distributed antenna system in a composite correlated rayleigh fading channel," *IET Communications*, vol. 9, no. 5, pp. 681–688, April 2015.

[16] D. Wang, X. You, J. Wang, Y. Wang, and X. Hou, "Spectral efficiency of distributed MIMO cellular systems in a composite fading channel," in *Proc. IEEE International Conference on Communications (ICC'08)*, Prague, Czech Repulic, May 2008, pp. 1259–1264.

[17] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, Feb. 2016.

[18] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Transactions on Wireless Communications*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.

[19] T. V. Chien, E. Björnson, and E. G. Larsson, "Joint power allocation and user association optimization for massive MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 9, pp. 6384–6399, Sept. 2016.

[20] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[21] A. Kammoun, A. Müller, E. Björnson, and M. Debbah, "Linear precoding based on polynomial expansion: Large-scale multi-cell MIMO systems," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 861–875, Oct. 2014.

[22] J. Li, D. Wang, P. Zhu, and X. You, "Downlink spectral efficiency of multi-cell multi-user large-scale DAS with pilot contamination," in *Proc. IEEE International Conference on Communications (ICC)*, London, UK, June 2015, pp. 2011–2016.

[23] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293–1308, Feb. 2016.

[24] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. K. Frenger, "How much do downlink pilots improve cell-free massive MIMO?" in *Proc. IEEE Global Communications Conference (GLOBECOM)*, 2016, to appear. [Online]. Available: https://arxiv.org/abs/1607.04753.

[25] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots," in *Proc. 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, Oct. 2013, pp. 293–298.

[26] G. Coluccia, E. Riegler, C. Mecklenbrauker, and G. Taricco, "Optimum MIMO-OFDM detection with pilot-aided channel state information," *IEEE Journal of Selected Topics in Signal Processing*, vol. 3, no. 6, pp. 1053–1065, Dec. 2009.

[27] A. Khansefid and H. Minn, "Achievable downlink rates of MRC and ZF precoders in massive MIMO with uplink and downlink pilot contamination," *IEEE Transactions on Communications*, vol. 63, no. 12, pp. 4849–4864, Dec. 2015.

[28] J. Zhang and J. G. Andrews, "Adaptive spatical intercell interference cancellation in multicell wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1455–1468, Dec. 2010.

[29] J. Zhang, M. Kountouris, J. G. Andrews, and R. W. H. Jr., "Multi-mode transmission for the MIMO broadcast channel with imperfect channel state information," *IEEE Transactions on Communications*, vol. 59, no. 3, pp. 803–814, Mar. 2011.

[30] N. Seifi, J. Zhang, R. W. H. Jr., T. Svensson, and M. Coldrey, "Coordinated 3D beamforming for interference management in cellular networks," *IEEE Transactions on Wireless Communications*, vol. 13, no. 10, pp. 5396–5410, Oct. 2014.

[31] R. W. Heath, T. Wu, Y. H. Kwon, and A. C. K. Soong, "Multiuser MIMO in distributed antenna systems with out-of-cell interference," *IEEE Transactions on Signal Processing*, vol. 59, no. 10, pp. 4885–4899, Oct. 2011.

[32] K. Hosseini, W. Yu, and R. S. Adve, "Modeling and analysis of ergodic capacity in network MIMO systems," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2014, pp. 808–814.

[33] ——, "A stochastic analysis of network MIMO systems," *IEEE Transactions on Signal Processing*, vol. 64, no. 16, pp. 4113–4126, Aug. 2016.

[34] N. Seifi, R. W. H. Jr., M. Coldrey, and T. Svensson, "Joint transmission mode and tilt adaptation in coordinated small-cell networks," in *Proc. IEEE International Conference on Communications Workshops (ICC)*, Sydney, Australia, June 2014, pp. 598–603.

[35] K. Hosseini, W. Yu, and R. S. Adve, "Large-scale MIMO versus network MIMO for multicell interference mitigation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 930–941, Oct. 2014.

[36] N. Seifi, R. W. H. Jr., M. Coldrey, and T. Svensson, "Adaptive multicell 3D beamforming in multi-antenna cellular networks," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 8, pp. 6217–6231, Aug. 2016.

[37] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, 7th ed., A. Jeffrey and D. Zwillinger, Eds. Academic Press, 2007.

[38] J. Wang and L. Dai, "Asymptotic rate analysis of downlink multi-user systems with co-located and distributed antennas," *IEEE Transactions on Wireless Communications*, vol. 14, no. 6, pp. 3046–3058, June 2015.

[39] ——, "Downlink rate analysis for virtual-cell based large-scale distributed antenna systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 3, pp. 1998–2011, March 2016.

[40] H. Zhu and J. Wang, "Chunk-based resource allocation in ofdma systems - part i: chunk allocation," *IEEE Transactions on Communications*, vol. 57, no. 9, pp. 2734–2744, Sept. 2009.

[41] ——, "Chunk-based resource allocation in ofdma systems - part ii: joint chunk, power and bit allocation," *IEEE Transactions on Communications*, vol. 60, no. 2, pp. 499–509, Feb. 2012.

[42] H. Zhu, "Radio resource allocation for ofdma systems in high speed environments," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 4, pp. 748–759, May 2012.

[43] F. Fernandes, A. Ashikhmin, and T. L. Marzetta, "Inter-cell interference in noncooperative TDD large scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 192–201, Feb. 2013.

[44] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, Feb. 2013.

[45] Q. Sun, J. Wang, S. Jin, C. Xu, X. Gao, and K. K. Wong, "Rate analysis and pilot reuse design for dense small cell networks," in *Proc. IEEE International Conference on Communications (ICC)*, London, UK, June 2015, pp. 2845–2850.

[46] R. R. Müller, L. Cottatellucci, and M. Vehkaperä, "Blind pilot decontamination," *EEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 773–786, Oct. 2014.

[47] A. Adhikary, A. Ashikhmin, and T. L. Marzetta, "Uplink interference reduction in large scale antenna systems," in *Proc. IEEE International Symposium on Information Theory*, Honolulu, HI, USA, Jun. 2014, pp. 2529–2533.

[48] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. New York, NY: Wiley, 1982.

[49] J. Zhang and J. G. Andrews, "Adaptive spatial intercell interference cancellation in multicell wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 28, no. 9, pp. 1455–1468, Dec. 2010.

[50] S. Nadarajah, "A review of results on sums of random variables," *Acta Appl. Math.*, vol. 103, no. 2, pp. 131–140, Sep. 2008.

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

[52] L. Lin, "Asymptotic formulas associated with psi function with applications," *Journal of Mathematical Analysis and Applications*, vol. 405, no. 1, pp. 52–56, Sep. 2013.

[53] Y. G. Lim, C. B. Chae, and G. Caire, "Performance analysis of massive MIMO for cell-boundary users," *IEEE Transactions on Wireless Communications*, vol. 14, no. 12, pp. 6827–6842, Dec. 2015.

[54] D. Wang, J. Wang, X. You, Y. Wang, M. Chen, and X. Hou, "Spectral efficiency of distributed MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 10, pp. 2112–2127, Oct. 2013.

[55] A. P. Prudnikov, Y. A. Brychkov, and O. I. Marichev, *Integrals and Series, Volume 3: More Special Functions*, 1st ed. New York, NY, 1990.