SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •

February 2019, Vol. 62 022303:1–022303:14
https://doi.org/10.1007/s11432-018-9413-6

Impacts of practical channel impairments on the downlink spectral efficiency of large-scale distributed antenna systems

Jiamin LI*, Dongming WANG, Pengcheng ZHU & Xiaohu YOU

National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096*, China*

Received 18 January 2018/Revised 11 March 2018/Accepted 29 March 2018/Published online 18 December 2018

Abstract Channel impairments are major limiting factors in the performance of large-scale antenna systems. In this paper, we analyze the impacts of practical channel impairments caused by pilot contamination, Doppler shift, and phase noise on the downlink spectral efficiency of large-scale distributed antenna systems (L-DASs) with maximum ratio transmission (MRT) and zero-forcing (ZF) beamforming, in which per user power normalization is considered. Using a joint channel model that allows study of the simultaneous impacts of these channel impairments, we derive accurate and tractable closed-form approximations for the ergodic achievable downlink rate, thereby enabling spectral efficiency analysis of L-DASs and an efficient evaluation of the impacts of the channel impairments. It is shown that channel impairments reduce the downlink spectral efficiency and have a significant impact on ZF beamforming. The asymptotic user rate limit is also determined, from which we analyze the asymptotic performance of L-DASs with channel impairments. The analytical results show that MRT and ZF beamforming achieve the same asymptotic performance limit even with channel impairments. It is also found that the use of a large-scale antenna array at the base station sides can weaken the impacts of channel impairments.

Keywords Doppler shift, phase noise, large-scale distributed antenna systems, spectral efficiency, pilot contamination

Citation Li J M, Wang D M, Zhu P C, et al. Impacts of practical channel impairments on the downlink spectral efficiency of large-scale distributed antenna systems. Sci China Inf Sci, 2019, 62(2): 022303, https://doi.org/10.1007/s11432-018-9413-6

1 Introduction

Large-scale antenna systems (LSASs), also called massive multi-input multi-output (MIMO) or large-scale MIMO, have emerged as one of the most promising technologies for the future of wireless networks [1–5]. The advantages of LSASs were initially validated by assuming ideal propagation conditions. However, understanding the performance limits of LSASs with practical channel impairments is imperative. The impacts of various channel impairments on LSASs have been previously studied by considering the pilot contamination caused by the reuse of the same pilot sequences in adjacent cells [6–8], Doppler shift resulting from the relative movement between users and base stations (BSs) [9–13], and phase noise due to imperfect local oscillators (LOs) [14–18]. Some important observations can be summarized as follows: (1) pilot contamination constitutes a performance bottleneck in LSASs and (2) channel aging induced by Doppler shift and phase noise, which refers to the mismatch between the channel state information (CSI)

* Corresponding author (email: lijiamin@seu.edu.cn)

when it is estimated vs. when it is used for detection or precoding, contributes to the further degradation of LSASs.

The large-scale antenna array at BSs can be co-located or geographically distributed in cells; thus describing the two extremes of the LSAS paradigm, i.e., co-located LSASs and large-scale distributed antenna systems (L-DASs) [19]. Refs. [20–24] have shown that L-DASs can provide larger performance gains compared with co-located LSASs. However, a critical issue potentially exists for an L-DAS if it is too sensitive to various channel impairments. As shown above, the impacts of channel impairments on co-located LSASs have received much attention in recent years, whereas only a few studies have considered distributed arrays [16, 25–27]. Refs. [16, 25] analyzed the impacts of hardware imperfections at BSs on L-DASs with maximum ratio combing (MRC) receiver and maximum ratio transmission (MRT) beamforming, respectively. It was shown that L-DASs are resilient to additive distortions but multiplicative phase noise is a limiting factor for both uplink and downlink processes in L-DASs. The drawbacks of the studies by [16, 25] are that their analyses did not account for the impact of Doppler shift and only MRT and MRC were considered. A realistic channel model accounting for the impacts of both Doppler shift and phase noise was recently proposed by [26]; this approach allows for a more complete characterization of channel aging. However, Ref. [26] did not consider the impact of pilot contamination. Moreover, considering that 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, average power normalization was assumed by [16, 26], rather than practical per user power normalization, to give analytical tractability. Since the difference in performance between the two normalization schemes is relatively large in distributed systems [28, 29], per user power normalization is considered practical. Recently, a joint channel model was provided in [27]. This model incorporated the impacts of channel impairments caused by pilot contamination, Doppler shift, and phase noise. Based on this channel model, the impacts of channel impairments on the spectral efficiency of L-DASs were analyzed. However, only uplink was considered.

Motivated by the aspects mentioned above, we investigated how practical channel impairments influence the downlink spectral efficiency of L-DASs with both MRT and zero-forcing (ZF) beamforming under per user power normalization. Specially, our main contributions can be outlined as follows.

- The ergodic achievable downlink rate of L-DASs after applying MRT and ZF beamforming under per user power normalization is given in closed form in the presence of practical channel impairments, thereby enabling spectral efficiency analysis of L-DASs with channel impairments and efficient evaluation of the impacts of channel impairments.
- The asymptotic user rate limit is also given, from which we analyze the asymptotic performance of L-DASs with practical channel impairments.
- We corroborate our theoretical analysis with simulations, and draw insightful conclusions from our analysis of the impacts of channel impairments on the downlink performance of L-DASs with MRT and ZF beamforming under per user power normalization.

The remainder of this paper is organized as follows. In Section 2, we introduce the system model including its channel configuration, channel model with channel impairments, downlink signal model and achievable rates. Section 3 describes the main analytical work where we provide the derivations of the ergodic achievable downlink rate and the asymptotic user rate limit after applying MRT and ZF beamforming under per user power normalization in the presence of practical channel impairments caused by pilot contamination, Doppler shift, and phase noise. In Section 4, we collocate our theoretical analysis with simulations. Section 5 concludes the paper. In the appendices, we provide proof for the main analytical results.

Notation. Vectors and matrices are denoted by boldface lower case and upper case letters. $\boldsymbol{I}_N$ is the size-N identity matrix. $(\cdot)^{\mathrm{H}}$, $\mathrm{tr}\{\cdot\}$ and $\mathrm{E}[\cdot]$ represent the conjugate transpose operator, trace operator and expectation operator, respectively. $\|\cdot\|$ denotes the spectral norm of a matrix and $|\cdot|$ denotes the absolute value of a scalar. The $\mathrm{diag}\{\cdot\}$ operator generates a square diagonal matrix with the elements of the given vector on the main diagonal. We use $\mathcal{N}(\mu, \sigma^2)$ to denote the Gaussian distribution with mean μ and variance σ^2, and we use $\mathcal{CN}(0, \sigma^2)$ to denote the circularly symmetric complex Gaussian distribution with mean zero and variance σ^2. $\Gamma(k, \theta)$ denotes the Gamma distribution with shape parameter k and scale parameter θ. Nakagami(m,Ω) denotes the Nakagami distribution with shape parameter m and controlling spread parameter Ω.

2 System model with channel impairments

In this section, we describe the considered multi-cell multi-user L-DASs with frequency-flat fading channels, time-division duplex (TDD) mode, pilot contamination, Doppler shift, phase noise, and linear precoding under per user power normalization.

2.1 System configuration and channel model

Consider an L-DAS with L cells and operating under TDD protocol. Each cell consists of K single-antenna users that communicate simultaneously with M remote antenna units (RAUs). Each RAU is equipped with N antennas. Moreover, a quasi-static block fading channel model [9] is considered in this paper. Then, we let $\bar{\boldsymbol{g}}_{i,l,k}[t]$ be the channel vector from user k in cell l to all of the RAUs in cell i at time t, which can be modeled as

$$\bar{\boldsymbol{g}}_{i,l,k}[t] = \boldsymbol{\Lambda}_{i,l,k}^{1/2}\boldsymbol{h}_{i,l,k}[t], \tag{1}$$

where

$$\boldsymbol{\Lambda}_{i,l,k} \triangleq \mathrm{diag}\{\lambda_{i,1,l,k},\ldots,\lambda_{i,M,l,k}\} \otimes \boldsymbol{I}_N, \tag{2}$$

$\boldsymbol{h}_{i,l,k}[t] = [\boldsymbol{h}_{i,1,l,k}^{\mathrm{T}}[t],\ldots,\boldsymbol{h}_{i,M,l,k}^{\mathrm{T}}[t]]^{\mathrm{T}}$, $\boldsymbol{h}_{i,m,l,k}[t] \sim \mathcal{CN}(0,\boldsymbol{I}_N)$ is the small-scale fading between user k in cell l and RAU m in cell i at time t, and $\lambda_{i,m,l,k}$ represents the corresponding large-scale effect including shadowing and pathloss, which changes slowly and can be learned over a long period of time.

2.2 Joint channel model with channel impairments

In this subsection, we present a joint channel model that incorporates the impacts of practical channel impairments caused by pilot contamination, Doppler shift, and phase noise. As in [26, 27], we assume that each frame of duration T_{c} symbols consists of the uplink training phase of K symbols (the minimum number of pilot symbols) with time indices $t = -K+1,\ldots,0$, followed by a downlink data transmission phase of $T_{\mathrm{c}} - K$ symbols with time indices $t = 1,\ldots,T_{\mathrm{c}}-K$; channel estimation takes place at time 0.

According to the autoregressive model of 1 [9, 30], the impact of channel aging induced by Doppler shift can be modeled as

$$\bar{\boldsymbol{g}}_{i,l,k}[t] = \alpha_t\bar{\boldsymbol{g}}_{i,l,k}[0] + \bar{\boldsymbol{e}}_{i,l,k}[t], \tag{3}$$

where $\alpha_t \triangleq J_0(2\pi f_{\mathrm{D}}T_{\mathrm{s}}t)$ represents the temporal correlation parameter modeled as Jakes model; $J_0(\cdot)$ denotes the zero-order Bessel function of the first kind; the maximum Doppler shift f_{D} is given by $f_{\mathrm{D}} = \frac{vf_{\mathrm{c}}}{\mathrm{c}}$ (v is the relative velocity of the user, f_{c} is the carrier frequency, and $\mathrm{c} = 3\times10^8$ mps is the speed of light); T_{s} denotes the symbol time; and $\bar{\boldsymbol{e}}_{i,l,k}[t] \sim \mathcal{CN}(0,(1-\alpha_t^2)\boldsymbol{\Lambda}_{i,l,k})$ represents the uncorrelated channel error vector.

Phase noise is a further source of channel aging and induces an extra loss caused by noisy LOs at BSs and users. The impact of phase noise can be modeled as [31]

$$\boldsymbol{g}_{i,l,k}[t] = \boldsymbol{\Theta}_{i,l,k}[t]\bar{\boldsymbol{g}}_{i,l,k}[t], \tag{4}$$

where $\boldsymbol{\Theta}_{i,l,k}[t] \triangleq \mathrm{e}^{\mathrm{j}\varphi_{l,k}[t]}\mathrm{diag}\{\mathrm{e}^{\mathrm{j}\phi_{i,1}[t]},\ldots,\mathrm{e}^{\mathrm{j}\phi_{i,M}[t]}\}\otimes\boldsymbol{I}_N$, $\varphi_{l,k}[t]$ and $\phi_{i,m}[t]$ are the noises at the LOs of user k in cell l and RAU m in cell i at time t, which follow the Wiener processes [16] with independent phase noise increment variances $\sigma_{\varphi_{l,k}}^2$ and $\sigma_{\phi_{i,m}}^2$, i.e., $\varphi_{l,k}[t] \sim \mathcal{N}(\varphi_{l,k}[t-1],\sigma_{\varphi_{l,k}}^2)$, $\phi_{i,m}[t] \sim \mathcal{N}(\phi_{i,m}[t-1],\sigma_{\phi_{i,m}}^2)$.

The impacts of channel aging induced by Doppler shift in (3) and phase noise in (4) can be incorporated as a multiplicative parameter matrix, and the channel vector at time t can be given by [26, 27]

$$\boldsymbol{g}_{i,l,k}[t] = \boldsymbol{\Psi}_{i,l,k}[t]\boldsymbol{g}_{i,l,k}[0] + \boldsymbol{e}_{i,l,k}[t], \tag{5}$$

where

$$\boldsymbol{\Psi}_{i,l,k}[t] \triangleq \alpha_t \mathrm{e}^{-\frac{\sigma^2_{\varphi_{l,k}}}{2}t} \mathrm{diag}\left\{\mathrm{e}^{-\frac{\sigma^2_{\phi_{i,1}}}{2}t}, \ldots, \mathrm{e}^{-\frac{\sigma^2_{\phi_{i,M}}}{2}t}\right\} \otimes \boldsymbol{I}_N,$$

which can be obtained by minimizing the mean square error defined as $\mathrm{tr}\{\mathrm{E}[\boldsymbol{e}_{i,l,k}[t]\boldsymbol{e}^{\mathrm{H}}_{i,l,k}[t]]\}$ [26, Theorem 1], and $\boldsymbol{e}_{i,l,k}[t] \sim \mathcal{CN}(0, \boldsymbol{\Lambda}_{i,l,k} - \boldsymbol{\Psi}_{i,l,k}[t]\boldsymbol{\Lambda}_{i,l,k}\boldsymbol{\Psi}_{i,l,k}[t])$ is the combined error vector, which depends on the impacts of Doppler shift and phase noise.

As seen from (5), the channel vector $\boldsymbol{g}_{i,l,k}[t]$ at time t is obtained in terms of the channel vector $\boldsymbol{g}_{i,l,k}[0]$ at time 0. However, $\boldsymbol{g}_{i,l,k}[0]$ is usually not available and pilot-assisted transmission is employed at BSs to perform channel estimation. When the channel coherence time is limited, non-orthogonal pilot sequences must be reused in adjacent cells, which results in pilot contamination, i.e., correlated interference from the users using the same pilot sequences. Note that in this paper, we assume that the users in each cell use orthogonal pilots, i.e., the length of the uplink training phase is equal to K, which results in only inter-cell pilot contamination. Otherwise, there is still intra-cell pilot contamination [32].

During the uplink training phase ($t = -K+1, \ldots, 0$), as in [10, 26, 33], we assume that both the phase noise and channel remain constant, and we neglect the channel estimation error resulting from the channel aging effect. Considering that the duration of the uplink training phase is short and the consequent variation of the channel is unnoticeable [33], this assumption is valid and yields a simple and tractable model that enables us to analyze the simultaneous impacts of practical channel impairments. Note that during the downlink data transmission phase ($t = 1, \ldots, T_{\mathrm{c}} - K$), the channels are supposed to vary from symbol to symbol due to the channel aging effect. At time 0, the pilot symbols received by the users in cell i are given by

$$\boldsymbol{Y}_i[0] = \sum_{l=1}^{L} \boldsymbol{G}_{i,l}[0]\boldsymbol{\Phi} + \boldsymbol{Z}_i[0], \tag{6}$$

where $\boldsymbol{G}_{i,l}[0] = [\boldsymbol{g}_{i,l,1}[0], \ldots, \boldsymbol{g}_{i,l,K}[0]]$ is the channel matrix between all of the users in cell l to all of the RAUs in cell i, $\boldsymbol{g}_{i,l,k}[0] = \boldsymbol{\Theta}_{i,l,k}[0]\bar{\boldsymbol{g}}_{i,l,k}[0]$ from (4), $\boldsymbol{\Phi} \in \mathbb{C}^{K\times K}$ is the pilot matrix with pairwise orthogonal rows satisfying $\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathrm{H}} = \boldsymbol{I}_K$, $\boldsymbol{Z}_l$ is the noise matrix with $\mathcal{CN}(0, 1/\gamma_{\mathrm{P}})$ elements, and γ_{P} is the training signal-to-noise ratio (SNR). After correlating the received pilot symbols with the pilot sequence assigned for user k, i.e., the k-th row of $\boldsymbol{\Phi}$, the channel vector $\boldsymbol{g}_{i,l,k}[0]$ can be estimated based on the observation $\boldsymbol{y}_{i,k}[0]$, given by

$$\boldsymbol{y}_{i,k}[0] = \boldsymbol{g}_{i,l,k}[0] + \sum_{j\neq i}^{L} \boldsymbol{g}_{i,j,k}[0] + \boldsymbol{z}_{i,k}[0], \tag{7}$$

where $\boldsymbol{y}_{i,k}[0]$ and $\boldsymbol{z}_{i,k}[0]$ are the k-th column of $\boldsymbol{Y}_i[0]$ and $\boldsymbol{Z}_i[0]$, respectively. The linear minimum mean square error (LMMSE) estimate [34, Subsection 12.3] of $\boldsymbol{g}_{i,l,k}[0]$ is given by

$$\hat{\boldsymbol{g}}_{i,l,k}[0] = \boldsymbol{\Lambda}_{i,l,k}\boldsymbol{\Xi}^{-1}_{i,k}\boldsymbol{y}_{i,k}[0], \tag{8}$$

where

$$\boldsymbol{\Xi}_{i,k} \triangleq \sum_{j=1}^{L} \boldsymbol{\Lambda}_{i,j,k} + 1/\gamma_{\mathrm{P}}\boldsymbol{I}_{MN}. \tag{9}$$

With the definition of

$$\hat{\boldsymbol{h}}_{i,k}[0] \triangleq \boldsymbol{\Xi}^{-1/2}_{i,k}\boldsymbol{y}_{i,k}[0], \tag{10}$$

which is distributed as $\hat{\boldsymbol{h}}_{i,k}[0] \sim \mathcal{CN}(0, \boldsymbol{I}_{MN})$ and hence can be regarded as the equivalent Rayleigh fading portion of the channel estimate, we can rewrite the LMMSE estimate $\hat{\boldsymbol{g}}_{i,l,k}[0]$ in (8) as

$$\hat{\boldsymbol{g}}_{i,l,k}[0] = \mathrm{diag}\left\{\kappa_{i,1,l,k}, \ldots, \kappa_{i,M,l,k}\right\} \otimes \boldsymbol{I}_N\hat{\boldsymbol{h}}_{i,k}[0], \tag{11}$$

where $\kappa_{i,m,l,k} \triangleq \lambda_{i,m,l,k}(\sum_{j=1}^{L} \lambda_{i,M,j,k} + 1/\gamma_{\rm P})^{-1/2}$. Due to the properties of the LMMSE estimate, the channel vector $\boldsymbol{g}_{i,l,k}[0]$ can be decomposed as

$$\boldsymbol{g}_{i,l,k}[0] = \hat{\boldsymbol{g}}_{i,l,k}[0] + \tilde{\boldsymbol{g}}_{i,l,k}[0], \tag{12}$$

where $\tilde{\boldsymbol{g}}_{i,l,k}[0] \sim \mathcal{CN}(0, \boldsymbol{\Lambda}_{i,l,k} - \boldsymbol{\Lambda}_{i,l,k}\boldsymbol{\Xi}_{i,k}^{-1}\boldsymbol{\Lambda}_{i,l,k})$ is the estimation error vector which is independent of the channel estimate $\hat{\boldsymbol{g}}_{i,l,k}[0]$.

Substituting (12) into (5) yields the following joint channel model:

$$\boldsymbol{g}_{i,l,k}[t] = \hat{\boldsymbol{g}}_{i,l,k}[t] + \tilde{\boldsymbol{g}}_{i,l,k}[t], \tag{13}$$

for $t \in \{1, \ldots, T_{\rm c} - K\}$, where

$$\hat{\boldsymbol{g}}_{i,l,k}[t] \triangleq \boldsymbol{\Psi}_{i,l,k}[t]\hat{\boldsymbol{g}}_{i,l,k}[0] = {\rm diag}\left\{\beta_{i,1,l,k}^{1/2}, \ldots, \beta_{i,1,l,k}^{1/2}\right\} \otimes \boldsymbol{I}_N \hat{\boldsymbol{h}}_{i,k}[0] \tag{14}$$

is the available CSI at BS i at time t,

$$\tilde{\boldsymbol{g}}_{i,l,k}[t] \triangleq \boldsymbol{\Psi}_{i,l,k}[t]\tilde{\boldsymbol{g}}_{i,l,k}[0] + \boldsymbol{e}_{i,l,k}[t] \sim \mathcal{CN}\left(0, {\rm diag}\left\{\eta_{i,1,l,k}[t], \ldots, \eta_{i,M,l,k}[t]\right\} \otimes \boldsymbol{I}_N\right) \tag{15}$$

is the joint channel error vector, which is independent of the available CSI $\hat{\boldsymbol{g}}_{i,l,k}[t]$, and

$$\beta_{i,m,l,k}[t] \triangleq \alpha_t^2 {\rm e}^{-\sigma_{\varphi_{l,k}}^2 t} {\rm e}^{-\sigma_{\phi_{i,m}}^2 t} \frac{\lambda_{i,m,l,k}^2}{\sum_{j=1}^{L} \lambda_{i,m,j,k} + 1/\gamma_{\rm P}}, \tag{16}$$

$$\eta_{i,m,l,k}[t] \triangleq \lambda_{i,m,l,k} - \beta_{i,m,l,k}[t]. \tag{17}$$

The joint channel model (13) incorporates the impacts of practical channel impairments caused by pilot contamination and channel aging induced by Doppler shift and phase noise. Based on (13), we can analyze the simultaneous impacts of these practical channel impairments on the spectral efficiency of L-DASs. The tractable joint channel model (13) is very general, with the conventional models in [26] ($L = 1$), in [31] ($f_{\rm D} = 0$) and in [9] ($\sigma_i^2 = 0$, for $i = \varphi_{l,k}$ or $\phi_{i,m}$) as special cases. Thus, the results and analysis herein hold for arbitrary parameters. Moreover, from (14), it can be seen that owing to the effect of pilot contamination, the achievable CSI $\hat{\boldsymbol{g}}_{i,l,k}[t]$ and $\hat{\boldsymbol{g}}_{i,j,k}[t]$ become correlated random vectors, although the channel vectors $\boldsymbol{g}_{i,l,k}[t]$ and $\boldsymbol{g}_{i,j,k}[t]$ are independent for $j \neq l$.

Remark 1. The joint model (13) is based on the model proposed in [27]. For simplicity, [27] assumed that the phase noise increment variances of all of the M RAUs in cell i are equal as in [16], i.e., $\sigma_{\phi_{i,1}}^2 = \cdots = \sigma_{\phi_{i,M}}^2 = \sigma_{\phi_i}^2$. In (13), we extend the model presented in [27] to cases with different phase noise increment variances at the RAUs, which makes spectral efficiency analysis of L-DASs with channel impairments more challenging.

2.3 Downlink signal model and achievable rates

During the downlink data transmission phase ($t = 1, \ldots, T_{\rm c} - K$), the channels vary from symbol to symbol due to the channel aging effect. At time $t \in \{1, \ldots, T_{\rm c} - K\}$, the received signal $y_{l,k}[t] \in \mathcal{C}$ of user k in cell l is written as

$$y_{l,k}[t] = \sum_{i=1}^{L}\sum_{j=1}^{K} \boldsymbol{g}_{i,l,k}^{\rm H}[t]\boldsymbol{w}_{i,j}[t]s_{i,j}[t] + z_{l,k}[t], \tag{18}$$

where $s_{i,j}[t] \sim \mathcal{CN}(0, 1)$ is the transmitted data symbol assigned for user j in cell i at time t, $\boldsymbol{w}_{i,j}[t]$ is the associated beamforming vector, and $z_{l,k}[t] \sim \mathcal{CN}(0, 1/\gamma_{\rm DL})$ indicates the receiver noise.

Assuming that users detect the transmitted signals with statistical CSI, i.e, ${\rm E}[\boldsymbol{g}_{l,l,k}^{\rm H}[t]\boldsymbol{w}_{l,k}[t]]$, and treating the remaining signal component and the interference plus noise as worst-case Gaussian distributed noise, the ergodic achievable rate of user k in cell l can be given by [35, Theorem 1]

$$R_{l,k}[t] = \log_2\left(1 + \frac{|{\rm E}[\boldsymbol{g}_{l,l,k}^{\rm H}[t]\boldsymbol{w}_{l,k}[t]]|^2}{{\rm var}\left[\boldsymbol{g}_{l,l,k}^{\rm H}[t]\boldsymbol{w}_{l,k}[t]\right] + \sum_{(i,j)\neq(l,k)} {\rm E}\left[|\boldsymbol{g}_{i,l,k}^{\rm H}[t]\boldsymbol{w}_{i,j}[t]|^2\right] + 1/\gamma_{\rm DL}}\right). \tag{19}$$

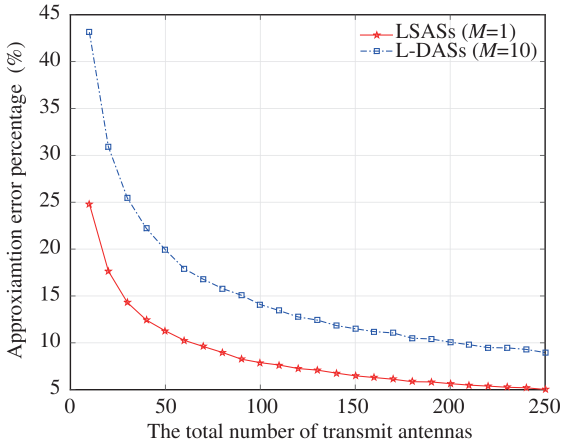

Figure 1 (Color online) Fluctuation of the norms of MRT beamforming vectors around their means versus the total number of transmit antennas.

Note that the ergodic achievable downlink rate analysis in this paper is performed under the condition of equal power allocation among users. However, it is straightforward to generalize the rate analysis to any given power allocation scheme by changing the data symbol from $s_{l,k}[t] \sim \mathcal{CN}(0,1)$ to $s_{l,k}[t] \sim \mathcal{CN}(0,p_{l,k})$ with power $p_{l,k} = \mathrm{E}[|s_{l,k}[t]|^2]$ and changing the noise from $z_{l,k}[t] \sim \mathcal{CN}(0,1/\gamma_{\mathrm{DL}})$ to $z_{l,k}[t] \sim \mathcal{CN}(0,\sigma^2)$; this makes no difference to the following derivation of the ergodic achievable downlink rate. Thus, the results obtained in this paper are applicable to an arbitrary power control strategy. Optimization of the power allocation among users requires a separate investigation.

In the present work, attention was restricted to MRT and ZF beamforming. With $\hat{\boldsymbol{g}}_{i,l,k}[t]$, the available CSI at BS i, and time t, the MRT and ZF beamforming vectors under practical per user power normalization are defined as

$$\boldsymbol{w}_{l,k}[t] = \begin{cases} \dfrac{\hat{\boldsymbol{g}}_{l,l,k}[t]}{\|\hat{\boldsymbol{g}}_{l,l,k}[t]\|}, & \text{for MRT,} \\ \dfrac{\boldsymbol{f}_{l,l,k}[t]}{\|\boldsymbol{f}_{l,l,k}[t]\|}, & \text{for ZF,} \end{cases} \tag{20}$$

respectively, where $\boldsymbol{f}_{l,l,k}[t]$ is the k-th column of $\boldsymbol{F}_l[t](\boldsymbol{F}_l^{\mathrm{H}}[t]\boldsymbol{F}_l[t])^{-1}$, and $\boldsymbol{F}_l[t] = [\hat{\boldsymbol{g}}_{l,l,1}[t],\ldots,\hat{\boldsymbol{g}}_{l,l,K}[t]]$.

Remark 2. Due to the effect of channel hardening [36], i.e.,

$$\frac{|\mathrm{E}[\|\boldsymbol{v}_{l,l,k}\|^2] - \|\boldsymbol{v}_{l,l,k}\|^2|}{\mathrm{E}[\|\boldsymbol{v}_{l,l,k}\|^2]} \xrightarrow{MN\to\infty} 0, \tag{21}$$

average power normalization instead of per user power normalization (divided by $\sqrt{\mathrm{E}[\|\boldsymbol{v}_{l,l,k}\|^2]}$ instead of $\|\boldsymbol{v}_{l,l,k}\|$ in (20)) was assumed in [37–41] to give analytical tractability, where $\boldsymbol{v}_{l,l,k} = \hat{\boldsymbol{g}}_{l,l,k}$ or $\boldsymbol{f}_{l,l,k}$. Figure 1 shows the difference in performance between the two power normalization schemes. As seen from the figure, with MRT beamforming, about 200 antennas are needed for L-DASs with $M = 10$ RAUs to achieve a 10% approximation error (defined in (21)), whereas only about 60 antennas are needed for co-located LSASs. This is because for L-DASs, the M RAUs are geographically distributed in cells; therefore, each user may be effectively served by only a portion of the RAUs, which leads to less channel hardening. Consequently, per user power normalization is more practical for L-DASs [28].

3 Downlink spectral efficiency analysis with channel impairments

In this section, we provide the derivations of the ergodic achievable downlink rates and asymptotic user rate limit after applying MRT and ZF beamforming under per user power normalization in the presence of practical channel impairments caused by pilot contamination and channel aging induced by Doppler shift and phase noise.

For the channel vector $\boldsymbol{g}_{i,l,k}[t]$, its strength can be expressed as

$$\boldsymbol{g}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{g}_{i,l,k}[t] = \sum_{m=1}^{M} \lambda_{i,m,l,k}\boldsymbol{h}_{i,m,l,k}^{\mathrm{H}}\boldsymbol{h}_{i,m,l,k}, \tag{22}$$

which is the summation of M independent but non-identically distributed terms and $\lambda_{i,m,l,k}\boldsymbol{h}_{i,m,l,k}^{\mathrm{H}}\boldsymbol{h}_{i,m,l,k} \sim \Gamma(N, \lambda_{i,m,l,k})$ [21]. In order to yield a mathematically tractable expression, we employ Lemma 1 to obtain an approximation of the distribution of $\boldsymbol{g}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{g}_{i,l,k}[t]$ in (22).

Lemma 1 ([42], Proposition 8)**.** Suppose that $\{x_i\}$ are independent $\Gamma(k_i, \theta_i)$ random variables. The Gamma distributed random variable $\Gamma(k, \theta)$ has the same first and second moments as $\sum_i x_i$, with the parameters given by

$$k = \frac{(\sum_i k_i\theta_i)^2}{\sum_i k_i\theta_i^2} \text{ and } \theta = \frac{\sum_i k_i\theta_i^2}{\sum_i k_i\theta_i}. \tag{23}$$

As a consequence of Lemma 1, we have

$$\boldsymbol{g}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{g}_{i,l,k}[t] \sim \Gamma(k_{i,l,k,\mathrm{a}}[t], \theta_{i,l,k,\mathrm{a}}[t]), \tag{24}$$

where

$$k_{i,l,k,\mathrm{a}}[t] = \frac{N(\sum_{m=1}^{M}\lambda_{i,m,l,k})^2}{\sum_{m=1}^{M}\lambda_{i,m,l,k}^2}, \tag{25}$$

$$\theta_{i,l,k,\mathrm{a}}[t] = \frac{\sum_{m=1}^{M}\lambda_{i,m,l,k}^2}{\sum_{m=1}^{M}\lambda_{i,m,l,k}}, \tag{26}$$

and a subscript "a" indicates approximation.

Given the distributions (24), the distributions of projection powers when channel vectors are projected onto the beamforming subspace can be obtained by applying Lemma 2.

Lemma 2 ([29], Lemma 3)**.** The projection power of the m-dimensional non-isotropic channel vector $\boldsymbol{g}_{i,l,k}[t]$ projected onto a s-dimensional subspace is distributed as $\Gamma(sk_{i,l,k,\mathrm{a}}[t]/m, \theta_{i,l,k,\mathrm{a}}[t])$ with $k_{i,l,k,\mathrm{a}}[t]$ and $\theta_{i,l,k,\mathrm{a}}[t]$ defined in (25) and (26).

Note that when characterizing the useful signal power and pilot contamination power, $s = MN$ with MRT beamforming and $s = MN - K + 1$ with ZF beamforming, whereas $s = 1$ for both MRT and ZF beamforming when characterizing interference power [43, 44].

Remark 3. The approximation accuracy of Lemma 2 depends on whether the path losses from a given user to the RAUs are similar. When the path losses vary drastically, we can increase the approximation accuracy by changing the dimension of the subspace from MN to $M_{\mathrm{s}}N$ for MRT beamforming or from $MN - K + 1$ to $M_{\mathrm{s}}N - K + 1$ for ZF beamforming, where M_{s} is the number of RAUs with similar and relatively less path losses [29].

Refs. [44–47] studied ergodic achievable rate approximation under the assumption of perfect CSI at BSs. In contrast, the present paper focuses on the practical case of imperfect CSI in the presence of channel impairments caused by pilot contamination, Doppler shift, and phase noise. Therefore, we need to further characterize the distributions of the powers of the non-isotropic available CSI $\hat{\boldsymbol{g}}_{i,l,k}[t]$ and the combined error vectors $\tilde{\boldsymbol{g}}_{i,l,k}[t]$ projected onto a s-dimensional beamforming subspace. We can first characterize the approximation distributions of $\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\hat{\boldsymbol{g}}_{i,l,k}[t]$ and $\tilde{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\tilde{\boldsymbol{g}}_{i,l,k}[t]$ based on Lemma 1, which can be given by

$$\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\hat{\boldsymbol{g}}_{i,l,k}[t] \sim \Gamma(\hat{k}_{i,l,k,\mathrm{a}}[t], \hat{\theta}_{i,l,k,\mathrm{a}}[t]), \tag{27}$$

$$\tilde{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\tilde{\boldsymbol{g}}_{i,l,k}[t] \sim \Gamma(\tilde{k}_{i,l,k,\mathrm{a}}[t], \tilde{\theta}_{i,l,k,\mathrm{a}}[t]), \tag{28}$$

where

$$\hat{k}_{i,l,k,\mathrm{a}}[t] = \frac{N(\sum_{m=1}^{M}\beta_{i,m,l,k}[t])^2}{\sum_{m=1}^{M}\beta_{i,m,l,k}^2[t]}, \tag{29}$$

$$\hat{\theta}_{i,l,k,\mathrm{a}}[t] = \frac{\sum_{m=1}^{M} \beta_{i,m,l,k}^2[t]}{\sum_{m=1}^{M} \beta_{i,m,l,k}[t]}, \tag{30}$$

$$\tilde{k}_{i,l,k,\mathrm{a}}[t] = \frac{N(\sum_{m=1}^{M} \eta_{i,m,l,k}[t])^2}{\sum_{m=1}^{M} \eta_{i,m,l,k}^2[t]}, \tag{31}$$

$$\tilde{\theta}_{i,l,k,\mathrm{a}}[t] = \frac{\sum_{m=1}^{M} \eta_{i,m,l,k}^2[t]}{\sum_{m=1}^{M} \eta_{i,m,l,k}[t]}, \tag{32}$$

and then, calculate the distributions of the projection powers of $\hat{\boldsymbol{g}}_{i,l,k}[t]$ and $\tilde{\boldsymbol{g}}_{i,l,k}[t]$ by applying Lemma 2.

Based on the ideas discussed above, we are able analyze the impacts of the practical channel impairments on the downlink spectral efficiency of L-DASs. Theorems 1 and 2, and Corollary 1 contain the main contributions of this paper and provide the derivations of the closed-form approximations for the ergodic achievable downlink rate (19) and the asymptotic user rate limit in L-DASs with MRT and ZF beamforming under per user power normalization in the presence of practical channel impairments.

Theorem 1. When MRT beamforming under per user power normalization is used, the closed-form approximation for the ergodic achievable downlink rate (19) in L-DASs with practical channel impairments is given by

$$R_{l,k}^{\mathrm{MRT}}[t] = \log_2\left(1 + \frac{\xi(\hat{k}_{l,l,k,\mathrm{a}}[t])\hat{\theta}_{l,l,k,\mathrm{a}}[t]}{\mathcal{I}_{l,k}^{\mathrm{MRT}}[t] + \sum_{i\neq l} \hat{k}_{i,l,k,\mathrm{a}}[t]\hat{\theta}_{i,l,k,\mathrm{a}}[t]}\right), \tag{33}$$

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

$$\xi(x) \triangleq \Gamma(x + 1/2)/\Gamma(x). \tag{34}$$

Proof. See Appendix 1.

Theorem 2. When ZF beamforming under per user power normalization is used, the closed-form approximation for the ergodic achievable downlink rate (19) in L-DASs with practical channel impairments is given by

$$R_{l,k}^{\mathrm{ZF}}[t] = \log_2\left(1 + \frac{\xi(\rho\hat{k}_{l,l,k,\mathrm{a}}[t])\hat{\theta}_{l,l,k,\mathrm{a}}[t]}{\mathcal{I}_{l,k}^{\mathrm{ZF}}[t] + \sum_{i\neq l} \hat{k}_{i,l,k,\mathrm{a}}[t]\hat{\theta}_{i,l,k,\mathrm{a}}[t]}\right), \tag{35}$$

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

$$\rho \triangleq \frac{MN - K + 1}{MN}. \tag{36}$$

Proof. See Appendix 2.

Based on the derived closed-form expressions (33) and (35), Corollary 1 provides the asymptotic user rate limit when $\frac{MN}{K} \to \infty$, from which we can investigate the asymptotic performance of L-DASs in the presence of practical channel impairments.

Corollary 1. At time t, as $\frac{MN}{K} \to \infty$, no matter with MRT or ZF beamforming, the ergodic achievable downlink rate (19) in the presence of practical channel impairments achieves the same asymptotic user rate limit given by

$$R_{l,k}^{\infty}[t] = \log_2\left(1 + \frac{\sum_{m=1}^{M} \mathrm{e}^{-\sigma_{\phi_{l,m}}^2 t}\mu_{l,m,l,k}}{\sum_{i\neq l}\sum_{m=1}^{M} \mathrm{e}^{-\sigma_{\phi_{i,m}}^2 t}\mu_{i,m,l,k}}\right). \tag{37}$$

Proof. See [21] for a similar proof.

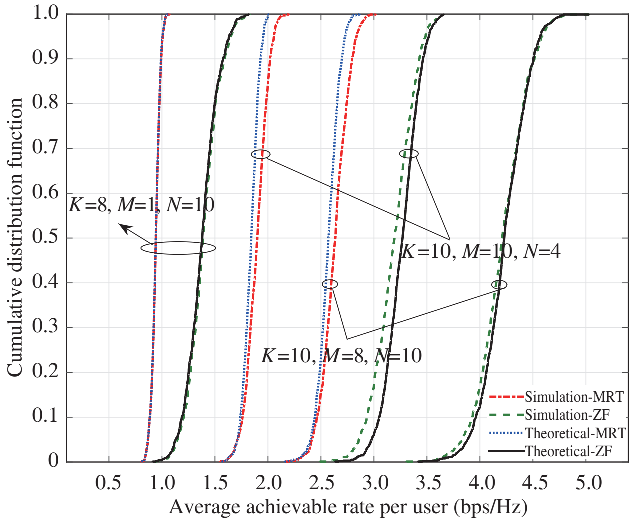

Figure 2 (Color online) Cumulative distribution function of the average achievable rate per user at time $t = 1$ for $L = 7$, $T_{\rm s} = 10^{-5}$ s, $T_{\rm c} = 100$, $\sigma_{\varphi_{l,k}} = \sigma_{\phi_{i,m}} = 0.72°$ with different K, M and N.

Corollary 1 demonstrates that the impacts of Doppler shift and phase noise at users vanish as $\frac{MN}{K} \to \infty$. This indicates that L-DASs can tolerate larger velocity of users and stronger phase noise at users. In particular, with common LOs or separate LOs, but with same phase noise variance, i.e., $\mathrm{e}^{-\sigma^2_{\phi_{l,m}}} = \mathrm{e}^{-\sigma^2_{\phi_{i,m}}}$ for $i \neq l$, the maximum achievable rate is not affected by the phase noise at BS sides. This insight is very important for practical deployment.

4 Numerical results

Our analytic results are validated in this section by studying the downlink of a hexagonal system with $L = 7$ cells. Each cell comprises M RAUs and K users, which are assumed to be uniformly distributed. We normalize the cell radius to 1 and set the minimum distance between RAUs and users to 0.01. For the large-scale fading $\lambda_{i,m,l,k}$, we consider the standard distance-based model given by $\lambda_{i,m,l,k} \triangleq cd^{-\alpha}_{i,m,l,k}$ [37, 39]. Here, $d_{i,m,l,k}$ denotes the distance between user k in cell l and RAU m in cell i; α represents the path loss exponent, which is set to 3.7; and $c = 1$ is the median of the mean path gain at $d_{i,m,l,k} = 1$. In all examples, $\gamma_{\rm DL} = 10$ dB and $\gamma_{\rm P} = K\gamma_{\rm DL}$.

In Figure 2, we verify the accuracy of Theorems 1 and 2 with practical channel impairments. The cumulative distribution function (CDF) curves of the average achievable rate per user at time $t \in \{1, T_{\rm c} - K\}$, defined by

$$R[t] = \frac{1}{LK} \sum_{l=1}^{L} \sum_{k=1}^{K} R_{l,k}[t] \tag{38}$$

obtained numerically and by Theorems 1 and 2 are presented in Figure 2 with different values of K, M and N. Without loss of generality, we set $t = 1$. The theoretical curves are averaged over 1000 channel realizations (different RAU and user locations) and the numerical curves are further averaged over small-scale channel fading. The carrier frequency is set to $f_{\rm c} = 2$ GHz. The Doppler spread is $f_{\rm D} = 250$ Hz, which corresponds to a user speed of 135 km/h. Then, the coherence time is $1/(4f_{\rm D}) = 1$ ms. Given that the symbol time $T_{\rm s} = 10^{-5}$ s, the coherence block includes $T_{\rm c} = 100$ symbols. For the parameters of channel impairments, the temporal correlation parameter can be calculated by $\alpha_t = J_0(2\pi f_{\rm D} T_{\rm s} t)$ and the increment standard deviation of the phase noise is set to $\sigma_{\varphi_{l,k}} = \sigma_{\phi_{i,m}} = 0.72°$ [15]. As can be seen in Figure 2, in the presence of practical channel impairments, the theoretical curves are almost indistinguishable from the numerical curves in co-located systems (Case 1) where all of the BS antennas are collocated at the origin. Although there is a small mismatch between the theoretical and numerical

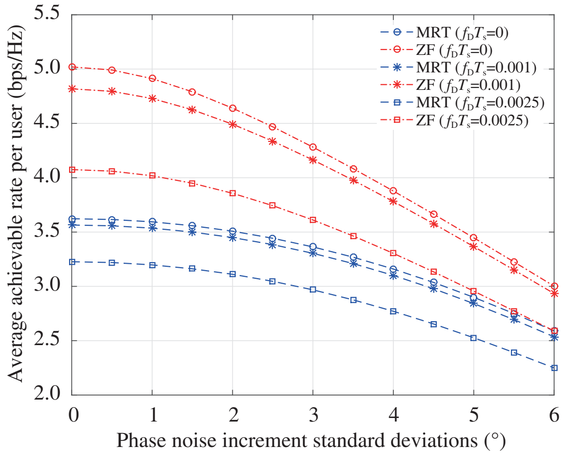

Figure 3 (Color online) Average achievable rate per user versus phase noise increment standard deviations for different values of $f_{\rm D}T_{\rm s}$, $L=7$, $M=10$, $N=20$, $K=8$, $T_{\rm s}=10^{-5}$ s, $T_{\rm c}=100$.

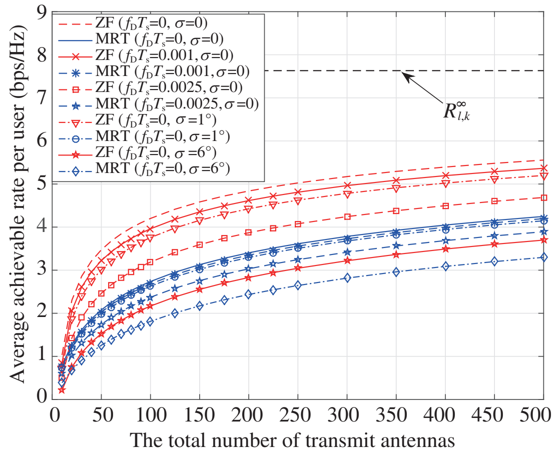

Figure 4 (Color online) Average achievable rate per user with channel impairments versus the total number of transmit antennas, $T_{\rm c}=100$, $L=7$, $M=5$, $K=8$.

curves in DASs (Cases 2 and 3) resulting from the approximation applied for the projection power of the non-isotropic channel vectors, they also match well. Note that the matches between the theoretical and numerical curves were preserved for different values of t, K, M, and N; we have omitted these results for the sake of brevity. In the following, we analyze the impacts of channel impairments on the spectral efficiency of L-DASs using these closed-form approximate expressions.

In the following, since the effective channels vary with t, we focus on the average achievable rate per user for each time instance of the downlink data transmission phase, which is defined by

$$R=\frac{1}{T_{\rm c}}\sum_{t=1}^{T_{\rm c}-K}R[t], \tag{39}$$

where the sum has $T_{\rm c}-K$ terms, which correspond to the number of symbols used for downlink data transmission and the pre-log factor $1/T_{\rm c}$ also accounts for the K symbols of pilot transmission.

Given the definition in (39), we investigated the impact of phase noise on average achievable rate per user (Figure 3) when $M=10$, $N=20$, $K=4$, $T_{\rm s}=10^{-5}$ s and $T_{\rm c}=100$ for different values of normalized Doppler shift $f_{\rm D}T_{\rm s}$. From Figure 3, we obtain the following findings. First, it is evident that as the value of the phase noise increment standard deviation increases, the average achievable rate per user loss becomes increasingly significant. Second, the larger is $f_{\rm D}T_{\rm s}$ (higher user mobility), the less important is the role of phase noise, and vice versa. As far as ZF beamforming is concerned, when the normalized Doppler shift $f_{\rm D}T_{\rm s}=0$ (no user mobility), the average achievable rate per user loss is about 40% with the phase noise increment standard deviations $\sigma_{\varphi_{l,k}}=\sigma_{\phi_{i,m}}=6^\circ$. However, when $f_{\rm D}T_{\rm s}=0.0025$ ($v=270$ km/h), the loss is about 36%. Meanwhile, when $\sigma_{\varphi_{l,k}}=\sigma_{\phi_{i,m}}=0^\circ$ (perfect LOs), the rate loss is about 18% with $f_{\rm D}T_{s}=0.0025$, and about 13% when $\sigma_{\varphi_{l,k}}=\sigma_{\phi_{i,m}}=6^\circ$ (strong phase noise scenario). Third, Doppler shift and phase noise have a larger impact on ZF beamforming and the gap between the average achievable rate per user performances of ZF and MRT beamforming decreases with increasing $f_{\rm D}T_{\rm s}$ and/or the phase noise increment standard deviations. This results from the poor interference cancellation capability of ZF beamforming when channel impairments are present.

In Figure 4, the average achievable rates per user with channel impairments are depicted as a function of the number of total transmit antennas MN when $K=8$. From this figure, we obtain the following findings. First, ZF beamforming performs worse (better) than MRT beamforming if the number of total transmit antennas MN is small (large) as it mitigates multiuser interference at a cost of decreasing the array gain from MN to $MN-K+1$. It should be noted that as $MN\to\infty$ for a given K, MRT beamforming exhibits the same multiuser interference suppression capability as ZF beamforming since the channel vectors between users become pairwise orthogonal [1] and the array gain of the two schemes become asymptotically equal; therefore, they approach the same asymptotic rate limit given in (37). Second, consistent with Corollary 1, as the number of total transmit antennas MN increases, the impacts of phase noise and Doppler shift decrease. Focusing on the case of $f_{\mathrm{D}}T_{\mathrm{s}} \approx 0.0025$ and $\sigma^2_{\varphi_{l,k}} = \sigma^2_{\phi_i} = 0°$, when $MN = 500$, the average rate per user loss decreases by 34% (from 50% to 16%) with ZF beamforming and by 10% (from 20% to 10%) with MRT beamforming, respectively. When $f_{\mathrm{D}}T_{\mathrm{s}} = 0$ and $\sigma^2_{\varphi_{l,k}} = \sigma^2_{\phi_i} = 6°$, the rate loss decreases by 54% (from 79% to 25%) with ZF beamforming and by 27% (from 49% to 22%) with MRT beamforming, respectively.

5 Conclusion

In this paper, we analyzed the impacts of practical channel impairments caused by pilot contamination, Doppler shift, and phase noise on the downlink spectral efficiency of L-DASs with MRT and ZF beamforming under per user power normalization. First, we derived closed-form approximate expressions for the ergodic achievable downlink rate with channel impairments, which proved to be accurate and tractable. Based on these expressions, we analyzed the spectral efficiency of L-DASs in the presence of practical channel impairments and efficiently evaluated the impacts of these channel impairments. Numerical results showed that channel impairments decreased the downlink achievable rate and had a larger impact on ZF beamforming. The asymptotic performance of L-DASs with channel impairments was also investigated. It was seen that the use of a large-scale antenna array at BS sides can weaken the impacts of channel impairments, and that the phase noise at the transmitter sides is dominant.

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

References

1 Marzetta T L. Noncooperative cellular wireless with unlimited numbers of base station antennas. IEEE Trans Wirel Commun, 2010, 9: 3590–3600
2 Wang D, Zhang Y, Wei H, et al. An overview of transmission theory and techniques of large-scale antenna systems for 5G wireless communications. Sci China Inf Sci, 2016, 59: 081301
3 Lu L, Li G Y, Swindlehurst A L, et al. An overview of massive MIMO: benefits and challenges. IEEE J Sel Top Signal Process, 2014, 8: 742–758
4 Rusek F, Persson D, Lau B K, et al. Scaling up MIMO: opportunities and challenges with very large arrays. IEEE Signal Process Mag, 2013, 30: 40–60
5 Zhang J, Wen C K, Jin S, et al. On capacity of large-scale MIMO multiple access channels with distributed sets of correlated antennas. IEEE J Sel Areas Commun, 2013, 31: 133–148
6 Fernandes F, Ashikhmin A, Marzetta T L. Inter-cell interference in noncooperative TDD large scale antenna systems. IEEE J Sel Areas Commun, 2013, 31: 192–201
7 Adhikary A, Ashikhmin A, Marzetta T L. Uplink interference reduction in large scale antenna systems. In: Proceedings of IEEE International Symposium on Information Theory, Honolulu, 2014. 2529–2533
8 Wen C K, Jin S, Wong K K, et al. Channel estimation for massive MIMO using gaussian-mixture Bayesian learning. IEEE Trans Wirel Commun, 2015, 14: 1356–1368
9 Truong K T, Heath R W. Effects of channel aging in massive MIMO systems. J Commun Netw, 2013, 15: 338–351
10 Papazafeiropoulos A K, Ngo H Q, Matthaiou M, et al. Uplink performance of conventional and massive MIMO cellular systems with delayed CSIT. In: Proceedings of IEEE International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC), Washington, 2014. 601–606
11 Papazafeiropoulos A K, Ratnarajah T. Deterministic equivalent performance analysis of time-varying massive MIMO systems. IEEE Trans Wirel Commun, 2015, 14: 5795–5809
12 Papazafeiropoulos A K. Impact of user mobility on optimal linear receivers in cellular networks. In: Proceedings of IEEE International Conference on Communications (ICC), London, 2015. 2239–2244
13 Guo K, Khodapanah B, Ascheid G. Performance analysis of downlink MMSE beamforming training in TDD MU-massive-MIMO. In: Proceedings of IEEE Wireless Communications and Networking Conference (WCNC), Doha, 2016
14 Pitarokoilis A, Moammed S K, Larsson E G. Achievable rates of ZF receivers in massive MIMO with phase noise impairments. In: Proceedings of Asilomar Conference on Signals, Systems and Computers, Pacific Grove, 2013. 1004–1008

15 Pitarokoilis A, Mohammed S K, Larsson E G. Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise. IEEE Trans Wirel Commun, 2015, 14: 711–723

16 Bjornson E, Matthaiou M, Debbah M. Massive MIMO with non-ideal arbitrary arrays: hardware scaling laws and circuit-aware design. IEEE Trans Wirel Commun, 2015, 14: 4353–4368

17 Corvaja R, Armada A G. Phase noise degradation in massive MIMO downlink with zero-forcing and maximum ratio transmission precoding. IEEE Trans Veh Technol, 2016, 65: 8052–8059

18 Zhu J, Schober R, Bhargava V K. Physical layer security for massive MIMO systems impaired by phase noise. In: Proceedings of the 17th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Edinburgh, 2016

19 Larsson E, Edfors O, Tufvesson F, et al. Massive MIMO for next generation wireless systems. IEEE Commun Mag, 2014, 52: 184–195

20 Lee S R, Moon S H, Kim J S, et al. Capacity analysis of distributed antenna systems in a composite fading channel. IEEE Trans Wirel Commun, 2012, 11: 1076–1086

21 Zhu P, You X, Li J, et al. Spectral efficiency analysis of large-scale distributed antenna system in a composite correlated Rayleigh fading channel. IET Commun, 2015, 9: 681–688

22 Wang D M, You X H, Wang J Z, et al. Spectral efficiency of distributed MIMO cellular systems in a composite fading channel. In: Proceedings of IEEE International Conference on Communications (ICC'08), Prague, 2008. 1259–1264

23 Wang J, Dai L. Asymptotic rate analysis of downlink multi-user systems with co-located and distributed antennas. IEEE Trans Wirel Commun, 2015, 14: 3046–3058

24 Wang J, Dai L. Downlink rate analysis for virtual-cell based large-scale distributed antenna systems. IEEE Trans Wirel Commun, 2016, 15: 1998–2011

25 Björnson E, Matthaiou M, Pitarokoilis A, et al. Distributed massive MIMO in cellular networks: impact of imperfect hardware and number of oscillators. In: Proceedings of European Signal Processing Conference (EUSIPCO), Nice, 2015. 2436–2440

26 Papazafeiropoulos A K. Impact of general channel aging conditions on the downlink performance of massive MIMO. IEEE Trans Veh Technol, 2017, 66: 1428–1442

27 Li J M, Wang D M, Zhu P C, et al. Uplink spectral efficiency analysis of distributed massive MIMO with channel impairments. IEEE Access, 2017, 5: 5020–5030

28 Interdonato G, Ngo H Q, Larsson E G, et al. How much do downlink pilots improve cell-free massive MIMO? In: Proceedings of IEEE Global Communications Conference (GLOBECOM), Washington, 2016

29 Li J, Wang D, Zhu P, et al. Downlink spectral efficiency of distributed massive MIMO systems with linear beamforming under pilot contamination. IEEE Trans Veh Technol, 2018, 67: 1130–1145

30 Jakes W C. Microwave Mobile Communications. New York: Wiley, 1974

31 Krishnan R, Khanzadi M R, Krishnan N, et al. Linear massive MIMO precoders in the presence of phase noise – a large-scale analysis. IEEE Trans Veh Technol, 2016, 65: 3057–3071

32 Carvalho E de, Björnson E, Larsson E G, et al. Random access for massive MIMO systems with intra-cell pilot contamination. In: Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Shanghai, 2016. 3361–3365

33 Truong K T, Lozano A, Heath R W, et al. Optimal training in continuous flat-fading massive MIMO systems. In: Proceedings of European Wireless Conference, Barcelona, 2014

34 Kay S M. Fundamental of Statistical Signal Processing: Estimation Theory. Englewood: Prentice-Hall, 1993

35 Jose J, Ashikhmin A, Marzetta T L, et al. Pilot contamination and precoding in multi-cell TDD systems. IEEE Trans Wirel Commun, 2011, 10: 2640–2651

36 Björnson E, Larsson E G, Marzetta T L. Massive MIMO: ten myths and one critical question. IEEE Commun Mag, 2016, 54: 114–123

37 Björnson E, Larsson E G, Debbah M. Massive MIMO for maximal spectral efficiency: how many users and pilots should be allocated? IEEE Trans Wirel Commun, 2016, 15: 1293–1308

38 Van Chien T, Bjornson E, Larsson E G. Joint power allocation and user association optimization for massive MIMO systems. IEEE Trans Wirel Commun, 2016, 15: 6384–6399

39 Hoydis J, Brink S ten, Debbah M. Massive MIMO in the UL/DL of cellular networks: how many antennas do we need? IEEE J Sel Areas Commun, 2013, 31: 160–171

40 Kammoun A, Muller A, Bjornson E, et al. Linear precoding based on polynomial expansion: large-scale multi-cell MIMO systems. IEEE J Sel Top Signal Process, 2014, 8: 861–875

41 Li J M, Wang D M, Zhu P C, et al. Downlink spectral efficiency of multi-cell multi-user large-scale DAS with pilot contamination. In: Proceedings of IEEE International Conference on Communications (ICC), London, 2015. 2011–2016

42 Heath J R W, Wu T, Kwon Y H, et al. Multiuser MIMO in distributed antenna systems with out-of-cell interference. IEEE Trans Signal Process, 2011, 59: 4885–4899

43 Zhang J, Andrews J G. Adaptive spatial intercell interference cancellation in multicell wireless networks. IEEE J Sel Areas Commun, 2010, 28: 1455–1468

44 Hosseini K, Yu W, Adve R S. Large-scale MIMO versus network MIMO for multicell interference mitigation. IEEE J Sel Top Signal Process, 2014, 8: 930–941

45 Hosseini K, Yu W, Adve R S. Modeling and analysis of ergodic capacity in network MIMO systems. In: Proceedings of IEEE Globecom Workshops (GC Wkshps), Austin, 2014. 808–814

46 Hosseini K, Yu W, Adve R S. A stochastic analysis of network MIMO systems. IEEE Trans Signal Process, 2016, 64: 4113–4126
47 Seifi N, Heath R W, Coldrey M, et al. Joint transmission mode and tilt adaptation in coordinated small-cell networks. In: Proceedings of IEEE International Conference on Communications Workshops (ICC), Sydney, 2016. 598–603

Appendix A Proof of Theorem 1

For the signal power term $|\mathrm{E}[\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]]|^2$, we have

$$\begin{aligned}\left|\mathrm{E}\left[\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]\right]\right|^2 &\overset{(\mathrm{a})}{=} \left|\mathrm{E}\left[\|\hat{\boldsymbol{g}}_{l,l,k}[t]\|\right]\right|^2 \\ &\overset{(\mathrm{b})}{=} \xi(\hat{k}_{l,l,k,\mathrm{a}}[t])\hat{\theta}_{l,l,k,\mathrm{a}}[t], \end{aligned} \tag{A1}$$

where (a) is obtained because $\hat{\boldsymbol{g}}_{l,l,k}[t]$ and $\tilde{\boldsymbol{g}}_{l,l,k}[t]$ are independent, (b) results from $\|\hat{\boldsymbol{g}}_{l,l,k}[t]\| \sim \mathrm{Nakagami}(\hat{k}_{l,l,k,\mathrm{a}}[t], \hat{k}_{l,l,k,\mathrm{a}}[t]\hat{\theta}_{l,l,k,\mathrm{a}}[t])$ since $\|\hat{\boldsymbol{g}}_{l,l,k}[t]\|^2 = |\hat{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]|^2 \sim \Gamma(\hat{k}_{l,l,k,\mathrm{a}}[t], \hat{\theta}_{l,l,k,\mathrm{a}}[t])$, which is obtained from Lemma 2 and (27).

Based on the independence of $\hat{\boldsymbol{g}}_{l,l,k}[t]$ and considering the effect of pilot contamination, we decompose the interference power terms $\mathrm{var}[\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]]$, and $\sum_{(i,j)\neq(l,k)}\mathrm{E}[|\boldsymbol{g}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{MRT}}[t]|^2]$ as follows:

$$\mathrm{var}\left[\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]\right] = \mathrm{E}\left[\left|\hat{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]\right|^2\right] + \mathrm{E}\left[\left|\tilde{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]\right|^2\right] - \left|\mathrm{E}\left[\hat{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]\right]\right|^2, \tag{A2}$$

and

$$\begin{aligned}\sum_{(i,j)\neq(l,k)}\mathrm{E}\left[\left|\boldsymbol{g}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{MRT}}[t]\right|^2\right] = &\sum_{j\neq k}\mathrm{E}\left[\left|\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,j}^{\mathrm{MRT}}[t]\right|^2\right] + \sum_{i\neq l}\sum_{j\neq k}\mathrm{E}\left[\left|\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{MRT}}[t]\right|^2\right] \\ &+ \sum_{i\neq l}\mathrm{E}\left[\left|\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,k}^{\mathrm{MRT}}[t]\right|^2\right] + \sum_{i\neq l}\sum_{j=1}^{K}\mathrm{E}\left[\left|\tilde{\boldsymbol{g}}_{i,l,k,}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{MRT}}[t]\right|^2\right]. \end{aligned} \tag{A3}$$

By applying Lemma 2 to approximate the distributions of the terms in (A2) and (A3), we have

$$\left|\hat{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]\right|^2 \sim \Gamma(\hat{k}_{l,l,k,\mathrm{a}}[t], \hat{\theta}_{l,l,k,\mathrm{a}}[t]), \tag{A4}$$

$$\left|\tilde{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{MRT}}[t]\right|^2 \sim \Gamma\left(\frac{1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}[t], \tilde{\theta}_{l,l,k,\mathrm{a}}[t]\right), \tag{A5}$$

$$\left|\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,j}^{\mathrm{MRT}}[t]\right|^2 \sim \Gamma\left(\frac{1}{MN}k_{l,l,k,\mathrm{a}}[t], \theta_{l,l,k,\mathrm{a}}[t]\right), \tag{A6}$$

$$\left|\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,k}^{\mathrm{MRT}}[t]\right|^2 \sim \Gamma(\hat{k}_{i,l,k,\mathrm{a}}[t], \hat{\theta}_{i,l,k,\mathrm{a}}[t]), \tag{A7}$$

$$\left|\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{MRT}}[t]\right|^2 \sim \Gamma\left(\frac{1}{MN}\hat{k}_{i,l,k,\mathrm{a}}[t], \hat{\theta}_{i,l,k,\mathrm{a}}[t]\right), \tag{A8}$$

$$\left|\tilde{\boldsymbol{g}}_{i,l,k,}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{MRT}}[t]\right|^2 \sim \Gamma\left(\frac{1}{MN}\tilde{k}_{i,l,k,\mathrm{a}}[t], \tilde{\theta}_{i,l,k,\mathrm{a}}[t]\right). \tag{A9}$$

Substituting (A1) and (A4)–(A9) into (19) concludes the proof.

Appendix B Proof of Theorem 2

First, given the distributions of $\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\hat{\boldsymbol{g}}_{i,l,k}[t]$ in (27), the non-isotropic achievable CSI $\hat{\boldsymbol{g}}_{i,l,k}[t]$ can be approximated as an isotropic vector $\hat{\boldsymbol{g}}_{i,l,k,\mathrm{a}}[t]$ with i.i.d. $\mathcal{CN}(0, \hat{\theta}_{i,l,k,\mathrm{a}})$ elements [29]. Then, with the definition of $\boldsymbol{F}_{l,\mathrm{a}}[t] \triangleq [\hat{\boldsymbol{g}}_{l,l,1,\mathrm{a}}[t], \ldots, \hat{\boldsymbol{g}}_{l,l,K,\mathrm{a}}[t]]$, the useful signal power term $|\mathrm{E}[\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]]|^2$ can be calculated by

$$\begin{aligned}\left|\mathrm{E}\left[\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]\right]\right|^2 &\overset{(\mathrm{a})}{=} \left|\mathrm{E}[1/\|\boldsymbol{f}_{l,l,k}[t]\|]\right|^2 \\ &\overset{(\mathrm{b})}{=} \left|\mathrm{E}\left[\left(\left[\left(\boldsymbol{F}_l^{\mathrm{H}}[t]\boldsymbol{F}_l[t]\right)^{-1}\right]_{k,k}\right)^{-1/2}\right]\right|^2 \\ &\overset{(\mathrm{c})}{=} \xi(\rho\hat{k}_{l,l,k,\mathrm{a}}[t])\hat{\theta}_{l,l,k,\mathrm{a}}[t], \end{aligned} \tag{B1}$$

where (a) is obtained because of the independence of $\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]$ and $\tilde{\boldsymbol{g}}_{l,l,k}[t]$ and $\hat{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t] = 1/\|\boldsymbol{f}_{l,l,k}[t]\|$, (b) results from $\|\boldsymbol{f}_{l,l,k}[t]\|^2 = [(\boldsymbol{F}_l^{\mathrm{H}}[t]\boldsymbol{F}_l[t])^{-1}]_{k,k}$, (c) results from $1/\|\boldsymbol{f}_{l,l,k}[t]\| \sim \mathrm{Nakagami}(\rho\hat{k}_{l,l,k,\mathrm{a}}[t], \rho\hat{k}_{l,l,k,\mathrm{a}}[t]\hat{\theta}_{l,l,k,\mathrm{a}}[t])$ where we have applied Lemma 2 to approximate the distribution of $[(\boldsymbol{F}_l^{\mathrm{H}}[t]\boldsymbol{F}_l[t])^{-1}]_{k,k}$ as $\Gamma(\rho\hat{k}_{l,l,k,\mathrm{a}}[t], \hat{\theta}_{l,l,k,\mathrm{a}}[t])$ since $[(\boldsymbol{F}_{l,\mathrm{a}}^{\mathrm{H}}[t]\boldsymbol{F}_{l,\mathrm{a}}[t])^{-1}]_{k,k} \sim \Gamma(MN-K+1, \hat{\theta}_{l,l,k,\mathrm{a}}[t])$[1)].

Similar to the analysis in the proof given for Theorem 1 and considering that $\hat{\boldsymbol{g}}_{l,l,k}[t]\boldsymbol{w}_{l,j}^{\mathrm{ZF}}[t] = 0$ for $j \neq k$, the interference power terms $\mathrm{var}[\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]]$ and $\sum_{(i,j)\neq(l,k)}\mathrm{E}[|\boldsymbol{g}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{ZF}}[t]|^2]$ can be decomposed as

$$\mathrm{var}\left[\boldsymbol{g}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]\right] = \mathrm{E}\left[\left|\hat{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]\right|^2\right] + \mathrm{E}\left[\left|\tilde{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]\right|^2\right] - \left|\mathrm{E}\left[\hat{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]\right]\right|^2, \tag{B2}$$

1) Khansefid A, Minn H. Achievable downlink rates of MRC and ZF precoders in massive MIMO with uplink and downlink pilot contamination. IEEE Trans Commun, 2015, 63: 4849–4864.

and

$$\sum_{(i,j)\neq(l,k)} \mathrm{E}\big[\big|\boldsymbol{g}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{ZF}}[t]\big|^2\big] = \sum_{j\neq k} \mathrm{E}\big[\big|\tilde{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,j}^{\mathrm{ZF}}[t]\big|^2\big] + \sum_{i\neq l}\sum_{j\neq k} \mathrm{E}\big[\big|\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{ZF}}[t]\big|^2\big]$$
$$+ \sum_{i\neq l} \mathrm{E}\big[\big|\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,k}^{\mathrm{ZF}}[t]\big|^2\big] + \sum_{i\neq l}\sum_{j=1}^{K} \mathrm{E}\big[\big|\tilde{\boldsymbol{g}}_{i,l,k,}[t]\boldsymbol{w}_{i,j}^{\mathrm{ZF}}\big|^2\big]. \tag{B3}$$

The distributions of the terms in (B2) and (B3) can be obtained by applying Lemma 2,

$$\big|\hat{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]\big|^2 \sim \Gamma(\rho\hat{k}_{l,l,k,\mathrm{a}}[t], \hat{\theta}_{l,l,k,\mathrm{a}}[t]), \tag{B4}$$

$$\big|\tilde{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,k}^{\mathrm{ZF}}[t]\big|^2 \sim \Gamma\left(\frac{1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}[t], \tilde{\theta}_{l,l,k,\mathrm{a}}[t]\right), \tag{B5}$$

$$\big|\tilde{\boldsymbol{g}}_{l,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{l,j}^{\mathrm{ZF}}[t]\big|^2 \sim \Gamma\left(\frac{1}{MN}\tilde{k}_{l,l,k,\mathrm{a}}[t], \tilde{\theta}_{l,l,k,\mathrm{a}}[t]\right), \tag{B6}$$

$$\big|\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,k}^{\mathrm{ZF}}[t]\big|^2 \sim \Gamma\big(\rho\hat{k}_{i,l,k,\mathrm{a}}[t], \hat{\theta}_{i,l,k,\mathrm{a}}[t]\big), \tag{B7}$$

$$\big|\hat{\boldsymbol{g}}_{i,l,k}^{\mathrm{H}}[t]\boldsymbol{w}_{i,j}^{\mathrm{ZF}}[t]\big|^2 \sim \Gamma\left(\frac{1}{MN}\hat{k}_{i,l,k,\mathrm{a}}[t], \hat{\theta}_{i,l,k,\mathrm{a}}[t]\right), \tag{B8}$$

$$\big|\tilde{\boldsymbol{g}}_{i,l,k}[t]\boldsymbol{w}_{i,j}^{\mathrm{ZF}}[t]\big|^2 \sim \Gamma\left(\frac{1}{MN}\tilde{k}_{i,l,k,\mathrm{a}}[t], \tilde{\theta}_{i,l,k,\mathrm{a}}[t]\right). \tag{B9}$$

Substituting (B1) and (B4)–(B9) into (19) yields the closed-form approximation (35). This completes the proof.