IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received May 12, 2022, accepted May 19, 2022, date of publication May 23, 2022, date of current version June 17, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3176946

Performance of Multiuser Downlink Cell-Free Massive MIMO Systems With Hard Deadlines

LE TAN[1], ZIYANG ZHANG[1,2], AND DONGMING WANG[2,3], (Member, IEEE)
[1]School of Information Science and Engineering, Southeast University, Nanjing 211189, China
[2]National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
[3]Purple Mountain Laboratories, Nanjing 211111, China

Corresponding author: Dongming Wang (wangdm@seu.edu.cn)

This work was supported in part by the Research and Development Project in Key Areas of Guangdong Province under Grant 2020B0101120003, and in part by the National Natural Science Foundation of China under Grant 61871122.

ABSTRACT Cell-free massive MIMO (mMIMO) is expected to be a novel network architecture to support massive connections, ultra-high transmission rate, and ultra-reliability low-latency transmission. Therefore, applications in ultra-reliable low-latency communications (URLLC) can be served by the cell-free mMIMO network in the future. This paper considers a multiuser downlink cell-free massive MIMO communication system and defines two events to evaluate the system's failure in latency and reliability performance of the transmission. The first kind of event is defined as *Transmission Error*, where the value of the downlink rate becomes too high to be supported by the channel between the access points (APs) and user equipments (UEs), while the second event is defined as *Time Overflow*, where the time of packet delivery violates the hard deadline. The probabilities of these events can help evaluate the URLLC performance of the system because the probability of *Transmission Error* reflects the reliability of the transmission while the probability of *Time Overflow* reflects the latency of the transmission. We derive the expressions of the probabilities of these two events and the final probability system outage, with maximum ratio transmission (MRT) and zero-forcing (ZF) precoding, based on the properties of Gamma distribution and the approximation for non-isotropic vectors. We finally provide the numerical results by system-level simulations to show that the expressions are accurate and discuss the influence of the noise variance, the number of UEs, and the precoding scheme on the URLLC performance.

INDEX TERMS URLLC, cell-free massive MIMO, outage performance.

I. INTRODUCTION

The Massive multiple-input multiple-output (mMIMO) is regarded as an essential 5G technology [1], [2]. The cellular network infrastructure is often deployed in conventional mMIMO. In such a scenario, each base station (BS) equipped with 128 or more antennas serves a set of user equipments (UEs) [3]. The spectral efficiency is improved via spatial multiplexing, and the data rates become far higher than before. However, the inter-cell interference and weak service at the cell edge become the main challenge [4].

The cell-free massive MIMO is considered an alternative network structure where the cell boundaries are eradicated. This method deploys plenty of access points (APs) serving the UEs by coherent joint transmission. All the APs connect to a controller, also called a central processing unit (CPU), edge-cloud processor [5], or cloud radio access network (C-RAN) data center [6], via fronthaul links. In a cell-free mMIMO system, a large number of antennas are spatially distributed rather than co-located at the BS. Each UE in the system is served by all APs and therefore has uniformly good service with each other in the same time and frequency resources. Another significant advantage of the cell-free mMIMO system is that it takes the benefit of spatial multiplexing to improve SE, as [7] shows that the cell-free mMIMO can achieve remarkable improvements in the median and 95%-likely spectral efficiency (SE).

A broad interest is growing in the cell-free mMIMO topic, as the framework is considered a candidate architecture for future networks. [8] investigates conjugate beamforming and the achievable rates with conjugate beamforming and zero-forcing (ZF) precoding. CSI acquisition and data sharing among the APs are analyzed in [9], while [10] researched the potentials of the cell-free mMIMO, which can provide

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

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

a high rate, reliability, and energy efficiency with simple signal processing. Four different implementations of cell-free mMIMO, characterized by different degrees of cooperation among the APs, are studied in [11]. Moreover, [12] provides a scalable framework and algorithm for the cell-free mMIMO system. [13] also proposes an optimization framework to improve the transmission accuracy for non-orthogonal wireless backhaul in the cell-free mMIMO with a deep neural network model. All the research improves the transmission reliability and reduces the transmission latency in the cell-free mMIMO. Hence, it enables the cell-free mMIMO better to serve ultra-reliable low-latency communications (URLLC) applications.

As the cell-free mMIMO is expected to be the key 6G technology to support massive connects, ultra-high data rate, and ultra-reliability low-latency transmission [14], [15], we notice that the research on ergodic capacity cannot accurately reflect the reliability and delay performance of the communication system. For applications in URLLC scene, a new metric to quantify the latency and reliability is needed. Inspired by [16], we define two events that will cause system outage, referring to the system's failure in latency and reliability, respectively. The lower probabilities of these outage events refer to a better performance in the reliability and latency of the transmission for URLLC applications. Therefore, the system outage probability can be treated as a metric to evaluate the URLLC performance.

This paper considers a multiuser downlink cell-free massive MIMO communication system to derive the expressions of the probabilities of the two events above, considering the multiuser interference. The first kind of event, referred to as *Transmission Error*, is caused by the chosen rate getting higher than the capacity of wireless channels between the UEs and APs such that the reliability of the transmission cannot be promised. The other kind of event, referred to as *Time Overflow*, is caused by the violation of transmission by an expected hard deadline. We find expressions for the probabilities of transmission error, time overflow, and system outage with maximum ratio transmission (MRT) and ZF precoding. Moreover, we simulate such a system to verify our expressions by comparing them with the numerical simulations. It shows that our expressions are accurate, and the noise variance, the number of UEs, and the precoding scheme affect the probabilities to varying degrees.

The rest of the paper is organized as follows. We introduce the downlink cell-free massive MIMO system model and define the two events that will lead to the system outage in Section II. Then, we analyze the corresponding expression for the probabilities of the events in Section III. Section IV presents the numerical analysis based on simulation results. Finally, our conclusion of the paper is presented in V.

Notations: We use boldface uppercase letters $\mathbf{A}$ to denote a matrix and $\mathbf{a}_i$ to denote the i-th column vector of matrix $\mathbf{A}$. We also use boldface lowercase letters $\mathbf{b}$ to denote a column vector, while b_j and $\|\mathbf{b}\|$ denote the j-th element and the Euclidean norm of vector $\mathbf{b}$, respectively. The operators $(\cdot)^{\mathrm{H}}$ and $(\cdot)^{\mathrm{T}}$ are used as the he transpose and Hermitian transpose operators, respectively, while the conjugate of a complex number z is denoted by z^*. We use $\mathcal{CN}(\cdot,\cdot)$ to refer to the circularly symmetric complex Gaussian distribution, while $\Gamma(\cdot,\cdot)$ refers to Gamma distribution. The expectations and covariances are denoted by $\mathbb{E}[\cdot]$ and $\mathrm{Var}[\cdot]$.

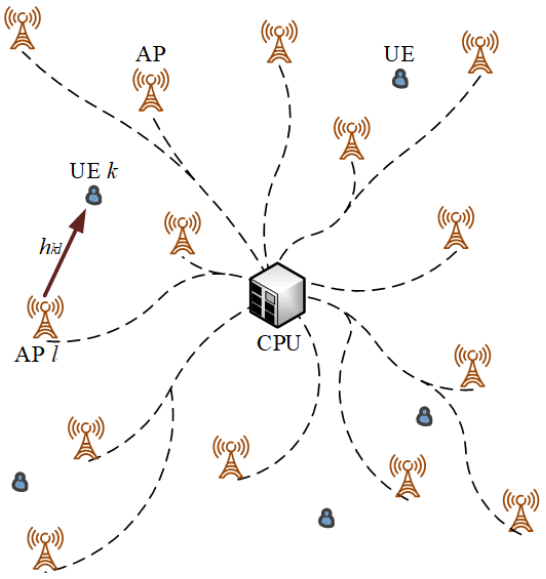

FIGURE 1. The multiuser downlink cell-free mMIMO system.

II. SYSTEM MODEL AND SYSTEM OUTAGE

A. SYSTEM SETUP

In this paper, we consider a periodic multiuser downlink cell-free massive MIMO communication system depicted in Fig. 1. There are L geographically distributed single-antenna APs and K single-antenna UEs. All APs are wired to a controller, called CPU [5], [17], in an arbitrary fashion. In addition, the channel state information (CSI) is shared among the APs for their cooperative transmission. We assume the system uses the standard time-division duplex (TDD) protocol in [18] and operates in cycles of period T over a bandwidth of W Hertz. The period can be divided into two phases: an uplink training phase of duration T_{U} and a downlink data phase of duration T_{D}, which means $T = T_{\mathrm{U}} + T_{\mathrm{D}}$. In addition, the system expects downlink data of fixed length, B bits, to be transmitted to each UE.

Let $h_{kl} \sim \mathcal{CN}(0, R_{kl})$ be the Rayleigh fading channel between AP l and UE k, where R_{kl} is the spatial channel correlation coefficient which describes the large-scale fading, including antenna gains, geometric pathloss, and shadowing [18]. Since the APs are spatially distributed in the system, the channels between the APs and the UEs are independently distributed, thus $\mathbb{E}\left[h_{kl}^* h_{kn}\right] = 0,\ l \neq n$. Furthermore, UE k's channel vector $\mathbf{h}_k = [h_{k1}, h_{k2}, \cdots, h_{kL}]^{\mathrm{T}} \in \mathbb{C}^{L\times 1}$ is independent from the channel vectors of other UEs. We assume the spatial correlation coefficients $\{R_{kl}\}$ are available wherever needed, and the practical correlation estimation methods are introduced in [19]–[21].

We also use a quasi-static fading model so that the fading value keeps constant during every circle. Furthermore, h_{kl} appears constant as the downlink packets are small enough.

B. CHANNEL ESTIMATION AND RATE SELECTION

We assume that there are τ_p mutually orthogonal τ_p-length pilot signals assigned to UEs in the system. We use $\mathcal{A}_t \subset \{1, 2, \cdots, K\}$ to denote the subset of UEs assigned to pilot t. Therefore, the received signal y_{tl} after correlating at AP l is

$$y_{tl} = \sum_{n \in \mathcal{A}_t} \sqrt{\tau_p p_n} h_{nl} + n_{tl} \tag{1}$$

where τ_p is the processing gain, p_n refers to UE n's transmit power, while $n_{tl} \sim \mathcal{CN}\left(0, \sigma^2\right)$ denotes the thermal noise, and σ^2 refers to the noise power. So the minimum mean-squared-error (MMSE) estimate of h_{kl} for $k \in \mathcal{A}_t$ is

$$\hat{h}_{kl} = \sqrt{\tau_p p_n} R_{kl} \Phi_{tl}^{-1} h_{tl} \tag{2}$$

where

$$\Phi_{tl} = \sum_{n \in \mathcal{A}_t} \tau_p p_n R_{nl} + \sigma^2 \tag{3}$$

is the correlation matrix of y_{tl}. Therefore, the channel estimate $\hat{h}_{kl}$ and the estimation error $\tilde{h}_{kl} = h_{kl} - \hat{h}_{kl}$ are independently distributed as $\hat{h}_{kl} \sim \mathcal{CN}(0, \beta_{kl})$ and $\tilde{h}_{kl} \sim \mathcal{CN}(0, \varepsilon_{kl})$, where

$$\beta_{kl} = \tau_p p_k \Phi_{tl}^{-1} R_{kl}^2 \tag{4a}$$

$$\varepsilon_{kl} = R_{kl} - \beta_{kl} \tag{4b}$$

In the paper, we assume that $\tau_p \geq K$, so that pilot contamination can be removed, Hence, the channel estimates $\hat{h}_{kl}$ become independent from each other. Let $\hat{\mathbf{h}}_k = \left[\hat{h}_{k1}, \hat{h}_{k2}, \cdots, \hat{h}_{kL}\right]^{\mathrm{T}} \in \mathbb{C}^{L \times 1}$ denote the estimated channel vector for UE k, while $\widehat{\mathbf{H}} = \left[\hat{\mathbf{h}}_1, \hat{\mathbf{h}}_2, \cdots, \hat{\mathbf{h}}_K\right] \in \mathbb{C}^{L \times K}$ refers to the estimated channel matrix. Then, two conventional low complexity precoding methods are considered in this paper as

$$\mathbf{G} = \begin{cases} \widehat{\mathbf{H}}, & \text{MRT} \\ \widehat{\mathbf{H}} \left(\widehat{\mathbf{H}}^{\mathrm{H}} \widehat{\mathbf{H}}\right)^{-1}, & \text{ZF} \end{cases} \tag{5}$$

The precoding vector $\mathbf{w}_k = \frac{\mathbf{g}_k}{\|\mathbf{g}_k\|}$, where $\mathbf{g}_k$ is the k-th column of matrix $\mathbf{G}$. Note that we use the normalized precoding vectors with both MRT and ZF precoding methods [12], which equivalently adds a constraint on the transmit power.

After channel estimation, the APs share the CSI and jointly select an appropriate downlink rate for the corresponding UE. We use R_k ro refer to the rate selected by the APs for the downlink transmission to UE k. Let T_k denote the time for UE k to receive the downlink signal. The APs choose the rate through transmit precoding, according to the MMSE estimation value $\hat{\mathbf{h}}_k$.

$$R_k = W \log_2 \left(1 + \frac{\left\|\hat{\mathbf{h}}_k^{\mathrm{H}} \mathbf{w}_k\right\|^2}{\sum_{i=1, i \neq k}^{K} \left\|\hat{\mathbf{h}}_k^{\mathrm{H}} \mathbf{w}_i\right\|^2 + \sigma^2}\right) \tag{6}$$

where $\mathbf{w}_k \in \mathbb{C}^{L \times 1}$ denotes the precoding vector that APs assign to UE k.

Then, the received signal at k-th UE is

$$\begin{aligned} y_k &= \hat{\mathbf{h}}_k^{\mathrm{H}} \left(\sum_{i=1}^{K} \mathbf{w}_k x_i\right) + n_k \\ &= \hat{\mathbf{h}}_k^{\mathrm{H}} \mathbf{w}_k x_k + \sum_{i=1, i \neq k}^{K} \hat{\mathbf{h}}_k^{\mathrm{H}} \mathbf{w}_i x_i + n_k \end{aligned} \tag{7}$$

where $x_i \sim \mathcal{CN}(0, 1)$ denotes the downlink data transmitted to UE i, while $n_k \sim \mathcal{CN}\left(0, \sigma^2\right)$ refers to the noise at the k-th UE.

C. SYSTEM OUTAGE

The system outage happens when the reliability of the following downlink data transmission cannot be guaranteed as the chosen rate is too large, or when the time of the corresponding downlink data transmission goes in excess of the intended budget T_{D} since the chosen rate is too low. According to [16], we define the former event as *transmission error* (TE) and refer to the latter event as *time overflow* (TO). To relate these above to the system model closely, we firstly specify these three events below:

Transmission Error (TE, [16]): When the chosen downlink rate between APs and the UE goes beyond the channel's capacity, also called the input-output mutual information, there will be a sharp increase in the UE's packet error probability as the UE cannot decode the downlink data correctly. The mutual information of the channel between APs and UE k is defined as in [22].

$$C_k = W \log_2 \left(1 + \frac{\left\|\mathbf{h}_k^{\mathrm{H}} \mathbf{w}_k\right\|^2}{\sum_{i=1, i \neq k}^{K} \left\|\mathbf{h}_k^{\mathrm{H}} \mathbf{w}_i\right\|^2 + \sigma^2}\right) \tag{8}$$

Note that the expression is raised for analysis under the infinite blocklength (IBL) assumption, while the finite blocklength (FBL) regime analysis will be deferred to future work. To describe the reliability of the whole system, we define the TE event as happening when one or more UEs fail to receive the corresponding data correctly. Therefore, we can express the probability of the TE event due to the device failure at UE k as

$$\begin{aligned} \mathbb{P}\left[\mathrm{TE}_k\right] &= \mathbb{P}\left[R_k > C_k\right] \\ &= \mathbb{P}\left[\frac{\left\|\hat{\mathbf{h}}_k^{\mathrm{H}} \mathbf{w}_k\right\|^2}{\sum_{i=1, i \neq k}^{K} \left\|\hat{\mathbf{h}}_k^{\mathrm{H}} \mathbf{w}_i\right\|^2 + \sigma^2} > \frac{\left\|\mathbf{h}_k^{\mathrm{H}} \mathbf{w}_k\right\|^2}{\sum_{i=1, i \neq k}^{K} \left\|\mathbf{h}_k^{\mathrm{H}} \mathbf{w}_i\right\|^2 + \sigma^2}\right] \end{aligned} \tag{9}$$

By deploying the quasi-static setting, the above probability can be written in terms of the relative values of the two signal-to-interference-and-noise ratio (SINR) terms in the expressions of R_k and C_k. Hence, the probability of the TE event is

expressed via the probability of the first kind of device failure as

$$\mathbb{P}[\mathtt{TE}] = 1 - \prod_{k=1}^{K} (1 - \mathbb{P}[\mathrm{TE}_k]) \tag{10}$$

where equation holds because of independent channel fades.

Time Overflow (TO): When one of the UEs is in deep fade, the chosen rate will be relatively low, leading to a long time for the UE to receive the complete downlink data. Considering that there are many standard requirements of delay duration in real-time applications, the TO event is defined to happen when one or more UEs fail to obtain the entire data before the cycle edge. Therefore, the probability of the TO event due to the device failure at UE k is

$$\mathbb{P}[\mathrm{TO}_k] = \mathbb{P}[T_k > T_D] = F_{r_k}\left(\frac{B}{T_D W}\right) \tag{11}$$

where r_k denotes the SE at UE k, while $F_{r_k}(\cdot)$ is the cumulative distribution function (CDF) of r_k. So the above probability can be written in terms of the required transmission duration T_k exceeding the downlink budget T_{D}. As the TO event refers to at least one UE's transmission time overflowing into a subsequent cycle, the probability of the TO event can be expressed via the probability of the second kind of device failure as

$$\mathbb{P}[\mathtt{TO}] = 1 - \prod_{k=1}^{K} (1 - \mathbb{P}[\mathrm{TO}_k]) \tag{12}$$

After choosing the downlink rates $\{R_k\}$, the CPU can judge whether the TO event will happen. All APs might not transmit any data when the TO event is detected at the CPU side.

System Outage (SO, [16]): When the TE event or the TO event happens, the SO event is defined to happen. So combining these two events, we can express the system outage probability as

$$\mathbb{P}[\mathtt{SO}] = 1 - \mathbb{P}\left[\bigcap_{k=1}^{K}\left\{\frac{B}{T_D} \le R_k \le C_k\right\}\right] \tag{13}$$

Let $\mathbb{P}\left[\frac{B}{T_D} \le R_k \le C_k\right]$ refers to the probability that the two kinds of device failure will not happen at UE k. Combining the equation (9) and (11), the two kinds of device failure happen independently at UE k. Therefore, the probability of device failure at UE k is $\mathbb{P}[\mathrm{TE}_k] + \mathbb{P}[\mathrm{TO}_k]$. Considering the independent channel fades, the device failure at each UE happens independently. Then, we define the final probability of device failure as

$$\mathbb{P}[\mathrm{DF}_k] = 1 - \mathbb{P}\left[\frac{B}{T_D} \le R_k \le C_k\right] = \mathbb{P}[\mathrm{TE}_k] + \mathbb{P}[\mathrm{TO}_k] \tag{14}$$

Due to the independent channel fades, the device failure happens independently at each UE. Therefore, the system outage probability can be rewritten as

$$\mathbb{P}[\mathtt{SO}] = 1 - \prod_{k=1}^{K} (1 - \mathbb{P}[\mathrm{DF}_k]) \tag{15}$$

So, we only need to analyze $\mathbb{P}[\mathtt{TE}]$, $\mathbb{P}[\mathtt{TO}]$, and $\mathbb{P}[\mathtt{SO}]$ by the analysis of the probabilities of the two kinds of device failure, $\mathbb{P}[\mathrm{TE}_k]$ and $\mathbb{P}[\mathrm{TO}_k]$. Moreover, the probability $\mathbb{P}\left[\frac{B}{T_D} \le R_k \le C_k\right]$ can be expressed via $\mathbb{P}[\mathrm{TE}_k]$ and $\mathbb{P}[\mathrm{TO}_k]$ as

$$\mathbb{P}\left[\frac{B}{T_D} \le R_k \le C_k\right] = 1 - \mathbb{P}[\mathrm{TE}_k] - \mathbb{P}[\mathrm{TO}_k] \tag{16}$$

III. OUTAGE ANALYSIS

In this section, we first characterize the probability of the probabilities of the two kinds of device failure $\mathbb{P}[\mathrm{TE}_k]$ and $\mathbb{P}[\mathrm{TO}_k]$ with MRT and ZF precoding. Then, we present the expressions of $\mathbb{P}[\mathtt{TE}]$, $\mathbb{P}[\mathtt{TO}]$, and $\mathbb{P}[\mathtt{SO}]$ based on the these device failure probabilities. We list some lemmas which helps to derive the expressions above all.

Lemma 1 [23]: If vector $\mathbf{a} \in \mathbb{C}^{M\times 1}$ consists of M i.i.d. $\mathcal{CN}\left(0, \sigma_a^2\right)$ elements, then $\mathbf{a}^{\mathrm{H}}\mathbf{a} \sim \Gamma\left(M, \sigma_a^2\right)$.

Lemma 2 [23]: If variable $Z \sim \Gamma(k, \theta)$, then $\mathbb{E}[Z] = k\theta$, $\mathrm{Var}[Z] = k\theta^2$, and $cZ \sim \Gamma(k, c\theta)$, where $c > 0$.

Lemma 3 [24]: If random variables $\{Z_i\}$ are independent and satisfy $\Gamma(k_i, \theta_i)$ respectively, the sum $\sum_i Z_i$ will satisfy

$$\begin{aligned}
\mathbb{E}\left[\sum_i Z_i\right] &= \sum_i k_i\theta_i \\
\mathbb{E}\left[\left(\sum_i Z_i\right)^2\right] &= \sum_i k_i\theta_i^2 + \left(\sum_i k_i\theta_i\right)^2 \\
\mathrm{Var}\left[\sum_i Z_i\right] &= \sum_i k_i\theta_i^2
\end{aligned} \tag{17}$$

where $\sum_i Z_i \sim \Gamma\left(\frac{\left(\sum_i k_i\theta_i\right)^2}{\sum_i k_i\theta_i^2}, \frac{\sum_i k_i\theta_i^2}{\sum_i k_i\theta_i}\right)$.

Lemma 4 [24]: We can project an m-dimensional non-isotropic channel vector $\mathbf{g}_k \sim \Gamma(k, \theta)$ onto an p dimensional subspace. Then, the projection power obeys $\Gamma\left(\frac{p}{m}k, \theta\right)$, where $p = L$ with MRT precoding and $p = L - K + 1$ with ZF precoding, while independent vectors will be projected onto a one-dimensional subspace.

Note that we can obtain a better approximation by *Lemma 4* when the path loss from UE k to each nearby AP is similar; otherwise, the degrees of freedom are overestimated [24]. However, we can adjust the number of effective dimension of the final subspace to increase the accuracy of our approximation. Then, based on the first three lemmas, we can derive the distributions of $\mathbf{h}_k^{\mathrm{H}}\mathbf{h}_k$, $\hat{\mathbf{h}}_k^{\mathrm{H}}\hat{\mathbf{h}}_k$, and $\tilde{\mathbf{h}}_k^{\mathrm{H}}\tilde{\mathbf{h}}_k$ as

$$\mathbf{h}_k^{\mathrm{H}}\mathbf{h}_k \sim \Gamma(k_k, \theta_k) \tag{18a}$$

$$\hat{\mathbf{h}}_k^{\mathrm{H}}\hat{\mathbf{h}}_k \sim \Gamma\left(\hat{k}_k, \hat{\theta}_k\right) \tag{18b}$$

$$\tilde{\mathbf{h}}_k^{\mathrm{H}}\tilde{\mathbf{h}}_k \sim \Gamma\left(\tilde{k}_k, \tilde{\theta}_k\right) \tag{18c}$$

where

$$k_k = \frac{\left(\sum_{l=1}^{L} R_{kl}\right)^2}{\sum_{l=1}^{L} R_{kl}^2}, \quad \theta_k = \frac{\sum_{l=1}^{L} R_{kl}^2}{\sum_{l=1}^{L} R_{kl}} \tag{19a}$$

$$\hat{k}_k = \frac{\left(\sum_{l=1}^{L} \beta_{kl}\right)^2}{\sum_{l=1}^{L} \beta_{kl}^2}, \quad \hat{\theta}_k = \frac{\sum_{l=1}^{L} \beta_{kl}^2}{\sum_{l=1}^{L} \beta_{kl}} \tag{19b}$$

$$\tilde{k}_k = \frac{\left(\sum_{l=1}^{L} \varepsilon_{kl}\right)^2}{\sum_{l=1}^{L} \varepsilon_{kl}^2}, \quad \tilde{\theta}_k = \frac{\sum_{l=1}^{L} \varepsilon_{kl}^2}{\sum_{l=1}^{L} \varepsilon_{kl}} \tag{19c}$$

A. TRANSMISSION ERROR

First of all, we calculate the probability of the first kind of device failure $\mathbb{P}[\mathrm{TE}_k]$ to prepare for the calculation of $\mathbb{P}[\mathrm{TE}]$. Calculating $\mathbb{P}[\mathrm{TE}_k]$ needs to divide the power terms in SINR of R_k and C_k into different independent variables to get the final expression. We will decide the detailed method according to the precoding vectors.

1) $\mathbb{P}[\mathrm{TE}_k]$ WITH MRT PRECODING

Considering the MRT precoding vector, we can rewrite the expression of $\mathbb{P}[\mathrm{TE}_k]$ as

$$\mathbb{P}\left[\frac{\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2}{\sum_{i=1,i\neq k}^{K}\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2+\sigma^2} > \frac{\left\|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2}{\sum_{i=1,i\neq k}^{K}\left\|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2+\sigma^2}\right] = \mathbb{P}\left[XV > \left(Y+\sigma^2\right)U\right] \tag{20}$$

where

$$\begin{aligned} X &= \left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2, \quad Y = \sum_{i=1,i\neq k}^{K}\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2, \\ U &= \left\|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2 - \left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2, \\ V &= \sum_{i=1,i\neq k}^{K}\left\|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2 - \sum_{i=1,i\neq k}^{K}\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2 \end{aligned} \tag{21}$$

These four variables become almost independent as the number of transmit antennas or APs grows. Moreover, it is evident that the signal term X and the interference term Y obey Gamma distribution with MRT precoding. Then, we need to analyze the distributions of the rest two variables. As for U, it refers to the sum of three terms

$$U = \left\|\tilde{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2 + \mathbf{w}_k^{\mathrm{H}}\hat{\mathbf{h}}_k\tilde{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k + \mathbf{w}_k^{\mathrm{H}}\tilde{\mathbf{h}}_k\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k \tag{22}$$

where the first term is a random variable obeying Gamma distribution. The sum of the rest two terms is a real number as the two terms are conjugated. The sum obeys Gaussian distribution according to [25]. The value of former term is mainly decided by ε_{kl}, while the value of latter terms is mainly influenced by β_{kl}. Therefore, the approximation distribution of variable U can be seen as Gaussian distribution when σ^2 is low, while it can be seen as Gamma distribution when σ^2 is large. Although V is similarly refers to the sum of these three kinds of terms, the sum of the pairs of conjugate terms is ignorable due to the independence of $\mathbf{w}_i$, $\hat{\mathbf{h}}_k$, and $\tilde{\mathbf{h}}_k$. Therefore, V can be seen as a Gamma variable.

Then, we use the probability density functions (PDFs) of X, Y, U, and V to calculate $\mathbb{P}[\mathrm{TE}_k]$ as

$$\mathbb{P}[\mathrm{TE}_k] = \int_0^{+\infty}\int_0^{+\infty}\int_0^{+\infty}\int_{-\infty}^{\frac{x\cdot v}{(y+\sigma^2)}} \times f_X(x) f_Y(y) f_V(v) f_U(u)\, du\, dv\, dy\, dx \tag{23}$$

Step 1: Calculate the first integration $\int_{-\infty}^{\frac{x\cdot v}{(y+\sigma^2)}} f_U(u)\, du$.

The CDF of variable U mainly decides the result $F_U\left(\frac{x\cdot v}{y+\sigma^2}\right)$. Since $U = \left\|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2 - \left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2$, the mean and variance of U can be expressed as $\mathbb{E}[U] = k_k\theta_k - \hat{k}_k\hat{\theta}_k$ and $\mathrm{Var}[U] = k_k\theta_k^2 - \hat{k}_k\hat{\theta}_k^2$. Therefore, we have

$$F_U\left(\frac{x\cdot v}{y+\sigma^2}\right) = \begin{cases} \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{\frac{x\cdot v}{y+\sigma^2}-\mu_u}{\sqrt{2}\sigma_u}\right), & \text{Gaussian} \\ \frac{1}{\Gamma(k_u)}\gamma\left[k_u, \frac{x\cdot v}{(y+\sigma^2)\theta_u}\right], & \text{Gamma} \end{cases} \tag{24}$$

where

$$\begin{aligned} \mu_u &= k_k\theta_k - \hat{k}_k\hat{\theta}_k = \sum_{l=1}^{L}\varepsilon_{kl} \\ \sigma_u^2 &= k_k\theta_k^2 - \hat{k}_k\hat{\theta}_k^2 = \sum_{l=1}^{L}\varepsilon_{kl}\left(R_{kl}+\beta_{kl}\right) \end{aligned} \tag{25a}$$

$$\begin{aligned} k_u &= \frac{\left(k_k\theta_k - \hat{k}_k\hat{\theta}_k\right)^2}{k_k\theta_k^2 - \hat{k}_k\hat{\theta}_k^2} = \frac{\left(\sum_{l=1}^{L}\varepsilon_{kl}\right)^2}{\sum_{l=1}^{L}\varepsilon_{kl}\left(R_{kl}+\beta_{kl}\right)} \\ \theta_u &= \frac{k_k\theta_k^2 - \hat{k}_k\hat{\theta}_k^2}{k_k\theta_k - \hat{k}_k\hat{\theta}_k} = \frac{\sum_{l=1}^{L}\varepsilon_{kl}\left(R_{kl}+\beta_{kl}\right)}{\sum_{l=1}^{L}\varepsilon_{kl}} \end{aligned} \tag{25b}$$

In addition, by using a taylor series approximation, we can make an asymptotic expansion of the results as

$$\begin{aligned} &F_U\left(\frac{x\cdot v}{y+\sigma^2}\right) \\ &= \int_{-\infty}^{\frac{x\cdot v}{(y+\sigma^2)}} f_U(u)\, du \\ &= \begin{cases} \frac{1}{2} + \frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\frac{(-1)^j\left(\frac{x\cdot v}{y+\sigma^2}-\mu_u\right)^{2j+1}}{2^j j!(2j+1)\sigma_u^{2j+1}}, & \text{Gaussian} \\ \frac{1}{\Gamma(k_u)}\sum_{j=0}^{\infty}(-1)^j\frac{\left(\frac{x\cdot v}{y+\sigma^2}\right)^{k_u+j}}{j!(k_u+j)\theta_u^{k_u+j}}, & \text{Gamma} \end{cases} \end{aligned} \tag{26}$$

Step 2: Calculate the rest three integrations

$$\int_0^{+\infty} f_X(x)\int_0^{+\infty} f_Y(y)\int_0^{+\infty} \times f_V(v)\cdot F_U\left(\frac{x\cdot v}{y+\sigma^2}\right) dv\, dy\, dx. \tag{27}$$

The rest integrations calculate the expectations of X, Y, and V. Therefore, by replacing the rest Gamma variables with their means, we have

$$\int_0^{+\infty}\int_0^{+\infty}\int_0^{+\infty}\int_{-\infty}^{\frac{x\cdot v}{(y+\sigma^2)}}$$

$$\times f_X(x) f_Y(y) f_V(v) f_U(u)\, du\, dv\, dy\, dx$$
$$= \begin{cases} \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{j=0}^{\infty} \frac{(-1)^j \left(\frac{k_x\theta_x k_y\theta_y}{k_y\theta_y+\sigma^2} - \mu_u\right)^{2j+1}}{2^j\, j!(2j+1)\sigma_u^{2j+1}}, & \text{Gaussian} \\ \frac{1}{\Gamma(k_u)} \sum_{j=0}^{\infty} (-1)^j \frac{\left(\frac{k_x\theta_x k_y\theta_y}{k_y\theta_y+\sigma^2}\right)^{k_u+j}}{j!(k_u+j)\theta_u^{k_u+j}}, & \text{Gamma} \end{cases} \tag{28}$$

where

$$k_x = \hat{k}_k, \theta_x = \hat{\theta}_k \tag{29a}$$
$$k_y = \frac{K-1}{L}\hat{k}_k, \theta_y = \hat{\theta}_k \tag{29b}$$
$$k_v = \frac{(K-1)\left(k_k\theta_k - \hat{k}_k\hat{\theta}_k\right)^2}{L\left(k_k\theta_k^2 - \hat{k}_k\hat{\theta}_k^2\right)}, \theta_v = \frac{k_k\theta_k^2 - \hat{k}_k\hat{\theta}_k^2}{k_k\theta_k - \hat{k}_k\hat{\theta}_k} \tag{29c}$$

Finally, we can calculate $\mathbb{P}[\mathrm{TE}_k]$ with MRT precoding according to equation (30), as shown at the bottom of the page, for UE k ($k = 1, \cdots, K$) in the systems.

2) $\mathbb{P}[\mathrm{TE}_k]$ WITH ZF PRECODING

Due to the definition, we can ignore the the interference term $\sum_{i=1,i\neq k}^{K} \left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2$ as it is far less than other terms. Therefore, $\mathbb{P}[\mathrm{TE}_k]$ can be rewritten as

$$\mathbb{P}\left[\frac{\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2}{\sigma^2} > \frac{\left\|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2}{\sum_{i=1,i\neq k}^{K}\left\|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2 + \sigma^2}\right] = \mathbb{P}\left[XY > \sigma^2 \cdot U\right] \tag{31}$$

where

$$X = \left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2,\ U = \left\|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2 - \left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2$$
$$Y = \sum_{i=1,i\neq k}^{K}\left\|\mathbf{h}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2 = \sum_{i=1,i\neq k}^{K}\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2 + \sum_{i=1,i\neq k}^{K}\left(\mathbf{w}_i^{\mathrm{H}}\tilde{\mathbf{h}}_k\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i + \mathbf{w}_i^{\mathrm{H}}\tilde{\mathbf{h}}_k\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right) + \sum_{i=1,i\neq k}^{K}\left\|\tilde{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2$$
$$\approx \sum_{i=1,i\neq k}^{K}\left\|\tilde{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2 \tag{32}$$

We simplify the expression of variable Y in the equation above for further computation, as the rest two sum terms in the original expression are negligible. These three variables become almost independent as the number of transmit antennas or APs grows. It is also evident that two of these variables, X and Y, obey Gamma distribution with ZF precoding.

Then, we need to analyze the distribution of U. Similar to analysis with MRT precoding, the approximation distribution of variable U can be seen as Gaussian distribution when σ^2 is low, while it can be seen as Gamma distribution when σ^2 is large. Therefore, we use the PDFs of X, Y, and U to express $\mathbb{P}[\mathrm{TE}_k]$ as

$$\mathbb{P}[\mathrm{TE}_k] = \int_0^{+\infty}\int_0^{+\infty}\int_{-\infty}^{\frac{x\cdot y}{\sigma^2}} f_X(x) f_Y(y) f_U(u)\, du\, dy\, dx \tag{33}$$

Step 1: Calculate the first integration $\int_{-\infty}^{\frac{x\cdot y}{\sigma^2}} f_U(u)\, du$.

Similar to *Step 1* with MRT precoding, we can get the result according to variable U's CDF. To further simplify the calculation, we also make an asymptotic expansion as

$$F_U\left(\frac{x\cdot y}{\sigma^2}\right) = \int_{-\infty}^{\frac{x\cdot y}{\sigma^2}} f_U(u)\, du = \begin{cases} \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{j=0}^{\infty} \frac{(-1)^j \left(\frac{x\cdot y}{\sigma^2} - \mu_u\right)^{2j+1}}{2^j\, j!(2j+1)\sigma_u^{2j+1}}, & \text{Gaussian} \\ \frac{1}{\Gamma(k_u)} \sum_{j=0}^{\infty} (-1)^j \frac{\left(\frac{x\cdot y}{\sigma^2}\right)^{k_u+j}}{j!(k_u+j)\theta_u^{k_u+j}}, & \text{Gamma} \end{cases} \tag{34}$$

where

$$\mu_u = \frac{L-K+1}{L}\sum_{l=1}^{L}\varepsilon_{kl}$$
$$\sigma_u^2 = \frac{L-K+1}{L}\sum_{l=1}^{L}\varepsilon_{kl}\left(R_{kl}+\beta_{kl}\right) \tag{35a}$$
$$k_u = \frac{(L-K+1)\left(\sum_{l=1}^{L}\varepsilon_{kl}\right)^2}{L\sum_{l=1}^{L}\varepsilon_{kl}\left(R_{kl}+\beta_{kl}\right)}$$
$$\theta_u = \frac{\sum_{l=1}^{L}\varepsilon_{kl}\left(R_{kl}+\beta_{kl}\right)}{\sum_{l=1}^{L}\varepsilon_{kl}} \tag{35b}$$

Step 2: Calculate the rest two integrations

$$\int_0^{+\infty} f_X(x)\int_0^{+\infty} f_Y(y)\cdot F_U\left(\frac{x\cdot v}{\sigma^2}\right) dy\, dx. \tag{36}$$

$$\mathrm{MRT}: \mathbb{P}[\mathrm{TE}_k] = \begin{cases} \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \sum_{j=0}^{\infty} \frac{(-1)^j}{2^j\, j!\,(2j+1)\,\sigma_u^{2j+1}} \left[\frac{\left(\sum_{l=1}^{L}\beta_{kl}\right)\left(\sum_{l=1}^{L}\varepsilon_{kl}\right)}{\sum_{l=1}^{L}\beta_{kl} + \frac{\sigma^2 L}{K-1}} - \mu_u\right]^{2j+1}, & \text{Gaussian } U \\ \frac{1}{\Gamma(k_u)} \sum_{j=0}^{\infty} \frac{(-1)^j}{j!\,(k_u+j)\,\theta_u^{k_u+j}} \left[\frac{\left(\sum_{l=1}^{L}\beta_{kl}\right)\left(\sum_{l=1}^{L}\varepsilon_{kl}\right)}{\sum_{l=1}^{L}\beta_{kl} + \frac{\sigma^2 L}{K-1}}\right]^{k_u+j}, & \text{Gamma } U \end{cases} \tag{30}$$

The rest integrations calculate the expectations of X and Y. Therefore, by replacing the two Gamma variables with their moments, we have

$$\int_0^{+\infty}\int_0^{+\infty}\int_{-\infty}^{\frac{x\cdot y}{\sigma^2}} f_X(x) f_Y(y) f_U(u)\, du\, dy\, dx = \begin{cases} \frac{1}{2}+\frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\frac{(-1)^j\left(\frac{k_x\theta_x k_y\theta_y}{\sigma^2}-\mu_u\right)^{2j+1}}{2^j j!(2j+1)\sigma_u^{2j+1}}, & \text{Gaussian} \\ \frac{1}{\Gamma(k_u)}\sum_{j=0}^{\infty}(-1)^j\frac{(k_x\theta_x k_y\theta_y)^{k_u+j}}{j!(k_u+j)\theta_u^{k_u+j}\sigma^{2(k_u+j)}}, & \text{Gamma} \end{cases} \tag{37}$$

where

$$k_x = \frac{L-K+1}{L}\hat{k}_k,\ \theta_x = \hat{\theta}_k \tag{38a}$$

$$k_y = \frac{K-1}{L}\tilde{k}_k,\ \theta_y = \tilde{\theta}_k \tag{38b}$$

Finally, we can calculate $\mathbb{P}[\mathrm{TE}_k]$ with MRT precoding according to equation (39), as shown at the bottom of the page, for UE k ($k = 1, \cdots, K$) in the systems. So the probability of transmission error $\mathbb{P}[\mathrm{TE}] = 1-\prod_{k=1}^{K}(1-\mathbb{P}[\mathrm{TE}_k])$ can be obtained with $\mathbb{P}[\mathrm{TE}_k]$ with both MRT and ZF precoding.

B. TIME OVERFLOW

Calculating the probability of the second kind of device failure $\mathbb{P}[\mathrm{TO}_k]$ is the prerequisite for the calculation of $\mathbb{P}[\mathrm{TO}]$. To get the final expression, we need to divide the terms in SINR of R_k into different almost independent variables. Therefore, we rewrite the expression of $\mathbb{P}[\mathrm{TO}_k]$ as

$$\mathbb{P}\left[\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2 < \left(2^{\frac{B}{T_D W}}-1\right)\left(\sum_{i=1,i\neq k}^{K}\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2+\sigma^2\right)\right] = \mathbb{P}\left[X < \left(2^{\frac{B}{T_D W}}-1\right)\left(Y+\sigma^2\right)\right] \tag{40}$$

where X refers to the signal term $\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2$, while Y refers to the interference term $\sum_{i=1,i\neq k}^{K}\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2$. Based on the PDFs of X and Y, we can express $\mathbb{P}[\mathrm{TO}_k]$ as

$$\mathbb{P}[\mathrm{TO}_k] = \int_0^{+\infty}\int_0^{\left(2^{\frac{B}{T_D W}}-1\right)(y+\sigma^2)} f_X(x) f_Y(y)\, dx\, dy \tag{41}$$

Step 1: Calculate the first integration

$$\int_0^{\left(2^{\frac{B}{T_D W}}-1\right)(y+\sigma^2)} f_X(x)\, dx. \tag{42}$$

The result is mainly decided by variable X's CDF. We also make an asymptotic expansion for low complexity as

$$F_X\left[\left(2^{\frac{B}{T_D W}}-1\right)\left(y+\sigma^2\right)\right] = \frac{1}{\Gamma(k_x)} \cdot \sum_{j=0}^{\infty}(-1)^j\frac{1}{j!\,(k_x+j)}\left[\frac{\left(2^{\frac{B}{T_D W}}-1\right)\left(y+\sigma^2\right)}{\theta_x}\right]^{k_x+j} \tag{43}$$

Step 2: Calculate the rest integration

$$\int_0^{+\infty} f_Y(y)\cdot F_X\left[\left(2^{\frac{B}{T_D W}}-1\right)\left(y+\sigma^2\right)\right] dy. \tag{44}$$

The rest integration calculate the expectation of variable Y. Therefore, by replacing the Gamma variable Y with its moment, we have

$$\int_0^{+\infty} f_Y(y)\cdot F_X\left[\left(2^{\frac{B}{T_D W}}-1\right)\left(y+\sigma^2\right)\right] dy = \frac{1}{\theta_x^{k_x}\Gamma(k_x)}\sum_{j=0}^{\infty}\frac{(-1)^j\left(2^{\frac{B}{T_D W}}-1\right)^{k_x+j}}{j!\,(k_x+j)\,\theta_x^j}\left[\left(k_y\theta_y+\sigma^2\right)^{k_x+j}\right] \tag{45}$$

Then we determine the parameters according to the precoding vectors.

1) $\mathbb{P}[\mathrm{TO}_k]$ WITH MRT PRECODING

When calculate $\mathbb{P}[\mathrm{TO}_k]$ with MRT precoding, we set the parameters as

$$k_x = \hat{k}_k,\ \theta_x = \hat{\theta}_k \tag{46a}$$

$$k_y = \frac{K-1}{L}\hat{k}_k,\ \theta_y = \hat{\theta}_k \tag{46b}$$

2) $\mathbb{P}[\mathrm{TO}_k]$ WITH ZF PRECODING

When calculate $\mathbb{P}[\mathrm{TO}_k]$ with ZF precoding, we set the parameters as

$$k_x = \frac{L-K+1}{L}\hat{k}_k,\ \theta_x = \hat{\theta}_k \tag{47a}$$

$$\mathrm{ZF}: \mathbb{P}[\mathrm{TE}_k] = \begin{cases} \frac{1}{2}+\frac{1}{\sqrt{2\pi}}\sum_{j=0}^{\infty}\frac{(-1)^j}{2^j j!\,(2j+1)\,\sigma_u^{2j+1}}\left[\frac{(L-K+1)(K-1)\left(\sum_{l=1}^{L}\beta_{kl}\right)\left(\sum_{l=1}^{L}\varepsilon_{kl}\right)}{\sigma^2 L^2}-\mu_u\right]^{2j+1}, & \text{Gaussian } U \\ \frac{1}{\Gamma(k_u)}\sum_{j=0}^{\infty}\frac{(-1)^j}{j!\,(k_u+j)\,\theta_u^{k_u+j}}\left[\frac{(L-K+1)(K-1)\left(\sum_{l=1}^{L}\beta_{kl}\right)\left(\sum_{l=1}^{L}\varepsilon_{kl}\right)}{\sigma^2 L^2}\right]^{k_u+j}, & \text{Gamma } U \end{cases} \tag{39}$$

TABLE 1. Simulation setup parameters.

Parameters	Values
Number of single-antenna APs L	400
Carrier frequency f_c	2 GHz
Number of mutually orthogonal pilot signals τ_p	20
Transmit bandwidth W	20 MHz
Noise variance at the UEs σ_0^2	-174 dBm
Transmit power P_k of UE k	100 mW
Height difference between an AP and a UE	10m
Angular standard deviation around the nominal angle	15°

$$k_y = \frac{K-1}{L}\hat{k}_k, \theta_y = \hat{\theta}_k \tag{47b}$$

Finally, we can calculate $\mathbb{P}[\mathrm{TO}_k]$ with both MRT and ZF precoding according to equation (48), as shown at the bottom of the page, for UE k ($k = 1, \cdots, K$) in the systems. Note that the expressions for k_x and θ_x are finally decided by the precoding methods. So the probability of time overflow $\mathbb{P}[\mathrm{TO}] = 1 - \prod_{k=1}^{K}(1 - \mathbb{P}[\mathrm{TO}_k])$ can be obtained with $\mathbb{P}[\mathrm{TO}_k]$. Then, we calculate the probability of system outage $\mathbb{P}[\mathrm{SO}]$ by the expression as

$$\mathbb{P}[\mathrm{SO}] = 1 - \prod_{k=1}^{K}(1 - \mathbb{P}[\mathrm{TE}_k] - \mathbb{P}[\mathrm{TO}_k]). \tag{49}$$

IV. NUMERICAL RESULTS

This section verifies our expressions above via a set of Monte-Carlo simulations. We plot the probability of transmission error as a function of the noise variance value and the probability of time overflow as the deadline value. In addition, we compare the theoretical expressions given above with the result of Monte-Carlo simulations to test the accuracy of the theoretical result.

We match the simulation setup with the 3GPP Urban Microcell model in [26]. In our simulation, the UEs are are distributed randomly and uniformly in a $1 \times 1\mathrm{km}^2$ area. The expression of the large-scale fading coefficient is

$$R_{kl}\,(\mathrm{dB}) = -30.5 - 36.7\log_{10}\left(\frac{d_{kl}}{1\ \mathrm{m}}\right) + F_{kl} \tag{50}$$

where d_{kl} denotes the distance between UE k and AP l and F_{kl} refers to the shadow fading which only appears when the distance is large. The rest simulation parameter values are summarized in Table 1 [11], [12].

Firstly, we verify that the expressions of $\mathbb{P}[\mathrm{TE}_k]$ with MRT and ZF precoding in Section III are accurate. We obtain the probability of transmission error of UE 1, $\mathbb{P}[\mathrm{TE}_1]$, through Monte-Carlo simulations and plot it as a function of the noise variance in Fig. 2 and Fig. 3. Then, we calculate the theoretical values of equation (30) and (39) and plot them in the corresponding figure. We pick the closest theoretical points and plot them as the fourth line. As seen from the two figures, compared with the simulation results, the theoretical values are more accurate by characterizing the distribution of variable U as Gaussian distribution when noise variance is low. However, when noise variance is large, the theoretical values are more accurate by characterizing the distribution of variable U as Gamma distribution compared with the simulation values.

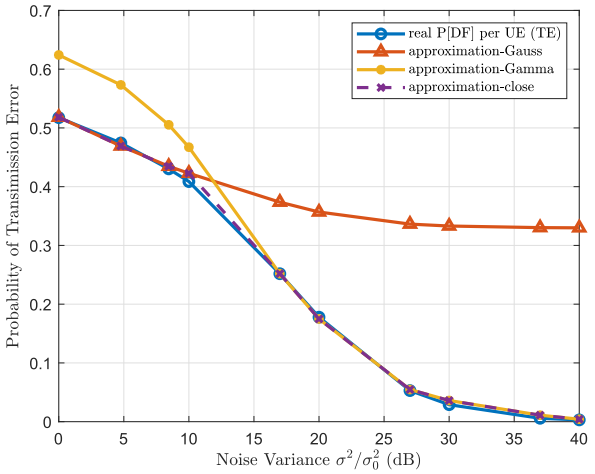

FIGURE 2. Simulation and theoretical results of $\mathbb{P}[\mathrm{TE}_1]$ with MRT percoding.

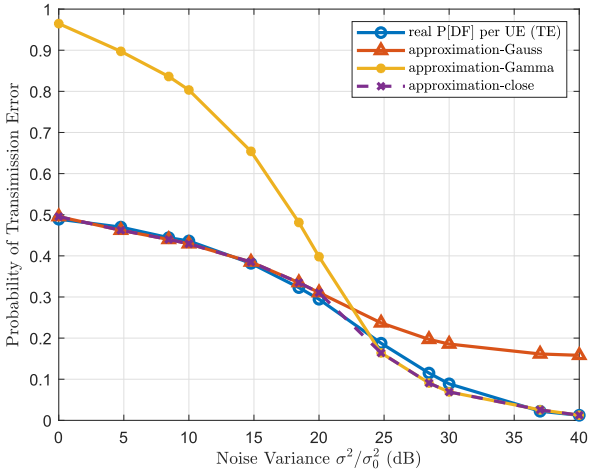

FIGURE 3. Simulation and theoretical results of $\mathbb{P}[\mathrm{TE}_1]$ with ZF percoding.

Secondly, we verify that the expressions of $\mathbb{P}[\mathrm{TO}_k]$ with MRT and ZF precoding in Section III are accurate whether the noise variance is large or low. We get the probability of time overflow of UE 1, $\mathbb{P}[\mathrm{TO}_1]$, through Monte-Carlo simulations and plot it as a function of the hard deadline value in Fig. 4 and Fig. 5. It can be seen from the two figures, the expression of

$$\mathbb{P}[\mathrm{TO}_k] = \frac{1}{\theta_x^{k_x}\Gamma(k_x)}\sum_{j=0}^{\infty}\frac{(-1)^j\left(2^{\frac{B}{TDW}}-1\right)^{k_x+j}}{j!\,(k_x+j)\,\theta_x^j}\left[\frac{(K-1)\sum_{l=1}^{L}\beta_{kl}}{L}+\sigma^2\right]^{k_x+j} \tag{48}$$

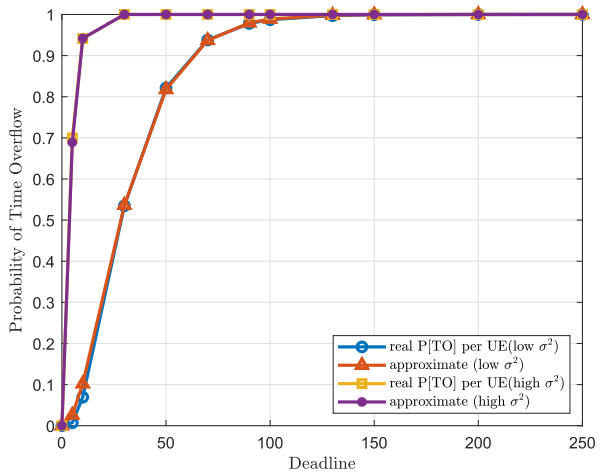

FIGURE 4. Simulation and theoretical result of $\mathbb{P}[\mathrm{TO}_1]$ with MRT percoding.

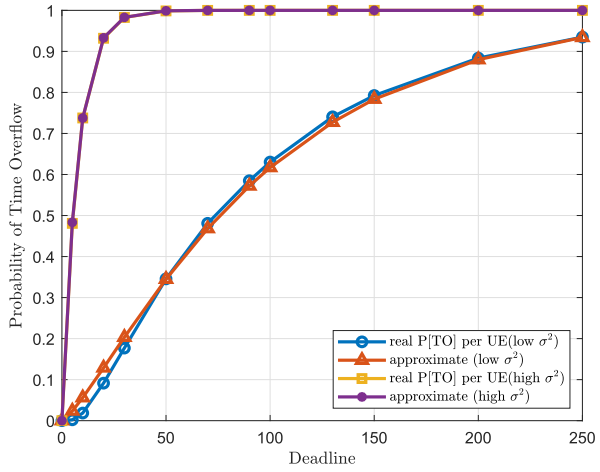

FIGURE 5. Simulation and theoretical result of $\mathbb{P}[\mathrm{TO}_1]$ with ZF percoding.

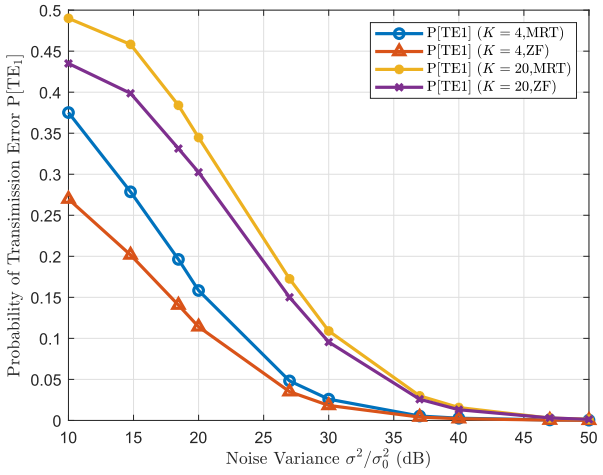

FIGURE 6. Simulation result of $\mathbb{P}[\mathrm{TE}_k]$ versus the noise variance with diffrent numbers of UE and diffrent precoding method.

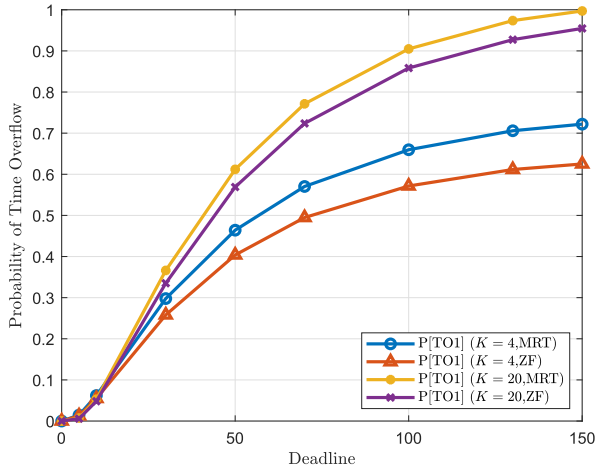

FIGURE 7. Simulation result of $\mathbb{P}[\mathrm{TO}_k]$ versus the noise variance with diffrent numbers of UE and diffrent precoding method.

$\mathbb{P}[\mathrm{TO}_k]$ in equation (45) with both MRT and ZF precoding is accurate compared with the Monte-Carlo simulation results. The figure also demonstrates that the increasing noise variance σ^2 brings about an increase in $\mathbb{P}[\mathrm{TO}_k]$, and the growth range is larger with ZF percoding.

Then, we plot the probabilities of transmission error and time overflow for different numbers of UE and different precoding methods to compare the outage performance of the systems. As can be seen from Fig. 6 and Fig. 7, systems with ZF precoding outperform systems with MRT precoding for having lower probabilities of transmission error and time overflow. Therefore, systems with ZF precoding have better outage performance. In addition, a more significant number of UEs leads to more interference and increasing probabilities of both transmission error and time overflow. Therefore, user grouping is necessary when the number of UEs grows to decrease the interference and probability of system outage.

For applications having more considerable noise variance, the probability of transmission error is low enough to satisfy the reliability required of URLLC, whatever the precoding method, according to Fig. 6. Then, the system can have a low outage probability as the length of downlink data is small enough to have a low probability of time overflow. However, transmission error probability converges to 0.5 when the noise variance is low, which will cause the system outage probability to get high. Therefore, a multiplicative *utilization factor*, $0 < \rho < 1$, can be introduced before the SINR term in the expressions of the chosen rate R_k to obtain a lower probability of transmission error when the noise variance decreases. So the chosen rate by the APs becomes

$$R_k = W\log_2\left(1+\frac{\rho\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_k\right\|^2}{\sum_{i=1,i\neq k}^{K}\left\|\hat{\mathbf{h}}_k^{\mathrm{H}}\mathbf{w}_i\right\|^2+\sigma^2}\right) \tag{51}$$

Similar ways can be found in [27], [28]. Moreover, a *back-off parameter* outside the log-function is proposed in [16] to circumvent this problem. Both the two parameters help decrease the probability of transmission error as the noise variance decreases. In addition, user grouping can also be introduced to obtain a lower probability of time overflow.

V. CONCLUSION

This paper modeled a multiuser, variable-rate downlink cell-free massive MIMO communication system. We defined two events that will lead to the system outage, to describe the

system's failure in latency and reliability performance of the transmission. The first kind of event was defined as *Transmission Error*, caused by the value of the downlink rate getting too high to be supported by the channel between the APs and UEs. The other event was defined as *Time Overflow*, caused by the violation of packet delivery against the hard deadline. We offered analytical methods, which used the properties of Gamma distribution and the approximation for non-isotropic vectors to derive the expressions of the probabilities of transmission error, time overflow, and system outage with MRT and ZF precoding considering the multiuser interference. The numerical results of system-level simulations presented that the expressions are accurate. Moreover, The noise variance, the number of UEs, and the precoding scheme influence the outage performance, or the URLLC performance, to varying degrees.

REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, Jun. 2014.

[2] S. Parkvall, E. Dahlman, A. Furuskar, and M. Frenne, "NR: The new 5G radio access technology," *IEEE Commun. Standards Mag.*, vol. 1, no. 4, pp. 24–30, Dec. 2017.

[3] V. H. M. Donald, "Advanced mobile phone service: The cellular concept," *Bell Syst. Tech. J.*, vol. 58, no. 1, pp. 15–41, Jan. 1979.

[4] O. D. Oyeleke, S. Thomas, I. Bismark, I. Muhammad, O. Nkemdilim, and S. U. Hussein, "Cell free massive MIMO-enabling technology in providing uniformly good service to users for 5G and B5G mobile communications," in *Proc. 15th Int. Conf. Electron., Comput. Comput. (ICECCO)*, Dec. 2019, pp. 1–7.

[5] A. G. Burr, M. Bashar, and D. Maryopi, "Ultra-dense radio access networks for smart cities: Cloud-RAN, fog-RAN and 'cell-free' massive MIMO," in *Proc. PIMRC, Int. Workshop CorNer*, 2018, pp. 1–5.

[6] S. Perlma and A. Forenza, "An introduction to pCell," Artemis Netw. LLC, San Francisco, CA, USA, White Paper, 2015. [Online]. Available: http://www.rearden.com/artemis/An-Introduction-to-pCell-White-Paper-150224.pdf

[7] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.

[8] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and H. Yang, "Cell-free massive MIMO systems," in *Proc. 49th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2015, pp. 695–699.

[9] J. Zhang, Y. Wei, E. Bjornson, Y. Han, and X. Li, "Spectral and energy efficiency of cell-free massive MIMO systems with hardware impairments," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2017, pp. 1–6.

[10] L. D. Nguyen, T. Q. Duong, H. Q. Ngo, and K. Tourki, "Energy efficiency in cell-free massive MIMO with zero-forcing precoding design," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1871–1874, Aug. 2017.

[11] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *IEEE Trans. Wireless Commun.*, vol. 19, no. 1, pp. 77–90, Jan. 2020.

[12] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 7, pp. 4247–4261, Jul. 2020.

[13] H. Yu, N. Ye, and A. Wang, "Non-orthogonal wireless backhaul design for cell-free massive MIMO: An integrated computation and communication approach," *IEEE Wireless Commun. Lett.*, vol. 10, no. 2, pp. 281–285, Feb. 2021.

[14] X. You, D. Wang, and J. Wang, *Fundamentals of Distributed MIMO and Cell-Free Mobile Communications*. Beijing, China: Science Press, 2020.

[15] X. You *et al.*, "Towards 6G wireless communication networks: Vision, enabling technologies, and new paradigm shifts," *Sci. China Inf. Sci.*, vol. 64, no. 1, pp. 1–74, 2021.

[16] R. Jurdi, S. R. Khosravirad, H. Viswanathan, J. G. Andrews, and R. W. Heath, Jr., "Outage of periodic downlink wireless networks with hard deadlines," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1238–1253, Feb. 2019.

[17] G. Interdonato, P. Frenger, and E. G. Larsson, "Scalability aspects of cell-free massive MIMO," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2019, pp. 1–6.

[18] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Found. Trends Signals Process.*, vol. 11, nos. 3–4, pp. 154–655, 2017.

[19] E. Bjornson, L. Sanguinetti, and M. Debbah, "Massive MIMO with imperfect channel covariance information," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 974–978.

[20] D. Neumann, M. Joham, and W. Utschick, "Covariance matrix estimation in massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 863–867, Jun. 2018.

[21] K. Upadhya and S. A. Vorobyov, "Covariance matrix estimation for massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 4, pp. 546–550, Apr. 2018.

[22] D. Wang, J. Wang, X. You, Y. Wang, M. Chen, and X. Hou, "Spectral efficiency of distributed MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 10, pp. 2112–2127, Oct. 2013.

[23] J. Li, D. Wang, P. Zhu, and X. You, "Uplink spectral efficiency analysis of distributed massive MIMO with channel impairments," *IEEE Access*, vol. 5, pp. 5020–5030, 2017.

[24] L. Zhang, P. Zhu, J. Li, and J. Cao, "Downlink ergodic rate analysis of DAS with linear beamforming under pilot contamination," in *Proc. 9th Int. Conf. Wireless Commun. Signal Process. (WCSP)*, Oct. 2017, pp. 1–6.

[25] S. Schiessl, J. Gross, M. Skoglund, and G. Caire, "Delay performance of the multiuser MISO downlink under imperfect CSI and finite-length coding," *IEEE J. Sel. Areas Commun.*, vol. 37, no. 4, pp. 765–779, Apr. 2019.

[26] *Further Advancements for E-UTRA Physical Layer Aspects (Release 9)*, 3GPP, document TS 36.814, Mar. 2017.

[27] S. Akoum, M. Kountouris, and R. W. Heath, Jr., "On imperfect CSI for the downlink of a two-tier network," in *Proc. IEEE Int. Symp. Inf. Theory Proc.*, Jul. 2011, pp. 553–557.

[28] J. Zhang and J. G. Andrews, "Distributed antenna systems with randomness," *IEEE Trans. Wireless Commun.*, vol. 7, no. 9, pp. 3636–3646, Sep. 2008.

LE TAN is with the School of Information Science and Engineering, Southeast University. Her research interests include cell-free massive MIMO communications and URLLC.

ZIYANG ZHANG received the B.S. degree in communication and information systems from Hohai University, China, in 2020. He is currently pursuing the Ph.D. degree with the National Mobile Communications Research Laboratory, Southeast University, China. His current research interests include cell-free massive MIMO systems and ultra-reliable low-latency communication design.

DONGMING WANG (Member, IEEE) received the B.S. degree from the Chongqing University of Posts and Telecommunications, in 1999, the M.S. degree from the Nanjing University of Posts and Telecommunications, in 2002, and the Ph.D. degree from Southeast University, China, in 2006. He joined at the National Mobile Communications Research Laboratory, Southeast University, in 2006, where he is currently a Professor. His research interests include signal processing for wireless communications and large-scale distributed MIMO systems (cell-free massive MIMO).