

This is a repository copy of *Uplink Spectral and Energy Efficiency of Cell-Free Massive MIMO with Optimal Uniform Quantization*.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/171842/>

Version: Accepted Version

Article:

Bashar, Manijeh, Ngo, Hien Quoc, Cumanan, Kanapathippillai orcid.org/0000-0002-9735-7019 et al. (4 more authors) (2021) Uplink Spectral and Energy Efficiency of Cell-Free Massive MIMO with Optimal Uniform Quantization. IEEE Transactions on Communications. pp. 223-245. ISSN 0090-6778

<https://doi.org/10.1109/TCOMM.2020.3028305>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Uplink Spectral and Energy Efficiency of Cell-Free Massive MIMO with Optimal Uniform Quantization

Manijeh Bashar, *Member, IEEE*, Hien Quoc Ngo, *Senior Member, IEEE*, Kanapathippillai Cumanan, *Senior Member, IEEE*, Alister G. Burr, *Senior Member, IEEE*, Pei Xiao, *Senior Member, IEEE*, Emil Björnson, *Senior Member, IEEE* and Erik G. Larsson, *Fellow, IEEE*

Abstract—This paper investigates the performance of limited-fronthaul cell-free massive multiple-input multiple-output (MIMO) taking account the fronthaul quantization and imperfect channel acquisition. Three cases are studied, which we refer to as Estimate&Quantize, Quantize&Estimate, and Decentralized, according to where channel estimation is performed and exploited. Maximum-ratio combining (MRC), zero-forcing (ZF), and minimum mean-square error (MMSE) receivers are considered. The Max algorithm and the Bussgang decomposition are exploited to model optimum uniform quantization. Exploiting the optimal step size of the quantizer, analytical expressions for spectral and energy efficiencies are presented. Finally, an access point (AP) assignment algorithm is proposed to improve the performance of the decentralized scheme. Numerical results investigate the performance gap between limited fronthaul and perfect fronthaul cases, and demonstrate that exploiting relatively few quantization bits, the performance of limited-fronthaul cell-free massive MIMO closely approaches the perfect-fronthaul performance.

Keywords: Bussgang decomposition, cell-free massive MIMO, C-RAN, energy efficiency, quantization, optimal uniform quantization.

I. INTRODUCTION

A. Why Cell-Free Massive MIMO?

The goal of the fifth generation (5G) wireless network is to provide greatly increased throughput while supporting very high density of users. A wireless network that can serve large numbers of users at the same time must become an ultra-dense network (UDN). In a UDN, the density of base stations (BSs) may exceed that of users, and distance between users and BSs is reduced to a few metres [2], [3]. The huge interference

M. Bashar and P. Xiao are with the Institute for Communication Systems, home of 5G Innovation Centre (5GIC) at the University of Surrey, Guildford, GU2 7XH, U.K., e-mail: {m.bashar, p.xiao}@surrey.ac.uk. H. Q. Ngo is with the School of Electronics, Electrical Engineering and Computer Science, Queen’s University Belfast, Belfast, T7 1NN, U.K., e-mail: hien.ngo@qub.ac.uk. K. Cumanan, and A. G. Burr are with the Department of Electronic Engineering, University of York, York, YO10 5NG, U.K. e-mail: {kanapathippillai.cumanan, alister.burr}@york.ac.uk. E. Björnson and E. G. Larsson are with the Department of Electrical Engineering, Linköping University, Linköping, SE-581 83, Sweden, e-mail: {emil.bjornson, erik.g.larsson}@liu.se.

The work of P. Xiao and M. Bashar was supported by the U.K. Engineering and Physical Sciences Research Council under Grant EP/P008402/2 and EP/R001588/1.

The work of H. Q. Ngo was supported by the UK Research and Innovation Future Leaders Fellowships under Grant MR/S017666/1.

The work of K. Cumanan and A. G. Burr was supported by H2020-MSCARISE-2015 under grant number 690750.

The work of E. G. Larsson was supported in part by the Swedish Research Council (VR) and ELLIIT.

Parts of this work was presented at the IEEE Asilomar 2018 [1].

from the neighbouring BSs is such that the capacity scaling law of the cellular paradigm no longer holds and hence interference becomes a limiting factor [2]–[4]. Exploiting coordination techniques, several neighbouring BSs can share data via fronthaul links: however, coordination among BSs results in a very large overhead in UDNs, which is costly and practically infeasible. The centralized or “cloud” radio access network (C-RAN) is a promising network architecture that can effectively address the interference and coordination issues in UDNs [2].

In C-RAN, the distributed remote radio heads (RRHs) service the users while the base-band signal processing is performed at a central base band unit (BBU) [5], [6]. C-RAN also allows implementation of a version of coordinated multipoint processing (CoMP), where the RRHs replace the coordinated BSs, and everything is processed at the “cloud” (the centralized processor). While C-RAN was created to reduce the hardware costs, by allowing multiple base stations to share a processor, it also enables the implementation of CoMP-based cooperation between base stations connected to the same “cloud”-based processor.

C-RAN and cell-free massive multiple-input multiple-output (MIMO) are not thus alternative technologies. Cell-free massive MIMO is very different since it is explicitly not based on the cellular principle. In cell-free Massive MIMO, we process the physical-layer signals in a different way and we deploy the access points (APs) in a different way from conventional cellular networks. Specifically, many APs are distributed across the area, removing the concepts of cell and cell-edge users. However, since cell-free massive MIMO requires a “central processing unit” (CPU), it can be implemented using a C-RAN-like architecture, and similarly the APs are connected to the CPU through fronthaul links. Nevertheless the software that we will run and the way that we deploy the APs will be different from that normally considered in C-RAN, reflecting the techniques used in massive MIMO. For example, some functional operations, such as channel estimation through uplink pilots, are carried out at the APs. Users are supported using time-division duplex (TDD), as in massive MIMO.

B. State-of-the-Art Limited-Fronthaul C-RAN

The common public radio interface (CPRI) specification was published by a union of radio equipment manufacturers, aiming to standardize the fronthaul communication of BBU and RRHs in C-RAN [7], [8]. In CPRI, each sample is

quantized exploiting a given large number of bits (around 15 bits) per sample. Based on [9], CPRI is not suitable for large-scale C-RAN and practical implementations using limited-fronthaul links are necessary. On the other hand, using a non-uniform quantizer, the quantization levels can be optimized as a function of the statistics of the signal using standard techniques such as the Lloyd-Max algorithm [10], [11]. Another approach in C-RAN is exploiting Wyner-Ziv coding. The coding scheme at the RRHs in order to leverage the side information at the receiver is known as Wyner-Ziv coding [12]. Using Wyner-Ziv quantization, the RRH can exploit a better quantizer with an improved resolution without any need to increase the fronthaul rate [8]. Another technique in C-RAN is compute-and-forward, which is originated from network information theory and nested lattice codes [8]. Finally, many of the network information theoretic problems for the design of C-RANs are open and improvement in this domain may provide a progress in the C-RAN technology [8], [13]. In this paper, we investigate the performance of limited-fronthaul cell-free massive using the Bussgang decomposition, which can be applied to the C-RAN technology.

C. The Limited Capacity of Fronthaul Links in Cell-Free Massive MIMO

The limited capacity links from the APs to the CPU constitute one of the most important challenges in cell-free massive MIMO [14]–[22]. This fronthaul limitation is a more crucial challenge on the uplink, as in the downlink mode the signals are sent as bit streams to the APs which then apply local modulation and coding whereas as the fronthaul links send the quantized version of the received signals at the APs to the CPU, which introduces additional self-interference to the signals at the CPU.

The total data rate required to transmit these quantized signals with sufficient precision to avoid performance loss is several times the total user data rate supported by those signals. In the C-RAN literature, this has been estimated as 20–50 times the corresponding data rate [23], implemented using the CPRI standard [7], typically over optical fiber.

The assumption of infinite fronthaul in [24] is not realistic in practice. It is reasonable to assume, however, that the fronthaul network will carry quantized signals, at least in the uplink direction, and that this will affect the network performance. This paper therefore provides an approach for the analysis of the effect of fronthaul quantization on the uplink of cell-free massive MIMO. While there has been significant work in the context of network MIMO on compression techniques such as Wyner-Ziv coding for interconnection of base stations, here for simplicity (and hence improved scalability) we assume simple uniform quantization. J. Max in [25] developed an algorithm to solve the problem of minimizing the mean-squared distortion (or mean-squared error (MSE)). In addition, P. Zillmann in [26] studied the problem of minimizing the MSE of the uniform quantizer exploiting the *Bussgang decomposition* [27]. Note that the *Max algorithm* and the scheme in [26] provide the same signal-to-distortion-plus-noise ratio (SDNR). In this paper, we exploit both the Max algorithm and the Bussgang decomposition to model optimal uniform quantization.

D. Different Implementations of Cell-Free Massive MIMO

In this paper, we study four cases for the implementation of cell-free massive MIMO networks with linear processing and present a fair comparison between the different schemes. By fair comparison, we refer to the fact that we assume the same fronthaul rate for all schemes. We investigate an uplink cell-free massive MIMO system with limited-fronthaul links and three different linear detection schemes, namely, maximum-ratio combining (MRC), zero-forcing (ZF), and minimum mean-square error (MMSE). We study the case in which the APs estimate the channel and the quantized versions of the estimated channels are transferred to the CPU through limited fronthaul links. Moreover, APs also send the quantized received signals to the CPU: we use the Bussgang decomposition to model the quantized signal. This case is referred to as Estimate&Quantize. Then, the case Quantize&Estimate is investigated, in which the quantized version of the received pilots is sent back to the CPU and the CPU performs the channel estimation. In this case, the APs also need to transfer the quantized versions of the received user data to the CPU. Next, we derive the fronthaul bit rate of limited-fronthaul cell-free massive MIMO. Finally, a decentralized scheme is investigated, in which the distributed APs multiply the received signal by the conjugate of the channel estimates, and a quantized version of this weighted signal is sent back to the CPU. Note that [28] considers the downlink of cell-free massive MIMO whereas in this paper, we investigate the uplink transmission. Moreover, the energy efficiency of cell-free massive MIMO with only error-free and unlimited capacity fronthaul links is investigated in [28].

E. Applications

Considering a limited-capacity fronthaul is practically important since this is the case that will appear in practice. For the first time, we investigate the performance of the system under realistic assumption of limited-capacity fronthaul links and using different detectors. The analytical results for the achievable rate and the energy efficiency can be used for power control optimization and other resource allocation tasks. Moreover, the analysis in this paper provides the means for system designers to investigate the outage probability, user assignment schemes, pilot assignment algorithms and bit error rate analysis with the realistic assumption of limited-capacity fronthaul links. Finally, the different implementations of cell-free massive MIMO (namely, Estimate&Quantize, Quantize&Estimate and the decentralized scheme) open the door to investigating flexible functional splits and signal compression techniques in this system model.

F. Contributions

The connection between the APs and the CPU is a challenge for the practical implementation of cell-free massive MIMO. One practical and feasible architecture of the cell-free massive MIMO is the radio stripe system [15], where the fronthaul links have limited capacity. The analysis in the current paper show that using only a few quantization bits (resulting in a

few tens of Mbit/s required fronthaul rate), the performance of limited-fronthaul cell-free massive MIMO closely approaches the perfect-fronthaul performance. The required capacity of fronthaul links obtained in this paper are compatible with the capacity of Ethernet links: a capacity of 100 Mbit/s can be achieved by using twisted pair Ethernet [29]. Moreover, energy-efficient techniques are considered as one of the key avenues for addressing these issues in the development of future wireless systems. As shown in the paper, the energy efficiency of the system is a function of the number of quantization bits. We accordingly find the optimal number of quantization bits to maximize the energy efficiency. As a result, the proposed uniform quantization scheme has low-power consumption and easy deployment which is very suitable for the practical implementation of cell-free massive MIMO. A “tradeoff” between the energy efficiency (as measured in bit/J) and spectral efficiency (as measured in bit/sec/Hz) of the cell-free massive MIMO system is an important performance metric [30]. However, the authors in [30] consider co-located massive MIMO whereas the investigation of the tradeoff between spectral efficiency and energy efficiency in cell-free massive MIMO is essential. Therefore, using the achievable rate obtained in the paper, we evaluate the tradeoff between the spectral efficiency and energy efficiency of the cell-free massive MIMO system. The contributions of the paper can be summarized as follows:

1. We exploit the Busgang decomposition and the Max algorithm to model the effect of uniform quantization with Estimate&Quantize, Quantize&Estimate and decentralized schemes. We derive analytical expressions for the user signal-to-interference-plus-noise ratio (SINR) for different schemes with three linear receivers, namely MRC, ZF and MMSE, taking into account the effects of limited fronthaul links, quantization errors and pilot contamination. We show that to analyse the performance of the aforementioned schemes with quantized channel gain at the receiver and non-Gaussian noise, we need to exploit the scheme in [31]. Note that this is the first paper which investigates the scheme in [31] and shows that the achievable rate can be obtained through a similar way as in [31, Section 2.3.5 and Table 2.3]. These analyses are novel and not presented anywhere else.
2. The required fronthaul rate and power consumption of the Estimate&Quantize, Quantize&Estimate and decentralized cases are investigated, which enables us to formulate the total power consumption of the system. Given the effect of the channel coherence time, the capacity of fronthaul links and the effect the quantization distortion, a closed form expression for an achievable spectral efficiency and the energy efficiency of the system are obtained. We show that very high spectral efficiency and energy efficiency can be obtained even with simple MRC processing. A tradeoff between the energy efficiency and spectral efficiency of the system is presented. It is shown that exploiting a moderately large number of APs can significantly enhance the spectral and energy efficiency of the system.

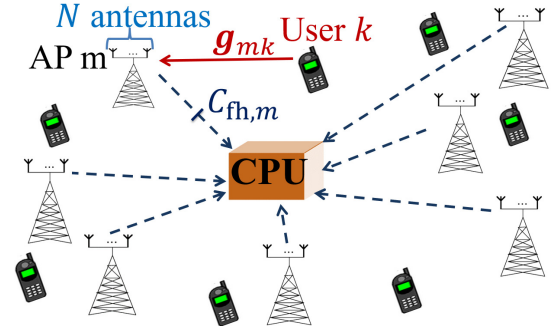


Figure 1. The uplink of a cell-free massive MIMO system with K single-antenna users and M APs. Each AP is equipped with N antennas. The solid lines denote the uplink channels and the dashed lines present the limited capacity fronthaul links between the APs and the CPU.

3. One of the aims of this work is to find a practical user assignment algorithm for cell-free massive MIMO. A novel and efficient user assignment algorithm is proposed, which is based on the capacity of fronthaul links and improves the performance of decentralized cell-free massive MIMO. Note that the proposed AP assignment is different from the AP selection schemes investigated in [15], [28], [32], as here we investigate the effect of the limited capacity of the fronthaul links on the total number of users which can be supported by each AP.
4. We present a performance comparison between the case of perfect fronthaul links, the Estimate&Quantize, Quantize&Estimate and the decentralized scheme. Next, we investigate the optimal number of AP antennas along with the optimal number of quantization bits and optimal number of APs to maximize the uplink energy efficiency of cell-free Massive MIMO. Numerical results confirm that relatively few quantization bits are needed for the limited-fronthaul case to meet the performance of the system with perfect fronthaul, which reveals that limited-fronthaul cell-free massive MIMO can be considered as a practical system for beyond 5G.
5. We investigate one of the most important and practical issues of cell-free massive MIMO: should we allocate more bits to quantize the estimated channel at the APs or more bits to quantize the received signal? It is interesting that the numerical results show that the same number of bits should be exploited to quantize the estimated channel at the APs or to quantize the received signal to maximize the spectral efficiency and energy efficiency of the system.

Outline: The rest of the paper is organized as follows. Section II describes the system model and Section III reviews the optimal uniform quantization model. Section IV describes the Estimate&Quantize scheme. Sections V presents the decentralized processing scheme whereas the Quantize&Estimate scheme is investigated in Section VI. In addition, Sections VII and VIII investigate the total energy efficiency and proposed AP assignment schemes, respectively. Numerical results are presented in Section IX, and finally Section X concludes the paper.

II. SYSTEM MODEL

We consider the uplink transmission in a cell-free massive MIMO system with M APs and K single-antenna users arbitrarily distributed in a large area, as shown in Fig. 1. Moreover, we assume each AP has N antennas. The channel coefficient vector between the k th user and the m th AP, $\mathbf{g}_{mk} \in \mathbb{C}^{N \times 1}$, is modeled as

$$\mathbf{g}_{mk} = \sqrt{\beta_{mk}} \mathbf{h}_{mk}, \quad (1)$$

where β_{mk} is the large-scale fading coefficient and $\mathbf{h}_{mk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ is a complex Gaussian random vector with covariance matrix \mathbf{I}_N which represents the small-scale fading [28].¹

A. Channel Estimation at the APs

All pilot sequences transmitted by the K users in the channel estimation phase are collected in a matrix $\Phi \in \mathbb{C}^{\tau_p \times K}$, where τ_p is the length of the pilot sequence for each user and the k th column, ϕ_k , represents the pilot sequence used by the k th user, where $\|\phi_k\|^2 = 1$. Let $\sqrt{\tau} \phi_k \in \mathbb{C}^{\tau \times 1}$, where $\|\phi_k\|^2 = 1$, be the pilot sequence assigned to the k th user. The MMSE estimate of the channel coefficient is given by [28]

$$\hat{\mathbf{g}}_{mk} = c_{mk} \left(\sqrt{\tau_p P_p} \mathbf{g}_{mk} + \sqrt{\tau_p P_p} \sum_{k' \neq k}^K \mathbf{g}_{mk'} \phi_k^H \phi_{k'} + \Omega_{p,m} \phi_k \right), \quad (2)$$

where $\Omega_{p,m}$ denotes the noise vector at the m th AP whose elements are independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$, p_p represents the normalized signal-to-noise ratio (SNR) of each pilot symbol, where it is assumed that \bar{p}_p denotes the power of pilot sequence where $p_p = \frac{\bar{p}_p}{p_n}$ and p_n is the noise power [28]. Moreover, we have $c_{mk} = \frac{\sqrt{\tau_p P_p} \beta_{mk}}{\tau_p P_p \sum_{k'=1}^K \beta_{mk'} |\phi_k^H \phi_{k'}|^2 + 1}$ and $\gamma_{mk} = \frac{\tau_p P_p \beta_{mk}^2}{\tau_p P_p \sum_{k'=1}^K \beta_{mk'} |\phi_k^H \phi_{k'}|^2 + 1}$ [28].

B. Uplink Transmission

In this subsection, we consider the uplink data transmission, where all users send their signals to the APs. The transmitted signal from the k th user is denoted by $x_k = \sqrt{\rho} q_k s_k$, where s_k ($\mathbb{E}\{|s_k|^2\} = 1$) and q_k denotes the transmitted symbol and the transmit power, respectively. Moreover, ρ refers to the normalized uplink SNR. The $N \times 1$ received signal at the m th AP is given by

$$\mathbf{y}_m = \sqrt{\rho} \sum_{k=1}^K \mathbf{g}_{mk} \sqrt{q_k} s_k + \mathbf{n}_m, \quad (3)$$

where $\mathbf{n}_m \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ is the noise vector at the m th AP. The baseline of cell-free massive MIMO with perfect fronthaul and linear receivers has been presented in [14], [37]–[39]. Hence, we only present the performance analysis for the limited fronthaul cases.

¹With multiple antennas at the users, if the users do not have channel state information, then a system with K N -antenna users have the same sum uplink spectral efficiency as a system having KN single-antenna users. However, in the case that users have channel state information, the effect of multiple antennas at the user terminals is investigated in [33] and is out of scope of the current paper. A proper and realistic channel model [34]–[36] is necessary. This is left aside for future research.

III. OPTIMAL UNIFORM QUANTIZATION MODEL

In this section, we provide an overview of the optimal uniform quantization. Note that J. Max in [25] developed an algorithm to define the necessary conditions to minimize the distortion of a scalar quantizer [40]. In addition, the Bussgang decomposition [27] is also used in this paper, enabling us to exploit the scheme proposed by P. Zillmann in [26] to model the quantization and hence find the optimum step size of the quantizer by maximizing the SDNR. Note that the Max algorithm and the scheme based on the Bussgang decomposition in [26] result in the same SDNR. The main difference between them is that using the Bussgang decomposition, the output of the quantizer can be represented by a scalar multiple of the input plus an uncorrelated distortion [26], [27] whereas exploiting the Max algorithm, the quantization distortion and the output of the quantizer are uncorrelated [25], [41], [42]. The details of the optimal uniform quantization models are provided in the following subsections.

A. Optimal Uniform Quantization with the Bussgang Decomposition

Based on the Bussgang decomposition [27], the output of a quantizer can be represented by a scalar multiple of the input plus uncorrelated distortion as follows [16], [26], [43]:

$$\mathcal{Q}(z) = h(z) = az + n_d, \quad \forall k, \quad (4)$$

where $\mathcal{Q}(\cdot)$ denotes the quantization operation, and $h(z)$ denotes a memoryless nonlinear function with the Gaussian-distributed input z , a is a constant, n_d refers to the distortion noise which is uncorrelated with the input of the quantizer, z [26]. The term a is given by $a = \frac{\mathbb{E}\{zh(z)\}}{\mathbb{E}\{z^2\}} = \frac{1}{p_z} \int_{\mathcal{Z}} zh(z) f_z(z) dz$, where $p_z = \mathbb{E}\{z^2\}$ denotes the power of z and we drop absolute value as z is a real number, and $f_z(z)$ represents the probability distribution function of z . In general, the terms a is a function of the power of the quantizer input, p_z . Similar to the methodology in [16], to remove this dependency, we normalize the input signal by dividing the input signal, z , by the square root of its power, $\sqrt{p_z}$, and then multiply the quantizer output by its square root, $\sqrt{p_z}$. Hence, by introducing a new variable $\tilde{z} = \frac{z}{\sqrt{p_z}}$, we have

$$\mathcal{Q}(z) = \sqrt{p_z} \mathcal{Q}(\tilde{z}) = \tilde{a} \sqrt{p_z} \tilde{z} + \sqrt{p_z} \tilde{n}_d = \tilde{a} z + \sqrt{p_z} \tilde{n}_d. \quad (5)$$

In other words, \tilde{a} is the constant term in the Bussgang decomposition for the case of normalized input. Finally, the optimum step size of the quantizer and the corresponding \tilde{a} are given in [16, Subsection II-B], which are summarized in Table I.

B. Max Algorithm for Optimal Uniform Quantization

Based on the analysis provided by J. Max in [25], the linear quantization can be modeled as:

$$\mathcal{Q}(z) = h(z) = z + n_d, \quad \forall k. \quad (6)$$

Note that based on the analysis in [42], [44, Chapter 3], the quantization model in (6) is a special case of (5), where $\tilde{a} =$

Table I
THE OPTIMAL STEP SIZE AND DISTORTION POWER OF A UNIFORM
QUANTIZER *with and without the Bussgang decomposition* AND UNIT
VARIANCE INPUT SIGNAL.

α	Δ_{opt}	$\sigma_{\tilde{n}_d}^2 = \tilde{b} - \tilde{a}^2 = \sigma_{\tilde{e},B}^2$	\tilde{a}	$\sigma_{\tilde{n}_d}^2 = \sigma_{\tilde{e}}^2$
1	1.596	0.2313	0.6366	0.3634 [25]
2	0.9957	0.10472	0.88115	0.1188 [25]
3	0.586	0.036037	0.96256	0.03744 [25]
4	0.3352	0.011409	0.98845	0.01154 [25]
5	0.1881	0.003482	0.996505	0.00349 [25]
6	0.1041	0.0010389	0.99896	-
7	0.0568	0.0003042	0.99969	-
8	0.0307	0.0000876	0.999912	-
9	0.0165	0.0000249	0.999975	-
10	0.0088	6.99696×10^{-6}	0.999993	-
11	0.004649	1.94441×10^{-6}	0.999998	-
12	0.0024484	5.35536×10^{-7}	0.999999	-
13	0.001283	1.46369×10^{-7}	0.9999998	-
14	0.001283	3.97394×10^{-8}	0.99999997	-
15	0.000349	1.0727×10^{-8}	1	-

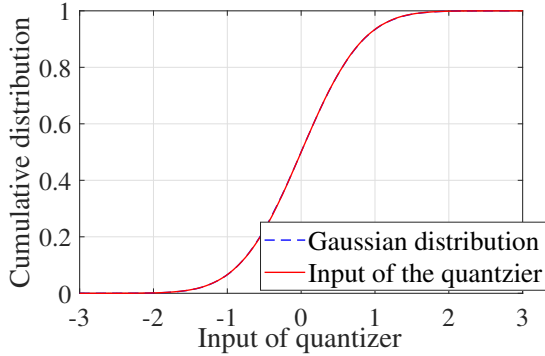


Figure 2. Cumulative distribution of the input of the quantizer with $K = 10$ and $N = 1$.

$1 - \sigma_{\tilde{e}}^2$. Using J. Max algorithm, the output of the quantizer and the distortion are uncorrelated [25], [41], [42]. For this case, to calculate the variance of the quantization error, we exploit the following scheme:

$$\sigma_{\tilde{n}_d}^2 = \begin{cases} \sigma_{\tilde{e}}^2, & \text{obtained in [25], } \alpha \leq 5, \\ \tilde{a}(1 - \tilde{a}), & [45], \alpha \geq 6, \end{cases} \quad (7)$$

where α denotes the number of quantization bits. In the following sections, we investigate the performance of limited-fronthaul cell-free massive MIMO with the aforementioned optimal uniform quantization schemes.

IV. ESTIMATE&QUANTIZE SCHEME

In this section, we present the Estimate&Quantize scheme. With this scheme, the m th AP quantizes the estimated channels, $\hat{\mathbf{g}}_{mk}$, $\forall k$, and the received signal, \mathbf{y}_m , using the optimal uniform quantization, and forwards the quantized channel and the quantized signal to the CPU. In the following subsections, we exploit the Bussgang decomposition [26], [27] and the Max algorithm [25] to quantize the received signal and the estimated channel, respectively. These enable us to

exploit the scheme in [30] to derive the SINR of the limited-fronthaul cell-free massive MIMO system with the considered Estimate&Quantize scheme.

A. Quantization of the Received Signal

Using the Bussgang decomposition [26], [27], the quantized signal can be obtained as:

$$\check{\mathbf{y}}_m = \tilde{a}\mathbf{y}_m + \mathbf{e}_m^y, \quad \forall m, \quad (8)$$

where \mathbf{y}_m is the received signal at the m th AP and is defined in (3). Exploiting the analysis in Section IV, the variance of the distortion noise is given by $\sigma_{[\mathbf{e}_m^y]_n}^2 = \sigma_{[\tilde{\mathbf{e}}_m^y]_n}^2 \mathbb{E}\{|\mathbf{y}_m|_n|^2\}$. Hence, we have

$$\begin{aligned} \sigma_{[\mathbf{e}_m^y]_n}^2 &= \sigma_{[\tilde{\mathbf{e}}_m^y]_n}^2 \left(\rho \sum_{k'=1}^K \beta_{mk'} q_{k'} + 1 \right) \\ &= \sigma_{\tilde{e}^y}^2 \left(\rho \sum_{k'=1}^K \beta_{mk'} q_{k'} + 1 \right), \quad \forall m, n, \end{aligned} \quad (9)$$

where $\sigma_{\tilde{e}^y}^2$ is the variance of the distortion noise with unit variance input for the given number of quantization bits. Note that using \mathbf{y}_m defined in (3), the elements of \mathbf{y}_m are uncorrelated, i.e., $\mathbb{E}\{\mathbf{y}_m \mathbf{y}_m^H\} = \mathbf{0}$. Moreover, in the second equality in (9) we used the same number of bits in all APs and all antennas to quantize the received signal and hence $\sigma_{[\tilde{\mathbf{e}}_m^y]_n}^2 = \sigma_{\tilde{e}^y}^2 = \sigma_{\tilde{e},B}^2, \forall m, n$. Note that subscript B denotes the Bussgang decomposition. The optimal values of $\sigma_{\tilde{e},B}^2$ for different numbers of quantization bits are given in Table I.

Remark 1. *J. Bussgang in his original work [27] assumes that the input of quantizer has Gaussian distribution. Since the input of quantizer is the sum of many random variates, from the central limit theorem, it has near Gaussian distribution. Therefore, we use the Bussgang decomposition, making the approximation that the input of the quantizer is Gaussian distributed. The Gaussian approximation can be verified numerically, as shown in Fig. 2. We can see that the cumulative distribution of the empirical distribution matches very well with that of the Gaussian distribution.*

Remark 2. *Using the Bussgang decomposition, the quantizer input is uncorrelated with the distortion noise. This implies that $\mathbb{E}\{\mathbf{y}_m \mathbf{e}_m^{yH}\} = \mathbf{0}, \forall k$.*

B. Quantization of the Estimated Channel

We quantize the estimated channel with the optimal quantizer obtained using the Max algorithm [25] as follows:

$$\check{\mathbf{g}}_{mk} = \hat{\mathbf{g}}_{mk} + \mathbf{e}_{mk}^g, \quad \forall k. \quad (10)$$

Using the analysis in Section IV, the variance of the quantization error is obtained as $\sigma_{[\mathbf{e}_{mk}^g]_n}^2 = \sigma_{[\tilde{\mathbf{e}}_{mk}^g]_n}^2 \mathbb{E}\{|\hat{\mathbf{g}}_{mk}|_n|^2\}$, which results in

$$\sigma_{[\mathbf{e}_{mk}^g]_n}^2 = \sigma_{[\tilde{\mathbf{e}}_{mk}^g]_n}^2 \gamma_{mk} = \sigma_{\tilde{e}^g}^2 \gamma_{mk}, \quad \forall m, k, n, \quad (11)$$

where for simplicity we use the same number of bits in all APs to quantize the estimated channels.

Remark 3. Note that J. Max investigated the case of an input signal with Gaussian distribution, and here the input signal, i.e., $\check{\mathbf{g}}_{mk}$ has Gaussian distribution.

Remark 4. Based on [25], [41], [42], the quantizer output is assumed to be uncorrelated with the distortion noise. Hence, we have

$$\mathbb{E} \left\{ \check{\mathbf{g}}_{mk} \mathbf{e}_{mk}^g H \right\} = \mathbf{0}, \forall k. \quad (12)$$

Note that the reason why we do not use the Bussgang decomposition to model the quantization of the estimated channel is explained later after the proof of Proposition 5.

Remark 5. If the probability density function of input of the quantizer is even, and exploiting the symmetrical quantizer, the distortion noise has zero mean [25], [41], [42]. Hence, we have:

$$\mathbb{E} \left\{ \mathbf{e}_{mk}^g \right\} = \mathbf{0} \ \& \ \mathbb{E} \left\{ \mathbf{e}_m^y \right\} = \mathbf{0}. \quad (13)$$

C. Data Detection

Let $\check{\mathbf{V}} \in \mathbb{C}^{M \times K}$ be the linear detector matrix depending on the side information at the receiver $\check{\mathbf{g}}_{mk}, \forall m, k$. We let $\check{\mathbf{v}}_k = [\check{\mathbf{v}}_{1k}^T, \dots, \check{\mathbf{v}}_{Mk}^T]^T$ refer to the k th column of the detector matrix $\check{\mathbf{V}}$, and $\check{\mathbf{v}}_{mk} \in \mathbb{C}^N$. The received signal after using the linear detector at the CPU is given by

$$\check{s}_k = \check{\mathbf{v}}_k^H [\check{\mathbf{y}}_1^T, \dots, \check{\mathbf{y}}_M^T], \quad (14)$$

where $\check{\mathbf{y}}_m$ is defined in (8). Next, the received signal for the k th user is re-written as:

$$\begin{aligned} \check{s}_k &= \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{y}}_m = \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H (\tilde{a} \mathbf{y}_m + \mathbf{e}_m^y) \\ &= \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \left(\tilde{a} \sqrt{\rho} \sum_{k=1}^K \check{\mathbf{g}}_{mk} \sqrt{q_k} s_k + \tilde{a} \mathbf{n}_m + \mathbf{e}_m^y \right) \\ &= \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \left(\tilde{a} \sqrt{\rho} \sum_{k=1}^K (\check{\mathbf{g}}_{mk} - \mathbf{e}_{mk}^g + \check{\mathbf{g}}_{mk}) \sqrt{q_k} s_k + \tilde{a} \mathbf{n}_m + \mathbf{e}_m^y \right) \\ &= \underbrace{\tilde{a} \sqrt{\rho} q_k \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk} s_k}_{A_1} + \underbrace{\tilde{a} \sqrt{\rho} \sum_{k' \neq k} \sqrt{q_{k'}} \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk'} s_{k'}}_{A_2} \\ &\quad + \underbrace{\tilde{a} \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \mathbf{n}_m}_{A_3} + \underbrace{\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \mathbf{e}_m^y}_{A_4} - \underbrace{\tilde{a} \sqrt{\rho} \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \mathbf{e}_{mk'}^g s_{k'}}_{A_5} \\ &\quad + \underbrace{\tilde{a} \sqrt{\rho} \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \check{\mathbf{g}}_{mk'} s_{k'}}_{A_6}. \end{aligned} \quad (15)$$

D. Capacity Bound with Quantized Channel and Non-Gaussian Noise

Note that the quantized version of the estimated channel is available as side information at the CPU to decode the signal, where we use the term $\Omega = \check{\mathbf{G}}$ to refer to it. Moreover, the received signal in (15) can be re-written as $\check{s}_k = \tilde{a} A_1 s_k + \tilde{a} A_2 + \tilde{a} A_3 + A_4 - \tilde{a} A_5 + \tilde{a} A_6$, and $W = A_2 + A_3 + A_4 + A_5 + A_6$. Note that here we use the scheme in [31, Section 2.3.5 and Table 2.3] for fading channel with additive non-Gaussian noise W and side information Ω . Let us suppose the following four conditions hold: **1.** Terms $A_1 s_k, A_2, A_3, A_4, A_5$ and A_6 are mutually uncorrelated, **2.** $\mathbb{E} \{W|\Omega\} \approx 0$, **3.** $\mathbb{E} \{s_k^* W|\Omega\} \approx 0$, and **4.** $\mathbb{E} \{A_1^* s_k^* W|\Omega\} \approx 0$. Then the closed-form expression for the achievable rate of the k th user is given by $R_k \approx \mathbb{E} \{\log_2(1 + \text{SINR}_k)\}$, where the SINR_k is the achievable SINR of the k th user and the expectation is taken over $\check{\mathbf{G}}$. Using the analysis in [31], the achievable SINR is obtained by (16) (defined at the bottom of this page). In the following lemmas, we prove the four aforementioned conditions.

Lemma 1. Terms $A_1 s_k, A_2, A_3, A_4, A_5$ and A_6 are mutually uncorrelated.

Proof: Please refer to Appendix A. ■

Lemma 2. We have $\mathbb{E} \{W|\Omega\} \approx 0$. Note that the approximation term in $\mathbb{E} \{W|\Omega\} \approx 0$ means that the term $\mathbb{E} \{W|\Omega\}$ is small enough and can be neglected.

Proof: Please refer to Appendix B. ■

Lemma 3. We have $\mathbb{E} \{s_k^* W|\Omega\} \approx 0$, where approximation sign indicates that the term $\mathbb{E} \{s_k^* W|\Omega\}$ is small enough and can be neglected.

Proof: Please refer to Appendix C. ■

Lemma 4. We have $\mathbb{E} \{A_1^* s_k^* W|\Omega\} \approx 0$.

Proof: Please refer to Appendix D. ■

Next, the achievable SINR of the k th user is obtained in the following theorem.

Theorem 1. The SINR of the k th user in cell-free massive MIMO for the case when APs send back the quantized versions of the estimated channels and the quantized version of the received signals to the CPU through limited fronthaul links is given by (17) (defined at the top of the next page), where in $\text{SINR}_k^{\text{E\&Q,lf}}$, the superscript “E&Q,lf” stands for Estimate&Quantize and limited fronthaul.

Proof: Using Lemma 1 and the analysis in [31], the achievable SINR is obtained by (16). Next, using (15) and (11), and after some mathematical manipulation, we have (18), where $\mathbf{W}_{k'}^{\text{E\&Q,lf}}$ and $\mathbf{F}^{\text{E\&Q,lf}}$ are defined in (20). It is straightforward to calculate the terms $\mathbb{E} \{|A_1 s_k| \check{\mathbf{G}}\}^2$, $\mathbb{E} \{|A_2| \check{\mathbf{G}}\}^2$, and

$$\text{SINR}_k^{\text{E\&Q,lf}} = \frac{|\mathbb{E} \{A_1| \check{\mathbf{G}}\}|^2}{\mathbb{E} \{|A_2|^2| \check{\mathbf{G}}\} + \mathbb{E} \{|A_3|^2| \check{\mathbf{G}}\} + \frac{1}{\tilde{a}^2} \mathbb{E} \{|A_4|^2| \check{\mathbf{G}}\} + \mathbb{E} \{|A_5|^2| \check{\mathbf{G}}\} + \mathbb{E} \{|A_6|^2| \check{\mathbf{G}}\}}. \quad (16)$$

$$\text{SINR}_k^{\text{E\&Q,If}} = \frac{\rho q_k \left| \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk} \right|^2}{\rho \sum_{k' \neq k}^K q_{k'} \left| \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk'} \right|^2 + \rho \sum_{m=1}^M \|\check{\mathbf{v}}_{mk}\|^2 \sum_{k'=1}^K q_{k'} \left[\beta_{mk'} \left(1 + \frac{\sigma_{\check{e}^y, \text{B}}^2}{\tilde{a}^2} \right) - \gamma_{mk'} (1 - \sigma_{\check{e}^g}^2) \right] + \left(1 + \frac{\sigma_{\check{e}^y, \text{B}}^2}{\tilde{a}^2} \right) \sum_{m=1}^M \|\check{\mathbf{v}}_{mk}\|^2} \quad (17)$$

$$\begin{aligned} \frac{1}{\tilde{a}^2} \mathbb{E} \left\{ |A_4| \check{\mathbf{G}} \right\}^2 + \mathbb{E} \left\{ |A_5| \check{\mathbf{G}} \right\}^2 + \mathbb{E} \left\{ |A_6| \check{\mathbf{G}} \right\}^2 &= \sum_{m=1}^M \|\check{\mathbf{v}}_{mk}\|^2 \sum_{k'=1}^K \rho q_{k'} \left[\beta_{mk'} \left(1 + \frac{\sigma_{\check{e}^y, \text{B}}^2}{\tilde{a}^2} \right) - \gamma_{mk'} (1 - \sigma_{\check{e}^g}^2) \right] \\ &+ \sum_{m=1}^M \|\check{\mathbf{v}}_{mk}\|^2 \frac{\sigma_{\check{e}}^2}{\tilde{a}^2} = \check{\mathbf{v}}_k^H \left(\rho \sum_{k'=1}^K q_{k'} \mathbf{W}_{k'}^{\text{E\&Q,If}} + \mathbf{F}^{\text{E\&Q,If}} \right) \check{\mathbf{v}}_k. \end{aligned} \quad (18)$$

$\mathbb{E} \left\{ |A_3| \check{\mathbf{G}} \right\}^2$. Finally, by substituting (18) into (16), we obtain (17), which completes the proof. ■

In this case, we will consider the following linear detectors.

$$\check{\mathbf{V}} = \begin{cases} \check{\mathbf{G}}, & \text{for MRC} \\ \check{\mathbf{G}} \left(\check{\mathbf{G}}^H \check{\mathbf{G}} \right)^{-1}, & \text{for ZF} \\ \left(\tilde{a}^2 \rho \sum_{k'=1}^K q_{k'} \check{\mathbf{g}}_{k'} \check{\mathbf{g}}_{k'}^H + \mathbf{R}^{\text{E\&Q,If}} \right)^{-1} \check{\mathbf{G}}, & \text{for MMSE} \end{cases} \quad (19)$$

where $\check{\mathbf{G}} = [\check{\mathbf{g}}_1, \dots, \check{\mathbf{g}}_K]$, and $\mathbf{R}^{\text{E\&Q,If}}$ is obtained as follows:

$$\mathbf{R}^{\text{E\&Q,If}} = \rho \sum_{k'=1}^K q_{k'} \mathbf{W}_{k'}^{\text{E\&Q,If}} + \mathbf{I}_{MN} + \mathbf{F}^{\text{E\&Q,If}}, \quad (20a)$$

$$\mathbf{W}_{k'}^{\text{E\&Q,If}} = \mathbf{S}_{k'}^{\text{E\&Q,If}} - \mathbf{T}_{k'}^{\text{E\&Q,If}} \quad (20b)$$

$$\mathbf{S}_{k'}^{\text{E\&Q,If}} = \left(\frac{\sigma_{\check{e}^y, \text{B}}^2}{\tilde{a}^2} + 1 \right) \text{diag} \left(\text{rep}(\beta_{1k'}, N) \cdots \text{rep}(\beta_{Mk'}, N) \right), \quad (20c)$$

$$\mathbf{T}_{k'}^{\text{E\&Q,If}} = \left(1 - \sigma_{\check{e}^g}^2 \right) \text{diag} \left(\text{rep}(\gamma_{1k'}, N) \cdots \text{rep}(\gamma_{Mk'}, N) \right), \quad (20d)$$

$$\mathbf{F}^{\text{E\&Q,If}} = \frac{\sigma_{\check{e}^y, \text{B}}^2}{\tilde{a}^2} \mathbf{I}_{MN}. \quad (20e)$$

Note that the MMSE detector is optimal as it maximizes the instantaneous SINR of each user.

V. DECENTRALIZED PROCESSING SCHEME

In this section, we investigate MRC, which due to its low complexity and can be designed in a distributed fashion where most processing is locally performed at the APs [24]. First, the received signal from the k th user at the m th AP, i.e., \mathbf{y}_m , is multiplied by the Hermitian of the local channel estimate $\hat{\mathbf{g}}_{mk}^H$. Following the terminology in [16], the m th AP quantizes the term $z_{m,k} = \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m, \forall k$. Next, the AP forwards the quantized signals, i.e., $z_{m,k}$, in each symbol duration to the CPU. Similar to [42], [46], a low resolution analog-to-digital converter (ADC) quantizes the real and imaginary parts of $z_{m,k}$ with α bits each, which introduces quantization errors to the received signals [42], [44], [45], [47]. Exploiting the Bussgang decomposition, the relation between z_{mk} and its quantized version, \tilde{z}_{mk} , is given by

$$\tilde{z}_{mk} = \tilde{a} z_{mk} + e_{mk}^z, \quad (21)$$

where the term e_{mk}^z is the distortion noise of the m th AP. Note that the input of quantizer, i.e., $z_{m,k} = \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m, \forall k$, is the sum of many random variates, from the central limit theorem, it has a near Gaussian distribution. This allows us to exploit Bussgang theorem [48, Section II-D]. The aggregated received signal at the CPU can be written as (22) (defined at the top of the next page), where DS_k and BU_k are the desired signal (DS) and beamforming uncertainty (BU) for the k th user, respectively, and IUI_k is the inter-user-interference (IUI) caused by the k' th user. Moreover, TN_k accounts for the total noise (TN) following the MRC detection, and finally TQE_k is to the total quantization error (TQE) at the k th user. Using the fact that the terms DS_k , BU_k , $\text{IUI}_{kk'}$, TN_k , and TQE_k are mutually uncorrelated, the achievable uplink rate of the k th user with the decentralized processing is given by (23) (defined at the top of the next page), where the index dec refers to the decentralized scheme [16, Section III].

VI. QUANTIZE&ESTIMATE SCHEME

In this section, we provide the Quantize&Estimate scheme. With this scheme, the APs quantize the received signal and received pilot, using the optimal uniform quantization, and forwards the quantized versions of the received signal and received pilot to the CPU. In the following subsections, we exploit the Bussgang decomposition [26], [27] to quantize the received signal and the received pilot.

A. Quantization of Received Pilot

The users simultaneously transmit their pilot sequences, and similar to [24] we assume the pilot sequence for the k th user is $\sqrt{\tau_p} \boldsymbol{\phi}_k \in \mathbb{C}^{\tau_p \times 1}$, where $\|\boldsymbol{\phi}_k\|^2 = 1$. Hence, the received signal at the m th AP is given by [24]

$$\mathbf{Y}_m^p = \sqrt{\tau_p p_p} \sum_{k=1}^K \mathbf{g}_{mk} \boldsymbol{\phi}_k^H + \mathbf{W}_m^p, \quad (24)$$

where matrix $\mathbf{W}_m^p \in \mathbb{C}^{N \times \tau_p}$ is the noise whose elements are i.i.d. $\mathcal{CN}(0, 1)$. Next each AP quantizes the received signal and sends the quantized versions of the received signals to the

$$\begin{aligned}
r_k &= \sum_{m=1}^M \left(\underbrace{\tilde{a} \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m + e_{mk}^z}_{z_{mk}} \right) = \underbrace{\tilde{a} \sqrt{\rho} \mathbb{E} \left\{ \sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\}}_{\text{DS}_k} s_k + \underbrace{\tilde{a} \sqrt{\rho} \left(\sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} - \mathbb{E} \left\{ \sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\} \right)}_{\text{BU}_k} s_k \\
&+ \underbrace{\sum_{k' \neq k}^K \tilde{a} \sqrt{\rho} \sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'}}}_{\text{IUI}_{kk'}} s_{k'} + \underbrace{\tilde{a} \sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m}_{\text{TN}_k} + \underbrace{\sum_{m=1}^M e_{mk}^z}_{\text{TQE}_k}. \tag{22}
\end{aligned}$$

$$\begin{aligned}
&\text{SINR}_k^{\text{dec}} \tag{23} \\
&= \frac{N^2 q_k \left(\sum_{m=1}^M \gamma_{mk} \right)^2}{N^2 \sum_{k' \neq k}^K q_{k'} \left(\sum_{m=1}^M \frac{\gamma_{mk} \beta_{mk'}}{\beta_{mk}} \right)^2 |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 + \frac{N^2 \sigma_{\tilde{e},B}^2}{\tilde{a}^2} \sum_{k'=1}^K q_{k'} \sum_{m=1}^M \gamma_{mk'}^2 |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 + N \left(\frac{\sigma_{\tilde{e},B}^2}{\tilde{a}^2} + 1 \right) \sum_{m=1}^M \gamma_{mk} \sum_{k'=1}^K q_{k'} \beta_{mk'} + \frac{N}{\rho} \left(\frac{\sigma_{\tilde{e},B}^2}{\tilde{a}^2} + 1 \right) \sum_{m=1}^M \gamma_{mk}}.
\end{aligned}$$

CPU. Using the Bussgang theorem, the quantized version of the received pilot signal at the m th AP is given by

$$\begin{aligned}
\check{\mathbf{Y}}_m^p &= \tilde{a} \mathbf{Y}_m^p + \mathbf{E}_m^{(y^p)} \\
&= \tilde{a} \sqrt{\tau_p p_p} \sum_{k=1}^K \mathbf{g}_{mk} \boldsymbol{\phi}_k^H + \tilde{a} \mathbf{W}_m^p + \mathbf{E}_m^{(y^p)}, \tag{25}
\end{aligned}$$

where $\mathbf{E}_m^{(y^p)}$ is the quantization distortion.

Remark 6. Note that \mathbf{Y}_m^p has Gaussian distribution as it is sum of Gaussian variables, i.e., $\sqrt{\tau_p p_p} \mathbf{g}_{m1} \boldsymbol{\phi}_1^H, \dots, \sqrt{\tau_p p_p} \mathbf{g}_{mK} \boldsymbol{\phi}_K^H, \mathbf{W}_m^p$. This enables us to exploit Bussgang's approach in [27]. In this case, the inputs are also independent which implies that the outputs will be independent (based on the analysis in [44, Section 2.2]).

B. Channel Estimation at the CPU

Next, the CPU exploits the pilot sequence $\boldsymbol{\phi}_k$ to correlate the received quantized signal from the m th AP with the pilot sequence as follows:

$$\begin{aligned}
\check{\mathbf{y}}_{m,k}^p &= \check{\mathbf{Y}}_m^p \boldsymbol{\phi}_k \\
&= \tilde{a} \sqrt{\tau_p p_p} \mathbf{g}_{mk} + \tilde{a} \sqrt{\tau_p p_p} \sum_{k' \neq k}^K \mathbf{g}_{mk'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + \tilde{a} \mathbf{w}_{mk}^p + \mathbf{e}_{mk}^{(y^p)}, \tag{26}
\end{aligned}$$

where $\mathbf{w}_{mk} \triangleq \mathbf{W}_m^p \boldsymbol{\phi}_k$, and $\mathbf{e}_{mk}^{(y^p)} \triangleq \mathbf{E}_m^{(y^p)} \boldsymbol{\phi}_k$.

Lemma 5. Linear minimum mean-square error (LMMSE) estimate of \mathbf{g}_{mk} given $\check{\mathbf{y}}_{m,k}^p$ is:

$$\check{\mathbf{g}}_{mk} = \frac{\mathbb{E} \left\{ \mathbf{g}_{mk}^H \check{\mathbf{y}}_{m,k}^p \right\}}{\mathbb{E} \left\{ \check{\mathbf{y}}_{m,k}^p H \check{\mathbf{y}}_{m,k}^p \right\}} \check{\mathbf{y}}_{m,k}^p = c_{mk} \check{\mathbf{y}}_{m,k}^p, \tag{27}$$

where $c_{mk} = \frac{\tilde{a} \sqrt{\tau_p p_p} \beta_{mk}}{\tilde{a}^2 \left(\tau_p p_p \sum_{k'=1}^K \beta_{mk'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 + 1 \right) + \sigma_{\tilde{e}}^2 \left(p_p \sum_{k=1}^K \beta_{mk} + 1 \right)}$.

Proof: It is easy to show that $\mathbb{E} \left\{ \mathbf{g}_{mk}^H \check{\mathbf{y}}_{m,k}^p \right\} = N \tilde{a} \sqrt{\tau_p p_p} \beta_{mk}$. In addition, using (26), the denominator of c_{mk} is obtained as follows:

$$\begin{aligned}
&\mathbb{E} \left\{ \check{\mathbf{y}}_{m,k}^p H \check{\mathbf{y}}_{m,k}^p \right\} \\
&= N \tilde{a}^2 \left(\tau_p p_p \sum_{k'=1}^K \beta_{mk'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 + 1 \right) + N \mathbb{E} \left\{ \left| \left[\mathbf{E}_m^{(y^p)} \right]_{n,t} \right|^2 \right\}, \tag{28}
\end{aligned}$$

where again $\check{\mathbf{Y}}_m^p$ is the input of the quantizer at the m th AP, given in (25), and $[\cdot]_{n,t}$ refers to the (n, t) -th element, and $n = 1, \dots, N$, and $t = 1, \dots, \tau_p$. Hence, the variance of the quantization error can be obtained as follows:

$$\begin{aligned}
&\mathbb{E} \left\{ \left| \left[\mathbf{E}_m^{(y^p)} \right]_{n,t} \right|^2 \right\} = \sigma_{\tilde{e}}^2 \mathbb{E} \left\{ \left| \left[\mathbf{Y}_m^p \right]_{n,t} \right|^2 \right\} \\
&= \sigma_{\tilde{e}}^2 \left(p_p \sum_{k'=1}^K \beta_{mk} + 1 \right), \tag{29}
\end{aligned}$$

which completes the proof.

The mean-square of the n th element of $\check{\mathbf{g}}_{mk}$ is denoted by γ_{mk} , and given by

$$\begin{aligned}
\gamma_{mk} &\triangleq \mathbb{E} \left\{ \left| \left[\check{\mathbf{g}}_{mk} \right]_n \right|^2 \right\} \\
&= \frac{\tilde{a}^2 \tau_p p_p \beta_{mk}^2}{\tilde{a}^2 \left(\tau_p p_p \sum_{k'=1}^K \beta_{mk'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 + 1 \right) + \sigma_{\tilde{e}}^2 \left(p_p \sum_{k=1}^K \beta_{mk} + 1 \right)}. \tag{30}
\end{aligned}$$

Note that the channel estimation error $\tilde{\mathbf{g}}_{mk}$ is given by

$$\tilde{\mathbf{g}}_{mk} = \mathbf{g}_{mk} - \check{\mathbf{g}}_{mk}, \tag{31}$$

where the elements of $\tilde{\mathbf{g}}_{mk} \sim \mathcal{CN}(0, \beta_{mk} - \gamma_{mk})$ are i.i.d. random variables and independent of $\check{\mathbf{g}}_{mk}$. This is obtained from the fact that the term in (25) is the sum of many random variates, from the central limit theorem, it has near Gaussian distribution [43]. Therefore, the MMSE estimation has Gaussian distribution. This fact is also observed in Fig. 3 with $M = 40$, $N = 1$, $K = 20$ and $\alpha_{m,2} = 1$ bit.

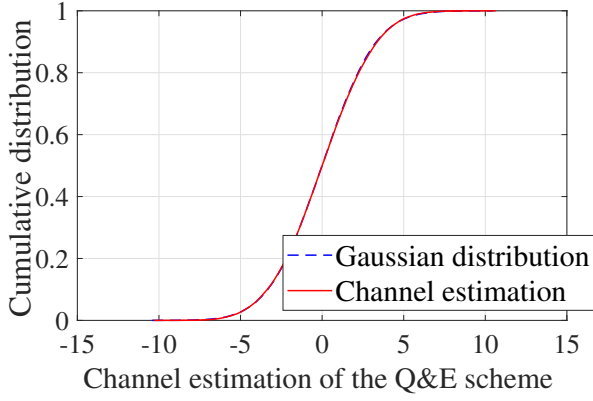


Figure 3. Cumulative distribution of the channel estimation with $M = 40$, $N = 1$, $K = 20$ and $\alpha_{m,2} = 1$ bit.

C. Quantization of Received Signal

The received signal at the m th AP from all users, i.e., \mathbf{y}_m , is given by (3). Exploiting the Bussgang decomposition, the quantized version of \mathbf{y}_m is

$$\check{\mathbf{y}}_m = \tilde{\alpha} \mathbf{y}_m + \mathbf{e}_m^y, \quad \forall m, \quad (32)$$

where the quantization noise is uncorrelated with the input of the quantizer, i.e., $\mathbb{E} \{ \mathbf{y}_m \mathbf{e}_m^{yH} \} = \mathbf{0}$, $\forall m$. Moreover, note that

$$\begin{aligned} \sigma_{[\mathbf{e}_m^y]_n}^2 &= \sigma_{[\tilde{\mathbf{e}}_m^y]_n}^2 \left(\rho \sum_{k'=1}^K \beta_{mk'} q_{k'} + 1 \right) \\ &= \sigma_{\tilde{\mathbf{e}}_m^y}^2 \left(\rho \sum_{k'=1}^K \beta_{mk'} q_{k'} + 1 \right), \end{aligned} \quad (33)$$

where in second equality of (33), we assumed the same number of bits at all APs, i.e., $\sigma_{[\tilde{\mathbf{e}}_m^y]_n}^2 = \sigma_{\tilde{\mathbf{e}}_m^y}^2, \forall m$, for simplicity. Note that in Remark 1, we prove that \mathbf{y}_m has Gaussian distribution, which enables us to use J. Bussgang's approach in [27].

D. Data Detection

Let $\check{\mathbf{V}} \in \mathbb{C}^{MN \times K}$ be the linear detector matrix, which depends on the side information at the receiver, $\check{\mathbf{g}}_{mk}, \forall m, k$. We assume $\check{\mathbf{v}}_k = [\check{\mathbf{v}}_{1k}^T, \dots, \check{\mathbf{v}}_{Mk}^T]^T$ refers to the k th column of the detector matrix $\check{\mathbf{V}}$, and $\check{\mathbf{v}}_{mk} \in \mathbb{C}^N$. The received signal for the k th user after using the detector at the CPU is given by

$$\check{s}_k = \check{\mathbf{v}}_k^H [\check{\mathbf{y}}_1^T, \dots, \check{\mathbf{y}}_M^T], \quad (34)$$

where $\check{\mathbf{y}}_m$ is defined in (32). Next, \check{s}_k is re-written as (35) (defined at the top of the next page), where $\check{\mathbf{v}}_k = [\check{\mathbf{v}}_{1k}^T, \dots, \check{\mathbf{v}}_{Mk}^T]^T$, and refer to the k th column of a $MN \times K$ detector matrix $\check{\mathbf{V}}$, and $\check{\mathbf{V}} = [\check{\mathbf{v}}_1, \dots, \check{\mathbf{v}}_K]$.

E. Capacity Bound with Quantized Channel and Non-Gaussian Noise

The CPU exploits the estimated channel as side information to detect the signal, where here we refer to the side

information at the CPU as $\mathcal{U} = \check{\mathbf{G}}$, and the received signal in (33) is re-written as $\check{s}_k = \tilde{\alpha} B_1 s_k + \tilde{\alpha} B_2 + \tilde{\alpha} B_3 + \tilde{\alpha} B_4 + B_5$, where $\Xi = B_2 + B_3 + B_4 + B_5$ is the additive non-Gaussian noise. Based on [31, Section 2.3.5 and Table 2.3], if the following four conditions hold: **1.** Terms $B_1 s_k$, B_2 , B_3 , B_4 , and B_5 are mutually uncorrelated, **2.** $\mathbb{E} \{ \Xi | \mathcal{U} \} \approx 0$, **3.** $\mathbb{E} \{ s_k^* \Xi | \mathcal{U} \} \approx 0$, and **4.** $\mathbb{E} \{ B_1^* s_k^* \Xi | \mathcal{U} \} \approx 0$. The closed-form expression for the achievable rate of the k th user is given by $R_k \approx \mathbb{E} \{ \log_2(1 + \text{SINR}_k) \}$, where the SINR_k is the achievable SINR of the k th user and the expectation is taken over $\check{\mathbf{G}}$. The achievable SINR is obtained by

$$\begin{aligned} \text{SINR}_k &= \frac{\mathbb{E} \{ |B_1 | \check{\mathbf{G}} \|^2 \}}{\mathbb{E} \{ |B_2|^2 | \check{\mathbf{G}} \} + \mathbb{E} \{ |B_3|^2 | \check{\mathbf{G}} \} + \mathbb{E} \{ |B_4|^2 | \check{\mathbf{G}} \} + \frac{1}{\tilde{\alpha}^2} \mathbb{E} \{ |B_5|^2 | \check{\mathbf{G}} \}}. \end{aligned} \quad (36)$$

Therefore, in the following lemmas, we show that the four above-mentioned required conditions hold.

Lemma 6. *Terms B_1 , B_2 , B_3 , B_4 , and B_5 are mutually uncorrelated.*

Proof: First, we use \check{s}_k given in (35)

$$\begin{aligned} \mathbf{y}_m &= \sqrt{\rho} \sum_{k=1}^K \mathbf{g}_{mk} \sqrt{q_k} s_k + \mathbf{n}_m \\ &= \sqrt{\rho} \sum_{k=1}^K \left(\check{\mathbf{g}}_{mk} + \tilde{\mathbf{g}}_{mk} \right) \sqrt{q_k} s_k + \mathbf{n}_m. \end{aligned} \quad (37)$$

As \mathbf{n}_m and $\tilde{\mathbf{g}}_{mk}$ are uncorrelated i.i.d. Gaussian noise and i.i.d. Gaussian MMSE error, respectively. Moreover, based on the analysis in Section III, using the Bussgang theorem, the quantization noise is uncorrelated with the input of the quantizer. Hence, the terms the terms B_1 and B_5 and the terms B_2 and B_5 are uncorrelated. Note that terms B_3 and B_4 include i.i.d. Gaussian noise and i.i.d. Gaussian MMSE error, respectively. Hence terms B_2 and B_6 are uncorrelated with other terms. These reveal that the terms B_1 , B_2 , B_3 , B_4 , and B_5 are mutually uncorrelated, which completes the proof of Lemma 6. ■

Lemma 7. *We have $\mathbb{E} \{ \Xi | \mathcal{U} \} \approx 0$. Note that the approximation term in $\mathbb{E} \{ \Xi | \mathcal{U} \} \approx 0$ means that the term $\mathbb{E} \{ \Xi | \mathcal{U} \}$ is small enough and can be neglected.*

Proof: Please refer to Appendix E. ■

Lemma 8. *We have $\mathbb{E} \{ s_k^* \Xi | \mathcal{U} \} \approx 0$.*

Proof: Please refer to Appendix F. ■

Lemma 9. *We have $\mathbb{E} \{ B_1^* s_k^* \Xi | \mathcal{U} \} \approx 0$.*

Proof: Please refer to Appendix G. ■

Finally, the achievable SINR of the k th user is obtained in the following theorem.

Theorem 2. *The SINR of the k th user in cell-free massive MIMO for the case when APs send back the quantized version of received pilots and the quantized version of the received*

$$\begin{aligned}
\check{s}_k &= \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{y}}_m = \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \left(\tilde{a} \sqrt{\rho} \sum_{k=1}^K \mathbf{g}_{mk} \sqrt{q_k} s_k + \tilde{a} \mathbf{n}_m + \mathbf{e}_m^y \right) = \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \left(\tilde{a} \sqrt{\rho} \sum_{k=1}^K \left(\check{\mathbf{g}}_{mk} + \tilde{\mathbf{g}}_{mk} \right) \sqrt{q_k} s_k + \tilde{a} \mathbf{n}_m + \mathbf{e}_m^y \right) \\
&= \underbrace{\tilde{a} \sqrt{q_k \rho} \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk} s_k}_{B_1} + \underbrace{\tilde{a} \sqrt{\rho} \sum_{k' \neq k} \sqrt{q_{k'}} \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk'} s_{k'}}_{B_2} + \underbrace{\tilde{a} \sqrt{\rho} \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \tilde{\mathbf{g}}_{mk'} s_{k'}}_{B_3} + \underbrace{\tilde{a} \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \mathbf{n}_m}_{B_4} + \underbrace{\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \mathbf{e}_m^y}_{B_5}. \quad (35)
\end{aligned}$$

$$\text{SINR}_k^{\text{Q\&E,lf}} = \frac{\rho q_k \left| \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk} \right|^2}{\rho \sum_{k' \neq k} q_{k'} \left| \sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk'} \right|^2 + \rho \sum_{m=1}^M \|\check{\mathbf{v}}_{mk}\|^2 \sum_{k'=1}^K q_{k'} \left(\left(1 + \frac{\sigma_{\tilde{e},B}^2}{\tilde{a}^2} \right) \beta_{mk'} - \gamma_{mk'} \right) + \left(1 + \frac{\sigma_{\tilde{e},B}^2}{\tilde{a}^2} \right) \sum_{m=1}^M \|\check{\mathbf{v}}_{mk}\|^2}. \quad (38)$$

$$\mathbb{E} \{ [\mathbf{Y}_m^p]_{nt}^* [\mathbf{Y}_m^p]_{n't'} \} = 0, \forall nt \neq n't', \quad \mathbb{E} \{ [\mathbf{Y}_m^p]_{nt}^* [\mathbf{y}_m]_{n'} \} = 0, \forall n \neq n', \quad \mathbb{E} \{ [\mathbf{Y}_m^p]_{nt}^* [\mathbf{y}_l]_{n'} \} = 0, \forall l \neq m, \quad (39a)$$

$$\mathbb{E} \{ \mathbf{y}_m \mathbf{y}_l^H \} = \mathbf{0}, \forall l \neq m, \quad \mathbb{E} \{ [\mathbf{y}_m]_n^* [\mathbf{y}_m]_{n'} \} = 0, \forall n \neq n'. \quad (39b)$$

$$\mathbb{E} \{ [\mathbf{Y}_m^p]_{nt}^* [\mathbf{Y}_m^p]_{n't'} \} = \mathbb{E} \left\{ \left(\sqrt{\tau_p p} \sum_{k=1}^K [\mathbf{g}_{mk} \boldsymbol{\phi}_k^H]_{nt} + [\mathbf{W}_m^p]_{nt} \right)^* \left(\sqrt{\tau_p p} \sum_{k=1}^K [\mathbf{g}_{mk} \boldsymbol{\phi}_k^H]_{n't'} + [\mathbf{W}_m^p]_{n't'} \right) \right\} = 0, \forall nt \neq n't', \quad (40a)$$

$$\mathbb{E} \{ [\mathbf{Y}_m^p]_{nt}^* [\mathbf{y}_m]_{n'} \} = \mathbb{E} \left\{ \left(\sqrt{\tau_p p} \sum_{k=1}^K [\mathbf{g}_{mk} \boldsymbol{\phi}_k^H]_{nt} + [\mathbf{W}_m^p]_{nt} \right)^* \left(\sqrt{\rho} \sum_{k=1}^K \sqrt{q_k} [\mathbf{g}_{mk}]_{n'} s_k + [\mathbf{n}_m]_{n'} \right) \right\} = 0, \forall n \neq n', \quad (40b)$$

$$\mathbb{E} \{ [\mathbf{Y}_m^p]_{nt}^* [\mathbf{y}_m]_{n'} \} = \mathbb{E} \left\{ \left(\sqrt{\tau_p p} \sum_{k=1}^K [\mathbf{g}_{mk} \boldsymbol{\phi}_k^H]_{nt} + [\mathbf{W}_m^p]_{nt} \right)^* \left(\sqrt{\rho} \sum_{k=1}^K \sqrt{q_k} [\mathbf{g}_{mk}]_{n'} s_k + [\mathbf{n}_m]_{n'} \right) \right\} = 0, \forall n \neq n', \quad (40c)$$

$$\mathbb{E} \{ [\mathbf{Y}_m^p]_{nt}^* [\mathbf{y}_l]_{n'} \} = \mathbb{E} \left\{ \left(\sqrt{\tau_p p} \sum_{k=1}^K [\mathbf{g}_{mk} \boldsymbol{\phi}_k^H]_{nt} + [\mathbf{W}_m^p]_{nt} \right)^* \left(\sqrt{\rho} \sum_{k=1}^K \sqrt{q_k} [\mathbf{g}_{lk}]_{n'} s_k + [\mathbf{n}_l]_{n'} \right) \right\} = 0, \forall m \neq l. \quad (40d)$$

signals to the CPU through the limited fronthaul links (and hence CPU estimates the channel) is given by (38) (defined at the top of this page), where in $\text{SINR}_k^{\text{Q\&E,lf}}$, the superscript ‘‘Q&E, lf’’ refers to Quantize&Estimate and limited fronthaul.

Proof: First, we present a proposition which helps us to prove Theorem 1. Note that the proof for ZF and MMSE follows the same steps as in the MRC case.

Proposition 1. *The input signals of the quantizers are uncorrelated.*

Proof: As explained in Subsections VI-A and VI-C, we need to quantize the received signals and the received pilots in the Quantize&Estimate scheme. To prove that the inputs of quantizers are uncorrelated, we need to show that the equalities in (39) hold. Note that using (3) and (24), we have the equalities in (40). Moreover, we have $\mathbb{E} \{ \mathbf{y}_m \mathbf{y}_l^H \} = \mathbf{0}$, $\forall l \neq m$ and $\mathbb{E} \{ [\mathbf{y}_m]_n^* [\mathbf{y}_m]_{n'} \} = 0$, $\forall n \neq n'$. This completes the proof of Proposition 1.

By following the proof of Theorem 1 and using Proposition 1, we can arrive at the desired result as in (38). ■

Note that the linear receiver is given by

$$\check{\mathbf{v}} = \begin{cases} \check{\mathbf{G}}, & \text{for MRC} \\ \check{\mathbf{G}} \left(\check{\mathbf{G}}^H \check{\mathbf{G}} \right)^{-1}, & \text{for ZF} \\ \left(\tilde{a}^2 \rho \sum_{k'=1}^K q_{k'} \check{\mathbf{g}}_{k'} \check{\mathbf{g}}_{k'}^H + \mathbf{R}^{\text{Q\&E,lf}} \right)^{-1} \check{\mathbf{G}}, & \text{for MMSE} \end{cases} \quad (41)$$

where $\check{\mathbf{G}} = [\check{\mathbf{g}}_1, \dots, \check{\mathbf{g}}_K]$, and $\mathbf{R}^{\text{Q\&E,lf}}$ is obtained as follows:

$$\mathbf{R}^{\text{Q\&E,lf}} = \rho \sum_{k'=1}^K q_{k'} \mathbf{W}_{k'}^{\text{Q\&E,lf}} + \mathbf{I}_{MN} + \mathbf{F}^{\text{Q\&E,lf}}, \quad (42a)$$

$$\mathbf{W}_{k'}^{\text{Q\&E,lf}} = \mathbf{S}_{k'}^{\text{E\&Q,lf}} - \mathbf{T}_{k'}^{\text{Q\&E,lf}} \quad (42b)$$

$$\mathbf{S}_{k'}^{\text{Q\&E,lf}} = \left(1 + \frac{\sigma_{\tilde{e},B}^2}{\tilde{a}^2} \right) \text{diag} \left(\text{rep}(\beta_{1k'}, N) \cdots \text{rep}(\beta_{Mk'}, N) \right), \quad (42c)$$

$$\mathbf{T}_{k'}^{\text{Q\&E,lf}} = \text{diag} \left(\text{rep}(\gamma_{1k'}, N) \cdots \text{rep}(\gamma_{Mk'}, N) \right), \quad (42d)$$

$$\mathbf{F}^{\text{Q\&E,lf}} = \frac{\sigma_{\tilde{e},B}^2}{\tilde{a}^2} \mathbf{I}_{MN}. \quad (42e)$$

VII. ENERGY EFFICIENCY

In this section, we investigate the energy efficiency performance of the limited-fronthaul cell-free massive MIMO

system, which is obtained by dividing the system throughput by the consumed power. In the following subsections, the required fronthaul capacities of Estimate&Quantize and Quantize&Estimate cases are investigated. Next, the power consumption model and energy efficiency analysis are presented.

A. Fronthaul Bit Rate

Let us assume the length of frame (which represents the length of the uplink data) is $\tau_f = \tau_c - \tau_p$, where τ_c denotes the number of samples for each coherence interval. Defining the number of the quantization bits as $\alpha_{m,i}$, for $i = 1, 2, 3$, corresponding to Estimate&Quantize, Quantize&Estimate and decentralized scheme, for $i = 1$, $i = 2$ and $i = 3$, respectively, where the index m denotes the m th AP. For Estimate&Quantize, the required number of bits for each AP to quantize the estimated channel and the uplink data during each coherence interval is $2\alpha_{m,1} \times (NK + N\tau_f)$ whereas Quantize&Estimate requires $2\alpha_{m,2} \times (N\tau_p + N\tau_f) = 2\alpha_{m,2} \times N\tau_c$ bits for each AP to quantize the received pilots and the uplink data during each coherence interval. Moreover, we need $2K\tau_f\alpha_{m,3}$ bits to quantize the signal during each coherence interval for the case of the decentralized scheme. Finally $R_{\text{fh},m}$ is the fronthaul rate of cell-free massive MIMO at the m th AP to the CPU, is given by

$$R_{\text{bh},m} = \begin{cases} \frac{2N(K+\tau_f)\alpha_{m,1}}{T_c}, & \text{Estimate\&Quantize,} \\ \frac{2N\tau_c\alpha_{m,2}}{T_c}, & \text{Quantize\&Estimate,} \\ \frac{2K\tau_f\alpha_{m,3}}{T_c}, & \text{Decentralized scheme,} \end{cases} \quad (43a-c)$$

where T_c (in sec.) refers to coherence time.

B. Power Consumption Model

The total power consumption can be defined as follows: $P_{\text{total}} = P_{\text{TX}} + P_{\text{CP}}$, where P_{TX} is the uplink power amplifiers (PAs) due to transmit power at the users and PA dissipation [49], and P_{CP} refers to the circuit power (CP) consumption [49]. The power consumption P_{TX} is given by $P_{\text{TX}} = \frac{1}{\zeta} \rho N_0 \sum_{k=1}^K q_k$, where ζ is the PA efficiency at each user. The power consumption P_{CP} is obtained as $P_{\text{CP}} = MP_{\text{fix}} + KP_U + \sum_{m=1}^M P_{\text{fh},m}$, where P_{fix} is a fixed power consumption (including control signals and fronthaul) at each AP, P_U denotes the required power to run circuit components at each user and finally, fronthaul power consumption from the m th AP to the CPU is obtained as follows [50], [51]:

$$P_{\text{fh},m} = P_{\text{BT},m} \frac{R_{\text{fh},m}}{C_{\text{fh},m}}, \quad (44)$$

where $P_{\text{BT},m}$ is the power required for fronthaul traffic (BT) at the m th AP, and $C_{\text{fh},m}$ introduces the capacity of the fronthaul link between the m th AP and the CPU.

C. Total Energy Efficiency

In this section, we formulate the total energy efficiency of cell-free massive MIMO uplink. The total energy efficiency is achieved by dividing the sum throughput by the total consumed power which is given by

$$E_e(\mathbf{V}^{(i)}, \alpha_{m,i}) = \frac{B \cdot S(\mathbf{V}^{(i)}, \alpha_{m,i})}{P_{\text{total}}(\alpha_{m,i})} \left(\frac{\text{bit}}{\text{Joule}} \right), \quad (45)$$

where B is the bandwidth, and

$$\mathbf{V}^{(i)} = \begin{cases} \check{\mathbf{V}}, & \text{if } i = 1 \text{ (Estimate\&Quantize)}, \\ \check{\check{\mathbf{V}}}, & \text{if } i = 2 \text{ (Quantize\&Estimate)}, \end{cases} \quad (46)$$

and the sum spectral efficiency is given by

$$S(\mathbf{V}^{(i)}, \alpha_{m,i}) = \sum_{k=1}^K S_k(\mathbf{V}^{(i)}, \alpha_{m,i}), \quad (47)$$

where the spectral efficiency (in bit/s/Hz) is defined by

$$S_k(\mathbf{V}^{(i)}, \alpha_{m,i}) = \left(1 - \frac{\tau_p}{\tau_c} \right) \mathbb{E} \left\{ \log_2 \left(1 + \text{SINR}_k^{(i)} \right) \right\}, \quad (48)$$

where $\text{SINR}_k^{(i)}$ refers to the SINR at the k th user and the term (i) refers to E&Q,lf and Q&E,lf for $i = 1$ and $i = 2$, respectively. Note that the expectation is taken over small-scale fading. Next, for the case of decentralized scheme, the total energy efficiency is given by

$$E_e(\alpha_{m,3}) = \frac{B \cdot S(\alpha_{m,3})}{P_{\text{total}}(\alpha_{m,3})} \left(\frac{\text{bit}}{\text{Joule}} \right), \quad (49)$$

where the sum spectral efficiency is given by

$$S(\alpha_{m,3}) = \sum_{k=1}^K S_k(\alpha_{m,3}), \quad (50)$$

where the spectral efficiency (in bit/s/Hz) is defined as follows:

$$S_k(\alpha_{m,3}) = \left(1 - \frac{\tau_p}{\tau_c} \right) \log_2 \left(1 + \text{SINR}_k^{\text{dec}} \right). \quad (51)$$

VIII. AP ASSIGNMENT

For the case of the decentralized scheme, the total fronthaul rate required between the m th AP and the CPU increases linearly with the total number of users served by the m th AP, which is obvious from (43c). This reveals the need to pick a proper set of active users for each AP, i.e., $|\mathcal{U}|_m, \forall m$. Using (43c) and the constraint $R_{\text{bh},m} \leq C_{\text{fh},m}$, we have

$$\alpha_{m,3} \times |\mathcal{U}|_m \leq \frac{C_{\text{fh},m} T_c}{2\tau_f}, \forall m. \quad (52)$$

The important result of (52) is that reducing the number of active users at each AP allows for using more bits to quantize each weighted signal at the APs. To find the proper number of active users for each AP, we modify the proposed antenna selection schemes in [28]. The proposed modified received power based AP assignment scheme is summarized in Algorithm 1, where σ_{th} is a predetermined threshold. We will investigate the effect of σ_{th} on the system performance in the next section. Moreover, inspired by the methodology

Algorithm 1 Modified received power based AP assignment

1. Initialize σ_{th} .
 2. Find set \mathcal{A}_k : First sort $\mathcal{P}^{\text{RP}}(m, k)$ as $\mathcal{P}^{\text{RP}}(k^{(1)}, k) \leq \dots \leq \mathcal{P}^{\text{RP}}(k^{(M)}, k)$ where we define $K^{(m)} \in \{1, \dots, M\}$. Suppose $\mathcal{A}_k = \{k^{(1)}\}$.
 for $i = 2 : M$
 if $\sum_{m \in \mathcal{A}_k} \mathcal{P}(m, k) \geq \sigma_{\text{th}}\%$, then stop,
 else $\mathcal{A}_k = \mathcal{A}_k \cup \{k^{(i)}\}$
 end
 end
 3. Calculate \mathcal{U}_m from \mathcal{A}_k . Next, set $\alpha_{m,3} = \left\lfloor \frac{T_c C_{\text{fh},m}}{2|\mathcal{U}_m|\tau_f} \right\rfloor$ and calculate \tilde{a} and $\sigma_{\tilde{e}_y}^2$ from Table I. Finally, set $\gamma_{mk} = 0$ when $k \notin \mathcal{U}_m$, and calculate the SINR given in (23).
-

in [28], we use $\mathcal{P}^{\text{RP}}(m, k) = \frac{\sqrt{q_k} \gamma_{mk}}{\sum_{m'=1}^M \sqrt{q_k} \gamma_{m'k}}$ in Algorithm

1. Next, we enhance the large-scale antenna selection scheme proposed in [28]. For the modified large-scale AP assignment, run Algorithm 1 with $\mathcal{P}^{\text{LS}}(m, k) = \frac{\beta_{mk}}{\sum_{m'=1}^M \beta_{m'k}}$. Note that in Algorithm 1, \mathcal{A}_k is the set of active APs for the k th user. Finally, we turn off the m th AP, if \mathcal{U}_m is empty.

IX. NUMERICAL RESULTS AND DISCUSSION

In this section, we provide numerical results to evaluate the performance of the cell-free massive MIMO system with different parameters and different beamforming schemes. A cell-free massive MIMO system with M APs and K single-antenna users is considered in a $D \times D$ simulation area, where both APs and users are uniformly distributed in random locations. In the following subsections, we define the simulation parameters and then present the corresponding simulation results. Throughout the rest of the paper, the index m is dropped from $\alpha_{m,i}$ as we consider the same number of bits to quantize the signal at all APs. Note that the term ‘‘orthogonal pilots’’ refer to the case where orthogonal pilots are assigned to all users, while in ‘‘random pilot assignment’’ each user is randomly assigned a pilot sequence from a set of orthogonal sequences of length τ_p ($< K$), following the approach of [24], [52].

A. Simulation Parameters

The channel coefficients between users and APs are modeled in (1) where the coefficient β_{mk} is given by [24] $\beta_{mk} = \frac{\sigma_{sh} z_{mk}}{\text{PL}_{mk} 10^{-\frac{\sigma_{sh} z_{mk}}{10}}}$, where PL_{mk} is the path loss from the k th user to the m th AP and $10^{-\frac{\sigma_{sh} z_{mk}}{10}}$ denotes the shadow fading with standard deviation $\sigma_{sh} = 8$ dB, and $z_{mk} \sim \mathcal{N}(0, 1)$. In the simulation, an uncorrelated shadowing model and a three-slope model for the path loss in [24] are considered. Moreover, it is assumed that that \bar{p}_p and $\bar{\rho}$ denote the pilot sequence and the uplink data powers, respectively, where $p_p = \frac{\bar{p}_p}{P_n}$ and $\rho = \frac{\bar{\rho}}{P_n}$. In simulations, we set $\bar{P}_p = 100$ mW and $\bar{\rho} = 100$ mW, $P_{\text{BT}} = 5$ Watt, and $D = 1$ km unless otherwise stated. Similar to [24], we assume that the simulation area is wrapped around

at the edges, and hence simulate an area without boundaries. Hence, the square simulation area has eight neighbours. We evaluate the average sum spectral and energy efficiency of the system over 300 random realizations of the locations of APs, users and shadow fading. Moreover, we set $\zeta = 0.3$, $P_U = 0.1$ Watt, $P_{\text{fix}} = 0.825$ Watt [49]–[51], [53].

B. Numerical Results

1) *Effect of Number of Quantization Bits*: This section investigates the spectral efficiency and energy efficiency performance of different cases of cell-free massive MIMO with fixed capacity of fronthaul links and different numbers of quantization bits. We assume that $M = 40$ APs each with $N = 10$ antennas are uniformly distributed in the area. Moreover, we assume $K = 40$ users and $\tau_p = 40$ as the length of pilot sequences. In Fig. 4a, the average sum spectral efficiency of the system is plotted versus the number of quantization bits while assuming a fixed $C_{\text{fh}} = 64$ Mbits/s as the capacity of fronthaul links and three different receivers, namely the MMSE, ZF and MRC receivers. As the figure shows only $\alpha_1 = \alpha_2 = 6$ bits for Case Q&E and E&Q and $\alpha_3 = 2$ bits for decentralized scheme are enough to closely achieve the performance of perfect fronthaul links. Note that exploiting (39), $\alpha_1 = \alpha_2 = 6$ and $\alpha_3 = 2$ bits lead to $R_{\text{fh}} = 24$ Mbit/s for all cases. In addition, the figure demonstrates that the performance of decentralized scheme can be significantly improved by using the proposed AP assignment schemes. Note that in the figures, legends RP-assignment and LS-assignment refer to the proposed modified received power based AP assignment and the large-scale based AP assignment, respectively. Note that for a fixed fronthaul rate, the performance gap between decentralized scheme and MRC with limited-fronthaul is due to the fact that in the decentralized scheme, the CPU does not have access to the quantized channel estimates and exploits only the statistics of the channel to decode the data. Next, the average total energy efficiency performance of the system is investigated in Fig. 6b. As the figure shows $\alpha_1 = \alpha_2 = 6$ and $\alpha_3 = 2$ bits are the optimal values to maximize the performance of energy efficiency of cell-free massive MIMO for Q&E and E&Q and decentralized schemes, respectively.

2) *Effect of the Capacity of fronthaul Links*: In this section, the effect of capacity of fronthaul links on the spectral efficiency and energy efficiency performance of cell-free massive MIMO is presented. We assume $M = 40$ distributed APs each equipped with $N = 10$ antennas serve $K = 40$ uniformly distributed users. We also consider $\tau_p = 40$ orthogonal pilots. Figs. 5a and 5b demonstrate the spectral efficiency and energy efficiency performance of the system versus the capacity of fronthaul links, respectively. To exploit the available fronthaul capacity we exploit different numbers of bits for different situations. For example, let us consider the first point, i.e., $C_{\text{fh}} = 12.8$ Mbits/s. Using the fronthaul rate given in (43) and setting $\alpha_1 = 3$, $\alpha_2 = 3$ and $\alpha_3 = 1$, we have $R_{\text{fh}} = 12$ Mbit/s, $R_{\text{fh}} = 12$ Mbit/s and $R_{\text{fh}} = 12.8$ Mbit/s, respectively. As expected the average sum spectral efficiency increases as the capacity of fronthaul links increase which is shown in

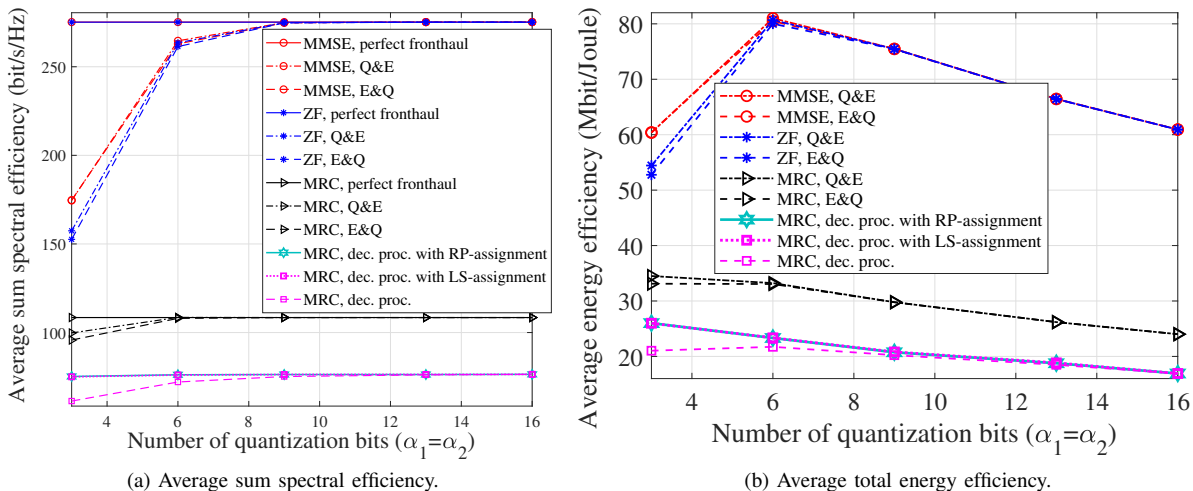


Figure 4. Here, we set $M = 40$, $N = 10$, $K = 40$, $\tau_p = 40$, $C_{\text{fh}} = 64$ Mbits/s, $P_{\text{BT}} = 1$ Watt, $\tau_c = 200$, and $T_c = 1$ ms. We set $\alpha_1 = \alpha_2 = \{3, 6, 9, 13, 16\}$, and $\alpha_3 = \{1, 2, 3, 4, 5\}$ which using (43) results in $R_{\text{fh}} = \{12, 24, 36, 52, 64\}$ Mbit/s for all three cases. Moreover for the case of AP assignment we use $\delta_{\text{th}} = \{96\%, 99\%, 99\%, 99\%, 100\%\}$.

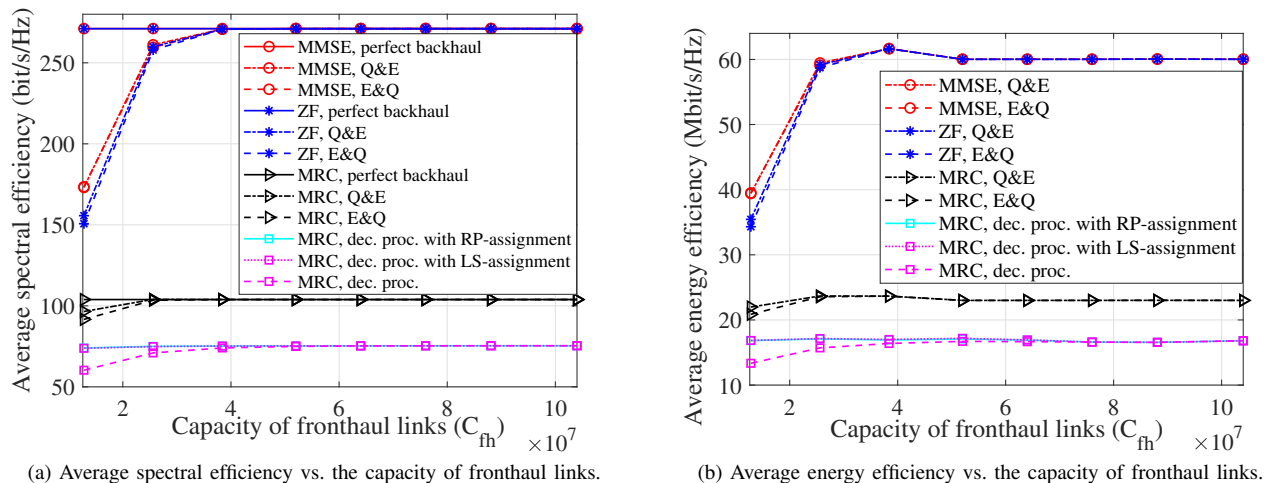


Figure 5. Here, we set $M = 40$, $N = 10$, $K = 40$, $\tau_p = 40$, $P_{\text{BT}} = 1$, $\tau_c = 200$, and $T_c = 1$ ms. We use $\alpha_1 = \alpha_2 = \{3, 6, 9, 13, 16, 19, 22, 26\}$ which using (43a) and (43b) results in $R_{\text{fh}} = \{12, 24, 36, 52, 64, 76, 88, 104\}$ Mbit/s, respectively whereas we set $\alpha_3 = \{1, 2, 3, 4, 5, 6, 7, 8\}$ which leads to $R_{\text{fh}} = \{12.8, 25.6, 38.4, 51.2, 64, 76, 89.6, 102.4\}$ Mbit/s using (43c). Moreover for the case of AP assignment we use $\delta_{\text{th}} = \{95\%, 99\%, 99.6\%, 99.6\%, 99.7\%, 100\%, 100\%, 100\%\}$.

Fig. 5a. Moreover, interestingly Fig. 5a reveals that using $R_{\text{fh}} = 38.4$ Mbits/s the spectral efficiency performance of the cell-free massive MIMO system is very close to the performance of the system with perfect fronthaul links. Next, Fig. 5b demonstrate the average total energy efficiency of the cell-free massive MIMO system. As it can be observed to maximize the energy efficiency of the system we need to set $R_{\text{fh}} = 38.4$ Mbits/s. Note that the proposed AP assignment algorithms can significantly improve the performance of cell-free massive MIMO in the low fronthaul rate regime. Note that both the spectral efficiency and the power consumption increase with the number of quantization bits. The increasing spectral efficiency will lead to an increasing energy efficiency as long as the quantization bits have a negligible impact on the total power consumption. After a certain point, the energy efficiency reduces instead since the power consumption increases faster than the spectral efficiency. Finally, Figs. 5a

and 5b demonstrate that the performance of the ZF receiver is closer to the performance of the MMSE receiver in the high fronthaul rate regime.

3) *Effect of Number of Antennas at APs*: This section investigates the effect of number of antennas per AP in cell-free massive MIMO, where we fix the total number of service antennas (i.e., MN), and change the number of antennas per AP, N , and total number of APs, M . Fig. 6b shows the average total energy efficiency of the system with $MN = 300$, $K = 40$, $\tau_p = 40$ and $C_{\text{fh}} = 122.4$ Mbits/s. For the sets of $M = \{2, 4, 5, 10, 15, 20, 25\}$ and $N = \{150, 75, 60, 30, 20, 15, 12\}$, we use $\alpha_1 = \alpha_2 = \{2, 4, 5, 10, 15, 20, 25\}$ and $\alpha_3 = 12$ which using (43) result in $R_{\text{fh}} = 120$ Mbit/s, $R_{\text{fh}} = 120$ Mbit/s $R_{\text{fh}} = 122.4$ Mbit/s, for Q&E and E&Q and decentralized schemes, respectively. We observe that to maximize the energy efficiency, one can determine optimal values of M and N . In addition, for this system set-up, the average sum spectral

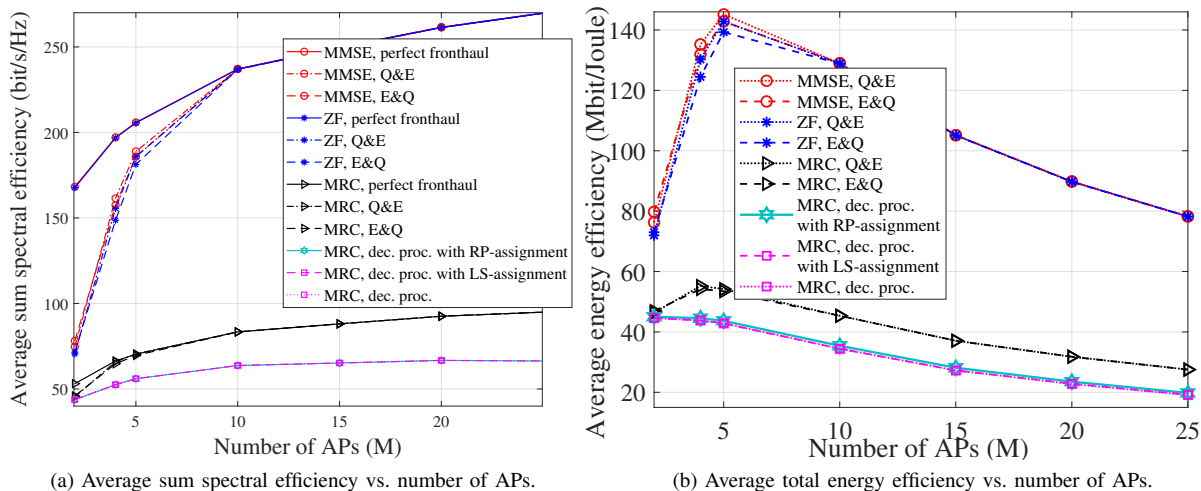


Figure 6. Here, we assume a fixed total service antennas $MN = 300$, $K = 40$, $\tau_p = 40$, $C_{\text{fh}} = 122.4$ Mbits/s and $\tau_c = 200$, $T_c = 1$ ms. For the sets of $M = \{2, 4, 5, 10, 15, 20, 25\}$ and $N = \{150, 75, 60, 30, 20, 15, 12\}$, we use $\alpha_1 = \alpha_2 = \{2, 4, 5, 10, 15, 20, 25\}$ and $\alpha_3 = 12$ which using (43) result in $R_{\text{fh}} = 120$ Mbit/s, $R_{\text{fh}} = 120$ Mbit/s, $R_{\text{fh}} = 122.4$ Mbit/s, for Q&E and E&Q and decentralized schemes, respectively. Finally, $\delta_{\text{th}} = \{100\%, 100\%, 100\%, 100\%, 100\%, 99.5\%, 99.5\%\}$.

efficiency performance is shown in Fig. 6a. As expected, the average sum spectral efficiency increases by distributing more antennas in the area, i.e., increasing M (as it will bring APs close to the users and improve the quality of the channel). Therefore, Figs. 6b and 6a present a tradeoff between M and N . Moreover, as expected, Figs. 6a and 6b reveal that the proposed AP assignment schemes are not able to improve the performance for the case of a few APs. Moreover, for the case of MRC, there is a performance gap between the decentralized scheme and the centralized scheme (Q&E and E&Q). This is due to the fact that in general the level of channel hardening in cell-free massive MIMO is small which is a result of [54]. Interestingly, Fig. 6a shows that by increasing the number of antennas per AP, the performance gap between the decentralized and the centralized schemes reduces.

4) *More Bits to Quantize the Estimated Channel or the Received Signal in Case E&Q?*: In Section IV, we investigate the case when the AP estimates the channel and sends back the quantized channel estimate as well as the quantized received signal to the CPU. Using (43a), the fronthaul rate for this case is obtained by $R_{\text{fh}} = \frac{2(NK\alpha_1^g + N\tau_f\alpha_{m,1}^y)}{T_c}$, when we exploit the same number of bits to quantize the channel estimate and the received signal. However, we could exploit different numbers of bits to quantize the received signal and the channel estimate. Let us assume $\alpha_{m,1}^y, \forall m$ and $\alpha_{m,1}^g, \forall m$ refer to the quantization bits to quantize the received signal and the channel estimate, respectively. Hence we have $R_{\text{fh}} = \frac{2(NK\alpha_{m,1}^g + N\tau_f\alpha_{m,1}^y)}{T_c}, \forall m$, where we drop the index m as we use the same number of bits at all APs and use $R_{\text{fh}} = \frac{2(NK\alpha_1^g + N\tau_f\alpha_1^y)}{T_c}$. First we consider a cell-free massive MIMO system with $M = 20$, $N = 20$, $K = 40$, $\tau_p = 40$, $\tau_c = 200$, $T_c = 1$ ms, and $C_{\text{fh}} = 40$ Mbits/s. Figs. 9a and 9b demonstrate the average sum spectral efficiency and energy efficiency, respectively. We assume pairs $(\alpha_1^y, \alpha_1^g) = \{(21, 1), (17, 2), (13, 3), (9, 4), (5, 5), (1, 6)\}$

where using $R_{\text{fh}} = \frac{2(NK\alpha_1^g + N\tau_f\alpha_1^y)}{T_c}$ results in $R_{\text{fh}} = 40$ Mbits/s for all pairs. Figs. 9a and 9b show that the pair for this network set-up, having the same number of bits to quantize the channel estimate and the signal, i.e., $(\alpha_1^y, \alpha_1^g) = (5, 5)$ is optimal.

Next, we assume a cell-free massive MIMO network with $M = 200$, $N = 2$, $K = 40$, $\tau_p = 40$, $\tau_c = 200$, $T_c = 1$ ms, and $C_{\text{fh}} = 12$ Mbits/s. To quantize the channel estimate and the received signal we consider pairs $(\alpha_1^y, \alpha_1^g) = \{(71, 1), (67, 2), (63, 3), (59, 4), (55, 5), (51, 6), (47, 7), (43, 8), (39, 9), (35, 10), (31, 11), (27, 12), (23, 13), (19, 14), (15, 15), (11, 16), (7, 17), (3, 18)\}$ where using $R_{\text{fh}} = \frac{2(NK\alpha_1^g + N\tau_f\alpha_1^y)}{T_c}$ results in $R_{\text{fh}} = 12$ Mbits/s for all pairs. The average sum spectral efficiency for this network set-up is presented in Fig. 8a. As the figure shows exploiting the same number of bits to quantize the received signal and the channel estimate is optimal and maximizes the spectral efficiency performance of the cell-free massive MIMO system.

To investigate the effect of coherence time on the system performance, we present the average sum spectral efficiency the cell-free massive MIMO system with $M = 200$, $N = 2$, $K = 50$, $\tau_p = 40$, $\tau_c = 750$, $T_c = 2.5$ ms, and $C_{\text{fh}} = 12$ Mbits/s in Fig. 8. The pairs $(\alpha_1^y, \alpha_1^g) = \{(80, 5), (66, 6), (52, 7), (38, 8), (24, 9), (10, 10)\}$ are considered where exploiting $R_{\text{fh}} = \frac{2(NK\alpha_1^g + N\tau_f\alpha_1^y)}{T_c}$ we have $R_{\text{fh}} = 12$ Mbits/s for all pairs. As the figure shows exploiting the same number of bits to quantize the channel estimate and the received signal, i.e., $(\alpha_1^y, \alpha_1^g) = (5, 5)$ maximizes the spectral efficiency of the system.

5) *The Optimal δ_{th} for the Proposed AP Assignment Schemes for the Centralized Scheme*: In this section, we investigate the optimal value of δ_{th} for the proposed AP assignment schemes. Figs. 9a and 9b present spectral efficiency and the energy efficiency of the cell-free massive MIMO, respectively, with $M = 40$, $N = 10$, $K = 40$, $C_{\text{fh}} = 12.8$ Mbits/s, $P_{\text{fh}} = 1$ Watt, and $\alpha_3 = 1$.

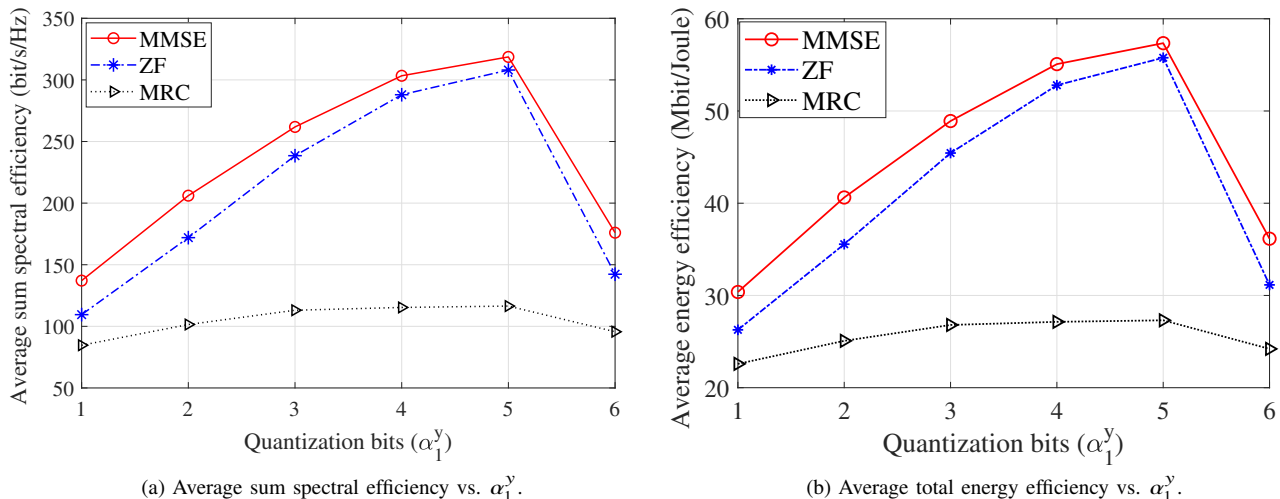
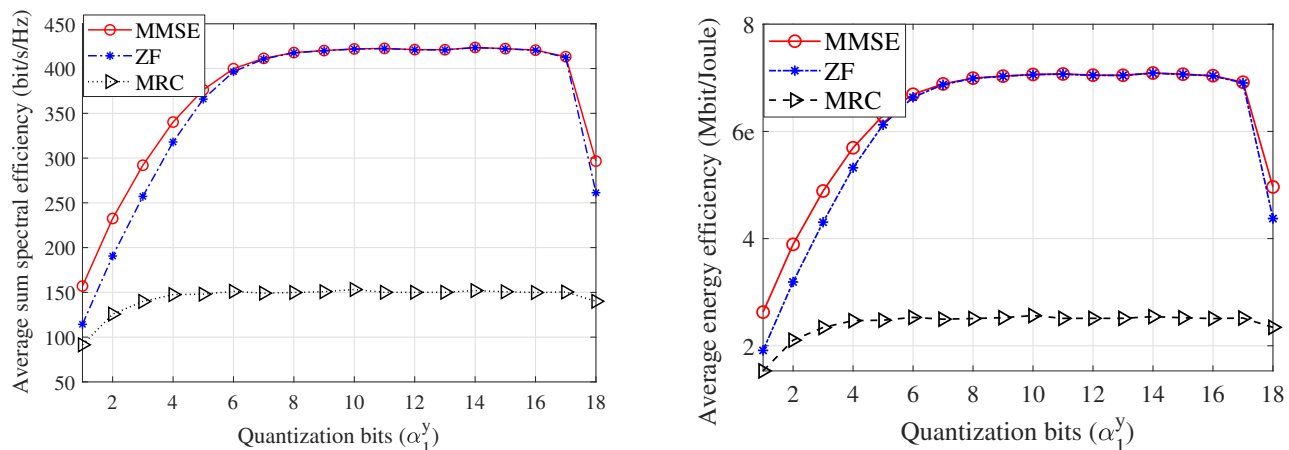


Figure 7. Here, we set $M = 20$, $N = 20$, $K = 40$, $\tau_P = 40$, $\tau_c = 200$, $T_c = 1$ ms, and $C_{fh} = 40$ Mbits/s. We assume the pairs $(\alpha_1^y, \alpha_1^g) = \{(21, 1), (17, 2), (13, 3), (9, 4), (5, 5), (1, 6)\}$ and $R_{fh} = 40$ Mbits/s.



(a) Here, we set $K = 40$, $\tau_P = 40$, $\tau_c = 200$, $T_c = 1$ ms, and $C_{fh} = 12$ Mbits/s. It is assumed $(\alpha_1^y, \alpha_1^g) = \{(71, 1), (67, 2), (63, 3), (59, 4), (55, 5), (51, 6), (47, 7), (43, 8), (39, 9), (35, 10), (31, 11), (27, 12), (23, 13), (19, 14), (15, 15), (11, 16), (7, 17), (3, 18)\}$ and $R_{fh} = 12$ Mbits/s for all pairs.

(b) Here, we set $K = 50$, $\tau_P = 40$, $\tau_c = 750$, $T_c = 2.5$ ms, and $C_{fh} = 12$ Mbits/s in Fig. . The pairs $(\alpha_1^y, \alpha_1^g) = \{(80, 5), (66, 6), (52, 7), (38, 8), (24, 9), (10, 10)\}$ are considered and $R_{fh} = 12$ Mbits/s for all pairs.

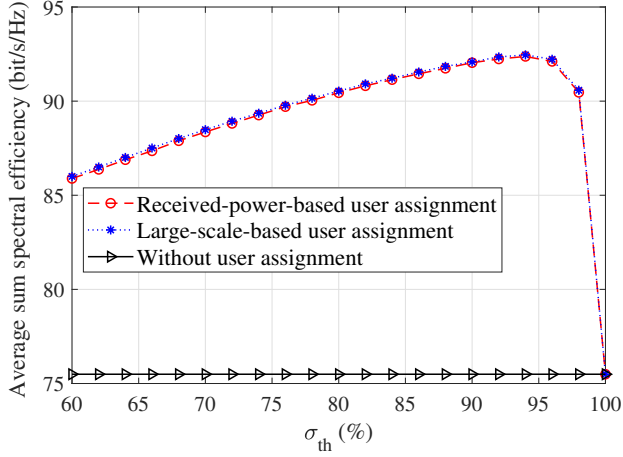
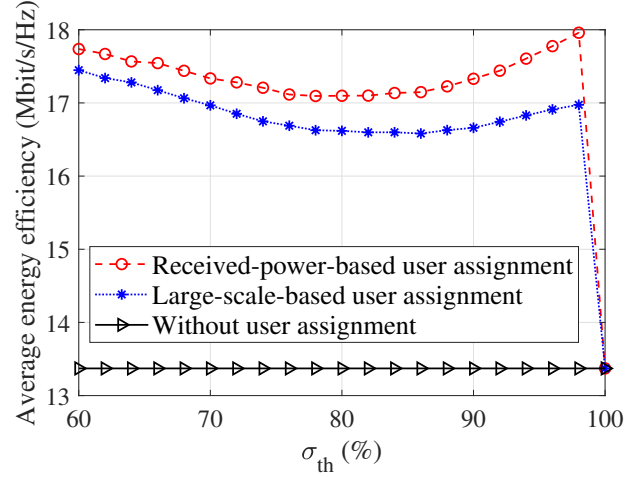
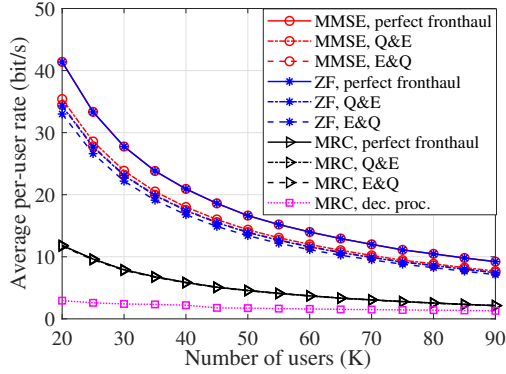
Figure 8. Average sum spectral efficiency versus α_1^y with $M = 200$, $N = 2$.

6) *The Effect of Total Number of Users*:: In this section, we investigate the effect of total number of users on the system performance. In Fig. 10 with $M = 40$, $N = 10$ and $C_{fh} = 30$ Mbit/s. As expected, by increasing total number of users, the average per-user rate of the cell-free massive MIMO system decreases. In this paper, we investigate the ergodic rates so there is no concept of outage performance. One possible application of the results of the current paper is power control where we could provide the exact rates that the users are requesting, instead of the rates that are obtained with a given uplink power. This is left aside for future research.

X. CONCLUSIONS

We have considered limited-fronthaul cell-free massive MIMO, and a performance comparison between different ways of implementing cell-free massive MIMO uplink has been

presented. First, we have investigated the performance of cell-free massive MIMO with perfect fronthaul links and linear receivers. Next, two other different cases have been studied: (i) Estimate&Quantize: when APs estimate the channel and transfer the quantized version of the estimated channel and the quantized versions of the received signals to the CPU and (ii) Quantize&Estimate, APs transfer the quantized version of the received uplink pilot and the quantized version of the received uplink data to the CPU, and hence CPU estimates the channels. We have made use of the Bussgang decomposition, which enables us to find a linear relationship between the input of the quantizer and the quantization noise. Spectral and energy efficiency of the different cases with different linear receivers have been derived. Numerical results have been provided and showed that the performance of limited-fronthaul cell-free massive MIMO is fairly close to the per-

(a) Average sum spectral efficiency vs. δ_{th} .(b) Average total energy efficiency vs. δ_{th} .Figure 9. Here, we set $M = 40$, $N = 10$, $K = 40$, $C_{fh} = 12.8$ Mbits/s, $P_{fh} = 1$ Watt, and $\alpha_3 = 1$.Figure 10. The average per-user rate of the system with $M = 40$, $N = 10$ and $C_{fh} = 30$ Mbits/s. For $K = \{20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90\}$, we set $\alpha_1 = \alpha_2 = 5$ (which comes from (43a) and (43b)) and $\alpha_3 = \{3, 2, 2, 2, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1\}$ (coming from (43c)).

formance of the perfect-fronthaul system while exploiting only a few bits for quantization. Moreover, the performance of Quantize&Estimate is only marginally better than that of Estimate&Quantize. Finally, it was shown that in the limited-fronthaul cell-free massive MIMO, to achieve the best performance the same number of bits should be used to quantize the estimated channel at the APs and to quantize the received signal.

APPENDIX A: PROOF OF LEMMA 1

We use the following proposition and remarks to prove Lemma 1 for the MRC case. Note that the proof for ZF and MMSE follows the same steps as in the MRC case.

Proposition 2. *The input signals of quantizers defined in (8) and (10) are uncorrelated.*

Proof: As explained in Subsections IV-A and IV-B, we need to quantize the received signal and the estimated channel in the Estimate&Quantize scheme. To prove that the inputs of

quantizers are uncorrelated, it is sufficient to prove that the following equalities hold:

$$\mathbb{E}\{\mathbf{y}_m \mathbf{y}_l^H\} = \mathbf{0}, \forall l \neq m, \quad \mathbb{E}\{[\mathbf{y}_m]_n^* [\mathbf{y}_m]_{n'}\} = 0, \forall n \neq n', \quad (53a)$$

$$\mathbb{E}\{\hat{\mathbf{g}}_{mk} \hat{\mathbf{g}}_{lk}^H\} = \mathbf{0}, \forall l \neq m, \quad (53b)$$

$$\mathbb{E}\{[\hat{\mathbf{g}}_{mk}]_n^* [\hat{\mathbf{g}}_{mk}]_{n'}\} = 0, \forall n \neq n', \quad \mathbb{E}\{\mathbf{y}_m \hat{\mathbf{g}}_{lk}^H\} = 0, \forall l \neq m, \quad (53c)$$

$$\mathbb{E}\{[\mathbf{y}_m]_n^* [\hat{\mathbf{g}}_{mk}]_{n'}\} = 0, \forall n \neq n', \quad (53d)$$

where $\mathbf{0}$ refers to a matrix with size $MN \times 1$ with all zero elements. Note that using (2) and (3) we have the equalities in (54) (defined at the top of the next page), which completes the proof of Lemma 1. \square

Proposition 3. *Using Remark 4, terms A_{1s_k} and A_5 are uncorrelated.*

Proposition 4. *Terms A_{1s_k} and A_4 are uncorrelated.*

Proof: We have

$$\begin{aligned} \mathbb{E}\{A_{1s_k}^* A_4\} &= \mathbb{E}\left\{\left(\tilde{a} \sqrt{\rho q_k} \sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H \check{\mathbf{g}}_{mk} s_k\right)^* \left(\sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H \mathbf{e}_m^y\right)\right\} \\ &= M \tilde{a} \sqrt{\rho q_k} \mathbb{E}\{\|\check{\mathbf{g}}_k\|^2 \hat{\mathbf{g}}_k^H \mathbf{e}^y s_k^*\} = 0, \end{aligned} \quad (55)$$

where $\mathbf{e}^y = [\mathbf{e}_1^T, \dots, \mathbf{e}_M^T]^T$, and the second equality is due to the following facts: $\mathbb{E}\{\check{\mathbf{g}}_k^H s_k\} = \mathbf{0}$, $\mathbb{E}\{\hat{\mathbf{g}}_k^H \mathbf{e}^y\} = 0$, and $\mathbb{E}\{\mathbf{e}^y s_k\} = \mathbf{0}$, where $\mathbf{0} = [0, \dots, 0]^T \in \mathbb{C}^{MN \times 1}$. This completes proof of Proposition 4. \square

Proposition 5. *Terms A_4 and A_5 are uncorrelated.*

Proof: Note that we have

$$\begin{aligned} &\mathbb{E}\{A_4^* A_5\} \\ &= \mathbb{E}\left\{\left(\sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H \mathbf{e}_m^y\right)^* \left(\tilde{a} \sqrt{\rho} \sum_{m=1}^M \hat{\mathbf{g}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \mathbf{e}_{mk'}^y s_{k'}\right)\right\} = 0, \end{aligned} \quad (56)$$

$$\mathbb{E} \{ \mathbf{y}_m \mathbf{y}_l^H \} = \mathbb{E} \{ (\mathbf{g}_{mk} s_k + \mathbf{n}_m) (\mathbf{g}_{lk} s_k + \mathbf{n}_l)^H \} = \sum_{k=1}^K \mathbb{E} \{ \mathbf{g}_{mk} \mathbf{g}_{lk}^H \} = \underline{\mathbf{0}}, \forall l \neq m, \quad (54a)$$

$$\mathbb{E} \{ [\mathbf{y}_m]_n^* [\mathbf{y}_m]_{n'} \} = \mathbb{E} \{ ([\mathbf{g}_{mk}]_n s_k + [\mathbf{n}_m]_n)^* ([\mathbf{g}_{mk}]_{n'} s_k + [\mathbf{n}_m]_{n'}) \} = 0, \forall n \neq n', \quad (54b)$$

$$\mathbb{E} \{ \hat{\mathbf{g}}_{mk} \hat{\mathbf{g}}_{lk}^H \} = \mathbb{E} \left\{ c_{mk} \left(\sqrt{\tau_p P_p} \sum_{k'=1}^K \mathbf{g}_{mk'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + \boldsymbol{\Omega}_{p,m} \boldsymbol{\phi}_k \right) = c_{lk} \left(\sqrt{\tau_p P_p} \sum_{k'=1}^K \mathbf{g}_{lk'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + \boldsymbol{\Omega}_{p,l} \boldsymbol{\phi}_k \right)^H \right\} = \underline{\mathbf{0}}, \forall l \neq m, \quad (54c)$$

$$\begin{aligned} \mathbb{E} \{ [\hat{\mathbf{g}}_{mk}]_n^* [\hat{\mathbf{g}}_{mk}]_{n'} \} &= \mathbb{E} \left\{ c_{mk} \left(\sqrt{\tau_p P_p} \sum_{k'=1}^K [\mathbf{g}_{mk'}]_n \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + [\boldsymbol{\Omega}_{p,m} \boldsymbol{\phi}_k]_n \right)^* \left(c_{mk} \left(\sqrt{\tau_p P_p} \sum_{k'=1}^K [\mathbf{g}_{mk'}]_{n'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + [\boldsymbol{\Omega}_{p,m} \boldsymbol{\phi}_k]_{n'} \right) \right) \right\} \\ &= 0, \forall n \neq n', \end{aligned} \quad (54d)$$

$$\mathbb{E} \{ \mathbf{y}_m \hat{\mathbf{g}}_{lk}^H \} = \mathbb{E} \left\{ (\mathbf{g}_{mk} s_k + \mathbf{n}_m) c_{lk} \left(\sqrt{\tau_p P_p} \sum_{k'=1}^K \mathbf{g}_{lk'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + \boldsymbol{\Omega}_{p,l} \boldsymbol{\phi}_k \right)^H \right\} = \underline{\mathbf{0}}, \forall l \neq m, \quad (54e)$$

$$\mathbb{E} \{ [\mathbf{y}_m]_n^* [\hat{\mathbf{g}}_{mk}]_{n'} \} = \mathbb{E} \left\{ ([\mathbf{g}_{mk}]_n s_k + [\mathbf{n}_m]_n)^* c_{mk} \left(\sqrt{\tau_p P_p} \sum_{k'=1}^K [\mathbf{g}_{mk'}]_{n'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + [\boldsymbol{\Omega}_{p,m} \boldsymbol{\phi}_k]_{n'} \right) \right\} = 0, \forall n \neq n', \quad (54f)$$

where the second equality is due to the following facts:

$$\begin{aligned} \mathbb{E} \{ \check{\mathbf{g}}_{mk}^H s_{k'} \} &= 0, \quad \mathbb{E} \{ \mathbf{e}_{mk'}^g s_{k'} \} = 0, \\ \mathbb{E} \{ \check{\mathbf{g}}_{mk}^H \mathbf{e}_{mk'}^y \} &= 0, \quad \mathbb{E} \{ \check{\mathbf{g}}_{mk}^H \mathbf{e}_m^y \} = 0, \end{aligned} \quad (57)$$

where the first two equalities are due to the fact that there is no correlation between the transmitted signal s_k and the quantized version of the estimated channel. Note that the third equality in (57) comes from Remark 4. This completes the proof of Proposition 5.

Note that in the proof of Proposition 4, we used the property of Max algorithm that the output of the quantizer and the distortion are uncorrelated (which is explained in Section III-B of the current paper). Otherwise, without using the Max algorithm, it is not possible to prove that the term A_4 and A_5 are uncorrelated in Proposition 4.

Proposition 6. *Using Remark 4, terms A_2 and A_5 are uncorrelated.*

Proposition 7. *As terms A_3 and A_6 include i.i.d. Gaussian noise and i.i.d. Gaussian MMSE error; respectively, A_2 and A_6 are uncorrelated with other terms.*

Finally, using Propositions 2- 7, it is easy to show that terms $A_1 s_k$, A_2 , A_3 , A_4 , A_5 and A_6 are mutually uncorrelated, which completes the proof of Lemma 1. ■

APPENDIX B: PROOF OF LEMMA 2

Firstly, note that A_2 , A_5 and A_6 include s_k , which has zero mean and is independent of Ω , hence, $\mathbb{E} \{ A_2 + A_5 + A_6 | \Omega \} = 0$. Similarly, A_3 includes \mathbf{n}_m which has zero mean and independent of Ω , resulting in $\mathbb{E} \{ A_3 | \Omega \} = 0$. Next, we need to show that $\mathbb{E} \{ A_4 | \Omega \} = 0$. We have $\mathbb{E} \{ A_4 | \Omega \} = \mathbb{E} \{ \check{\mathbf{v}}_{mk}^H \mathbf{e}_m^y | \Omega \} = \check{\mathbf{v}}_{mk}^H \mathbb{E} \{ \mathbf{e}_m^y | \Omega \}$. Therefore, we need to prove that $\mathbb{E} \{ \mathbf{e}_m^y | \Omega \} = 0$. Note that $\mathbb{E} \{ \mathbf{e}_m^y | \check{\mathbf{v}}_{mk} \} = 0$ is equivalent to $\mathbb{E} \{ \mathbf{e}_m^y | \check{\mathbf{g}}_{mk} \} = 0$,

as $\check{\mathbf{v}}_{mk}$ is a function of only $\check{\mathbf{g}}_{mk}$. Next, approximating the Bayesian estimator with a linear estimator, we have [42], [45]

$$\mathbb{E} \{ \mathbf{e}_m^y | \check{\mathbf{g}}_{mk} \} \stackrel{(a)}{\approx} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E} \{ \mathbf{e}_m^y \check{\mathbf{g}}_{mk}^H \} \check{\mathbf{g}}_{mk}. \quad (58)$$

Note that the approximation in step (a) is widely used in literature in the concept of uniform quantization [42], [44]–[46]. Next, we aim to calculate the term $\mathbb{E} \{ \mathbf{e}_m^y \check{\mathbf{g}}_{mk}^H \}$. As a remind, the term \mathbf{e}_m^y denotes the quantization distortion while we quantize the received signal, i.e., \mathbf{y}_m , in (8), and the term $\check{\mathbf{g}}_{mk}$ is the output of quantizer where we quantize the estimated channel, i.e., $\hat{\mathbf{g}}_{mk}$, in (10). It is easy to show that the input of these two quantizers are uncorrelated. More precisely, using (2) and (3), we have

$$\begin{aligned} \mathbb{E} \{ \mathbf{y}_m \hat{\mathbf{g}}_{mk}^H \} &= c_{mk} \mathbb{E} \left\{ \left(\sqrt{\rho} \sum_{k=1}^K \mathbf{g}_{mk} \sqrt{q_k} s_k + \mathbf{n}_m \right) \right. \\ &\quad \left. \left(\sqrt{\tau_p P_p} \sum_{k'=1}^K \mathbf{g}_{mk'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + \boldsymbol{\Omega}_{p,m} \boldsymbol{\phi}_k \right)^H \right\} = \underline{\mathbf{0}}. \end{aligned} \quad (59)$$

This shows that the inputs of quantizers when we separately quantize the estimated channel and the received signal are uncorrelated. Moreover, using the analysis in [44, Section 2.2], when we quantize two uncorrelated signals x_1 and x_2 using the uniform quantizer (i.e., $\mathcal{Q}(x_1) = x_1 + n_1$, and $\mathcal{Q}(x_2) = x_2 + n_2$), we have $\mathbb{E} \{ \mathcal{Q}(x_1) n_2 \} = \mathbb{E} \{ \mathcal{Q}(x_2) n_1 \} = 0$. Exploiting this fact and (59), we have

$$\mathbb{E} \{ \mathbf{e}_m^y \check{\mathbf{g}}_{mk}^H \} = \underline{\mathbf{0}}. \quad (60)$$

Finally, substituting (60) into (58), we get $\mathbb{E} \{ \mathbf{e}_m^y | \check{\mathbf{v}}_{mk} \} \approx 0$, which completes the proof. ■

APPENDIX C: PROOF OF LEMMA 3

Firstly, note that $\mathbb{E} \{ s_k^* (A_2 + A_3 + A_4) | \Omega \} = 0$, as none of the terms A_2 , A_3 and A_4 include s_k , which has zero mean and is independent of Ω . In the following, we prove that $\mathbb{E} \{ s_k^* (A_5 + A_6) | \Omega \} = 0$. Next, approximating the Bayesian

estimator with a linear estimator we show that $\mathbb{E}\{s_k^* A_5 | \check{\mathbf{g}}_{mk}\} \approx 0$, which is given as follows:

$$\begin{aligned} \mathbb{E}\{s_k^* A_5 | \check{\mathbf{g}}_{mk}\} &\approx \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\{s_k^* A_5 \check{\mathbf{g}}_{mk}^H\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{s_k^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \mathbf{e}_{mk',s_{k'}}^g\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(a_1)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\sqrt{q_k} \sum_{m=1}^M \left(\check{\mathbf{v}}_{mk}^H \mathbf{e}_{mk}^g\right) \check{\mathbf{g}}_{mk}^H |s_k|^2\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(a_2)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \sqrt{q_k} \mathbb{E}\left\{\sum_{m=1}^M \left(\check{\mathbf{v}}_{mk}^H \mathbf{e}_{mk}^g\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \stackrel{(a_3)}{=} 0, \quad (61) \end{aligned}$$

where in step (a₁) we used the fact that the terms s_k s are independent variables. Moreover, step (a₂) is due to the fact that s_k has zero mean and independent of Ω , and $\mathbb{E}\{|s_k|^2\} = 1$ whereas step (a₃) comes from (12). Next we have:

$$\begin{aligned} \mathbb{E}\{s_k^* A_6 | \check{\mathbf{g}}_{mk}\} &\approx \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\{s_k^* A_6 \check{\mathbf{g}}_{mk}^H\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{s_k^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \tilde{\mathbf{g}}_{mk',s_{k'}}\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \sqrt{q_k} \mathbb{E}\left\{\sum_{m=1}^M \left(\check{\mathbf{v}}_{mk}^H \tilde{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H |s_k|^2\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(b_1)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \sqrt{q_k} \mathbb{E}\left\{\underbrace{\sum_{m=1}^M \left(\check{\mathbf{v}}_{mk}^H \tilde{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H}_{J_1}\right\} \check{\mathbf{g}}_{mk} \stackrel{(b_2)}{\approx} 0, \quad (62) \end{aligned}$$

where in step (b₁) we used the fact that s_k s are independent variables. Moreover, note that the elements of $\check{\mathbf{g}}_{mk}$ are i.i.d. $CN(0, \beta_{mk} - \gamma_{mk})$ and are uncorrelated and independent of the input of the quantizer, i.e., $\mathbb{E}\{\check{\mathbf{g}}_{mk} \check{\mathbf{g}}_{mk}^H\} = \mathbf{0}$, and hence uncorrelated with the output of the quantizer, i.e., $\mathbb{E}\{\check{\mathbf{g}}_{mk} \tilde{\mathbf{g}}_{mk}^H\} = \mathbf{0}$. The step (b₂) comes from the fact that the output of the quantizer $\check{\mathbf{g}}_{mk}$ and the channel estimation error $\tilde{\mathbf{g}}_{mk}$ are uncorrelated and independent. This independence is shown in Figs. 11a, 11b, 11c, 11d, where the results reveal that $P(\mu, \xi) \approx P(\mu) \times P(\xi)$ while ξ and μ represent the channel estimation error and the output of the quantizer, respectively. This completes the proof. ■

APPENDIX D: PROOF OF LEMMA 4

It can be easily shown that $\mathbb{E}\{A_1^* s_k^* (A_2 + A_3 + A_4) | \Omega\} = 0$, as none of the terms A_1, A_2, A_3 and A_4 include s_k , which has zero mean and is independent of Ω . In the following, we show that $\mathbb{E}\{A_1^* s_k^* (A_5 + A_6) | \Omega\} = 0$. First, we aim to show

that $\mathbb{E}\{A_1^* s_k^* A_5 | \check{\mathbf{g}}_{mk}\} \approx 0$, as follows:

$$\begin{aligned} \mathbb{E}\{A_1^* s_k^* A_5 | \check{\mathbf{g}}_{mk}\} &\approx \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\{A_1^* s_k^* A_5 \check{\mathbf{g}}_{mk}^H\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* s_k^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \mathbf{e}_{mk',s_{k'}}^g\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(c_1)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sqrt{q_k} \mathbf{e}_{mk}^g\right) |s_k|^2 \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(c_2)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sqrt{q_k} \mathbf{e}_{mk}^g\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\underbrace{\left(\check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\check{\mathbf{v}}_{mk}^H \sqrt{q_k} \mathbf{e}_{mk}^g\right) \check{\mathbf{g}}_{mk}^H}_{I_1}\right\} \check{\mathbf{g}}_{mk} \\ &\quad + \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\underbrace{\sum_{n=1}^M \sum_{m \neq n}^M \left(\check{\mathbf{v}}_{nk}^H \check{\mathbf{g}}_{nk}\right)^* \left(\check{\mathbf{v}}_{mk}^H \sqrt{q_k} \mathbf{e}_{mk}^g\right) \check{\mathbf{g}}_{mk}^H}_{I_2}\right\} \check{\mathbf{g}}_{mk} = 0, \quad (63) \end{aligned}$$

where in step (c₁) we use the fact that the terms s_k s are independent variables. In addition, step (c₂) is due to the fact that s_k has zero mean and independent of Ω , and $\mathbb{E}\{|s_k|^2\} = 1$. Moreover, we have $I_1 = 0$, due to (12), and $I_2 = 0$ (as $\hat{\mathbf{g}}_{mk} \hat{\mathbf{g}}_{nk}^H = \mathbf{0}, \forall n \neq m$, based on the analysis in [44, Section 2.2]), we end up with $\mathbb{E}\{\check{\mathbf{g}}_{mk} \mathbf{e}_{nk}^g\} = \mathbf{0}, \forall n \neq m$. Next, approximating the Bayesian estimator with a linear estimator, we have:

$$\begin{aligned} \mathbb{E}\{A_1^* s_k^* A_6 | \check{\mathbf{g}}_{mk}\} &\approx \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\{A_1^* s_k^* A_6 \check{\mathbf{g}}_{mk}^H\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* s_k^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \tilde{\mathbf{g}}_{mk',s_{k'}}\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(d_1)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sqrt{q_k} \tilde{\mathbf{g}}_{mk}\right) |s_k|^2 \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(d_2)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sqrt{q_k} \tilde{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \sqrt{q_k} \mathbb{E}\left\{\underbrace{\left(\check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\check{\mathbf{v}}_{mk}^H \tilde{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H}_{I_3}\right\} \check{\mathbf{g}}_{mk} \\ &\quad + \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\underbrace{\sum_{n=1}^M \sum_{m \neq n}^M \left(\check{\mathbf{v}}_{nk}^H \check{\mathbf{g}}_{nk}\right)^* \left(\check{\mathbf{v}}_{mk}^H \sqrt{q_k} \tilde{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H}_{I_4}\right\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \sqrt{q_k} \mathbb{E}\left\{\left(\check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\check{\mathbf{v}}_{mk}^H \tilde{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \stackrel{(d_3)}{\approx} 0, \quad (64) \end{aligned}$$

where in step (d₁) we used the fact that s_k s are independent variables. Moreover, step (d₂) is due to the fact that s_k has zero mean and independent of Ω and $\mathbb{E}\{|s_k|^2\} = 1$. Again, note that the elements of $\check{\mathbf{g}}_{mk}$ are i.i.d. $CN(0, \beta_{mk} - \gamma_{mk})$ and are uncorrelated and independent of the input of the quantizer, i.e., $\mathbb{E}\{\check{\mathbf{g}}_{mk} \check{\mathbf{g}}_{mk}^H\} = \mathbf{0}$, and hence uncorrelated with the output of the quantizer, i.e., $\mathbb{E}\{\check{\mathbf{g}}_{mk} \tilde{\mathbf{g}}_{mk}^H\} = \mathbf{0}$. Next, $I_4 = 0$, which

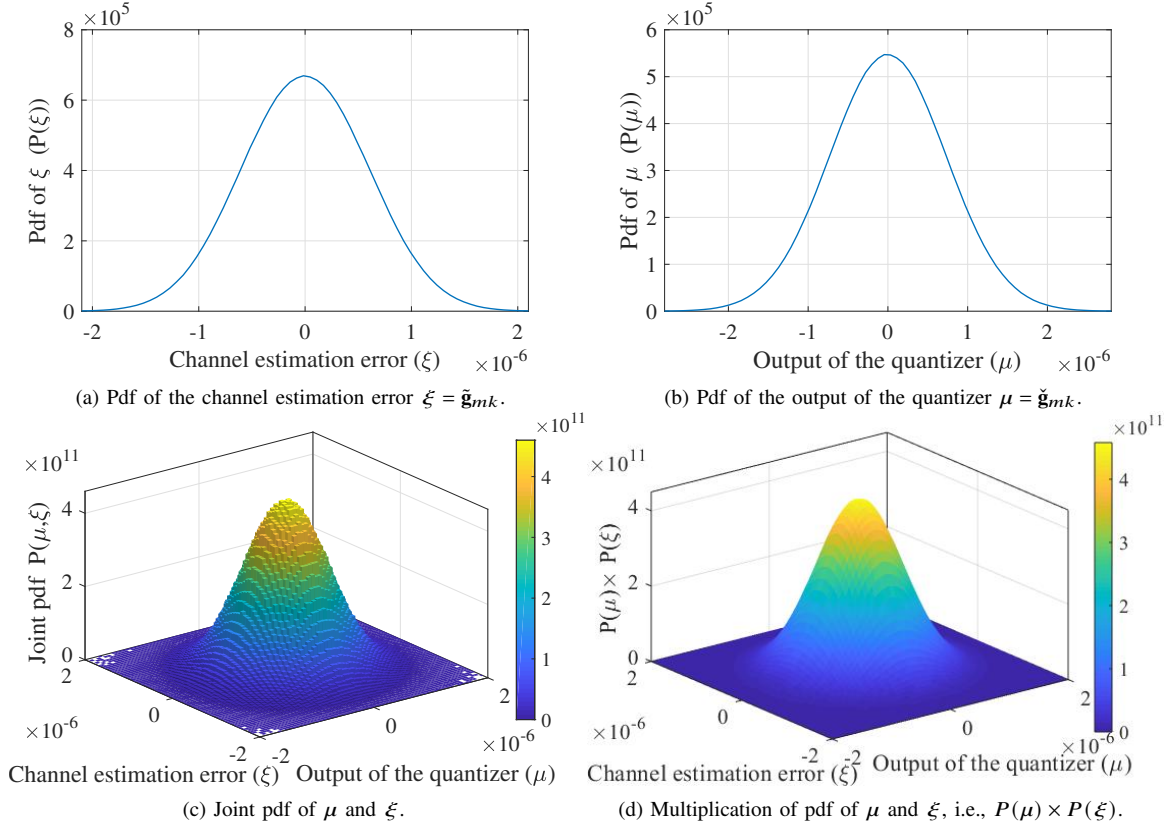


Figure 11. Pdf of $\mu = \check{\mathbf{g}}_{mk}$ and $\xi = \hat{\mathbf{g}}_{mk}$ with $\alpha = 1$ quantization bit.

is due to the fact that the equality $\mathbb{E}\{\hat{\mathbf{g}}_{nk}\hat{\mathbf{g}}_{nk}^H\} = \mathbf{0}, \forall n \neq m$ results in $\mathbb{E}\{\check{\mathbf{g}}_{nk}\check{\mathbf{g}}_{nk}^H\} = \mathbf{0}, \forall n \neq m$ (based on the analysis in [44, Section 2.2]), we end up with $\mathbb{E}\{\check{\mathbf{g}}_{mk}\mathbf{e}_{nk}^g{}^H\} = \mathbf{0}, \forall n \neq m$. Finally, the step (d_3) is due to the fact that the output of the quantizer $\check{\mathbf{g}}_{mk}$ and the channel estimation error $\hat{\mathbf{g}}_{mk}$ are uncorrelated and independent. The independence shown in Figs. 11a, 11b, 11c, 11d where the results show that $P(\mu, \xi) \approx P(\mu) \times P(\xi)$ while ξ and μ are the channel estimation error and the output of the quantizer, respectively. This completes the proof. \blacksquare

APPENDIX E: PROOF OF LEMMA 7

Firstly, note that B_2 and B_3 include s_k , which has zero mean and is independent of \mathcal{U} , hence, $\mathbb{E}\{B_2 + B_3|\mathcal{U}\} = 0$. Similarly, B_4 includes \mathbf{n}_m which has zero mean and independent of \mathcal{U} , resulting in $\mathbb{E}\{B_4|\mathcal{U}\} = 0$. Next, we need to show that $\mathbb{E}\{B_5|\mathcal{U}\} = 0$. We have $\mathbb{E}\{B_5|\mathcal{U}\} = \mathbb{E}\{\check{\mathbf{v}}_{mk}^H \mathbf{e}_m^y|\mathcal{U}\} = \check{\mathbf{v}}_{mk}^H \mathbb{E}\{\mathbf{e}_m^y|\mathcal{U}\}$. Therefore, we need to prove that $\mathbb{E}\{\mathbf{e}_m^y|\mathcal{U}\} = 0$. Note that $\mathbb{E}\{\mathbf{e}_m^y|\check{\mathbf{v}}_{mk}\} = 0$ is equivalent to $\mathbb{E}\{\mathbf{e}_m^y|\check{\mathbf{g}}_{mk}\} = 0$, as $\check{\mathbf{v}}_{mk}$ is a function of only $\check{\mathbf{g}}_{mk}$. Next, approximating the Bayesian estimator with a linear estimator, we have [42], [45]

$$\mathbb{E}\{\mathbf{e}_m^y|\check{\mathbf{g}}_{mk}\} \stackrel{(e)}{\approx} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\{\mathbf{e}_m^y \check{\mathbf{g}}_{mk}^H\} \check{\mathbf{g}}_{mk}. \quad (65)$$

Note that the approximation in step (e) is widely used in literature in the concept of uniform quantization [42], [44]–[46]. Next, we aim to calculate the term $\mathbb{E}\{\mathbf{e}_m^y \check{\mathbf{g}}_{mk}^H\}$. As a remind, the term \mathbf{e}_m^y denotes the quantization distortion while we quantize the received signal, i.e., \mathbf{y}_m , in (32), and the term

$\check{\mathbf{g}}_{mk}$ is the output of quantizer where we quantize the estimated channel, i.e., $\hat{\mathbf{g}}_{mk}$, in (31). It is easy to show that the input of these two quantizers are uncorrelated. More precisely, using (2) and (3), we have

$$\mathbb{E}\{\mathbf{y}_m \hat{\mathbf{g}}_{mk}^H\} = c_{mk} \mathbb{E}\left\{\left(\sqrt{\rho} \sum_{k=1}^K \mathbf{g}_{mk} \sqrt{q_k} s_k + \mathbf{n}_m\right) \left(\sqrt{\tau_p \rho} \sum_{k'=1}^K \mathbf{g}_{mk'} \boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k + \boldsymbol{\Omega}_{p,m} \boldsymbol{\phi}_k\right)^H\right\} = \mathbf{0}. \quad (66)$$

This shows that the inputs of quantizers when we separately quantize the estimated channel and the received signal are uncorrelated. Moreover, using the analysis in [44, Section 2.2], when we quantize two uncorrelated signals x_1 and x_2 using the uniform quantizer (i.e., $Q(x_1) = x_1 + n_1$, and $Q(x_2) = x_2 + n_2$), we have $\mathbb{E}\{Q(x_1)n_2\} = \mathbb{E}\{Q(x_2)n_1\} = 0$. Exploiting this fact and (59), we have

$$\mathbb{E}\{\mathbf{e}_m^y \check{\mathbf{g}}_{mk}^H\} = \mathbf{0}. \quad (67)$$

Finally, substituting (60) into (58), we get $\mathbb{E}\{\mathbf{e}_m^y|\check{\mathbf{v}}_{mk}\} \approx 0$, which completes the proof. \blacksquare

APPENDIX F: PROOF OF LEMMA 8

Firstly, note that $\mathbb{E}\{s_k^* (B_2 + B_4 + B_5)|\mathcal{U}\} = 0$, as none of the terms B_4 , and B_5 include s_k , which has zero mean

and is independent of \mathcal{U} . In the following, we prove that $\mathbb{E}\{s_k^*(B_3)|\mathcal{U}\} = 0$. Next we have:

$$\begin{aligned} & \mathbb{E}\{s_k^* B_3 | \check{\mathbf{g}}_{mk}\} \approx \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\{s_k^* B_3 \check{\mathbf{g}}_{mk}^H\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{s_k^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \check{\mathbf{g}}_{mk'} s_{k'}\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \sqrt{q_k} \mathbb{E}\left\{\sum_{m=1}^M \left(\check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H |s_k|^2\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(f_1)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \sqrt{q_k} \mathbb{E}\left\{\sum_{m=1}^M \left(\check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \stackrel{(f_2)}{=} 0, \quad (68) \end{aligned}$$

where in step (f₁) we used the fact that s_k s are independent variables. The elements of $\check{\mathbf{g}}_{mk}$ are uncorrelated and independent of the input of the quantizer and $\mathbb{E}\{\check{\mathbf{g}}_{mk} \check{\mathbf{g}}_{mk}^H\} = \mathbf{0}$. The step (f₂) comes from the fact that the channel estimate $\check{\mathbf{g}}_{mk}$ and the channel estimation error $\check{\mathbf{g}}_{mk}$ are uncorrelated and independent. This completes the proof. ■

APPENDIX G: PROOF OF LEMMA 9

It can be easily shown that $\mathbb{E}\{B_1^* s_k^* (B_2 + B_4 + B_5) | \mathcal{U}\} = 0$, as none of the terms B_2 , B_4 and B_5 include s_k , which has zero mean and is independent of \mathcal{U} . In the following, we show that $\mathbb{E}\{B_1^* s_k^* (B_3) | \mathcal{U}\} = 0$. First, we aim to show that $\mathbb{E}\{B_1^* s_k^* B_3 | \check{\mathbf{g}}_{mk}\} \approx 0$, as follows. Approximating the Bayesian estimator with a linear estimator, we have:

$$\begin{aligned} & \mathbb{E}\{B_1^* s_k^* B_3 | \check{\mathbf{g}}_{mk}\} \approx \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\{B_1^* s_k^* B_3 \check{\mathbf{g}}_{mk}^H\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* s_k^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sum_{k'=1}^K \sqrt{q_{k'}} \check{\mathbf{g}}_{mk'} s_{k'}\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(g_1)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sqrt{q_k} \check{\mathbf{g}}_{mk}\right) |s_k|^2 \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &\stackrel{(g_2)}{=} \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \sqrt{q_k} \check{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \\ &= \sigma_{\check{\mathbf{g}}_{mk}}^{-2} \sqrt{q_k} \mathbb{E}\left\{\left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right)^* \left(\sum_{m=1}^M \check{\mathbf{v}}_{mk}^H \check{\mathbf{g}}_{mk}\right) \check{\mathbf{g}}_{mk}^H\right\} \check{\mathbf{g}}_{mk} \stackrel{(g_3)}{\approx} 0, \quad (69) \end{aligned}$$

where in step (g₁) we used the fact that s_k s are independent variables. Moreover, step (g₂) is due to the fact that s_k has zero mean and independent of Ω and $\mathbb{E}\{|s_k|^2\} = 1$. The step (g₃) comes from the fact that the channel estimate $\check{\mathbf{g}}_{mk}$ and the channel estimation error $\check{\mathbf{g}}_{mk}$ are uncorrelated and independent. This completes the proof. ■

REFERENCES

- [1] M. Bashar, H. Q. Ngo, A. G. Burr, D. Maryopi, K. Cumanan, and E. G. Larsson, "On the performance of backhaul constrained cell-free massive MIMO with linear receivers," in *Proc. IEEE Asilomar*, Oct. 2018, pp. 1–7.
- [2] C. Pan, M. Elkashlan, J. Wang, J. Yuan, and L. Hanzo, "User-centric C-RAN architecture for ultra-dense 5G networks: Challenges and methodologies," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 14–20, Jun. 2018.
- [3] A. 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. IEEE PIMRC*, Sep. 2018, pp. 1–5.
- [4] M. Ding, P. Wang, D. López-Pérez, G. Mao, and Z. Lin, "Performance impact of LoS and NLoS transmissions in dense cellular networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 2365–2380, Mar. 2016.
- [5] M. Peng, C. Wang, J. Li, H. Xiang, and V. Lau, "Recent advances in underlay heterogeneous networks: Interference control, resource allocation, and self-organization," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, p. 700–729, Jun. 2015.
- [6] M. Peng, S. Yan, K. Zhang, and C. Wang, "Fog-computing-based radio access networks: issues and challenges," *IEEE Network*, vol. 30, no. 4, pp. 46–53, Jul. 2016.
- [7] *Common Public Radio Interface (CPRI)*. Interface Specification V6.0, Aug 2013.
- [8] T. Quek, M. Peng, O. Simeone, and W. Yu, *Cloud Radio Access Networks*. Cambridge University Press, Feb 2017.
- [9] A. Checko, H. L. Christiansen, Y. Yan, L. Scolari, G. Kardaras, M. S. Berger, and L. Dittmann, "Cloud RAN for mobile networks- a technology overview," *Commun. Surveys Tuts.*, vol. 17, no. 1, pp. 405–426, Mar. 2015.
- [10] D. Samardzija, J. Pastalan, M. MacDonald, S. Walker, and R. Valenzuela, "Compressed transport of baseband signals in radio access networks," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3216–3225, Jun. 2012.
- [11] M. Bashar, *Cell-free Massive MIMO and Millimeter Wave Channel Modelling for 5G and Beyond*. Ph.D. dissertation, University of York, United Kingdom, 2019.
- [12] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Trans. Inf. Theory*, vol. 22, no. 1, pp. 1–10, Jan. 1976.
- [13] N. Liu and W. Kang, "A new achievability scheme for downlink multicell processing with finite backhaul capacity," in *Proc. IEEE ISIT*, Jul. 2014, pp. 1–5.
- [14] 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.
- [15] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO communications," *EURASIP 1022 J. Wireless Commun. Netw.*, pp. 1–13, Aug. 2019.
- [16] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, M. Debbah, and P. Xiao, "Max-min SINR of cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6796–6814, Oct. 2019.
- [17] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, and M. Debbah, "Cell-free massive MIMO with limited backhaul," in *Proc. IEEE ICC*, May 2018, pp. 1–7.
- [18] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "On the energy efficiency of limited-backhaul cell-free massive MIMO," in *Proc. IEEE ICC*, Jul. 2019, pp. 1–7.
- [19] M. Bashar, A. Akbari, K. Cumanan, H. Q. Ngo, A. G. Burr, P. Xiao, M. Debbah, and J. Kittler, "Exploiting deep learning in limited-fronthaul cell-free massive MIMO uplink," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1678–1697, Jun. 2020.
- [20] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "Enhanced max-min SINR for uplink cell-free massive MIMO systems," in *Proc. IEEE ICC*, May 2018, pp. 1–6.
- [21] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, L. Hanzo, and P. Xiao, "On the performance of cell-free massive mimo relying on adaptive noma/oma mode-switching," *IEEE Trans. Commun.*, vol. 68, no. 2, pp. 792–810, Nov. 2019.
- [22] M. Bashar, A. Akbari, K. Cumanan, H. Q. Ngo, A. G. Burr, P. Xiao, and M. Debbah, "Deep learning-aided finite-capacity fronthaul cell-free massive MIMO with zero forcing," in *Proc. IEEE ICC*, Jun. 2020, pp. 1–6.
- [23] T. Rep, *C-RAN: the road towards green RAN*. China Mobile Research Institute, Oct 2011.
- [24] 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.
- [25] J. Max, "Quantizing for minimum distortion," *IEEE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.
- [26] P. Zillmann, "Relationship between two distortion measures for memoryless nonlinear systems," *IEEE Signal Process. Lett.*, vol. 17, no. 11, pp. 917–920, Feb. 2010.
- [27] J. Bussgang, *Crosscorrelation functions of amplitude-distorted gaussian signals*. Research Laboratory of Electronics, MIT, Mar 1952.

- [28] H. Q. Ngo, L. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. and Net.*, vol. 2, no. 1, pp. 25–39, Mar. 2017.
- [29] C. E. Spurgeon and J. Zimmerman, *Ethernet: The Definitive Guide Paperback*. O'Reilly Media, Apr. 2014.
- [30] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [31] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of massive MIMO*. Cambridge University Press, Nov. 2016.
- [32] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, E. G. Larsson, and P. Xiao, "Energy efficiency of the cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Trans. Green Commun. and Net.*, vol. 3, no. 4, pp. 971–987, Dec. 2019.
- [33] A. Burr, S. F. Islam, J. Zhao, and M. Bashar, "Cell-free massive MIMO with multi-antenna access points and user terminals," in *IEEE Asilomar*, Nov. 2020, pp. 1–5.
- [34] M. Bashar, A. G. Burr, K. Haneda, K. Cumanan, M. M. Molu, M. Khalily, and P. Xiao, "Evaluation of low complexity massive MIMO techniques under realistic channel conditions," *IEEE Trans. Veh. Technol.*, vol. 68, no. 2, pp. 9297–9302, Sep. 2019.
- [35] M. Bashar, A. Burr, K. Haneda, and K. Cumanan, "Robust user scheduling with cost 2100 channel model for massive mimo networks," in *Proc. IEEE GLOBECOM*, Dec. 2018, pp. 1786–1792.
- [36] M. Bashar, A. G. Burr, K. Haneda, and K. Cumanan, "Robust user scheduling with COST 2100 channel model for Massive MIMO networks," *IET microwave antenna and propagation*, vol. 12, pp. 1786 – 1792, Aug. 2018.
- [37] E. Nayebi, A. Ashikhmin, T. L. Marzetta, and B. D. Rao, "Performance of cell-free massive MIMO systems with MMSE and LSFD receivers," in *IEEE Asilomar*, Nov. 2016, pp. 1–6.
- [38] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "On the uplink max-min SINR of cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2021–2036, Jan. 2019.
- [39] M. Bashar, A. G. Burr, K. Haneda, and K. Cumanan, "Mixed quality of service in cell-free massive MIMO," *IEEE Commun. Letters*, vol. 22, no. 7, pp. 1494–1497, Jul. 2018.
- [40] R. G. Gallager, *Principles of Digital Communication*. Cambridge University Press, Jan 2008.
- [41] J. G. Proakis, *Digital Communications*. McGraw-Hill, 1995.
- [42] A. Kakkavas, J. Munir, A. Mezghani, H. Brunner, and J. A. Nossek, "Weighted sum rate maximization for multiuser MISO systems with low resolution digital to analog converters," in *Proc. IEEE WSA*, Mar. 2016, pp. 1–8.
- [43] A. G. Burr, M. Bashar, and D. Maryopi, "Cooperative access networks: Optimum fronthaul quantization in distributed massive MIMO and cloud RAN," in *Proc. IEEE VTC*, Jun. 2018, pp. 1–7.
- [44] A. Mezghani, *Information theoretic Analysis and Signal Processing Techniques for Quantized MIMO Communications*. Ph.D. dissertation, Technical University of Munich, Germany, Dec. 2015.
- [45] A. Kakkavas, J. Munir, A. Mezghani, and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," in *Proc. IEEE ISIT*, Jul. 2012, pp. 1–5.
- [46] Y. Li, C. Tao, G. Seco-Granados, A. Mezghani, A. L. Swindlehurst, and L. Liu, "Channel estimation and performance analysis of one-bit massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 15, pp. 4075–4089, Aug. 2017.
- [47] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-time signal processing*. Prentice-hall Englewood Cliffs, 1989.
- [48] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, and E. Larsson, "Energy efficiency of the cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Trans. Green Commun. and Net.*, vol. 3, no. 4, pp. 971–987, Dec. 2019.
- [49] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.
- [50] A. J. Fehske, P. Marsch, and G. P. Fettweis, "Bit per joule efficiency of cooperating base stations in cellular networks," in *Proc. IEEE Globecom Workshops*, Dec. 2010, pp. 1406–1411.
- [51] O. Onireti, F. Heliot, and M. A. Imran, "On the energy efficiency-spectral efficiency trade-off of distributed MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 9, pp. 3741–3753, Sep. 2013.
- [52] H. Ahmadi, A. Farhang, N. Marchetti, and A. MacKenzie, "A game theoretic approach for pilot contamination avoidance in massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 5, no. 1, pp. 12–15, Feb. 2016.
- [53] O. Onireti, F. Heliot, and M. A. Imran, "On the energy efficiency-spectral efficiency trade-off in the uplink of CoMP system," *IEEE Trans. Commun.*, vol. 11, no. 2, pp. 556–561, Feb. 2012.
- [54] Z. Chen and E. Björnson, "Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry," *IEEE Trans. Commun.*, pp. 1–14, Jun. 2018.