



**QUEEN'S
UNIVERSITY
BELFAST**

Max-Min Rate of Cell-Free Massive MIMO Uplink with Optimal Uniform Quantization

Bashar, M., Cumanan, K., Burr, A. G., Ngo, H. Q., Debbah, M., & Xiao, P. (2019). Max-Min Rate of Cell-Free Massive MIMO Uplink with Optimal Uniform Quantization. *IEEE Transactions on Communications*, 67(10), 6796-6815. [8756286]. <https://doi.org/10.1109/TCOMM.2019.2926706>, <https://doi.org/10.1109/TCOMM.2019.2926706>

Published in:

IEEE Transactions on Communications

Document Version:

Peer reviewed version

Queen's University Belfast - Research Portal:

[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

© 2019 IEEE.

This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Max-Min Rate of Cell-Free Massive MIMO Uplink with Optimal Uniform Quantization

Manijeh Bashar, *Student Member, IEEE*, Kanapathippillai Cumanan, *Member, IEEE*, Alister G. Burr, *Senior Member, IEEE*, Hien Quoc Ngo, *Member, IEEE*, Merouane Debbah, *Fellow, IEEE*, and Pei Xiao, *Senior Member, IEEE*

Abstract—Cell-free Massive multiple-input multiple-output (MIMO) is considered, where distributed access points (APs) multiply the received signal by the conjugate of the estimated channel, and send back a quantized version of this weighted signal to a central processing unit (CPU). For the first time, we present a performance comparison between the case of perfect fronthaul links, the case when the quantized version of the estimated channel and the quantized signal are available at the CPU, and the case when only the quantized weighted signal is available at the CPU. The Bussgang decomposition is used to model the effect of quantization. The max-min problem is studied, where the minimum rate is maximized with the power and fronthaul capacity constraints. To deal with the non-convex problem, the original problem is decomposed into two sub-problems (referred to as receiver filter design and power allocation). Geometric programming (GP) is exploited to solve the power allocation problem whereas a generalized eigenvalue problem is solved to design the receiver filter. An iterative scheme is developed and the optimality of the proposed algorithm is proved through uplink-downlink duality. A user assignment algorithm is proposed which significantly improves the performance. Numerical results demonstrate the superiority of the proposed schemes.

Keywords: Cell-free Massive MIMO, generalized eigenvalue, geometric programming, limited fronthaul.

I. INTRODUCTION

Cell-free Massive multiple-input multiple-output (MIMO) has been recognized as a potential technology for 5th Generation (5G) systems, where large number of distributed access points (APs) serve a much smaller number of users, and hence, uniformly good service performance for all users is ensured [1]–[5]. Interestingly, in [2], it is shown that the system performance of cell-free Massive MIMO depends only on large-scale fading, i.e., the small-scale fading and noise can

be averaged out when number of APs is large. In [6] a user-centric approach is proposed where each user is served by a small number of APs. Cell-free Massive MIMO effectively implements a user-centric approach [7]. In [8], the authors consider distributed Massive MIMO in a multi-cell manner, which is different from cell-free massive MIMO (as there is no cell concept).

One of the main issues of cell-free Massive MIMO systems which requires more investigation is the limited-capacity fronthaul links from the APs to a central processing unit (CPU). The assumption of infinite fronthaul in [1], [2], [9] is not realistic in practice. The fronthaul requirements for Massive MIMO systems, including small-cell and macro-cell base stations (BSs) have been investigated in [10]. The fronthaul load is the main challenge in any distributed antenna systems [10], [11]. First, we consider the case where all APs send back the quantized version of the minimum mean-square error (MMSE) estimate of the channel from each user and the quantized version of the received signal to the CPU. We next study the case when each AP multiplies the received signal by the conjugate of the estimated channel from each user, and sends back a quantized version of this weighted signal to the CPU. We derive the total number of bits for both cases and show that given the same fronthaul capacity for both cases, the relative performance of the aforementioned cases depends on the number of antennas at each AP, the total number of APs and the channel coherence time. A new approach is provided to 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 distributed base stations, here for simplicity (and hence improved scalability) we assume simple uniform quantization. We exploit the Bussgang decomposition [12] to model the effect of quantization.

In [1], [2], [13] the authors propose that the APs design the linear receivers based on the estimated channels, and that this is carried out locally at the APs. Hence, the CPU exploits only the statistics of the channel for data detection. However, in this paper, we propose to exploit a new receiver filter at the CPU to improve the performance of cell-free Massive MIMO systems. The coefficients of the proposed receiver filter are designed based on only the statistics of the channel, which is different from the linear receiver at the APs. The proposed receiver filter provides more freedom in the design parameters and hence, significantly improves the performance of the uplink

M. Bashar, K. Cumanan and A. G. Burr are with the Department of Electronic Engineering, University of York, Heslington, York, UK. e-mail: {mb1465, kanapathippillai.cumanan, alister.burr}@york.ac.uk. M. Bashar is also with home of the 5G Innovation Centre, Institute for Communication Systems, University of Surrey, UK. e-mail: m.bashar@surrey.ac.uk. H. Q. Ngo is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, UK. e-mail: hien.ngo@qub.ac.uk. M. Debbah is with the Large Networks and Systems Group (LANEAS), CentraleSupélec, Université Paris-Saclay, Gif-sur-Yvette 91192, France, and also with the Mathematical and Algorithmic Sciences Lab, Huawei Technologies Co., Ltd., Boulogne-Billancourt 92100, France. e-mail: m.debbah@centralesupelec.fr and merouane.debbah@huawei.com. Pei Xiao is with home of the 5G Innovation Centre, Institute for Communication Systems, University of Surrey, UK. e-mail: p.xiao@surrey.ac.uk.

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

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

of cell-free Massive MIMO. The work in [14] presents a large scale fading decoding (LSFD) postcoding vector and power allocation scheme to solve max-min signal-to-interference-plus-noise ratio (SINR) problem. However, note that the work in [14] does not present any iterative algorithm to jointly solve power minimization problem and LSFD postcoding vector design. In [15], the authors use a bisection search approach to solve the power allocation problem. Next, MMSE receiver is exploited to determine the LSFD postcoding vectors. However, in our work, we exploit geometric programming (GP) to optimally solve the power allocation problem. Moreover, we prove that the proposed algorithm is optimal whereas the authors in [14] does not present any proof of optimality. In addition, the work in [14] does not consider any quantization errors whereas our work investigates the realistic assumption of limited-capacity fronthaul links.

We next investigate an uplink max-min rate problem with limited fronthaul links. In particular, the receiver filter coefficients and power allocation are optimized in the proposed scheme whereas the work in [2] only considered user power allocations. In particular, we propose a new approach to solve this max-min problem. A similar max-min rate problem based on SINR known as *SINR balancing* in the literature has been considered for cognitive radio networks in [16], [17]. In [18], [19], the authors consider MIMO systems and study the problem of max-min user rate to maximize the smallest user rate. The problem of uplink-downlink duality has been investigated in [20], [21]. Note that none of the previous works on uplink-downlink duality consider Massive MIMO and the SINR formula in single-cell does not include any pilot contamination, channel estimation and quantization errors. To tackle the non-convexity of the original max-min rate problem, we propose to decouple the original problem into two sub-problems, namely, receiver filter coefficient design, and power allocation. We next show that the receiver filter coefficient design problem may be solved through a generalized eigenvalue problem [22]. Moreover, the user power allocation problem is solved through standard GP [23]. We present an iterative algorithm to alternately solve each sub-problem while one of the design parameters is fixed. Next an uplink-downlink duality for cell-free Massive MIMO system with limited fronthaul links is established to validate the optimality of the proposed scheme. We finally propose an efficient user assignment algorithm and show that further improvement is achieved by the proposed user assignment algorithm.

The idea of exploiting an iterative algorithm to design the receiver filter and power coefficients in cell-free Massive MIMO system has been proposed in [24]. However, in [24], the authors investigate a cell-free Massive MIMO with single-antenna APs and perfect fronthaul links whereas in the this work we exploit a cell-free Massive MIMO system with multiple-antenna APs and limited-capacity fronthaul links. Furthermore, in this work, unlike [24], user assignment is investigated. The contributions of the paper are summarized as follows:

1. We consider two cases: i) the quantized versions of the channel estimates and the received signals at the APs are available at the CPU and ii) the quantized versions of

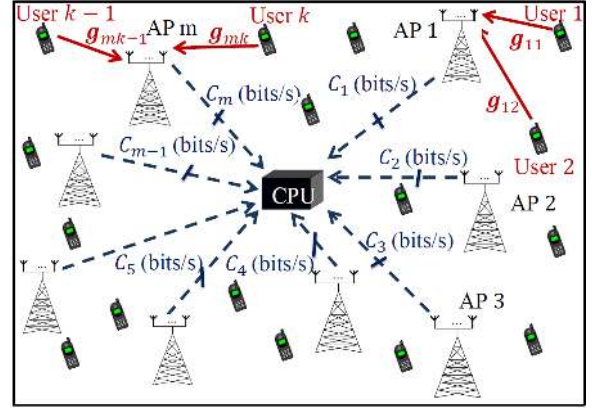


Figure 1. The uplink of a cell-free Massive MIMO system with K single-antenna users and M APs. Each AP is equipped with N antennas. The solid lines denote the uplink channels and the dashed lines present the limited-capacity fronthaul links from the APs to the CPU.

processed signals at the APs are available at the CPU. The corresponding achievable rates are derived by using the Use-and-then-Forget (UaF) bounding technique taking into account the effects of channel estimation error and quantization error.

2. We make use of the Bussgang decomposition to model the effect of quantization and present the analytical solution to find the optimal step size of the quantizer.
3. We propose a max-min fairness power control problem which maximizes the smallest of all user rates under the per-user power and fronthaul capacity constraints. To solve this problem, the original problem is decomposed into two sub-problems and an iterative algorithm is developed. The optimality of the proposed algorithm is proved through establishing the uplink-downlink duality for the cell-free Massive MIMO system with limited fronthaul link capacities.
4. A novel and efficient user assignment algorithm based on the capacity of fronthaul links is proposed which results in significant performance improvement.

The rest of the paper is organized as follows. Section II describes the system model and Section III provides performance analysis. The proposed max-min rate scheme is presented in Section IV and the convergence is provided in Section V. The optimality of the proposed scheme is proved in Section VI. Section VII investigates the proposed user assignment algorithm. Numerical results are presented in Section VIII, and finally Section IX concludes the paper.

II. SYSTEM MODEL

We consider uplink transmission in a cell-free Massive MIMO system with M APs and K single-antenna users randomly distributed in a large area. 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 $\mathbf{g}_{mk} = \sqrt{\beta_{mk}} \mathbf{h}_{mk}$, where β_{mk} denotes the large-scale fading, the elements of \mathbf{h}_{mk} are independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ random variables, and represents the small-scale fading [2].

A. Uplink Channel Estimation

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 for the k th user. After performing a de-spreading operation, the MMSE estimate of the channel coefficient between the k th user and the m th AP is given by [2]

$$\hat{\mathbf{g}}_{mk} = c_{mk} \left(\sqrt{\tau_p p_p} \mathbf{g}_{mk} + \sqrt{\tau_p p_p} \sum_{k' \neq k}^K \mathbf{g}_{mk'} \phi_{k'}^H \phi_k + \mathbf{W}_{p,m} \phi_k \right), \quad (1)$$

where $\mathbf{W}_{p,m} \in \mathbb{C}^{M \times K}$ denotes the noise sequence at the m th AP whose elements are i.i.d. $CN(0, 1)$, p_p represents the normalized SNR of each pilot sequence (which we define in Section VIII), and $c_{mk} = \frac{\sqrt{\tau_p p_p} \beta_{mk}}{\tau_p p_p \sum_{k'=1}^K \beta_{mk'} |\phi_{k'}^H \phi_k|^2 + 1}$. Note that, as in [2], we assume that the large-scale fading, β_{mk} , is known.¹

B. Optimal Quantization Model

Based on Busgang's theorem [12], a nonlinear output of a quantizer can be represented as a linear function as follows:

$$Q(z) = h(z) = az + n_d, \quad \forall k, \quad (2)$$

where a is a constant value and n_d refers to the distortion noise which is uncorrelated with the input of the quantizer, z . The term a is given by

$$a = \frac{\mathbb{E}\{zh(z)\}}{\mathbb{E}\{z^2\}} = \frac{1}{p_z} \int_{\mathcal{Z}} zh(z) f_z(z) dz, \quad (3)$$

where $p_z = \mathbb{E}\{|z|^2\} = \mathbb{E}\{z^2\}$ is the power of z and we drop absolute value as z is a real number, and $f_z(z)$ is the probability distribution function of z . Denote by²

$$b = \frac{\mathbb{E}\{h^2(z)\}}{\mathbb{E}\{z^2\}} = \frac{1}{p_z} \int_{\mathcal{Z}} h^2(z) f_z(z) dz. \quad (4)$$

Then, the signal-to-distortion noise ratio (SDNR) is

$$\text{SDNR} = \frac{\mathbb{E}\{(az)^2\}}{\mathbb{E}\{n_d^2\}} = \frac{p_z a^2}{p_z (b - a^2)} = \frac{a^2}{b - a^2}, \quad (5)$$

According to [12], [27], [28], the midrise uniform quantizer function $h(z)$ is given by

$$h(z) = \begin{cases} -\frac{L-1}{2}\Delta & z \leq -\left(\frac{L}{2} + 1\right)\Delta, \\ \left(l + \frac{1}{2}\right)\Delta & l\Delta \leq z \leq (l+1)\Delta, l = -\frac{L}{2} + 1, \dots, \frac{L}{2} - 2, \\ \frac{L-1}{2}\Delta & z \geq \left(\frac{L}{2} - 1\right)\Delta, \end{cases} \quad (6)$$

where Δ is the step size of the quantizer and $L = 2^\alpha$, where α is number of quantization bits.

¹The large-scale fading β_{mk} changes very slowly with time. Compared to the small-scale fading, the large-scale fading changes much more slowly, some 40 times slower according to [25], [26]. Therefore, β_{mk} can be estimated in advance. One simple way is that the AP takes the average of the power level of the received signal over a long time period. A similar technique for collocated Massive MIMO is discussed in Section III-D of [26].

²Equations (2)-(4) come from [12] but we include them here for completeness, and to define the terms we used.

Lemma 1. The terms a and b are obtained as follows:

$$a = \Delta \sqrt{\frac{2}{\pi p_z} \left(\sum_{l=1}^{\frac{L}{2}-1} e^{-\frac{l^2 \Delta^2}{2p_z}} + 1 \right)}, \quad b = \frac{\Delta^2}{p_z} \left(\frac{1}{4} + 4 \sum_{l=1}^{\frac{L}{2}-1} l Q\left(\frac{l\Delta}{\sqrt{p_z}}\right) \right), \quad (7)$$

where $Q(x)$ is the Q -function and is given by $Q(x) = \frac{1}{2} \text{erfc}\left(\frac{x}{\sqrt{2}}\right)$, where erfc refers to the complementary error function [29].

Proof: Please refer to Appendix A. \blacksquare

In general, terms a and b are functions of the power of the quantizer input, p_z . 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 $\dot{z} = \frac{z}{\sqrt{p_z}}$, we have

$$Q(z) = \sqrt{p_z} Q(\dot{z}) = \dot{a} \sqrt{p_z} \dot{z} + \sqrt{p_z} \dot{n}_d = \dot{a} z + \sqrt{p_z} \dot{n}_d. \quad (8)$$

Note that (8) enables us to find the optimum step size of the quantizer and the corresponding \dot{a} . Note that for the case of $\dot{\Delta} = \frac{1}{\sqrt{p_z}} \Delta$, we have $\dot{a} = a$, $\dot{b} = b$. The optimal step size of the quantizer is obtained by solving the following maximization problem:

$$\begin{aligned} \Delta_{\text{opt}} &= \arg \max_{\Delta} \text{SDNR} = \arg \max_{\Delta} \frac{a^2}{b - a^2} \stackrel{I_1}{=} \arg \max_{\dot{\Delta}} \frac{\dot{a}^2}{\dot{b} - \dot{a}^2} \\ &= \arg \max_{\dot{\Delta}} \frac{\dot{a}^2}{\dot{b}} \stackrel{I_2}{=} \arg \max_{\dot{\Delta}} \left(\frac{2\dot{\Delta}^2}{\pi} \left(\sum_{l=1}^{\frac{L}{2}-1} \exp\left(-\frac{l^2 \dot{\Delta}^2}{2}\right) + 1 \right)^2 \right) \\ &= \arg \max_{\dot{\Delta}} \left(\frac{\left(\sum_{l=1}^{\frac{L}{2}-1} 2 \exp\left(-\frac{l^2 \dot{\Delta}^2}{2}\right) + 1 \right)^2}{\frac{1}{4} + 4 \sum_{l=1}^{\frac{L}{2}-1} l Q(l\dot{\Delta})} \right), \end{aligned} \quad (9)$$

where in step I_1 , we have used (8) and step I_2 comes from results in Lemma 1. Moreover, note that $\dot{\Delta} = \frac{\Delta}{\sqrt{p_z}}$. The maximization problem in (9) can be solved through a one-dimensional search over $\dot{\Delta}$ for a given L in a symbolic mathematics tool such as Mathematica. For the input \dot{z} with $p_z = 1$, the optimal step size of the quantizer $\dot{\Delta}_{\text{opt}}$, the resulting distortion noise power, $p_{\dot{n}_d} = \mathbb{E}\{\dot{n}_d^2\} = \dot{b} - \dot{a}^2$, and the resulting \dot{a} are summarized in Table I.

Remark 1. Interestingly, the optimal values for quantization step size, $\dot{\Delta}_{\text{opt}}$, given in Table I, are exactly the same as the optimal values of quantization step size in [30]. In [30], J. Max did not provide any analytical solution to solve the problem of minimizing the mean-squared distortion (or mean-squared error (MSE)) and to obtain the optimal quantization step size. Moreover, J. Max only calculates the optimal step size and the resulting distortion power for $\alpha = 1, \dots, 5$ whereas Lemma 1 enables us to calculate the optimal step size and the resulting distortion power for any quantization resolution. Values for α up to 10 are listed in Table I.

Table I
THE OPTIMAL STEP SIZE AND DISTORTION POWER OF A UNIFORM
QUANTIZER WITH BUSSGANG DECOMPOSITION.

α	$\hat{\Delta}_{\text{opt}}$	$\sigma_e^2 = p n_d = b - \hat{a}^2$	\hat{a}
1	1.596	0.2313	0.6366
2	0.9957	0.10472	0.88115
3	0.586	0.036037	0.96256
4	0.3352	0.011409	0.98845
5	0.1881	0.003482	0.996505
6	0.1041	0.0010389	0.99896
7	0.0568	0.0003042	0.99969
8	0.0307	0.0000876	0.999912

C. 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 represented 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 from the k th user, respectively, where ρ represents the normalized uplink SNR (see Section VIII for more details). The $N \times 1$ received signal at the m th AP from all users is given by

$$\mathbf{y}_m = \sqrt{\rho} \sum_{k=1}^K \mathbf{g}_{mk} \sqrt{q_k} s_k + \mathbf{n}_m, \quad (10)$$

where each element of $\mathbf{n}_m \in \mathbb{C}^{N \times 1}$, $n_{n,m} \sim \mathcal{CN}(0, 1)$ is the noise at the m th AP.

III. PERFORMANCE ANALYSIS

In this section, the performance analysis for two cases is presented. First we consider the case when the quantized versions of the channel estimates and the received signals are available at the CPU. Next, it is assumed that only the quantized versions of the weighted signals are available at the CPU. The corresponding achievable rates are derived by exploiting the UaF bounding technique.

Case 1. Quantized Estimate of the Channel and Quantized Signal Available at the CPU: The m th AP quantizes the terms $\hat{\mathbf{g}}_{mk}$, $\forall k$, and \mathbf{y}_m , and forwards the quantized channel state information (CSI) and the quantized signals in each symbol duration to the CPU. The quantized signal can be obtained as:

$$Q([\mathbf{y}_m]_n) = \hat{a}[\mathbf{y}_m]_n + [\mathbf{e}_m^y]_n = [\zeta_m]_n + j[\mathbf{v}_m]_n, \quad \forall m, n, \quad (11)$$

where $[\mathbf{e}_m^y]_n$ refers to the quantization error, and $[\zeta_m]_n$ and $[\mathbf{v}_m]_n$ are the real and imaginary parts of the output of the quantizer, respectively. Note that we separately quantize the imaginary and real parts of the input of the quantizer. Note that $[\mathbf{x}]_n$ represents the n th element of vector \mathbf{x} . The analog-to-digital converter (ADC) quantizes the real and imaginary parts of $[\mathbf{y}_m]_n$ with α bits each, which introduces quantization errors $[\mathbf{e}_m^y]_n$ to the received signals [31], [32]. In addition, the ADC quantizes the MMSE estimate of CSI as:

$$Q([\hat{\mathbf{g}}_{mk}]_n) = \hat{a}[\hat{\mathbf{g}}_{mk}]_n + [\mathbf{e}_{mk}^g]_n = [\varrho_{mk}]_n + j[\kappa_{mk}]_n, \quad \forall k, n, \quad (12)$$

where $[\varrho_{mk}]_n$ and $[\kappa_{mk}]_n$ denote the real and imaginary parts of the output of the quantizer, respectively. Again, note that the real and imaginary parts of the input of the quantizer are separately quantized. For simplicity, we

assume all APs use the same number of bits to quantize the received signal, \mathbf{y}_m , and the estimated channel, $\hat{\mathbf{g}}_{mk}$. Therefore, $[\mathbf{e}_m^y]_n = \mathbb{E}\{[|\mathbf{y}_m]_n|^2\}[\hat{\mathbf{e}}_m^y]_n$ and $[\mathbf{e}_{mk}^g]_n = \mathbb{E}\{[|\hat{\mathbf{g}}_{mk}]_n|^2\}[\hat{\mathbf{e}}_m^g]_n$, where $\mathbb{E}\{[|\hat{\mathbf{e}}_m^y]_n|^2\} = \mathbb{E}\{[|\hat{\mathbf{e}}_m^g]_n|^2\} = \sigma_e^2$. Note that $\mathbb{E}\{[|\hat{\mathbf{e}}_m^y]_n|^2\}$ and $\mathbb{E}\{[|\hat{\mathbf{e}}_m^g]_n|^2\}$ are quantization errors of a quantizer with normalized input $[\hat{\mathbf{y}}_m]_n = \frac{[\mathbf{y}_m]_n}{\sqrt{\mathbb{E}\{[|\mathbf{y}_m]_n|^2\}}}$ and $[\hat{\mathbf{g}}_{mk}]_n = \frac{[\hat{\mathbf{g}}_{mk}]_n}{\sqrt{\mathbb{E}\{[|\hat{\mathbf{g}}_{mk}]_n|^2\}}}$, respectively. Note that due to power normalization, \hat{a} , b , and optimal step size for (11) and (12) are the same and provided in Table I. The received signal for the k th user after using the maximum ratio combining (MRC) detector at the CPU is given by

$$\begin{aligned} r_k &= \sum_{m=1}^M u_{mk} (Q(\hat{\mathbf{g}}_{mk}))^H Q(\mathbf{y}_m) = \sum_{m=1}^M u_{mk} (\hat{a}\hat{\mathbf{g}}_{mk} + \mathbf{e}_{mk}^g)^H (\hat{a}\mathbf{y}_m + \mathbf{e}_m^y) \\ &= \sum_{m=1}^M u_{mk} (\hat{a}\hat{\mathbf{g}}_{mk} + \mathbf{e}_{mk}^g)^H \left(\hat{a}\sqrt{\rho} \sum_{k=1}^K \mathbf{g}_{mk} \sqrt{q_k} s_k + \hat{a}\mathbf{n}_m + \mathbf{e}_m^y \right) \\ &= \hat{a}^2 \sqrt{\rho} \underbrace{\mathbb{E}\left\{ \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\}}_{\text{DS}_k} s_k + \hat{a}^2 \underbrace{\sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m}_{\text{TN}_k} \\ &\quad + \hat{a}^2 \sqrt{\rho} \underbrace{\left(\sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} - \mathbb{E}\left\{ \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\} \right)}_{\text{BU}_k} s_k + \hat{a}^2 \\ &\quad \underbrace{\sum_{k' \neq k}^K \sqrt{\rho} \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'}} s_{k'}}_{\text{IUI}_{kk'}} + \underbrace{\sum_{k'=1}^K \hat{a} \sqrt{\rho} \sum_{m=1}^M u_{mk} (\mathbf{e}_{mk}^g)^H \mathbf{g}_{mk'} \sqrt{q_{k'}} s_{k'}}_{\text{TQE}_{kk'}} \\ &\quad + \hat{a} \underbrace{\sum_{m=1}^M u_{mk} (\mathbf{e}_{mk}^g)^H \mathbf{n}_m}_{\text{TQE}_k^g} + \hat{a} \underbrace{\sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{e}_m^y}_{\text{TQE}_k^y} + \underbrace{\sum_{m=1}^M u_{mk} (\mathbf{e}_{mk}^g)^H \mathbf{e}_m^y}_{\text{TQE}_k^{gy}}. \quad (13) \end{aligned}$$

where DS_k and BU_k denote the desired signal (DS) and beamforming uncertainty (BU) for the k th user, respectively, and $\text{IUI}_{kk'}$ represents the inter-user-interference (IUI) caused by the k' th user. In addition, TN_k accounts for the total noise (TN) following the MRC detection, and finally the terms TQE_k^g , TQE_k^y , $\text{TQE}_{kk'}$ refer to the total quantization error (TQE) at the k th user due to the quantization errors at the channel and signal. Moreover, by collecting all the coefficients u_{mk} , $\forall m$, corresponding to the k th user, we define $\mathbf{u}_k = [u_{1k}, u_{2k}, \dots, u_{Mk}]^T$ and without loss of generality, it is assumed that $\|\mathbf{u}_k\| = 1$. The optimal values of u_{mk} are investigated in Section IV.

Proposition 1. *Terms DS_k , BU_k , $\text{IUI}_{kk'}$, $\text{TQN}_{kk'}$, TQN_k^g , TQN_k^y , TQN_k^{gy} are mutually uncorrelated.*

Proof: Please refer to Appendix B. ■

To obtain an achievable rate, we use the UaF bounding technique as in [2]. This technique is commonly used in massive MIMO [33], [34] since it yields a simple and tight achievable rate which enables us to further design the

$$\text{SINR}_k^{\text{Case 1}} = \frac{\dot{a}^4 |\text{DS}_k|^2}{\dot{a}^4 \mathbb{E}\{|\text{BU}_k|^2\} + \dot{a}^4 \mathbb{E}\{|\text{TN}_k|^2\} + \dot{a}^4 \sum_{k' \neq k}^K \mathbb{E}\{|\text{IUI}_{kk'}|^2\} + \dot{a}^2 \mathbb{E}\{|\text{TQE}_k^y|^2\} + \dot{a}^2 \mathbb{E}\{|\text{TQE}_k^g|^2\} + \dot{a}^2 \sum_{k'=1}^K \mathbb{E}\{|\text{TQE}_{kk'}|^2\} + \mathbb{E}\{|\text{TQE}_k^{\text{sy}}|^2\}}. \quad (14)$$

$$\text{SINR}_k^{\text{Case 1}} = \frac{N^2 q_k \left(\sum_{m=1}^M u_{mk} \gamma_{mk} \right)^2}{N^2 \sum_{k' \neq k}^K q_{k'} \left(\sum_{m=1}^M u_{mk} \gamma_{mk} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2 \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 + N \left(\frac{C_{\text{tot}}}{\dot{a}^4} + 1 \right) \sum_{m=1}^M u_{mk} \gamma_{mk} \sum_{k'=1}^K q_{k'} \beta_{mk'} + \frac{N}{\rho} \left(\frac{C_{\text{tot}}}{\dot{a}^4} + 1 \right) \sum_{m=1}^M u_{mk} \gamma_{mk}}. \quad (15)$$

systems. The tightness of this bound for cell-free Massive MIMO is presented in [2]. Using Proposition 1 and the UaF bounding technique in [2], we can obtain an achievable rate as $R_k^{\text{Case 1}} = \log_2(1 + \text{SINR}_k^{\text{Case 1}})$, where $\text{SINR}_k^{\text{Case 1}}$ is given by (14). The closed-form expression for the achievable uplink rate of the k th user is given in the following theorem.

Theorem 1. *Having the quantized CSI and the quantized signal at the CPU and employing MRC detection at the CPU, the closed-form expression for the achievable rate of the k th user is given by $R_k^{\text{Case 1}} = \log_2(1 + \text{SINR}_k^{\text{Case 1}})$, where the $\text{SINR}_k^{\text{Case 1}}$ is given by (15) (defined at the top of this page), where $\gamma_{mk} = \sqrt{\tau_p P_p} \beta_{mk} c_{mk}$ and $C_{\text{tot}} = 2\dot{a}^2 \sigma_e^2 + \sigma_e^4$.*

Proof: The power of quantization errors can be obtained as

$$\begin{aligned} \mathbb{E}\{|\mathbf{e}_m^y|^2\} &= \mathbb{E}\{|\mathbf{e}_m^y|^2\} \left(\rho \sum_{k'=1}^K q_{k'} \beta_{mk'} + 1 \right), \\ \mathbb{E}\{|\mathbf{e}_m^g|^2\} &= \mathbb{E}\{|\mathbf{e}_m^g|^2\} \gamma_{mk}. \end{aligned} \quad (16)$$

Since $\mathbb{E}\{|\mathbf{e}_m^y|^2\} = \mathbb{E}\{|\mathbf{e}_m^g|^2\} = \sigma_e^2$, we have:

$$\begin{aligned} \mathbb{E}\{|\mathbf{e}_m^y|^2\} &= \sigma_e^2 \left(\rho \sum_{k'=1}^K q_{k'} \beta_{mk'} + 1 \right), \\ \mathbb{E}\{|\mathbf{e}_m^g|^2\} &= \sigma_e^2 \gamma_{mk}. \end{aligned} \quad (17)$$

Using (16) and the fact that quantization error is independent with the input of the quantizer, after some mathematical manipulations, we have:

$$\begin{aligned} &\dot{a}^2 \mathbb{E}\{|\text{TQE}_k^y|^2\} + \dot{a}^2 \mathbb{E}\{|\text{TQE}_k^g|^2\} + \dot{a}^2 \sum_{k'=1}^K \mathbb{E}\{|\text{TQE}_{kk'}|^2\} \\ &+ \mathbb{E}\{|\text{TQE}_k^{\text{sy}}|^2\} = N C_{\text{tot}} \sum_{m=1}^M u_{mk} \gamma_{mk} \left(\rho \sum_{k'=1}^K q_{k'} \beta_{mk'} + 1 \right). \end{aligned} \quad (18)$$

Note that the terms $|\text{DS}_k|^2$, $\mathbb{E}\{|\text{BU}_k|^2\}$, and $\mathbb{E}\{|\text{IUI}_{kk'}|^2\}$ are derived in (50), (51) and (56), respectively. Finally substituting (18), (50), (51) and (56) into (14) results in (15), which completes the proof of Theorem 1. ■

Case 2. *Quantized Weighted Signal Available at the CPU:* The m th AP quantizes the terms $z_{m,k} = \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m$, $\forall k, m$, and forwards the quantized signals in each symbol duration to the CPU as

$$z_{mk} = \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m = r_{mk} + j s_{mk}, \quad \forall k, m, \quad (19)$$

where r_{mk} and s_{mk} represent the real and imaginary parts of z_{mk} , respectively. An ADC quantizes the real and imaginary

parts of $z_{m,k}$ with α bits each, which introduces quantization errors to the received signals [31]. Let us consider the term e_{mk}^z as the quantization error of the m th AP. Hence, using the Bussgang decomposition, the relation between z_{mk} and its quantized version, \hat{z}_{mk} , can be written as

$$Q(z_{mk}) = \dot{a} z_{mk} + e_{mk}^z. \quad (20)$$

Note that given the fact that the input of quantizer, i.e., $z_{mk} = \hat{\mathbf{g}}_{mk}^H \mathbf{y}_m$, is the summation of many terms, it can be approximated as a Gaussian random variable. This enables us to exploit the values given in Table I, which are obtained for Gaussian input. The aggregated received signal at the CPU can be written as³

$$\begin{aligned} r_k &= \sum_{m=1}^M u_{mk} \left(\dot{a} \underbrace{\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m}_{z_{mk}} + e_{mk}^z \right) = \dot{a} \sqrt{\rho} \sum_{k'=1}^K \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_{k'}} s_{k'} \\ &+ \dot{a} \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m + \sum_{m=1}^M u_{mk} e_{mk}^z = \dot{a} \sqrt{\rho} \underbrace{\mathbb{E}\left\{ \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\}}_{\text{DS}_k} s_k \\ &+ \dot{a} \sqrt{\rho} \left(\underbrace{\sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k}}_{\text{BU}_k} - \mathbb{E}\left\{ \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\} \right) s_k + \sum_{k' \neq k}^K \dot{a} \\ &\underbrace{\sqrt{\rho} \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'}} s_{k'}}_{\text{IUI}_{kk'}} + \underbrace{\dot{a} \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m}_{\text{TN}_k} + \underbrace{\sum_{m=1}^M u_{mk} e_{mk}^z}_{\text{TQE}_k}, \end{aligned} \quad (21)$$

where TQE_k refers to the total quantization error (TQE) at the k th user. Note that in cell-free Massive MIMO with $M \rightarrow \infty$, due to the channel hardening property, detection using only the channel statistics is nearly optimal. This is shown in [2] (see Fig. 2 of reference [2] and its discussion). Moreover, in [2] the authors show that in cell-free Massive MIMO with $M \rightarrow \infty$, the received signal includes only the desired signal plus interference from the pilot sequence non-orthogonality. Finally, using the analysis in [2], the corresponding SINR of the received signal in (21) can be defined by considering the worst-case of the uncorrelated Gaussian noise is given by (22)

³Note that for both Case 1 and Case 2, the AP estimates the channel. Therefore the total complexity is the same for Case 1 and Case 2. However, in Case 1 the CPU performs N^2 multiplications and $M-1$ additions whereas in Case 2 the APs establish N multiplications and the CPU performs N multiplications and combines the transmitted signals from M APs (via fronthaul links) which requires $M-1$ additions.

$$\text{SINR}_k^{\text{Case 2}} = \frac{|\text{DS}_k|^2}{\mathbb{E}\{|\text{BU}_k|^2\} + \sum_{k' \neq k}^K \mathbb{E}\{|\text{IUI}_{kk'}|^2\} + \mathbb{E}\{|\text{TN}_k|^2\} + \frac{1}{\hat{a}^2} \mathbb{E}\{|\text{TQE}_k|^2\}}. \quad (22)$$

$$\text{SINR}_k^{\text{Case 2}} \approx \frac{N^2 q_k \left(\sum_{m=1}^M u_{mk} \gamma_{mk} \right)^2}{N^2 \sum_{k' \neq k}^K q_{k'} \left(\sum_{m=1}^M u_{mk'} \gamma_{mk'} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2 \left| \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 + N \sum_{m=1}^M u_{mk} \left(\frac{\sigma_e^2 (2\beta_{mk} - \gamma_{mk})}{\hat{a}^2} + \gamma_{mk} \right) \sum_{k'=1}^K q_{k'} \beta_{mk'} + \frac{N}{\rho} \left(\frac{\sigma_e^2}{\hat{a}^2} + 1 \right) \sum_{m=1}^M u_{mk} \gamma_{mk}}. \quad (23)$$

(defined at the top of the next page). Based on the SINR definition in (22), the achievable uplink rate of the k th user is given in the following theorem.

Theorem 2. *Having the quantized weighted signal at the CPU and employing MRC detection at the CPU, the achievable uplink rate of the k th user in the cell-free Massive MIMO system is $R = \log_2(1 + \text{SINR}^{\text{Case 2}})$, where $\text{SINR}^{\text{Case 2}}$ is given by (23) (defined at the top of this page).*

Proof: Please refer to Appendix C. \blacksquare

A. Required Fronthaul Capacity

Let τ_f be the length of the uplink payload data transmission for each coherence interval, i.e., $\tau_f = \tau_c - \tau_p$, where τ_c denotes the number of samples for each coherence interval and τ_p represents the length of pilot sequence. Defining the number of quantization bits as $\alpha_{m,i}$, for $i = 1, 2$, corresponding to Cases 1 and 2, and m refers to the m th AP. For Case 1, the required number of bits for each AP during each coherence interval is $2\alpha_{m,1} \times (NK + N\tau_f)$ whereas Case 2 requires $2\alpha_{m,2} \times (K\tau_f)$ bits for each AP during each coherence interval. Hence, the total fronthaul capacity required between the m th AP and the CPU for all schemes is defined as

$$C_m = \begin{cases} \frac{2(NK + N\tau_f) \alpha_{m,1}}{T_c}, & \text{Case 1,} \\ \frac{2(K\tau_f) \alpha_{m,2}}{T_c}, & \text{Case 2,} \end{cases} \quad (24)$$

where T_c (in sec.) refers to coherence time.⁴ In the following, we present a comparison between two cases of uplink transmission. To make a fair comparison between Case 1 and Case 2, we use the same total number of fronthaul bits for both cases, that is $2(NK + N\tau_f)\alpha_{m,1} = 2(K\tau_f)\alpha_{m,2}$.⁵

IV. PROPOSED MAX-MIN RATE SCHEME

In this section, we formulate the max-min rate problem for Case 2 of uplink transmission in cell-free Massive MIMO system, where the minimum uplink rates of all users is maximized while satisfying the transmit power constraint at each user and the fronthaul capacity constraint. Note that the same

⁴Exploiting the constraint $C_m \leq C_{\text{fb}}$, the largest number of quantization level is equal to $\alpha_{m,1} = \left\lfloor \frac{T_c C_{\text{fb}}}{2(NK + N\tau_f)} \right\rfloor$ and $\alpha_{m,2} = \left\lfloor \frac{T_c C_{\text{fb}}}{2K\tau_f} \right\rfloor$.

⁵Future work is needed to investigate the performance analysis for different numbers of quantization bits as well as the numbers of APs. This will be presented in [15].

approach can be used to investigate the max-min rate problem for Case 1. The achievable user SINR for the system model considered in the previous section is obtained by following a similar approach to that in [2]. Note that the main difference between the proposed approach and the scheme in [2] is the new set of receiver coefficients which are introduced at the CPU to improve the achievable user rates. The benefits of the proposed approach in terms of the achieved user uplink rate is demonstrated through numerical simulation results in Section V. In deriving the achievable rates of each user, it is assumed that the CPU exploits only the knowledge of channel statistics between the users and APs to detect data from the received signal in (21). Using the SINR given in (23), the achievable rate is obtained $R_k^{\text{UP}} = \log_2(1 + \text{SINR}_k^{\text{Case 2}})$. Defining $\mathbf{u}_k = [u_{1k}, u_{2k}, \dots, u_{Mk}]^T$, $\boldsymbol{\Gamma}_k = [\gamma_{1k}, \gamma_{2k}, \dots, \gamma_{Mk}]^T$, $\boldsymbol{\Upsilon}_{kk'} = \text{diag}\left[\beta_{1k'} \left(\frac{\sigma_e^2 (2\beta_{1k} - \gamma_{1k})}{\hat{a}^2} + \gamma_{1k} \right), \dots, \beta_{Mk'} \left(\frac{\sigma_e^2 (2\beta_{Mk} - \gamma_{Mk})}{\hat{a}^2} + \gamma_{Mk} \right)\right]$, $\boldsymbol{\Lambda}_{kk'} = \left[\frac{\gamma_{1k} \beta_{1k'}}{\beta_{1k}}, \frac{\gamma_{2k} \beta_{2k'}}{\beta_{2k}}, \dots, \frac{\gamma_{Mk} \beta_{Mk'}}{\beta_{Mk}} \right]^T$ and $\mathbf{R}_k = \text{diag}\left[\left(\frac{\sigma_e^2}{\hat{a}^2} + 1\right) \gamma_{1k}, \dots, \left(\frac{\sigma_e^2}{\hat{a}^2} + 1\right) \gamma_{Mk}\right]$, the achievable uplink rate of the k th user is given by

$$P_1 : \max_{q_k, \mathbf{u}_k, \alpha_2} \min_{k=1, \dots, K} R_k^{\text{UP}} \quad (26a)$$

$$\text{subject to } \|\mathbf{u}_k\| = 1, \forall k, 0 \leq q_k \leq p_{\text{max}}^{(k)}, \forall k, \quad (26b)$$

$$C_m \leq C_{\text{fb}}, \forall m, \quad (26c)$$

where $p_{\text{max}}^{(k)}$ and C_{fb} refer to the maximum transmit power available at user k and the capacity of fronthaul link between each AP and the CPU, respectively. Note that using (24), C_m is given as $C_m = \frac{2(K\tau_f)\alpha_{m,2}}{T_c}, \forall m$. Throughout the rest of the paper, the index m is dropped from $\alpha_{m,i}$, $i = 1, 2$, as we consider the same number of bits to quantize the signal at all APs. Problem P_1 is a discrete optimization with integer decision variables and it is obvious that the achievable user rates monotonically increase with the capacity of the fronthaul link between the m th AP and the CPU. Hence, the optimal solution is achieved when $C_m = C_{\text{fb}}, \forall m$, which leads to fixed values for the number of quantization bits. As a result, the max-min based max-min rate problem can be re-formulated as follows:

$$P_2 : \max_{q_k, \mathbf{u}_k} \min_{k=1, \dots, K} R_k^{\text{UP}} \quad (27a)$$

$$\text{subject to } \|\mathbf{u}_k\| = 1, \forall k, \quad (27b)$$

$$0 \leq q_k \leq p_{\text{max}}^{(k)}, \forall k. \quad (27c)$$

Problem P_2 is not jointly convex in terms of \mathbf{u}_k and power allocation $q_k, \forall k$. Therefore, it cannot be directly solved through existing convex optimization software. To tackle this non-convexity issue, we decouple Problem P_2 into two sub-problems: receiver coefficient design (i.e. \mathbf{u}_k) and the power allocation problem. The optimal solution for Problem P_2 , is obtained through alternately solving these sub-problems, as explained in the following subsections.

A. Receiver Filter Coefficient Design

In this subsection, the problem of designing the receiver coefficients is considered. We solve the max-min rate problem for a given set of allocated powers at all users, $q_k, \forall k$, and fixed values for the number of quantization levels, $Q_m, \forall m$. These coefficients (i.e., $\mathbf{u}_k, \forall k$) are obtained by independently maximizing the uplink SINR of each user. Therefore, the optimal receiver filter coefficients can be determined by solving the following optimization problem:

$$P_3 : \max_{\mathbf{u}_k} \quad (28a)$$

$$\frac{N^2 \mathbf{u}_k^H (q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H z) \mathbf{u}_k}{\mathbf{u}_k^H \left(N^2 \sum_{k' \neq k}^K q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Lambda}_{kk'} \mathbf{\Lambda}_{kk'}^H + N \sum_{k'=1}^K q_{k'} \mathbf{\Upsilon}_{kk'} + \frac{N \mathbf{R}_k}{\rho} \right) \mathbf{u}_k} \quad (28b)$$

subject to $\|\mathbf{u}_k\| = 1, \quad \forall k.$

Problem P_3 is a generalized eigenvalue problem [22], [24], [35], where the optimal solutions can be obtained by determining the generalized eigenvector of the matrix pair $\mathbf{A}_k = N^2 q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H$ and $\mathbf{B}_k = N^2 \sum_{k' \neq k}^K q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Lambda}_{kk'} \mathbf{\Lambda}_{kk'}^H + N \sum_{k'=1}^K q_{k'} \mathbf{\Upsilon}_{kk'} + \frac{N \mathbf{R}_k}{\rho}$ corresponding to the maximum generalized eigenvalue.

B. Power Allocation

In this subsection, we solve the power allocation problem for a given set of fixed receiver filter coefficients, $\mathbf{u}_k, \forall k$, and fixed values of quantization levels, $Q_m, \forall m$. The optimal transmit power can be determined by solving the following max-min problem:

$$P_4 : \max_{q_k} \min_{k=1, \dots, K} \text{SINR}_k^{\text{UP}} \quad (29a)$$

$$\text{subject to} \quad 0 \leq q_k \leq p_{\max}^{(k)}. \quad (29b)$$

Without loss of generality, Problem P_4 can be rewritten by introducing a new slack variable as

$$P_5 : \max_{t, q_k} t \quad (30a)$$

$$\text{subject to} \quad 0 \leq q_k \leq p_{\max}^{(k)}, \quad \forall k, \text{SINR}_k^{\text{UP}} \geq t, \quad \forall k. \quad (30b)$$

Proposition 2. *Problem P_5 can be formulated into a standard GP.*

Algorithm 1 Proposed algorithm to solve Problem P_2

1. Initialize $\mathbf{q}^{(0)} = [q_1^{(0)}, q_2^{(0)}, \dots, q_K^{(0)}], i = 1$
 2. Repeat steps 3-5 until $\frac{\text{SINR}_k^{\text{UP},(i)} - \text{SINR}_k^{\text{UP},(i-1)}}{\text{SINR}_k^{\text{UP},(i-1)}} \leq \epsilon, \forall k$
 3. Determine the optimal receiver coefficients $\mathbf{U}^{(i)} = [\mathbf{u}_1^{(i)}, \mathbf{u}_2^{(i)}, \dots, \mathbf{u}_K^{(i)}]$ through solving the generalized eigenvalue Problem P_3 in (28) for a given $\mathbf{q}^{(i-1)}$,
 4. Compute $\mathbf{q}^{(i)}$ through solving Problem P_5 in (30) for a given $\mathbf{U}^{(i)}$
 5. $i = i + 1$
-

Proof: Please refer to Appendix D. ■

Therefore, Problem P_5 is efficiently solved through existing convex optimization software. Based on these two sub-problems, an iterative algorithm has been developed by alternately solving both sub-problems at each iteration. The proposed algorithm is summarized in Algorithm 1. Note that ϵ in Step 2 of Algorithm 1 refers to a small predetermined value.

V. CONVERGENCE

In this section, we present the convergence of the proposed Algorithm 1. We propose to alternatively solve two sub-problems to find the solution of the original Problem P_2 , where at each iteration, one of the design parameters is determined by solving the corresponding sub-problem while other design variable is fixed. We showed that each sub-problem provides an optimal solution for the other given design variable. Let us assume at the $i - 1$ th iteration, that the receiver filter coefficients $\mathbf{u}_k^{(i-1)}, \forall k$ are obtained for a given power allocation $\mathbf{q}^{(i-1)}$ and similarly, the power allocation $\mathbf{q}^{(i)}$ is determined for a fixed set of receiver filter coefficients $\mathbf{u}_k^{(i-1)}, \forall k$. Note that, the optimal power allocation $\mathbf{q}^{(i)}$ determined for a given $\mathbf{u}_k^{(i-1)}$ achieves an uplink rate greater than or equal to that of the previous iteration. In addition, the power allocation $\mathbf{q}^{(i-1)}$ is a feasible solution to find $\mathbf{q}^{(i)}$ as the receiver filter coefficients $\mathbf{u}_k^{(i)}, \forall k$ are determined for a given $\mathbf{q}^{(i-1)}$. Note that the uplink rate of the system monotonically increases with the power. As a result, the achievable uplink rate of the system monotonically increases at each iteration. Note that the achievable uplink max-min rate is bounded from above for a given set of per-user power constraints and fronthaul link capacity constraint. Hence the proposed algorithm converges to a specific solution. Note that to the best of our knowledge and referring to [1] this is a common way to show the convergence. In the next section, we prove the optimality of the proposed Algorithm 1 through the principle of uplink-downlink duality.

$$R_k^{\text{UP}} \approx \log_2 \left(1 + \frac{\mathbf{u}_k^H (N^2 q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H) \mathbf{u}_k}{\mathbf{u}_k^H \left(N^2 \sum_{k' \neq k}^K q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \mathbf{\Lambda}_{kk'} \mathbf{\Lambda}_{kk'}^H + N \sum_{k'=1}^K q_{k'} \mathbf{\Upsilon}_{kk'} + \frac{N \mathbf{R}_k}{\rho} \right) \mathbf{u}_k} \right). \quad (25)$$

VI. OPTIMALITY OF THE PROPOSED MAX-MIN RATE ALGORITHM

In this section, we present a method to prove the optimality of the proposed Algorithm 1. The proof is based on two main observations: we first demonstrate that the original max-min Problem P_2 with per-user power constraint is equivalent to an uplink problem with an equivalent total power constraint. We next prove that the same SINRs can be achieved in both the uplink and the virtual downlink with an equivalent total power constraint, which enables us to establish an uplink-downlink duality. Finally, we show that the virtual downlink problem is quasi-convex and can be optimally solved through a bisection search [36]. Note that the uplink Problem P_1 and the equivalent virtual downlink problem achieve the same SINRs and the solution of the virtual downlink problem is optimal. As a result, the optimality of the proposed Algorithm 1 is guaranteed. The details of the proof are provided in the following subsections.

A. Equivalent Max-Min Uplink Problem

We aim to show the equivalence of Problem P_2 with a per-user power constraint and the uplink max-min rate problem with a total power constraint. Note that in the total power constraint, the maximum available transmit power is defined as the sum of all users' transmit power from the solution of Problem P_2 , which is formulated as:

$$P_6 : \quad \max_{q_k, \mathbf{u}_k} \quad \min_{k=1, \dots, K} \quad R_k^{\text{UP}} \quad (31a)$$

$$\text{subject to} \quad \|\mathbf{u}_k\| = 1, \quad \forall k, \quad (31b)$$

$$\sum_{k=1}^K q_k \leq P_{\text{tot}}^c. \quad (31c)$$

Problem P_6 is not convex in terms of receiver filter coefficients \mathbf{u}_k and power allocation $q_k, \forall k$. To deal with this non-convexity, similar to the proposed method to solve problem P_2 , we propose to modify Algorithm 1 to incorporate the total power constraint in Problem P_6 . Hence, we decompose Problem P_6 into receiver filter coefficient design and power allocation sub-problems. The same generalized eigenvalue problem in Problem P_3 is solved to determine the receiver filter coefficients whereas the GP formulation in P_5 is modified to incorporate the total power constraint (31c). Note that, the total power constraint is a convex constraint (posynomial function in terms of power allocation) and GP with the equivalent total power constraint can be used to find the optimum solution.

Lemma 2. *The original Problem P_2 (with per-user power constraint) and the equivalent Problem P_6 (with the equivalent total power constraint) have the same optimal solution.*

Proof: Please refer to Appendix E. ■

B. Uplink-Downlink Duality for Cell-free Massive MIMO

This subsection demonstrates an uplink-downlink duality for cell-free Massive MIMO systems. In particular, it is shown that the same SINRs (or rate regions) can be realized for all users in the uplink and the virtual downlink with the

equivalent total power constraints [21], [37], respectively. In other words, based on the principle of uplink-downlink duality, the same set of filter coefficients can be utilized in the uplink and the downlink to achieve the same SINRs for all users with different user power allocations. The following theorem defines the achievable virtual downlink rate for cell-free Massive MIMO systems:

Theorem 3. *By employing conjugate beamforming at the APs, the achievable virtual downlink rate of the k th user in the cell-free Massive MIMO system with K randomly distributed single-antenna users, M APs where each AP is equipped with N antennas and limited-capacity fronthaul links is given by (32) (defined at the top of this page).*

Proof: This can be derived by following the same approach as for uplink transmission in Theorem 2. ■

Note that in (32), $p_k, \forall k$ denotes the downlink power allocation for the k th user and the following equalities hold: $\mathbf{\Gamma}_k = [\gamma_{1k}, \gamma_{2k}, \dots, \gamma_{Mk}]^T$, $\mathbf{F}_{k'k} = \text{diag} \left[\beta_{1k} \left(\frac{\hat{a}^2 (2\beta_{1k'} - \gamma_{1k'})}{\sigma_e^2} + \gamma_{1k'} \right), \dots, \beta_{Mk} \left(\frac{\hat{a}^2 (2\beta_{Mk'} - \gamma_{Mk'})}{\sigma_e^2} + \gamma_{Mk'} \right) \right]$ and $\Delta_{k'k} = \left[\frac{\gamma_{1k'} \beta_{1k}}{\beta_{1k'}}, \frac{\gamma_{2k'} \beta_{2k}}{\beta_{2k'}}, \dots, \frac{\gamma_{Mk'} \beta_{Mk}}{\beta_{Mk'}} \right]^T$. The following Theorem provides the required condition to establish the uplink-downlink duality for cell-free Massive MIMO systems with limited-capacity fronthaul links:

Theorem 4. *By employing MRC detection in the uplink and conjugate beamforming in the virtual downlink, to realize the same SINR tuples in both the uplink and the virtual downlink of a cell-free Massive MIMO system, with the same fronthaul loads, the same filter coefficients and different transmit power allocations, the following condition should be satisfied:*

$$N \sum_{m=1}^M \sum_{k=1}^K \left(\frac{\sigma_e^2}{\hat{a}^2} + 1 \right) \gamma_{mk} |w_{mk}|^2 = \sum_{k=1}^K q_k^* = P_{\text{tot}}^c, \quad (34)$$

where $q_k^*, \forall k$ refer to the optimal solution of Algorithm 1, and w_{mk} denotes the (m, k) -th entry of matrix \mathbf{W} which is defined as follows:

$$\mathbf{W} = [\sqrt{p_1} \mathbf{u}_1, \sqrt{p_2} \mathbf{u}_2, \dots, \sqrt{p_K} \mathbf{u}_K]. \quad (35)$$

Proof: Please refer to Appendix F. ■

C. Equivalent Max-Min Downlink Problem

In this subsection, we provide an optimal solution for the max-min rate downlink problem with the equivalent total power constraint. This problem can be written as follows:

$$P_7 : \quad \max_{p_k, \mathbf{u}_k} \quad \min_{k=1, \dots, K} \quad R_k^{\text{DL}} \quad (36a)$$

$$\text{subject to} \quad \|\mathbf{u}_k\| = 1, \quad \forall k, \quad \sum_{k=1}^K p_k \leq P_{\text{tot}}^c, \quad (36b)$$

where $R_k^{\text{DL}} = \log_2(1 + \text{SINR}_k^{\text{DL}})$, and $\text{SINR}_k^{\text{DL}}$ is defined in (32). This problem is difficult to jointly solve in terms of transmit filter coefficients \mathbf{u}_k 's and power allocations p_k 's. However,

$$\text{SINR}_k^{\text{DL}}(\mathbf{U}, \mathbf{p}) = \frac{\mathbf{u}_k^H (N^2 p_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H) \mathbf{u}_k}{N^2 \sum_{k' \neq k}^K \mathbf{u}_{k'}^H p_{k'} |\boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k|^2 \Delta_{k'k} \Delta_{k'k}^H \mathbf{u}_{k'} + N \sum_{k'=1}^K \mathbf{u}_{k'}^H p_{k'} \mathbf{F}_{k'k} \mathbf{u}_{k'} + \frac{N}{\rho}}. \quad (32)$$

$$\text{SINR}_k^{\text{UP}}(\mathbf{U}, \mathbf{q}) = \frac{\mathbf{u}_k^H (N^2 q_k \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H) \mathbf{u}_k}{\mathbf{u}_k^H \left(N^2 \sum_{k' \neq k}^K q_{k'} |\boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k|^2 \Lambda_{kk'} \Lambda_{kk'}^H + N \sum_{k'=1}^K q_{k'} \Upsilon_{kk'} + \frac{N}{\rho} \mathbf{R}_k \right) \mathbf{u}_k}. \quad (33)$$

it can be represented by introducing a new variable \mathbf{W} to decouple the variables \mathbf{U} and \mathbf{q} as follows:

$$P_8 : \max_{\mathbf{W}} \min_{k=1, \dots, K} R_k^{\text{DL}} \quad (37a)$$

$$\text{subject to } N \sum_{m=1}^M \sum_{k=1}^K \left(\frac{\sigma_{\dot{e}}^2}{\dot{a}^2} + 1 \right) \gamma_{mk} |w_{mk}|^2 \leq P_{\text{tot}}^c. \quad (37b)$$

It is easy to show that Problem P_8 is quasi-convex. Hence, a bisection [36] approach can be used to obtain the optimal solution for the original Problem P_8 by sequentially solving the following power minimization problem for a given target SINR t at all users:

$$P_9 : \min_{\mathbf{W}} \sum_{m=1}^M \sum_{k=1}^K \gamma_{mk} |w_{mk}|^2 \quad (38a)$$

subject to

$$\frac{\mathbf{w}_k^H (N^2 \mathbf{\Gamma}_k \mathbf{\Gamma}_k^H) \mathbf{w}_k}{N^2 \sum_{k' \neq k}^K \mathbf{w}_{k'}^H |\boldsymbol{\phi}_{k'}^H \boldsymbol{\phi}_k|^2 \Delta_{k'k} \Delta_{k'k}^H \mathbf{w}_{k'} + N \sum_{k'=1}^K \mathbf{w}_{k'}^H \mathbf{F}_{k'k} \mathbf{w}_{k'} + \frac{N}{\rho}} \geq t, \quad (38c)$$

$$N \sum_{m=1}^M \sum_{k=1}^K \left(\frac{\sigma_{\dot{e}}^2}{\dot{a}^2} + 1 \right) \gamma_{mk} |w_{mk}|^2 \leq P_{\text{tot}}^c,$$

where \mathbf{w}_k represents the k th column of the matrix \mathbf{W} defined in (35). Problem P_9 can be reformulated by exploiting a second order cone programming (SOCP). Note that the objective function in (38) refers to the total transmit power. As a result, the optimal solution for Problem P_7 can be obtained by extracting the normalized transmit filter coefficients \mathbf{u}_k 's and power allocations p_k 's as

$$p_k^* = \|\mathbf{w}_k^*\|^2, \quad \forall k, \quad \& \quad \mathbf{u}_k^* = \frac{\mathbf{w}_k^*}{\|\mathbf{w}_k^*\|}, \quad \forall k, \quad (39)$$

where the \mathbf{w}_k^* 's refer to the optimal solution of Problem P_8 . Note that constraint (38c) is an equivalent total power constraint to the per-user power constraint in the original Problem P_2 , which is a more relaxed constraint than (27c). However, it is already shown in the previous sub-section that the same SINRs can be realized in both the uplink and the virtual downlink with per-user and the equivalent total power constraints.

D. Prove of Optimality of Algorithm 1

In Lemma 2, we prove that Problems P_2 and P_6 are equivalent, and have the same solution. Next, in Proposition 1, using uplink-downlink duality, we prove that Problem P_6 and the virtual downlink Problem P_7 are equivalent. Note that the SINR achieved by solving Problem P_7 are optimal (the optimal solution is obtained by a bisection search approach).

This confirms that the proposed algorithm to solve Problem P_2 is optimal.

VII. USER ASSIGNMENT

Exploiting (24), it is obvious that the total fronthaul capacity required between the m th AP and the CPU increases linearly with the total number of users served by the m th AP. This motivates the need to pick a proper set of active users for each AP. Using (24), we have

$$\alpha_2 \times K_m \leq \frac{C_{\text{th}} T_c}{2\tau_f}, \quad (40)$$

where K_m denotes the size of the set of active users for the m th AP. From (40), it can be seen that decreasing the size of the set of active users allows for a larger number of quantization levels. Motivated by this fact, and to exploit the capacity of fronthaul links more efficiently, we investigate all possible combinations of α_2 and K_m . First, for a fixed value of α_2 , we find an upper bound on the size of the set of active users for each AP. In the next step, we propose for all APs that the users are sorted according to $\beta_{mk}, \forall k$, and find the K_m users which have the highest values of β_{mk} among all users. If a user is not selected by any AP, we propose to find the AP which has the best link to this user (in Algorithm 2, $\pi(j) = \arg \max_m \beta_{mj}$ determines best link to the j th user, i.e., the index of the AP which is closest to the j th user). Note that to only consider the users that have links to other APs, we use $k|_{\mathcal{S}_k \pi_j \neq \emptyset}$, where \emptyset refers to empty set. Then we drop the user which has the lowest $\beta_{mk}, \forall k$, among the set of active users for that AP, which has links to other APs as well. Finally, we add the user which is not selected by any AP to the set of active users for this AP. We next solve the virtual downlink problem to maximize the minimum uplink rate of the users as follows

$$P_{10} : \max_{\mathbf{W}} \min_{k=1, \dots, K} R_k^{\text{DL}}(\vec{\gamma}_{mk}), \quad (41a)$$

$$\text{subject to } N \sum_{m=1}^M \sum_{k=1}^K \left(\frac{\sigma_{\dot{e}}^2}{\dot{a}^2} + 1 \right) \vec{\gamma}_{mk} |w_{mk}|^2 \leq P_{\text{tot}}^c, \quad (41b)$$

where

$$\vec{\gamma}_{mk} = \begin{cases} \gamma_{mk}, & m \in \mathcal{S}_k \\ 0, & \text{otherwise} \end{cases} \quad (42)$$

where \mathcal{S}_k refers to the set of active APs for the k th user. The proposed algorithm is summarized in Algorithm 2.

VIII. NUMERICAL RESULTS AND DISCUSSION

In this section, we provide numerical simulation results to validate the performance of the proposed max-min rate scheme with different parameters. A cell-free Massive MIMO system with M APs and K single-antenna users is considered in a

Algorithm 2 User Assignment

1. Using (40), find the maximum possible integer value for $K_m, \forall m$
 2. Sort users according to the ascending channel gain: $\beta_{m1} \geq \beta_{m2} \geq \dots \geq \beta_{mK}, \forall m$
 3. Assign K_m users with the highest values of $\beta_{mk}, \forall m$ to each AP, i.e., $\mathcal{T}_m \leftarrow \{k^{(1)}, k^{(2)}, \dots, k^{(K_m)}\}, \forall m$
 4. Find set of active APs for each user; $\mathcal{S}_k \leftarrow \{m^{(1)}, m^{(2)}, \dots, m^{(M_k)}\}, \forall k$
 5. for $j = 1 : K$
 if $\text{size} \{ \mathcal{S}_j \} = 0$
 $\pi(j) = \text{argmax}_m \beta_{mj}, \delta(j) = \text{argmin}_k \beta_{\pi(j)k}, k | \mathcal{S}_k \pi_j \neq \emptyset, \mathcal{T}_{\pi(j)} \leftarrow \mathcal{T}_{\pi(j)} \setminus \delta(j), \mathcal{T}_{\pi(j)} \leftarrow \mathcal{T}_{\pi(j)} \cup j$
 end
 end
 6. If $m \in \mathcal{S}_k$, then $\vec{\gamma}_{mk} \leftarrow \gamma_{mk}$, otherwise $\vec{\gamma}_{mk} = 0$ and solve the max-min rate problem P_2
-

$D \times D$ simulation area, where both APs and users are uniformly distributed at random. In the following subsections, we define the simulation parameters and then present the corresponding simulation results. The channel coefficients between users and APs are modeled in Section II where the coefficient β_{mk} is given by $\beta_{mk} = \text{PL}_{mk} 10^{\frac{\sigma_{sh} z_{mk}}{10}}$, where PL_{mk} is the path loss from the k th user to the m th AP and the second term $10^{\frac{\sigma_{sh} z_{mk}}{10}}$, denotes the shadow fading with standard deviation $\sigma_{sh} = 8$ dB, and $z_{mk} \sim \mathcal{N}(0, 1)$ [2]. In the simulation, an uncorrelated shadowing model is considered and a three-slope model for the path loss similar to [2]. The noise power is given by $p_n = \text{BW} \times k_B \times T_0 \times W$, where $\text{BW} = 20$ MHz denotes the bandwidth, $k_B = 1.381 \times 10^{-23}$ represents the Boltzmann constant, and $T_0 = 290$ (Kelvin) denotes the noise temperature. Moreover, $W = 9$ dB, and denotes the noise figure. It is assumed that that \bar{p}_p and $\bar{\rho}$ denote the power of 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 $\bar{p}_p = 200$ mW and $\bar{\rho} = 200$ mW. Similar to [2], we assume that the simulation area is wrapped around at the edges which can simulate an area without boundaries. Hence, the square simulation area has eight neighbours. We evaluate the average rate of the system over 300 random realizations of the locations of APs, users and shadow fading. Similar to the model in [38], the fronthaul links establish communications through wireless microwave links with limited capacity. Hence, we use $C_{\text{fh}} = 100$ Mb/s [38], unless otherwise it is indicated. In this paper, the term ‘‘orthogonal pilots’’ refers to the case where unique orthogonal pilots are assigned to all users, while in ‘‘random pilot assignment’’ each user is randomly assigned a pilot sequence from a set of τ_p orthogonal sequences of length $\tau_p (< K)$, following the approach of [2].

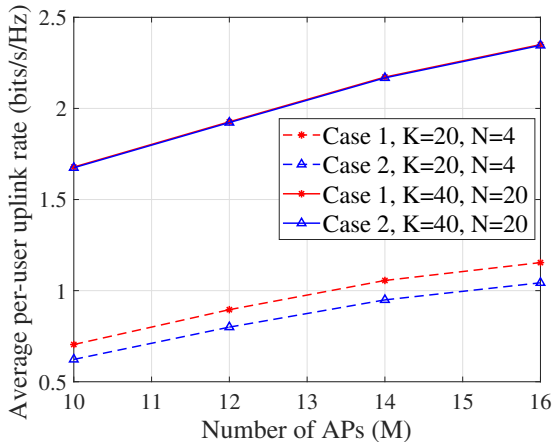
1) Performance of Different Cases of Uplink Transmission:

Fig. 2a presents the average per-user uplink rate, where the per-user uplink rate is obtained by solving Problem P_4 , given by (29) for Cases 1 and 2. The values of $\alpha_1 = 9$ and $\alpha_2 = 2$

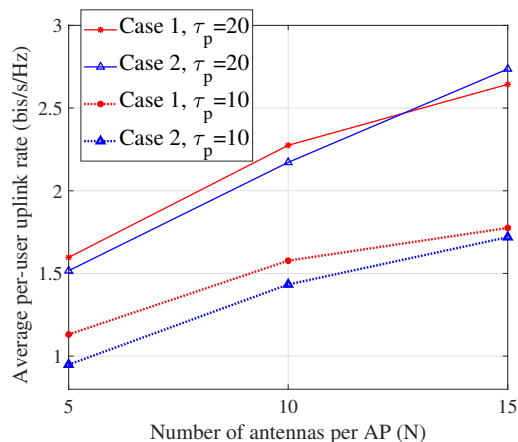
correspond to a total number of 14,400 bits for each AP during each coherence time (or frame). In addition, similar to [31] we use a uniform quantizer with fixed step size. As Fig 2a shows the performance of Case 1 is slightly better than Case 2 for $K = 20$. Next, the performance of the cell-free Massive MIMO system is evaluated for a system with $K = 40$ in which each AP is equipped with $N = 20$ antennas. Fig. 2a shows the average rate of the cell-free Massive MIMO system, where for Case 1 and Case 2, we set $\alpha_1 = 3$ and $\alpha_2 = 8$, respectively which leads to a total number of 64,000 fronthaul bits per AP per frame. Fig. 2a shows that the performances of Case 1 and Case 2 depend on the values of N , K and τ_f . Next, we investigate the effect of number of antennas per AP and τ_f for $K = 20$. Fig. 2b shows the average per-user uplink rate of cell-free Massive MIMO versus number of antennas per AP and two cases of $\tau_p = 20$ ($\tau_f = 180$) and $\tau_p = 10$ ($\tau_f = 190$). Moreover, we consider $(\alpha_1 = 18, \alpha_2 = 5)$, $(\alpha_1 = 18, \alpha_2 = 10)$, $(\alpha_1 = 18, \alpha_2 = 15)$ for the cases of $N = 5$, $N = 10$, $N = 15$, respectively, resulting 18,000 bits for all values of N . As the figure shows the difference between Case 1 and Case 2 decreases as N increases. Moreover, for the case of orthogonal pilots and $N = 15$, the performance of Case 2 is better than the performance of Case 1. Since in case 1, the CPU knows the quantized channel estimates, other signal processing techniques (e.g., zero-forcing processing) can be implemented to improve the system performance and can be considered in future work.

2) *Performance of the Proposed User Max-Min Rate Algorithm:* In this subsection, we evaluate the performance of the proposed uplink max-min rate scheme. To assess the performance, a cell-free Massive MIMO system is considered with 70 APs ($M = 70$) where each AP is equipped with $N = 4$ antennas and 40 users ($K = 40$) which are randomly distributed over the simulation area of size 1×1 km meters. Moreover, we consider the case $\{M = 50, N = 4, K = 30\}$ Fig. 3 presents the cumulative distribution of the achievable uplink rates for the proposed Algorithm 1 in the case similar to [2], without defining the coefficients \mathbf{u}_k , (i.e., $u_{mk} = 1 \forall m, k$) and solving Problem P_4 , with random pilot sequences with length $\tau_p = 30$. As seen in Fig. 3, the performance (i.e. the 10%-outage rate, R_{out} , refers to the case when $P_{\text{out}} = \text{Pr}(R_k < R_{\text{out}}) = 0.1$, where Pr refers to the probability function) of the proposed scheme is almost three times than that of the case with $u_{mk} = 1 \forall m, k$.

3) *Convergence:* Next, we provide simulation results to validate the convergence of the proposed algorithm for a set of different random realizations of the locations of APs, users and shadow fading. These results are generated over the simulation area of size 1×1 km² with random and orthogonal pilot sequences. Fig. 4a investigates the convergence of the proposed Algorithm 1 with 70 APs ($M = 70$) and 40 users ($K=40$) and random pilot sequences with length $\tau_p = 30$, whereas Fig. 4b demonstrates the convergence of the proposed Algorithm 1 for the case of $M = 30$ APs and $K = 50$ with orthogonal pilot sequences. The figures confirm that the proposed algorithm converges after a few iterations, while the minimum rate of the users increases with the iteration number.



(a) Average per-user uplink rate for cases 1 and 2, with $(N = 4, K = 20, \tau_p = 20, \alpha_1 = 9, \alpha_2 = 2)$, and $(N = 20, K = 40, \tau_p = 40, \alpha_1 = 8, \alpha_2 = 5)$ with $D = 1$ km and $\tau_c = 200$. Note that here $\tau_f = \tau_c - \tau_p = 160$.



(b) Average per-user uplink rate for cases 1 and 2, for $M = 20, K = 20, \tau_p = 20, \tau_p = 10, D = 1$ km and $\tau_c = 200$ versus number of antennas per AP. Note that we consider $(\alpha_1 = 18, \alpha_2 = 5), (\alpha_1 = 18, \alpha_2 = 10), (\alpha_1 = 18, \alpha_2 = 15)$ for the cases of $N = 5, N = 10, N = 15$, respectively.

Figure 2. Performance of different cases of uplink transmission

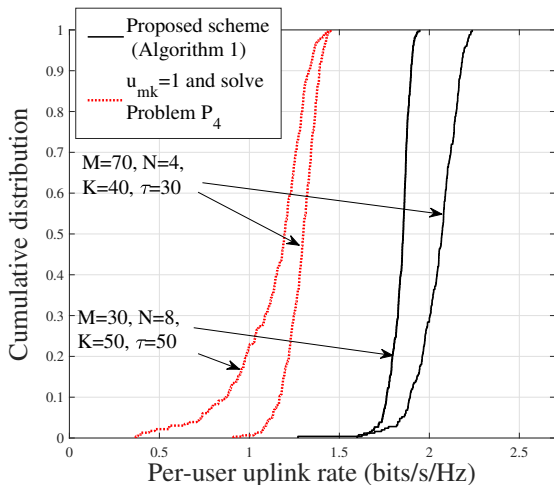


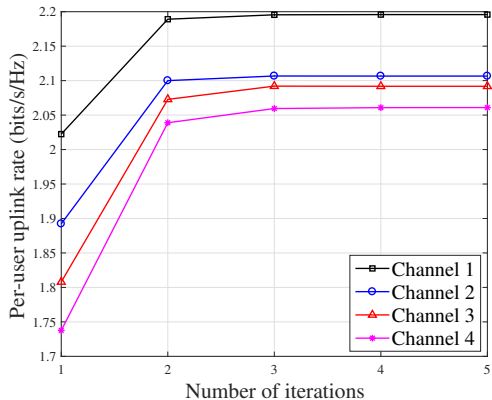
Figure 3. The cumulative distribution of the per-user uplink rate, for $\{M = 70, N = 4, K = 40\}$, $\{M = 50, N = 4, K = 30\}$, and $\tau_p = 30, \alpha_1 = 1$ and $D = 1$ km.

4) *Uplink-Downlink Duality in Cell-Free Massive MIMO System*: Here, the simulation results are provided to support the theoretical derivations of the uplink-downlink duality and the optimality of Algorithm 1. It is assumed that users are randomly distributed through the simulation area of size 1×1 km. Figs. 5 compares the cumulative distribution of the achievable uplink rates between the original uplink max-min problem (Problem P_1), the equivalent uplink problem (Problem P_6) and the equivalent downlink problem (Problem P_7). In Fig. 5, the minimum uplink rate is obtained for a system with 30 APs ($M = 30$) where each is equipped with $N = 8$ antennas and has 50 users ($K = 50$) for two cases of orthogonal pilot sequences and random pilot sequences with length $\tau_p = 30$. Moreover, Fig. 5 demonstrates the same results for 70 APs ($M = 70$), $N = 4$, 40 users ($K = 40$), and $\tau_p = 30$. The simulation results provided in Fig. 5 validate our result that the problem formulations P_1, P_6 and P_7 are equivalent and

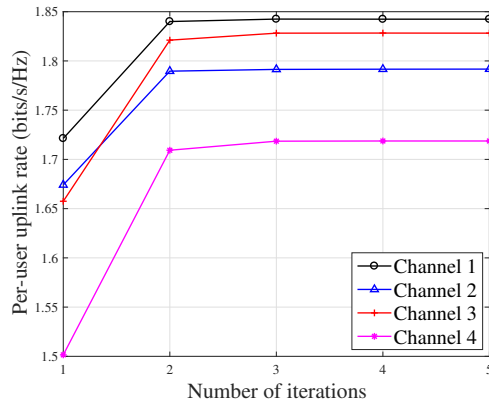
achieve the same minimum user rate. In addition, these results support our result on the uplink-downlink duality for cell-free Massive MIMO in Section VI and the proof of optimality of Algorithm 1.

5) *Performance of the Proposed User Assignment Algorithm 2*: This subsection investigates the performance of the proposed user assignment Algorithm 2. In Fig. 6a, the average per-user uplink rate is presented with $M = 120, N = 2, K = 50$, orthogonal pilot sequences and random pilot assignment with $D = 1$ km, versus the total number of active users per AP. Here, we used inequality (40) and set $\alpha_2 \times K_m = 100$ for all curves in Fig. 6a. The optimum value of K_m , (K_m^{opt}), depends on the system parameters and as Fig. 6a shows for both cases of $\tau_p = 50$ and $\tau_p = 30$, the optimum value is achieved by $K_m^{\text{opt}} = 20$. As a result, the proposed user assignment scheme can efficiently improve the performance of cell-free Massive MIMO systems with limited fronthaul capacity. For instance, using the proposed user assignment scheme for the case of $\tau_p = 50$ in Fig. 6a, one can achieve per-user uplink rate of 2.442 bits/s/Hz by setting $K_m^{\text{opt}} = 20$, instead of quantizing the signals of all $K = 40$ users and achieving per-user uplink rate of 2.3 bits/s/Hz, which indicates more than 5.2% in the performance of cell-free Massive MIMO systems with limited fronthaul capacity.

6) *Effect of the Capacity of Fronthaul Links*: What is the optimal capacity of fronthaul links in cell-free Massive MIMO systems to approach the performance of the system with perfect and error-free fronthaul links? The aim of this subsection is to answer this fundamental question. In this subsection, we evaluate the performance of the cell-free Massive MIMO system with two cases of perfect and limited fronthaul links. To assess the performance, a cell-free Massive MIMO system is considered with $M = 120, K = 50, N = 2, D = 1$ km, $\tau_p = 30$ and $\tau_p = 50$. To improve the performance of the system, we exploit the proposed user assignment algorithm. Fig. 6b presents average per-user uplink rate with the proposed



(a) The convergence of the proposed max-min rate approach (Algorithm 1) for $M = 70$, $N = 4$, $K = 40$, $\tau_P = 30$, $\alpha_2 = 1$ and $D = 1$ km.



(b) The convergence of the proposed max-min rate approach (Algorithm 1) for $M = 30$, $N = 8$, $K = 50$, $\tau_P = 50$, $\alpha_2 = 1$ and $D = 1$ km.

Figure 4. Convergence

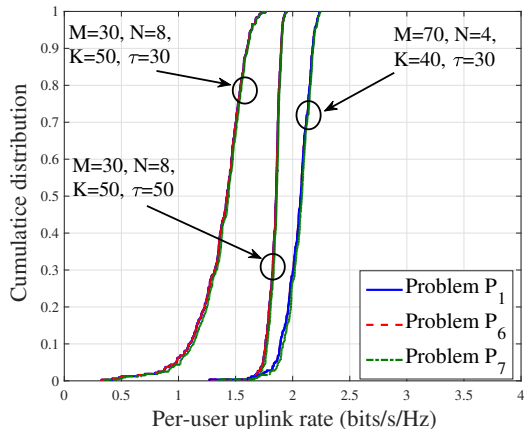


Figure 5. The cumulative distribution of the per-user uplink rate for the original problem with per-user power constraint (Problem P_1), the equivalent uplink problem with total power constraint (Problem P_5), and the equivalent downlink problem (Problem P_6), with $\alpha_1 = 1$ and $D = 1$ km.

max-min rate algorithm versus number of quantization bits, α_1 with the use of proposed user assignment algorithm. As Fig. 6b shows, for both cases of random and orthogonal pilots to closely approach the performance of perfect fronthaul links, we need to set $\alpha_1 \geq 8$.

IX. CONCLUSIONS

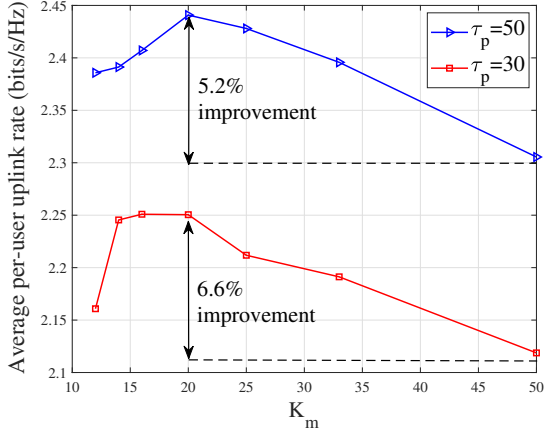
We have studied the uplink max-min rate problem in cell-free Massive MIMO with the realistic assumption of limited-capacity fronthaul links, and have proposed an optimal solution to maximize the minimum user rate. The original max-min problem was divided into two sub-problems which were iteratively solved by formulating them into generalized eigenvalue problem and GP. The optimality of the proposed solution has been validated through establishing an uplink-downlink duality. Numerical results have been provided to demonstrate the optimality of the proposed scheme in comparison with the existing schemes. In addition, these results confirmed that the proposed max-min rate algorithm can increase the median of the CDF of the minimum uplink rate of the users by more than two times, compared to existing algorithms. We finally showed that further improvement (more than three times) in

minimum rate of the users can be achieved by the proposed user assignment algorithm.

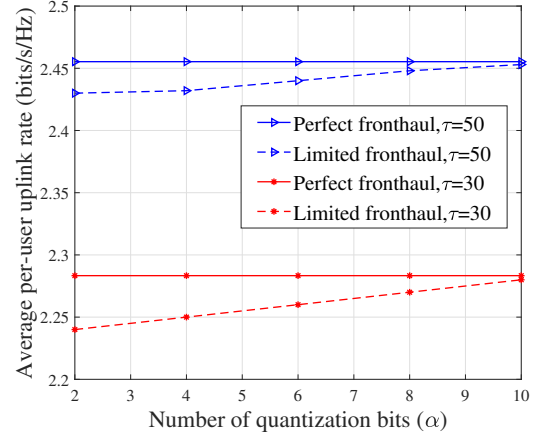
APPENDIX A: PROOF OF LEMMA 1

We exploit (3) and (6) to find a and b for uniform quantizer as follows:

$$\begin{aligned}
 a &= \frac{1}{p_z} \int_{-\frac{L}{2}+1}^{\infty} z h(z) f_z(z) dz = \frac{1}{p_z} \left(\int_{-\frac{L}{2}+1}^{-\frac{L-1}{2}} -z \frac{L-1}{2} \Delta f_z(z) dz \right. \\
 &\quad \left. + \sum_{l=-\frac{L}{2}+1}^{l=\frac{L}{2}-1} \int_{l\Delta}^{(l+1)\Delta} x \left(l + \frac{1}{2} \right) \Delta f_z(z) dz + \int_{\frac{L}{2}-1}^{\infty} z \frac{L-1}{2} \Delta f_z(z) dz \right) \\
 &\stackrel{a_1}{=} \frac{(L-1)\Delta}{2\sqrt{2\pi p_z}} \exp\left(-\frac{(\frac{L}{2}-1)^2}{2}\right) + \sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{(l+\frac{1}{2})\Delta}{\sqrt{2\pi p_z}} \\
 &\quad \left(\exp\left(-\frac{l^2\Delta^2}{2}\right) - \exp\left(-\frac{(l+1)^2\Delta^2}{2}\right) \right) + \frac{(L-1)\Delta}{2\sqrt{2\pi p_z}} \exp\left(-\frac{(\frac{L}{2}-1)^2}{2}\right) \\
 &\stackrel{a_2}{=} \sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{(l+\frac{1}{2})\Delta}{\sqrt{2\pi p_z}} \exp\left(-\frac{l^2\Delta^2}{2}\right) - \sum_{l'=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{(l'+\frac{1}{2})\Delta}{\sqrt{2\pi p_z}} \exp\left(-\frac{l'^2\Delta^2}{2}\right) \\
 &= \sum_{l=-\frac{L}{2}+1}^{\frac{L}{2}-2} \frac{\Delta}{\sqrt{2\pi p_z}} \exp\left(-\frac{l^2\Delta^2}{2}\right) = \sum_{l=1}^{\frac{L}{2}-1} \frac{2\Delta}{\sqrt{2\pi}} \exp\left(-\frac{l^2\Delta^2}{2}\right) + \frac{\Delta}{\sqrt{2\pi p_z}} \\
 &= \Delta \sqrt{\frac{2}{\pi p_z}} \left(\sum_{l=1}^{\frac{L}{2}-1} \exp\left(-\frac{l^2\Delta^2}{2}\right) + 1 \right), \tag{43}
 \end{aligned}$$



(a) Average per-user uplink rate versus the total number of active users for each AP with $M = 120$, $N = 2$, $K = 50$ and $\alpha_2 \times K_m = 100$.



(b) Average per-user uplink rate versus the number of quantization bits, α_2 , with limited and perfect fronthaul links and $M = 120$, $K = 50$, $N = 2$, $D = 1$ km, $\tau_p = 30$ and $\tau_p = 50$.

Figure 6. Average per-user uplink rate.

$$\begin{aligned}
 b &= \frac{\mathbb{E}\{h^2(z)\}}{\mathbb{E}\{z^2\}} = \frac{1}{p_z} \int_{-\infty}^{\infty} h^2(z) f_z(z) dz = \frac{2}{p_z} \left(\sum_{l=1}^{\frac{l}{2}-1} \int_{(l-1)\Delta}^{l\Delta} \left(l - \frac{1}{2}\right)^2 \Delta^2 f_z(z) dz \right. \\
 &+ \left. \int_{(\frac{l}{2}-1)\Delta}^{l\Delta} \left(\frac{L-1}{2}\right)^2 \Delta^2 f_z(z) dz \right) \stackrel{a_1}{=} \frac{2\Delta^2}{p_z} \left(\sum_{l=1}^{\frac{l}{2}-1} \left(l - \frac{1}{2}\right)^2 \left(\mathcal{Q}\left(\frac{(l-1)\Delta}{\sqrt{p_z}}\right) \right. \right. \\
 &- \left. \left. \mathcal{Q}\left(\frac{l\Delta}{\sqrt{p_z}}\right) \right) + \left(\frac{L-1}{2}\right)^2 \mathcal{Q}\left(\frac{(L-2)\Delta}{2p_z}\right) \right) \stackrel{a_2}{=} \frac{2\Delta^2}{p_z} \left(\sum_{l'=0}^{\frac{l}{2}-1} \left(l' + \frac{1}{2}\right)^2 \right. \\
 &\left. \mathcal{Q}\left(\frac{l'\Delta}{p_z}\right) - \sum_{l=1}^{\frac{l}{2}-1} \left(l - \frac{1}{2}\right)^2 \mathcal{Q}\left(\frac{l\Delta}{p_z}\right) \right) = \frac{2\Delta^2}{p_z} \left(\left(\frac{1}{2}\right)^2 \mathcal{Q}(0) \right. \\
 &\left. + \sum_{l=1}^{\frac{l}{2}-1} \left(\left(l + \frac{1}{2}\right)^2 - \left(l - \frac{1}{2}\right)^2 \right) \mathcal{Q}\left(\frac{l\Delta}{p_z}\right) \right) = \frac{\Delta^2}{p_z} \left(\frac{1}{4} + 4 \sum_{l=1}^{\frac{l}{2}-1} l \mathcal{Q}\left(\frac{l\Delta}{p_z}\right) \right), \quad (44)
 \end{aligned}$$

where the steps a_1 and a_2 come from the property that the input of the quantizer has the Gaussian distribution, and $l' = l+1$, respectively. ■

APPENDIX B: PROOF OF PROPOSITION 1

Terms \mathbf{e}_m^y and \mathbf{e}_{mk}^g have i.i.d. random variables with zero mean [31]. The value of the quantization error is uncorrelated with the input of the quantizer. This can be achieved by exploiting the Bussgang decomposition [12]. In this paper, we do not address the details of Bussgang decomposition and it can be considered an interesting future direction. As a result, we have $\mathbb{E}\{\mathbf{e}_m^y\} = 0$ & $\mathbb{E}\{\mathbf{e}_{mk}^g\} = 0$, $\mathbb{E}\{\mathbf{e}_m^y H \mathbf{e}_{mk}^g\} = 0$, $\mathbb{E}\{\mathbf{g}_{mk}^H \mathbf{e}_{mk}^g\} = 0$, $\mathbb{E}\{\mathbf{y}_m^H \mathbf{e}_{mk}^g\} = 0$, $\mathbb{E}\{\mathbf{y}_m^H \mathbf{e}_m^y\} = 0$, and $\mathbb{E}\{\mathbf{g}_{mk}^H \mathbf{e}_m^y\} = 0$. In addition, based on [1], we have $\mathbf{g}_{mk} = \hat{\mathbf{g}}_{mk} + \tilde{\mathbf{g}}_{mk}$, where $\tilde{\mathbf{g}}_{mk}$ has i.i.d. $CN(0, 1)$ elements. Hence, $\mathbb{E}\{\mathbf{g}_{mk}^H \mathbf{e}_m^y\} = 0$ & $\mathbb{E}\{\mathbf{g}_{mk}^H \mathbf{e}_{mk}^g\} = 0$. These result in

$$\mathbb{E}\{\text{TQN}_{kk'}\} = 0, \mathbb{E}\{\text{TQN}_k^g\} = 0, \quad (45)$$

$$\mathbb{E}\{\text{TQN}_k^y\} = 0, \mathbb{E}\{\text{TQN}_k^{gy}\} = 0. \quad (46)$$

Moreover, note that as the term DS_k is a constant, we have $\mathbb{E}\{\text{DS}_k^H \text{TQN}_k^y\} = \text{DS}_k^H \mathbb{E}\{\text{TQN}_k^y\} = 0$, and similarly $\mathbb{E}\{\text{DS}_k^H \text{TQN}_k^g\} = 0$, $\mathbb{E}\{\text{DS}_k^H \text{TQN}_k^{gy}\} = 0$, and $\mathbb{E}\{\text{DS}_k^H \text{TQN}_{kk'}\} = 0$. In addition, we have

$$\begin{aligned}
 \mathbb{E}\{\text{BU}_k^H \text{TQN}_{kk'}\} &= \mathbb{E}\left\{ \left(\sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right. \right. \\
 &- \left. \left. \underbrace{\left(\sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right)}_{A_1} \sum_{m=1}^M u_{mk} (\mathbf{e}_{mk}^g)^H \mathbf{g}_{mk'} \sqrt{q_{k'}} \right) \right\} \\
 &= \mathbb{E}\left\{ \left(\sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right)^H \left(\sum_{m=1}^M u_{mk} (\mathbf{e}_{mk}^g)^H \mathbf{g}_{mk'} \sqrt{q_{k'}} \right) \right\} \\
 &- \mathbb{E}\left\{ A_1^H \left(\sum_{m=1}^M u_{mk} (\mathbf{e}_{mk}^g)^H \mathbf{g}_{mk'} \sqrt{q_{k'}} \right) \right\}. \quad (47)
 \end{aligned}$$

For the first term of (47), we have

$$\begin{aligned}
 &\mathbb{E}\left\{ \left(\sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right)^H \left(\sum_{m=1}^M u_{mk} (\mathbf{e}_{mk}^g)^H \mathbf{g}_{mk'} \sqrt{q_{k'}} \right) \right\} \\
 &= \sqrt{q_k q_{k'}} \mathbb{E}\left\{ \sum_{m=1}^M \sum_{n=1}^M u_{mk} u_{nk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} (\mathbf{e}_{nk}^g)^H \mathbf{g}_{nk'} \right\} = 0, \quad (48)
 \end{aligned}$$

where the last equality is due to $\mathbb{E}\{\mathbf{g}_{mk}^H \mathbf{e}_m^y\} = 0$, $\mathbb{E}\{\mathbf{g}_{mk}^H \mathbf{e}_{mk}^g\} = 0$, and $\mathbb{E}\{\hat{\mathbf{g}}_{mk}^H \mathbf{e}_{mk}^g\} = 0$. For the second term of (47), as A_1 is a constant, and using $\mathbb{E}\{\mathbf{g}_{mk}^H \mathbf{e}_{mk}^g\} = 0$, we have

$$\mathbb{E}\left\{ A_1^H \left(\sum_{m=1}^M u_{mk} (\mathbf{e}_{mk}^g)^H \mathbf{g}_{mk'} \sqrt{q_{k'}} \right) \right\} = 0. \quad (49)$$

Finally, using (48) and (49), we have $\mathbb{E}\{\text{BU}_k^H \text{TQN}_{kk'}\} = 0$. Using the same approach, it is easy to show that the terms DS_k , BU_k , $\text{IUI}_{kk'}$, $\text{TQN}_{kk'}$, TQN_k^g , TQN_k^y , and TQN_k^{gy} are mutually uncorrelated, which completes the proof of Proposition 1. ■

APPENDIX C: PROOF OF THEOREM 2

The desired signal for the user k is given by

$$\text{DS}_k = \sqrt{\rho} \mathbb{E} \left\{ \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \sqrt{q_k} \right\} = N \sqrt{\rho q_k} \sum_{m=1}^M u_{mk} \gamma_{mk}. \quad (50)$$

Hence, $|\text{DS}_k|^2 = \rho q_k \left(N \sum_{m=1}^M u_{mk} \gamma_{mk} \right)^2$. Moreover, the term $\mathbb{E}\{|\text{BU}_k|^2\}$ can be obtained as

$$\begin{aligned} \mathbb{E}\{|\text{BU}_k|^2\} &= \rho \sum_{m=1}^M q_k u_{mk}^2 \left(\mathbb{E}\left\{ \left| \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} - \mathbb{E}\left\{ \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk} \right\} \right|^2 \right\} \right) \\ &= \rho N \sum_{m=1}^M q_k u_{mk}^2 \gamma_{mk} \beta_{mk}, \end{aligned} \quad (51)$$

where the last equality comes from the analysis in [2, Appendix A], and using $\gamma_{mk} = \sqrt{\tau_p p_p} \beta_{mk} c_{mk}$. The term $\mathbb{E}\{|\text{IU}_{kk'}|^2\}$ is obtained as

$$\begin{aligned} \mathbb{E}\{|\text{IU}_{kk'}|^2\} &= \rho q_{k'} \mathbb{E} \left\{ \underbrace{\left| \sum_{m=1}^M c_{mk} u_{mk} \mathbf{g}_{mk}^H \tilde{\mathbf{w}}_{mk} \right|^2}_A \right\} \\ &+ \rho \tau_p p_p \mathbb{E} \left\{ \underbrace{q_{k'} \left| \sum_{m=1}^M c_{mk} u_{mk} \left(\sum_{i=1}^K \mathbf{g}_{mi} \boldsymbol{\phi}_k^H \boldsymbol{\phi}_i \right)^H \mathbf{g}_{mk'} \right|^2}_B \right\}, \end{aligned} \quad (52)$$

where the third equality in (52) is due to the fact that for two independent random variables X and Y and $\mathbb{E}\{X\} = 0$, we have $\mathbb{E}\{|X+Y|^2\} = \mathbb{E}\{|X|^2\} + \mathbb{E}\{|Y|^2\}$ [2]. Since $\tilde{\mathbf{w}}_{mk} = \boldsymbol{\phi}_k^H \mathbf{W}_{p,m}$ is independent from the term $\mathbf{g}_{mk'}$ similar to [2, Appendix A], the term A in (52) immediately is given by $A = N q_{k'} \sum_{m=1}^M c_{mk}^2 u_{mk}^2 \beta_{mk'}$. The term B in (52) can be obtained as

$$\begin{aligned} B &= \tau_p p_p q_{k'} \mathbb{E} \left\{ \underbrace{\left| \sum_{m=1}^M c_{mk} u_{mk} \|\mathbf{g}_{mk'}\|^2 \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2}_C \right\} \\ &+ \tau_p p_p q_{k'} \mathbb{E} \left\{ \underbrace{\left| \sum_{m=1}^M c_{mk} u_{mk} \left(\sum_{i \neq k'}^K \mathbf{g}_{mi} \boldsymbol{\phi}_k^H \boldsymbol{\phi}_i \right)^H \mathbf{g}_{mk'} \right|^2}_D \right\}. \end{aligned} \quad (53)$$

The first term in (53) is given by

$$\begin{aligned} C &= \tau_p p_p q_{k'} \mathbb{E} \left\{ \left| \sum_{m=1}^M c_{mk} u_{mk} \|\mathbf{g}_{mk'}\|^2 \boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'} \right|^2 \right\} \\ &= N \tau_p p_p q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \sum_{m=1}^M c_{mk}^2 u_{mk}^2 \beta_{mk'}^2 \\ &+ N^2 q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \left(\sum_{m=1}^M u_{mk} \gamma_{mk} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2, \end{aligned} \quad (54)$$

where the last equality is derived based on the fact $\gamma_{mk} = \sqrt{\tau_p p_p} \beta_{mk} c_{mk}$. The second term in (53) can be obtained as

$$\begin{aligned} D &= \tau_p p_p q_{k'} \mathbb{E} \left\{ \left| \sum_{m=1}^M c_{mk} u_{mk} \left(\sum_{i \neq k'}^K \mathbf{g}_{mi} \boldsymbol{\phi}_k^H \boldsymbol{\phi}_i \right)^H \mathbf{g}_{mk'} \right|^2 \right\} \\ &= N \sqrt{\tau_p p_p} q_{k'} \sum_{m=1}^M u_{mk}^2 c_{mk} \beta_{mk'} \beta_{mk} - N q_{k'} \sum_{m=1}^M u_{mk}^2 c_{mk}^2 \beta_{mk'} \\ &- N \tau_p p_p q_{k'} \sum_{m=1}^M u_{mk}^2 c_{mk}^2 \beta_{mk'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2. \end{aligned} \quad (55)$$

Finally by substituting (54) and (55) into (53), and substituting (53) into (52), we obtain

$$\begin{aligned} \mathbb{E}\{|\text{IU}_{kk'}|^2\} &= N \rho q_{k'} \left(\sum_{m=1}^M u_{mk}^2 \beta_{mk'} \gamma_{mk} \right) \\ &+ N^2 \rho q_{k'} |\boldsymbol{\phi}_k^H \boldsymbol{\phi}_{k'}|^2 \left(\sum_{m=1}^M u_{mk} \gamma_{mk} \frac{\beta_{mk'}}{\beta_{mk}} \right)^2. \end{aligned} \quad (56)$$

The total noise for the user k is given by

$$\mathbb{E}\{|\text{TN}_k|^2\} = \mathbb{E} \left\{ \left| \sum_{m=1}^M u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m \right|^2 \right\} = N \sum_{m=1}^M u_{mk}^2 \gamma_{mk}, \quad (57)$$

where the last equality is due to the fact that the terms $\hat{\mathbf{g}}_{mk}$ and \mathbf{n}_m are uncorrelated. Based on the analysis in [39], we have

$$\mathbf{R}_{e_k^z e_k^z} \stackrel{(a)}{=} \sigma_e^2 \text{diag}(\mathbf{R}_{z_k z_k}), \quad (58)$$

where $\mathbf{R}_{e_k^z e_k^z}$ and $\mathbf{R}_{z_k z_k}$ refer to the covariance matrix of the quantization error and the covariance matrix of the input of the quantizer, respectively. Moreover, note that in step (a), we exploit the analysis in [39, Section V]. Thus, the power of the quantization error for user k is given by:

$$\mathbb{E}\{|\text{TQE}_k|^2\} = \mathbb{E} \left\{ \left| \sum_{m=1}^M u_{mk} e_{mk}^z \right|^2 \right\} = \sum_{m=1}^M u_{mk}^2 \mathbb{E}\{|e_{mk}^z|^2\}, \quad (59)$$

Finally, the power of the quantization error is obtained as the following:

$$\mathbb{E}\{|e_{mk}^z|^2\} = \mathbb{E}\{|z_{mk}^z|^2\} \sigma_{z_{mk}}^2 = \sigma_e^2 \sigma_{z_{mk}}^2, \quad (60)$$

where we used the fact that all APs use the same number of bits to quantize the weighted signal z_{mk} in (21). Next, the term $\sigma_{z_{mk}}^2$ is obtained as

$$\begin{aligned} \sigma_{z_{mk}}^2 &= \mathbb{E} \left\{ \left(\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m \right)^H \left(\hat{\mathbf{g}}_{mk}^H \mathbf{y}_m \right) \right\} = \mathbb{E} \left\{ \left(\sqrt{\rho} \sum_{k'=1}^K u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'} s_{k'}} \right. \right. \\ &\left. \left. + u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m \right)^H \left(\sqrt{\rho} \sum_{k'=1}^K u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'} s_{k'}} + u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m \right) \right\} \\ &\approx \rho \mathbb{E} \left\{ \left| \sum_{k'=1}^K u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'} s_{k'}} \right|^2 \right\} + \mathbb{E} \left\{ |u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{n}_m|^2 \right\}, \end{aligned} \quad (61)$$

where the approximation (61) is obtained byignoring the correlation between the terms $\hat{\mathbf{g}}_{mk}^H \mathbf{n}_m$ and $\sum_{k'=1}^K \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'} s_{k'}}$. Note that the simulation results confirm that this approximation is very tight [40]. The first term in (61) can be obtained as

$$\begin{aligned} &\mathbb{E} \left\{ \left| \sum_{k'=1}^K u_{mk} \hat{\mathbf{g}}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'} s_{k'}} \right|^2 \right\} \\ &= \mathbb{E} \left\{ \left| \sum_{k'=1}^K u_{mk} (\mathbf{g}_{mk} - \boldsymbol{\epsilon}_{mk})^H \mathbf{g}_{mk'} \sqrt{q_{k'} s_{k'}} \right|^2 \right\} \\ &= \mathbb{E} \left\{ \underbrace{\left| \sum_{k'=1}^K u_{mk} \mathbf{g}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'} s_{k'}} \right|^2}_I + \mathbb{E} \left\{ \underbrace{\left| \sum_{k'=1}^K u_{mk} \boldsymbol{\epsilon}_{mk}^H \mathbf{g}_{mk'} \sqrt{q_{k'} s_{k'}} \right|^2}_II \right\} \right\}, \end{aligned} \quad (62)$$

where each element of ϵ is given by $\epsilon_{mk} = \mathcal{CN}(0, \beta_{mk} - \gamma_{mk})$. The terms I and II in (62) are given as following: I = $N\beta_{mk} \sum_{k'=1}^K q_{k'}\beta_{mk'}$, and II = $N(\beta_{mk} - \gamma_{mk}) \sum_{k'=1}^K q_{k'}\beta_{mk'}$. Finally, we have

$$\begin{aligned} & \mathbb{E} \{ |\text{TQE}_k|^2 \} \\ & \approx N\sigma_{\check{e}}^2 \sum_{m=1}^M u_{mk}^2 \left[\sqrt{\rho}(2\beta_{mk} - \gamma_{mk}) \sum_{k'=1}^K q_{k'}\beta_{mk'} + \gamma_{mk} \right], \end{aligned} \quad (63)$$

By substituting (50), (51), (56) and (57) into (22), the corresponding SINR of the k th user is obtained by (23), which completes the proof of Theorem 2. ■

APPENDIX D: PROOF OF PROPOSITION 2

The standard form of GP is defined as follows [36]:

$$P_{12}: \quad \min \quad f_0(\mathbf{x}), \quad (64a)$$

$$\text{subject to} \quad f_i(\mathbf{x}) \leq 1, i = 1, \dots, m, g_i(\mathbf{x}) = 1, i = 1, \dots, p, \quad (64b)$$

where f_0 and f_i are posynomial and g_i are monomial functions. Moreover, $\mathbf{x} = \{x_1, \dots, x_n\}$ represent the optimization variables. The SINR constraint in (64) is not a posynomial function in this form, however it can be rewritten as the following posynomial function:

$$\frac{\mathbf{u}_k^H \left(N^2 \sum_{k' \neq k}^K q_{k'} |\phi_k^H \phi_{k'}|^2 \Lambda_{kk'} \Lambda_{kk'}^H + N \sum_{k'=1}^K q_{k'} \Upsilon_{kk'} + \frac{N\mathbf{R}_k}{\rho} \right) \mathbf{u}_k}{\mathbf{u}_k^H (N^2 q_k \Gamma_k \Gamma_k^H) \mathbf{u}_k} < \frac{1}{t}, \forall k. \quad (65)$$

By applying a simple transformation, (65) is equivalent to the following inequality:

$$q_k^{-1} \left(\sum_{k' \neq k}^K a_{kk'} q_{k'} + \sum_{k'=1}^K b_{kk'} q_{k'} + c_k \right) < \frac{1}{t}, \forall k, \quad (66)$$

where $a_{kk'} = \frac{\mathbf{u}_k^H (|\phi_k^H \phi_{k'}|^2 \Lambda_{kk'} \Lambda_{kk'}^H) \mathbf{u}_k}{\mathbf{u}_k^H (\Gamma_k \Gamma_k^H) \mathbf{u}_k}$, $b_{kk'} = \frac{\mathbf{u}_k^H \Upsilon_{kk'} \mathbf{u}_k}{\mathbf{u}_k^H (\Gamma_k \Gamma_k^H) \mathbf{u}_k}$ and $c_k = \frac{\mathbf{u}_k^H \mathbf{R}_k \mathbf{u}_k}{\rho \mathbf{u}_k^H (\Gamma_k \Gamma_k^H) \mathbf{u}_k}$. The transformation in (66) shows that the left-hand side of (65) is a posynomial function. Hence, the power allocation Problem P_4 is a GP (convex problem), where the objective function and constraints are monomial and posynomial, respectively, which completes the proof of Proposition 2. ■

APPENDIX E: PROOF OF LEMMA 2

This lemma is proven by exploiting the unique optimal solution of the uplink max-min SINR problem with total power limitation through an eigensystem [20]. This problem is iteratively solved and the optimal receiver filter coefficients $\check{\mathbf{U}}$ are determined by solving Problem P_3 . Next, we scale the power allocation at each user such that the per-user power constraints are satisfied. Let us consider the following optimization problem for a given receiver filter coefficients $\check{\mathbf{U}}$:

$$P_{11}: C^{\text{UP}}(\check{\mathbf{U}}, P_{\text{tot}}) = \max_{q_k} \min_{k=1, \dots, K} \text{SINR}_k^{\text{UP}}(\check{\mathbf{U}}, \mathbf{q}), \quad (67a)$$

$$\text{subject to} \quad \sum_{k=1}^K q_k \leq P_{\text{tot}}. \quad (67b)$$

The optimal solution of Problem P_{11} can be determined by finding the unique eigenvector associated with unique positive eigenvalue of an eigensystem and the power allocation $\check{\mathbf{q}}$ that satisfies the following condition [20]:

$$\sum_{k=1}^K \check{q}_k = P_{\text{tot}}. \quad (68)$$

The SINRs of all users can be collectively written as

$$\check{\mathbf{q}} \frac{1}{C_k^{\text{UP}}(\check{\mathbf{U}}, P_{\text{tot}})} = \mathbf{D}\Psi(\check{\mathbf{U}})\check{\mathbf{q}} + \mathbf{D}\sigma(\check{\mathbf{U}}), \quad (69)$$

where $\sigma(\check{\mathbf{U}}) \in \mathbb{C}^{K \times 1}$, $\sigma_k(\mathbf{u}_k) = \frac{N}{\rho} \left(\frac{\sigma_{\check{e}}^2}{\hat{d}^2} + 1 \right) \sum_{m=1}^M \check{u}_{mk} \gamma_{mk}$ and \mathbf{D} and $\Psi(\check{\mathbf{U}})$ are defined as

$$\mathbf{D} = \text{diag} \left[\frac{1}{\check{\mathbf{u}}_1^H \check{\mathbf{D}}_1 \check{\mathbf{u}}_1}, \dots, \frac{1}{\check{\mathbf{u}}_K^H \check{\mathbf{D}}_K \check{\mathbf{u}}_K} \right], \quad (70)$$

$$\left[\Psi(\check{\mathbf{U}}) \right]_{kk'} = \begin{cases} \check{\mathbf{u}}_k^H \check{\mathbf{R}}_{kk} \check{\mathbf{u}}_k, & k = k', \\ \check{\mathbf{u}}_k^H \check{\mathbf{R}}_{kk'} \check{\mathbf{u}}_k + \check{\mathbf{u}}_k^H \check{\mathbf{R}}_{k'k} \check{\mathbf{u}}_k, & k \neq k', \end{cases} \quad (71)$$

where $\check{\mathbf{D}}_k$, $\check{\mathbf{R}}_{kk'}$ and $\check{\mathbf{R}}_{k'k}$ are defined as

$$\text{SINR}_k^{\text{UP}} = \quad (72)$$

$$\frac{q_k \mathbf{u}_k^H \left(N^2 \Gamma_k \Gamma_k^H \right) \mathbf{u}_k}{\mathbf{u}_k^H \left(\sum_{k' \neq k}^K q_{k'} \underbrace{N^2 |\phi_k^H \phi_{k'}|^2 \Lambda_{kk'} \Lambda_{kk'}^H}_{\check{\mathbf{R}}_{kk'}} + \sum_{k'=1}^K q_{k'} \underbrace{N \Upsilon_{kk'}}_{\check{\mathbf{R}}_{k'k}} + \frac{N\mathbf{R}_k}{\rho} \right) \mathbf{u}_k}.$$

Having both sides of (69) multiplied by $\mathbf{1}^T = [1, \dots, 1]$, we obtain $\frac{1}{C_k^{\text{UP}}(\check{\mathbf{U}}, P_{\text{tot}})} = \frac{1}{P_{\text{tot}}} \mathbf{1}^T \check{\mathbf{D}}\Psi(\check{\mathbf{U}})\check{\mathbf{q}} + \frac{1}{P_{\text{tot}}} \mathbf{1}^T \mathbf{D}\sigma(\check{\mathbf{U}})$, which can be combined with (69) to define the following eigensystem:

$$\Lambda(\check{\mathbf{U}}, P_{\text{tot}})\check{\mathbf{q}}_{\text{ext}} = \frac{1}{C_k^{\text{UP}}(\check{\mathbf{U}}, P_{\text{tot}})} \check{\mathbf{q}}_{\text{ext}}, \quad (73)$$

$$\Lambda(\check{\mathbf{U}}, P_{\text{tot}}) = \begin{bmatrix} \mathbf{D}\Psi^T(\check{\mathbf{U}}) & \mathbf{D}\sigma(\check{\mathbf{U}}) \\ \frac{1}{P_{\text{tot}}} \mathbf{1}^T \mathbf{D}\Psi^T(\check{\mathbf{U}}) & \frac{1}{P_{\text{tot}}} \mathbf{1}^T \mathbf{D}\sigma(\check{\mathbf{U}}) \end{bmatrix}. \quad (74)$$

The optimal power allocation $\check{\mathbf{q}}$ is obtained by determining the eigenvector corresponding to the maximum eigenvalue of $\Lambda(\check{\mathbf{U}}, P_{\text{tot}})$ and scaling the last element to one as follows:

$$\Lambda(\check{\mathbf{U}}, P_{\text{tot}})\check{\mathbf{q}}_{\text{ext}} = \lambda_{\max}(\Lambda(\check{\mathbf{U}}, P_{\text{tot}}))\check{\mathbf{q}}_{\text{ext}}, \quad \check{\mathbf{q}}_{\text{ext}} = \begin{bmatrix} \check{\mathbf{q}} \\ 1 \end{bmatrix}, \quad (75)$$

Note that to find the optimal power allocation $\check{\mathbf{q}}$, the elements of eigenvector of $\Lambda(\check{\mathbf{U}}, P_{\text{tot}})$ should be scaled such that the

last element is one to satisfy the total power constraint. In particular, the element of the eigenvector that needs to be scaled depends on the type of power constraint in the problem. For example, to meet the total power constraint, the last element is scaled to one. Similarly, to meet the other types of power constraints (for example, per-user power constraint), the components of this eigenvector can be scaled by any positive value to satisfy a given condition as follows:

$$\Lambda(\check{\mathbf{U}}, P_{\text{tot}}) \delta_{\text{cons}} \check{\mathbf{q}}_{\text{ext}} = \lambda_{\text{max}} \left(\Lambda(\check{\mathbf{U}}, P_{\text{tot}}) \right) \delta_{\text{cons}} \check{\mathbf{q}}_{\text{ext}}, \quad (76)$$

where δ_{cons} is a positive constant. This is the key fact that exploited to show that both Problems P_2 and P_6 provide the same optimal solution. We further scale the power allocation $\check{\mathbf{q}}$ to satisfy the per-user power constraints which is performed through carrying out the following two steps:

$$\bar{\mathbf{q}} = \begin{bmatrix} \check{q}_1 \\ \frac{(1)}{P_{\text{max}}} \\ \vdots \\ \check{q}_K \\ \frac{(K)}{P_{\text{max}}} \end{bmatrix}. \quad (77)$$

Next, we find the maximum value among the elements of $\bar{\mathbf{q}}$, i.e., $\max(\bar{\mathbf{q}})$, and divide all elements of $\bar{\mathbf{q}}$ by it. Hence the power allocation $\check{\check{\mathbf{q}}}$ is defined as follows:

$$\check{\check{\mathbf{q}}} = \begin{bmatrix} \check{q}_1 \\ \frac{\check{q}_1}{\max(\bar{\mathbf{q}})} \\ \vdots \\ \check{q}_K \\ \frac{\check{q}_K}{\max(\bar{\mathbf{q}})} \end{bmatrix}. \quad (78)$$

In the next iteration, the same max-min problem is solved with a new total power constraint obtained by summing up the allocated power to all users in the previous iteration:

$$L_1 : \mathcal{C}^{\text{UP}}(\check{\mathbf{U}}, P_{\text{tot}}^{\text{new}}) = \max_{q_k} \min_{k=1, \dots, K} \text{SINR}_k^{\text{UP}}(\check{\mathbf{U}}, \mathbf{q}), \quad (79a)$$

$$\text{subject to } \sum_{k=1}^K q_k \leq P_{\text{tot}}^{\text{new}}, \quad \text{where } P_{\text{tot}}^{\text{new}} = \sum_{k=1}^K \check{\check{q}}_k. \quad (79b)$$

At the convergence of the algorithm, the per-user power constraints are satisfied with achieving the same uplink SINR for each user. Interestingly, if this max-min problem is solved with the corresponding total power constraint, then it will converge to the same optimal solution of max-min problem with per-user power constraints. This is due to the property that the eigensystem exploited to obtain the power allocation in (74) has a unique positive eigenvalue and a corresponding unique eigenvector. Furthermore, in both Problems P_2 and P_6 , different elements of the same eigenvector are scaled to meet the corresponding constraints on the power allocation. In other words, the last element is scaled to meet the total power constraint in P_6 whereas the element with the highest ratio as in (76) is scaled to meet the per-user power constraint. As the equivalent total power P_{tot}^c for Problem P_6 chosen from the solution of the original P_2 , both of them will converge to the same solution whose optimality is proven later by considering an equivalent problem related to the virtual downlink SINR. Therefore, Problems P_2 and P_6 are equivalent and have the same optimal solution. ■

APPENDIX F: PROOF OF THEOREM 4

To achieve the same SINR tuples in both the uplink and the downlink, we need:

$$\text{SINR}_k^{\text{DL}}(\mathbf{U}, \mathbf{p}) = \text{SINR}_k^{\text{UP}}(\mathbf{U}, \mathbf{q}), \quad \forall k. \quad (80)$$

By substituting uplink and downlink SINRs, in (33) and (32), respectively, in equation (80) and summing all equations by both sides, we have

$$\begin{aligned} p_1 N \sum_{m=1}^M \left(\frac{\sigma_e^2}{d^2} + 1 \right) u_{m1}^2 \gamma_{m1} + \dots + p_K N \sum_{m=1}^M \left(\frac{\sigma_e^2}{d^2} + 1 \right) u_{mK}^2 \gamma_{mK} \\ = \sum_{k=1}^K q_k. \end{aligned} \quad (81)$$

Therefore, this condition between the total transmit power on the uplink and the equivalent total transmit power on the downlink should be satisfied to realize the same SINRs for all users. ■

REFERENCES

- [1] 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.
- [2] 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.
- [3] 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.
- [4] 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.
- [5] M. Bashar, K. Cumanan, A. Burr, H. Q. Ngo, L. Hanzo, and P. Xiao, "NOMA/OMA mode selection-based cell-free Massive MIMO," in *Proc. IEEE ICC*, May 2019.
- [6] S. Buzzi and C. DAndrea, "Cell-free Massive MIMO: User-centric approach," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 1–4, Aug. 2017.
- [7] G. Interdonato, E. Bjornson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free Massive MIMO communications," *IEEE Commun. Mag.*, pp. 1–19, submitted.
- [8] J. Li, D. Wang, P. Zhu, J. Wang, and X. You, "Downlink spectral efficiency of distributed Massive MIMO systems with linear beamforming under pilot contamination," *IEEE Trans. Veh. Technol.*, vol. 67, no. 2, pp. 1130–1145, Feb. 2018.
- [9] 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.
- [10] Z. Gao, L. Dai, D. Mi, Z. Wang, M. A. Imran, and M. Z. Shakir, "MmWave Massive-MIMO-based wireless backhaul for the 5G ultra-dense network," *IEEE Trans. Wireless Commun.*, vol. 22, no. 5, pp. 13–21, Oct. 2015.
- [11] M. Bashar, K. Cumanan, A. Burr, H. Q. Ngo, E. Larsson, and P. Xiao, "On the energy efficiency of limited-backhaul cell-free Massive MIMO," in *Proc. IEEE ICC*, May 2019.
- [12] P. Zillmann, "Relationship between two distortion measures for memoryless nonlinear systems," *IEEE Signal Process. Lett.*, vol. 17, no. 11, pp. 917–920, Feb. 2010.
- [13] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and power optimization in cell-free Massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, Jul. 2017.
- [14] 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.
- [15] M. Bashar, H. Q. Ngo, K. Cumanan, A. G. Burr, D. Maryopi, and E. G. Larsson, "On the performance of backhaul constrained cell-free Massive MIMO with linear receivers," in *IEEE Asilomar*, 2018, pp. 1–6.

- [16] K. Cumanan, R. Krishna, L. Musavian, and S. Lambotharan, "Joint beamforming and user maximization techniques for techniques for cognitive radio networks based on branch and bound method," *IEEE Trans. Wireless Commun.*, vol. 9, no. 10, pp. 3082–3092, Oct. 2010.
- [17] K. Cumanan, L. Musavian, S. Lambotharan, and A. B. Gershman, "SINR balancing technique for downlink beamforming in cognitive radio networks," *IEEE Signal Process. Lett.*, vol. 17, no. 2, pp. 133–136, Feb. 2010.
- [18] A. Wiesel, Y. C. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 1, pp. 161 – 176, Jan. 2006.
- [19] D. W. H. Cai, T. Q. S. Quek, and C. W. Tan, "A unified analysis of max-min weighted SINR for MIMO downlink system," *IEEE Trans. Signal Process.*, vol. 59, no. 8, pp. 3850–3862, Aug. 2011.
- [20] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Veh. Technol.*, vol. 53, no. 1, pp. 18 – 28, Jan. 2004.
- [21] D. N. C. Tse and P. Viswanath, "Downlink-uplink duality and effective bandwidths," in *Proc. IEEE ISIT*, Jul. 2002.
- [22] G. Golub and C. V. Loan, *Matrix Computations*, 2nd ed. Baltimore, MD: The Johns Hopkins Univ. Press, 1996.
- [23] M. Chiang, C. W. Tan, D. P. Palomar, D. O'Neill, and D. Julian, *Power Control by Geometric Programming, in Resource Allocation in Next Generation Wireless Networks*. W. Li, Y. Pan, Editors, Nova Sciences Publishers, 2006.
- [24] 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, Apr. 2019.
- [25] T. S. Rappaport, *Wireless Communications: Principles and Practice*. Englewood Cliffs, NJ, USA: Prentice-Hall, 2002.
- [26] A. Ashikhmin, T. L. Marzetta, and L. Li, "Interference reduction in multi-cell Massive MIMO systems I: Large-scale fading precoding and decoding," Available: <http://arxiv.org/abs/1411.4182>, Submitted.
- [27] 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.
- [28] D. Maryopi, M. Bashar, and A. Burr, "On the uplink throughput of zero-forcing in cell-free Massive MIMO with coarse quantization," *IEEE Trans. Veh. Technol.*, pp. 1–5, Jun. 2019.
- [29] L. C. Andrews, *Special Functions of Mathematics for Engineers*, 2nd ed. Oxford University Press, 1997.
- [30] J. Max, "Quantizing for minimum distortion," *IEEE Trans. Inf. Theory*, vol. 6, no. 1, pp. 7–12, Mar. 1960.
- [31] A. V. Oppenheim, R. W. Schaffer, and J. R. Buck, *Discrete-time signal processing*. Prentice-hall Englewood Cliffs, 1989.
- [32] M. Bashar, *Cell-free Massive MIMO and Millimeter Wave Channel Modelling for 5G and Beyond*. Ph.D. dissertation, University of York, United Kingdom, 2019.
- [33] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.
- [34] A. Pizzo, D. Verenzuela, L. Sanguinetti, and E. Björnson, "Network deployment for maximal energy efficiency in uplink with multislope path loss," *IEEE Trans. Green Commun. and Net.*, pp. 1–30, submitted.
- [35] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, and H. V. Poor, "Mixed quality of service in cell-free Massive MIMO," *IEEE Commun. Lett.*, vol. 22, no. 7, pp. 706–709, Jul. 2018.
- [36] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [37] M. Schubert and H. Boche, "Iterative multiuser uplink and downlink beamforming under sinr constraints," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2324–2334, Jul. 2005.
- [38] 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.
- [39] 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.
- [40] 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 Netw.*, Accepted.



Manijeh Bashar received the B.Sc. degree in electrical engineering from the University of Gilan, Rasht, Iran, in 2009, and the M.Sc. degree in communication systems engineering (with honors) from the Shiraz University of Technology, Shiraz, Iran, in 2013. She received the Ph.D. degree in communications engineering from the University of York, York, U.K. in 2019. In 2017, she was an Academic Visitor with the Department of Electronics and Nanoengineering, Aalto University, Espoo, Finland, with a Short Term Scientific Mission (STSM) Scholarship Award from European COST-IC1004 "Cooperative Radio Communications for Green Smart Environments".

Manijeh Bashar is currently a research fellow at the Institute for Communication Systems, home of 5G Innovation Centre (5GIC) at the University of Surrey. Her current research interests include cooperative communications for 5G networks including distributed Massive MIMO, Cloud-RAN, Fog-RAN, NOMA, deep learning, resource allocation, and also millimetre-wave channel modelling.

She received the K. M. Stott Prize for excellent research in electronics engineering from the University of York, U.K. in 2019. She has been awarded First place (based on jury) in the IEEE WCNC'18 three-minute PhD thesis competition for her research in cell-free Massive MIMO.



Dr. Cumanan received the BSc degree with first class honors in electrical and electronic engineering from the University of Peradeniya, Sri Lanka in 2006 and the PhD degree in signal processing for wireless communications from Loughborough University, Loughborough, UK, in 2009.

He is currently a lecturer at the Department of Electronic Engineering, The University of York, UK. From March 2012 to November 2014, he was working as a research associate at School of Electrical and Electronic Engineering, Newcastle University, UK. Prior to this, he was with the School of Electronic, Electrical and System Engineering, Loughborough University, UK. In 2011, he was an academic visitor at Department of Electrical and Computer Engineering, National University of Singapore, Singapore. From January 2006 to August 2006, he was a teaching assistant with Department of Electrical and Electronic Engineering, University of Peradeniya, Sri Lanka. His research interests include non-orthogonal multiple access (NOMA), cell-free massive MIMO, physical layer security, cognitive radio networks, convex optimization techniques and resource allocation techniques. He has published more than 80 journal articles and conference papers which attracted more than 1200 Google scholar citations. He has been also recently appointed as an associate editor for IEEE Access journal.

Dr. Cumanan was the recipient of an overseas research student award scheme (ORSAS) from Cardiff University, Wales, UK, where he was a research student between September 2006 and July 2007.



Alister G. Burr was born in London, U.K., in 1957. He received the B.Sc. degree in electronic engineering from the University of Southampton, Southampton, U.K., in 1979, and the Ph.D. degree from the University of Bristol, Bristol, U.K., in 1984. Between 1975 and 1985, he was with Thorn-EMI Central Research Laboratories, London. In 1985, he joined the Department of Electronics (now Electronic Engineering), University of York, York, U.K., where he has been Professor of communications since 2000. His research interests include wireless

communication systems, especially multiple-input multiple-output, cooperative systems, physical layer network coding, and iterative detection and decoding techniques. He has authored or coauthored around 250 papers in refereed international conferences and journals. He has authored the book entitled *Modulation and Coding for Wireless Communications* (Englewood Cliffs, NJ, USA: PrenticeHall/PHEI, 2001), and coauthored the book entitled *Wireless Physical-Layer Network Coding* (Cambridge, U.K.: Cambridge Univ. Press, 2018). He was a recipient of a Senior Research Fellowship by the U.K. Royal Society in 1999 and the J. Langham Thompson Premium from the Institution of Electrical Engineers in 2002. He has also given more than 15 invited presentations, including three keynote presentations. He was the Chair, working group 2, of a series of European COST programs including IC1004 “Cooperative Radio Communications for Green Smart Environments” (which have been influential in 3GPP standardization). He was an Associate Editor for the *IEEE Communications Letters*, the Workshop Chair for the 2016 IEEE International Conference on Communications, and the Technical Program Committee Co-Chair for the 2018 Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications.



Hien Quoc Ngo received the B.S. degree in electrical engineering from the Ho Chi Minh City University of Technology, Vietnam, in 2007, the M.S. degree in electronics and radio engineering from Kyung Hee University, South Korea, in 2010, and the Ph.D. degree in communication systems from Linköping University (LiU), Sweden, in 2015. In 2014, he visited the Nokia Bell Labs, Murray Hill, New Jersey, USA. From January 2016 to April 2017, Hien Quoc Ngo was a VR researcher at the Department of Electrical Engineering (ISY), LiU.

He was also a Visiting Research Fellow at the School of Electronics, Electrical Engineering and Computer Science, Queen’s University Belfast, UK, funded by the Swedish Research Council.

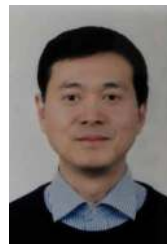
Hien Quoc Ngo is currently a Lecturer at Queen’s University Belfast, UK. His main research interests include massive (large-scale) MIMO systems, cell-free massive MIMO, physical layer security, and cooperative communications. He has co-authored many research papers in wireless communications and co-authored the Cambridge University Press textbook *Fundamentals of Massive MIMO* (2016).

Dr. Hien Quoc Ngo received the IEEE ComSoc Stephen O. Rice Prize in Communications Theory in 2015, the IEEE ComSoc Leonard G. Abraham Prize in 2017, and the Best PhD Award from EURASIP in 2018. He also received the IEEE Sweden VT-COM-IT Joint Chapter Best Student Journal Paper Award in 2015. He was an *IEEE Communications Letters* exemplary reviewer for 2014, an *IEEE Transactions on Communications* exemplary reviewer for 2015, and an *IEEE Wireless Communications Letters* exemplary reviewer for 2016. Dr. Hien Quoc Ngo currently serves as an Editor for the *IEEE Wireless Communications Letters*, *Digital Signal Processing*, and *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*. He was a Guest Editor of *IET Communications*, special issue on “Recent Advances on 5G Communications” and a Guest Editor of *IEEE Access*, special issue on “Modelling, Analysis, and Design of 5G Ultra-Dense Networks”, in 2017. He has been a member of Technical Program Committees for several IEEE conferences such as ICC, GLOBECOM, WCNC, and VTC.



Mérouane Debbah (S’01–M’04–SM’08–F’15) received the M.Sc. and Ph.D. degrees from the Ecole Normale Supérieure Paris-Saclay, France. He was with Motorola Labs, Saclay, France, from 1999 to 2002, and also with the Vienna Research Center for Telecommunications, Vienna, Austria, until 2003. From 2003 to 2007, he was an Assistant Professor with the Mobile Communications Department, Institut Eurecom, Sophia Antipolis, France. From 2007 to 2014, he was the Director of the Alcatel-Lucent Chair on Flexible Radio. Since 2007, he has

been a Full Professor with CentraleSupélec, Gif-sur-Yvette, France. Since 2014, he has been a Vice-President of the Huawei France Research Center and the Director of the Mathematical and Algorithmic Sciences Lab. He has managed 8 EU projects and more than 24 national and international projects. His research interests lie in fundamental mathematics, algorithms, statistics, information, and communication sciences research. He is an IEEE Fellow, a WWRF Fellow, and a Membre émérite SEE. He was a recipient of the ERC Grant MORE (Advanced Mathematical Tools for Complex Network Engineering) from 2012 to 2017. He was a recipient of the Mario Boella Award in 2005, the IEEE Glavieux Prize Award in 2011, and the Qualcomm Innovation Prize Award in 2012. He received 20 best paper awards, among which the 2007 IEEE GLOBECOM Best Paper Award, the Wi-Opt 2009 Best Paper Award, the 2010 Newcom++ Best Paper Award, the WUN CogCom Best Paper 2012 and 2013 Award, the 2014 WCNC Best Paper Award, the 2015 ICC Best Paper Award, the 2015 IEEE Communications Society Leonard G. Abraham Prize, the 2015 IEEE Communications Society Fred W. Ellersick Prize, the 2016 IEEE Communications Society Best Tutorial Paper Award, the 2016 European Wireless Best Paper Award, the 2017 Eurasip Best Paper Award, the 2018 IEEE Marconi Prize Paper Award, the 2019 IEEE Communications Society Young Author Best Paper Award and the Valuetools 2007, Valuetools 2008, CrownCom 2009, Valuetools 2012, SAM 2014, and 2017 IEEE Sweden VT-COM-IT Joint Chapter best student paper awards. He is an Associate Editor-in-Chief of the journal *Random Matrix: Theory and Applications*. He was an Associate Area Editor and Senior Area Editor of the *IEEE TRANSACTIONS ON SIGNAL PROCESSING* from 2011 to 2013 and from 2013 to 2014, respectively.



Pei Xiao (SM’11) is a professor of Wireless Communications at the Institute for Communication Systems, home of 5G Innovation Centre (5GIC) at the University of Surrey. He is the technical manager of 5GIC, leading the research team in the new physical layer work area, and coordinating/supervising research activities across all the work areas within 5GIC (www.surrey.ac.uk/5gic/research). Prior to this, he worked at Newcastle University and Queen’s University Belfast. He also held positions at Nokia Networks in Finland. He has published extensively

in the fields of communication theory and signal processing for wireless communications.