Scalable User Rate and Energy-Efficiency Optimization in Cell-Free Massive MIMO

H. D. Tuan, *Member, IEEE*, A. A. Nasir, H. Q. Ngo, *Senior Member, IEEE*, E. Dutkiewicz, *Senior Member, IEEE*, and H. V. Poor, *Life Fellow, IEEE*

***Abstract*— This paper considers a cell-free massive multiple-input multiple-output network (cfm-MIMO) with a massive number of access points (APs) distributed across an area to deliver information to multiple users. Based on only local channel state information, conjugate beamforming is used under both proper and improper Gaussian signalings. To accomplish the mission of cfm-MIMO in providing fair service to all users, the problem of power allocation to maximize the geometric mean (GM) of users' rates (GM-rate) is considered. A new scalable algorithm, which iterates linear-complex closed-form expressions and thus is practical regardless of the scale of the network, is developed for its solution. The problem of quality-of-service (QoS) aware network energy-efficiency is also addressed via maximizing the ratio of the GM-rate and the total power consumption, which is also addressed by iterating linear-complex closed-form expressions. Intensive simulations are provided to demonstrate the ability of the GM-rate based optimization to achieve multiple targets such as a uniform QoS, a good sum rate, and a fair power allocation to the APs.**

***Index Terms*— Cell-free massive MIMO (cfm-MIMO), conjugate beamforming (CB), energy efficiency, geometric mean, nonconvex optimization, scalable algorithms.**

I. Introduction

FOR massive access beyond 5G (B5G), cell-free massive MIMO (cfm-MIMO) [1]–[3], under which a massive number of access points (APs) are distributed across an area to provide uniform service to multiple users, has attracted significant attention. Resource allocation problems for spectral efficiency (SE) and energy efficiency (EE) maximization in a cfm-MIMO system are of paramount importance [4], [5] but are computationally demanding due to the network size.

To mitigate the computational burden caused by the large scale of cmf-MIMO, zero-forcing beamforming (ZFB) and conjugate beamforming (CB) are used [1], [6] with the preference going for the latter due to its low computational complexity yet substantial performance gain as the number of APs increases. Note that when the number of APs grows large, the channels become favorable, and hence, even with CB, the inter-user interference can be reduced significantly. More importantly, CB can be implemented in a distributed manner [1]. Other linear processing techniques such as MMSE/regularized ZFB may enhance the system performance, which however require high computational complexity (due to taking the expectation of random matrix inversions) and are subject to backhaul specifications. In addition, ZFB is very sensitive in scenarios where the channel matrix is rank deficient such as double-scattering channels, or correlated channels. This is a major reason why CB has continued to attract attention [7]–[11]. The conventional SE index is the minimum users' rate, which has been incorporated into constraints [12]–[15] or optimization objective functions [13], [15], [16]. However all algorithms proposed in these works must be based on iterating large-scale convex problems, which is computationally demanding.

Against the above background, this paper offers the following contributions:

- We show that the geometric mean of users' rates (GM-rate) is an appropriate SE index. As GM-rate is a function of the product of the all users' rates, its maximization gives rise to rates for users suffering from poor channel conditions without enforcing QoS constraints, maintaining a good sum rate (SR). Analogously, we show that GM-EE as the ratio of the GM-rate to the total power consumption is a meaningful EE index for quantifying the EE as its maximization also still maintains the QoS to all users. This is in contrast to the conventional SR-EE as the ratio of the SR to the total power consumption, which is a meaningful EE index only under additional QoS constraints.
- We develop a sophisticated but scalable optimization algorithms for these GM-rate and GM-EE maximization problems, which iterate linear-complex closed-form

Manuscript received 1 September 2021; revised 12 January 2022 and 31 March 2022; accepted 19 July 2022. Date of publication 26 July 2022; date of current version 16 September 2022. The work was supported in part by the Australian Research Council's Discovery Projects under Grant DP190102501, in part by the Deanship of Research Oversight and Coordination (DROC) at KFUPM for funding under the Interdisciplinary Research Center for Communication Systems and Sensing through project No. INCS2111, in part by the U.K. Research and Innovation Future Leaders Fellowships under Grant MR/S017666/1, and in part by the U.S. National Science Foundation under Grant CNS-2128448. The associate editor coordinating the review of this article and approving it for publication was S. Affes. *(Corresponding author: H. D. Tuan.)*

H. D. Tuan and E. Dutkiewicz are with the School of Electrical and Data Engineering, University of Technology Sydney, Sydney, Ultimo, NSW 2007, Australia (e-mail: tuan.hoang@uts.edu.au; eryk.dutkiewicz@uts.edu.au).

A. A. Nasir is with the Department of Electrical Engineering and the Center for Communication Systems and Sensing, King Fahd University of Petroleum & Minerals (KFUPM), Dhahran 31261, Saudi Arabia (e-mail: anasir@kfupm.edu.sa).

H. Q. Ngo is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast BT3 9DT, U.K. (e-mail: hien.ngo@qub.ac.uk).

H. V. Poor is with the Department of Electrical and Computer Engineering, Princeton University, Princeton, NJ 08544 USA (e-mail: poor@princeton.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCOMM.2022.3194046.

Digital Object Identifier 10.1109/TCOMM.2022.3194046

expressions. They are in sharp contrast to those algorithms developed in [12]–[16], which iterate large scale convex problems of third order polynomial complexity at best and thus are not practical for large-size networks. It is revealed that GM-rate maximization leads to very fair rate allocations to all users and achieves a better SR than that achieved by the max-min rate optimization. Moreover, the transmit powers at the APs are also rationally distributed without imposing any constraint.

- CB does not force the multi-user interference (MUI) to zero, which would require global CSI, but rather it manages MUI by exploiting the local CSI (each AP requires only its own channel estimates between itself and the users [17]). It is understood that under CB, cmf-MIMO is an interference-limited network, for which proper Gaussian signaling (PGS) relying on circularly symmetric Gaussian signals for carrying information is not necessarily the best signaling. Recent studies such as [18]–[23] and references therein have shown that improper Gaussian signaling (IGS) [24], which relies on improper/noncircular symmetric Gaussian signals for carrying information, can manage the interference more effectively. We thus use IGS to improve both the GM-rate and the GM-EE. Accordingly, we also propose scalable algorithms to solve GM-rate and GM-EE maximization under IGS. Compared to PGS, IGS not only improves the GM-rate and the GM-EE, but it also promises fairer rate distribution with smaller variance among the users.

The paper is organized as follows. Section II is devoted to PGS for cfm-MIMO, where the basic modeling and communication protocol are recalled, the problem of maximizing GM-rate or GM-EE is formulated and scalable algorithms (Alg. 1 and Alg. 2, respectively) are developed for its solution. Section III introduces IGS for payload data transmission and accordingly develops scalable optimization algorithms (Alg. 3 and Alg. 4) for the IGS-based GM-rate and GM-EE problems, respectively. Section IV provides simulations to confirm the analytical results of Sections II and III. Section V concludes the paper.

Notation: Bold-faced upper-case and lower-case letters, e.g., $\mathbf{X}$ and $\mathbf{x}$, are used for matrices and vectors, respectively, while lower-case letters, e.g., x, are used for scalars. $\mathbf{x}^H$, $\mathbf{x}^T$, and $\mathbf{x}^*$ denote Hermitian transpose, normal transpose, and conjugate of the vector $\mathbf{x}$, respectively. $\|\cdot\|$ denotes a vector's Euclidean norm and $|\cdot|$ stands for the absolute value of a complex/real scalar number. $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^H \mathbf{y}$ for the vectors $\mathbf{x}$ and $\mathbf{y}$ and $\langle \mathbf{X}, \mathbf{Y} \rangle = \text{trace}(\mathbf{X}^H \mathbf{Y})$ for the matrices $\mathbf{X}$ and $\mathbf{Y}$. We also use $\langle \mathbf{X} \rangle = \text{trace}(\mathbf{X})$ for the matrix $\mathbf{X}$. Also $\mathbf{X} \succ 0$ means the Hermitian symmetric matrix $\mathbf{X}$ is positive definite. $\mathbb{C}$ and $\mathbb{R}$, and $\mathbb{R}_+$ denote the sets of all complex numbers, real numbers, and positive numbers, respectively. $\Re\{\mathbf{x}\}$ and $\Im\{\mathbf{x}\}$ denote the real and imaginary parts of a complex vector $\mathbf{x}$, and $\mathbf{I}_n$ is the identity matrix of size $n \times n$. $\mathcal{C}(0,1)$ is the set of circular (proper) Gaussian random variables with zero means and unit variances, while $\mathcal{CN}(0,1)$ is the set of non-circular (improper) Gaussian random variables with zero means and unit variances. Their fundamental difference is that $\mathbb{E}(x^2) = 0$ for $x \in \mathcal{CN}(0,1)$ but $\mathbb{E}(x^2) \neq 0$ for $x \in \mathcal{C}(0,1)$.

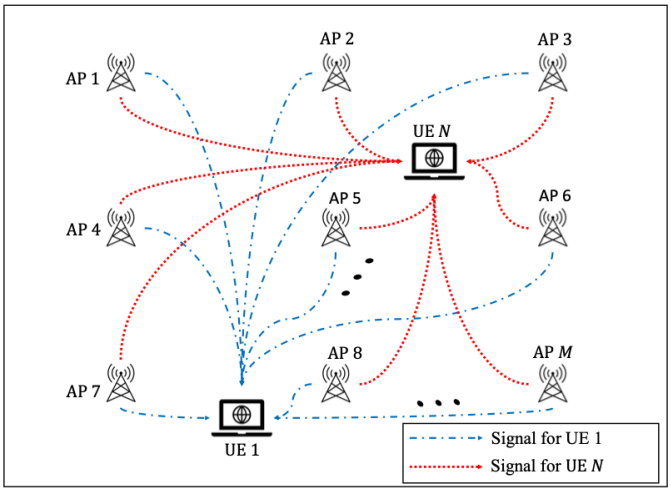

Fig. 1. An illustration of a cell-free massive MIMO system. Later, the simulation results also consider a user-centric cell-free network, where the users are served by a set of selected APs.

The following inequalities, which were proved in [25], are frequently used in the theoretical derivations:

$$\ln\left(1+\frac{x^2}{y}\right) \geq \ln\left(1+\frac{\bar{x}^2}{\bar{y}}\right) - \frac{\bar{x}^2}{\bar{y}} + 2\frac{\bar{x}}{\bar{y}}x - \frac{\bar{x}^2}{\bar{y}(\bar{x}^2+\bar{y})}(x^2+y), \quad (1)$$

for all $x \in \mathbb{R}$, $y > 0$, and $\bar{x} \in \mathbb{R}$, $\bar{y} > 0$, and

$$\begin{aligned} &\ln\left|\mathbf{I}_2 + [\mathbf{V}]^2(\mathbf{Y})^{-1}\right| \\ &\quad \geq \ln\left|\mathbf{I}_2 + [\bar{\mathbf{V}}]^2(\bar{\mathbf{Y}})^{-1}\right| - \langle (\bar{\mathbf{Y}})^{-1}, [\bar{\mathbf{V}}]^2 \rangle \\ &\qquad + 2\Re\{\langle (\bar{\mathbf{Y}})^{-1}\bar{\mathbf{V}}, \mathbf{V} \rangle\} \\ &\qquad - \langle (\bar{\mathbf{Y}})^{-1} - (\bar{\mathbf{Y}} + [\bar{\mathbf{V}}]^2)^{-1}, [\mathbf{V}]^2 + \mathbf{Y} \rangle, \end{aligned} \quad (2)$$

for all matrices $\mathbf{V}$, $\mathbf{Y} \succ 0$, $\bar{\mathbf{V}}$, and $\bar{\mathbf{Y}} \succ 0$ of size 2×2.

II. Basic Modeling and Proper Gaussian Signaling

Consider a cfm-MIMO system, which consists of M APs, indexed by $m \in \mathcal{M} \triangleq \{1, \ldots, M\}$, serving N single-antenna users (UEs) indexed by $n \in \mathcal{N} \triangleq \{1, \ldots, N\}$. Each AP is equipped with N_t antennas. The APs, which are linked to a central processing unit through backhauls, are distributed across the area to make them closer to the UEs. Under time-division duplexing, there are two phases of communication within each coherence interval [1]: uplink training for CSI acquisition, followed by payload data transmission.

A. Uplink Training for CSI Acquisition

Following [1], the channel vector $\mathbf{h}_{mn} \in \mathbb{C}^{N_t \times 1}$ between AP m and UE n is modeled as

$$\mathbf{h}_{mn} = \sqrt{\beta_{mn}}\tilde{\mathbf{h}}_{mn}, \quad (3)$$

where β_{mn} represents the large-scaling fading, which can be assume to be known a priori, and $\tilde{\mathbf{h}}_{mn}$ is an $N_t \times 1$ vector

of small-scale fading coefficients, the elements of which are unknown and independent and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$ random variables. Therefore, $\mathbf{h}_{mn}$ is unknown and needs to be estimated.

We assume channel reciprocity to estimate the channel $\mathbf{h}_{mn}$ between AP m and UE n. The standard block fading model is used, hence the channel $\mathbf{h}_{mn}$ is constant in time-frequency blocks of t channel uses, where t is the length of the coherence interval in samples defined by the product of the coherence time and coherence bandwidth B. Let t_{u} be the uplink training interval, so $t_{\mathrm{u}} < t$. For CSI acquisition, all UEs send their pilot sequences $\sqrt{\tau_{\mathrm{u}}}\boldsymbol{\psi}_n \in \mathbb{C}^{t_{\mathrm{u}}\times 1}$ with $\|\boldsymbol{\psi}_n\|^2 = 1$, $n \in \mathcal{N}$ and $\tau_{\mathrm{u}} \triangleq t_{\mathrm{u}} p_{\mathrm{u}}$, where p_{u} is the power allocated to pilots $\boldsymbol{\psi}_n$.

The signal received at AP $m \in \mathcal{M}$ is

$$\mathbf{Y}_m = \sqrt{\tau_{\mathrm{u}}} \sum_{n'=1}^{N} \mathbf{h}_{mn'} \boldsymbol{\psi}_{n'}^H + \mathbf{W}_m, \tag{4}$$

where $\mathbf{W}_m$ is an $N_t \times t_u$ noise matrix whose elements are i.i.d. $\mathcal{CN}(0,\sigma^2)$ random variables. Applying a matched filter to the received signal $\mathbf{Y}_m$ yields

$$\begin{aligned}\bar{\mathbf{y}}_{mn} &= \mathbf{Y}_m \boldsymbol{\psi}_n \\ &= \sqrt{\tau_{\mathrm{u}}}\mathbf{h}_{mn} + \sqrt{\tau_{\mathrm{u}}} \sum_{n' \in \mathcal{N}\setminus\{n\}} \mathbf{h}_{mn'} \varphi_{n'n} + \mathbf{W}_m \boldsymbol{\psi}_n, \end{aligned} \tag{5}$$

where $\varphi_{nn'} \triangleq \langle \boldsymbol{\psi}_n, \boldsymbol{\psi}_{n'} \rangle$. Thus, a minimum-mean-square-error (MMSE) estimate of $\mathbf{h}_{mn}$ is

$$\begin{aligned}\hat{\mathbf{h}}_{mn} &= \mathbb{E}\left\{\mathbf{h}_{mn}\bar{\mathbf{y}}_{mn}^H\right\} \left(\mathbb{E}\left\{\bar{\mathbf{y}}_{mn}\bar{\mathbf{y}}_{mn}^H\right\}\right)^{-1} \bar{\mathbf{y}}_{mn} \\ &= \alpha_{mn}\bar{\mathbf{y}}_{mn}, \end{aligned} \tag{6}$$

where $\alpha_{mn} \triangleq \dfrac{\sqrt{\tau_{\mathrm{u}}}\beta_{mn}}{\tau_{\mathrm{u}} \sum_{n' \in \mathcal{N}} \beta_{mn'} |\varphi_{n'n}|^2 + \sigma_{\mathrm{u}}^2}$, which is dependent on the large scale fading β_{mn}, training length t_u, uplink power allocation p_u, and $\varphi_{nn'}$. The mean-square of the i-th element of the channel estimate vector, for all $i = \{1, \ldots, N_t\}$, is given by [12]

$$\begin{aligned}\xi_{mn} &\triangleq \mathbb{E}\left\{\left|[\hat{\mathbf{h}}_{mn}]_i\right|^2\right\} \\ &= \frac{\tau_{\mathrm{u}}\beta_{mn}^2}{\tau_{\mathrm{u}} \sum_{n' \in \mathcal{N}} \beta_{mn'} |\varphi_{n'n}|^2 + \sigma_{\mathrm{u}}^2}. \end{aligned} \tag{7}$$

The channel estimation error $\hat{\mathbf{h}}_{mn} - \mathbf{h}_{mn}$ is independent of the channel estimate $\hat{\mathbf{h}}_{mn}$, and its elements are i.i.d. $\mathcal{CN}\left(0, \beta_{mn} - \xi_{mn}\right)$ random variables [12]. Thus, the mean-square error of the elements of channel estimation error vector is given by

$$\begin{aligned}\varepsilon_{mn} &= \mathbb{E}\left\{\left|[\hat{\mathbf{h}}_{mn}]_i - [\mathbf{h}_{mn}]_i\right|^2\right\} \\ &= \beta_{mn} - \xi_{mn} \\ &= \beta_{mn} - \frac{\tau_{\mathrm{u}}\beta_{mn}^2}{\tau_{\mathrm{u}} \sum_{n' \in \mathcal{N}} \beta_{mn'} |\varphi_{n'n}|^2 + \sigma_{\mathrm{u}}^2}. \end{aligned} \tag{8}$$

By using (8) and the definition of α_{mn}, defined below (6), we can express $\alpha_{mn} = \xi_{mn}/(\sqrt{\tau_{\mathrm{u}}}\beta_{mn})$. It should be realized that while $\hat{\mathbf{h}}_{mn}$ is available by (6), the scaling parameter α_{nm}, mean-square estimation error ε_{mn} and mean-square estimate ξ_{mn}, are deterministic that are defined beforehand.

B. Payload Data Transmission

We adopt the following CB at AP m [15],

$$\mathbf{x}_m = \sum_{n' \in \mathcal{N}} p_{n'} \hat{\mathbf{h}}_{mn'}^* s_{n'}, \quad p_{n'} \in \mathbb{R}, \tag{9}$$

where $p_{n'}$ is the square-root of power allocation to UE n', i.e., $p_{n'}^2$ expresses the power allocation across all the APs for the symbol $s_{n'}$ intended for UE n'.[1] The design of this CB involves only N decision variables instead of MN decision variables for the conventional CB [1], [12]. For instance, for $M = 128$ and $N = 32$, the design of CB (9) involves only 32 decision variables while that CB [1], [12] involves already 4096 decision variables. More importantly, the computational complexity of each iteration in all algorithms proposed in this paper is scalable in N while that of each iteration in those algorithms proposed in [12] is third order polynomial in NM at best. Using (9), the power transmitted by the AP m is given by

$$\begin{aligned}\mathbb{E}\left\{\|\mathbf{x}_m\|^2\right\} &= \sum_{n' \in \mathcal{N}} \mathbb{E}\left\{\|\hat{\mathbf{h}}_{mn'}\|^2\right\} p_{n'}^2 \\ &= N_t \sum_{n' \in \mathcal{N}} \xi_{mn'} p_{n'}^2 \\ &\triangleq \pi_m(\mathbf{p}), \end{aligned} \tag{10}$$

which is a convex quadratic function in $\mathbf{p} \triangleq \{p_n, n = 1, \ldots, N\}$.

The signal received at UE n is

$$\begin{aligned}y_n &= \sum_{m \in \mathcal{M}} \mathbf{h}_{mn}^T \mathbf{x}_m + w_n \\ &= p_n \sum_{m \in \mathcal{M}} \mathbf{h}_{mn}^T \hat{\mathbf{h}}_{mn}^* s_n + \sum_{n' \in \mathcal{N}\setminus\{n\}} p_{n'} T_{nn'} s_{n'} + w_n, \end{aligned} \tag{11}$$

where $w_n \sim \mathcal{CN}(0,\sigma^2)$ is the additive noise at UE n, and

$$T_{nn'} \triangleq \sum_{m \in \mathcal{M}} \mathbf{h}_{mn}^T \hat{\mathbf{h}}_{mn'}^*, \quad n' \in \mathcal{N} \setminus \{n\}. \tag{12}$$

Cfm-MIMO offers channel hardening property [1], especially when multiple antennas are considered at the APs [26]. The channel hardening property ensures that with high probability, the instantaneous effective channel gain of the desired signal $\sum_{m \in \mathcal{M}} \mathbf{h}_{mn}^T \hat{\mathbf{h}}_{mn}^*$ is very close to its mean value $S_n \triangleq \mathbb{E}\left\{\sum_{m \in \mathcal{M}} \mathbf{h}_{mn}^T \hat{\mathbf{h}}_{mn}^*\right\}$. Thus, we can rewrite (11) as

$$y_n = p_n S_n s_n + \underbrace{p_n U_n s_n + \sum_{n' \in \mathcal{N}\setminus\{n\}} p_{n'} T_{nn'} s_{n'} + w_n}_{\triangleq \tilde{\nu}_n}, \tag{13}$$

where the first term is considered as the desired signal part with deterministic channel gain S_n, and the sum of remaining terms, denoted by $\tilde{\nu}_n$, is considered as the effective noise, with

$$U_n \triangleq \sum_{m \in \mathcal{M}} \mathbf{h}_{mn}^T \hat{\mathbf{h}}_{mn}^* - \mathbb{E}\left(\sum_{m \in \mathcal{M}} \mathbf{h}_{mn}^T \hat{\mathbf{h}}_{mn}^*\right). \tag{14}$$

[1]The use of $p_{n'}^2$ instead of a simple symbol $p_{n'}$ helps later in handling non-convex problems.

The corresponding effective SINR at UE n is

$$g_n(\mathbf{p}) = \frac{p_n^2|S_n|^2}{\mathbb{E}(|\tilde{\nu}_n|^2)} = \frac{p_n^2|S_n|^2}{p_n^2\mathbb{E}(|U_n|^2) + \sum_{n'\in\mathcal{N}\setminus\{n\}} p_{n'}^2\mathbb{E}(|T_{nn'}|^2) + \sigma^2}, \tag{15}$$

By using a similar argument as that used for deriving [1, eq. (24)], we can obtain the following analytical expressions:

$$S_n = N_t \sum_{m\in\mathcal{M}} \xi_{mn}, \tag{16}$$

$$\mathbb{E}(|U_n|^2) = N_t \sum_{m\in\mathcal{M}} \beta_{mn}\xi_{mn}, \tag{17}$$

and

$$\mathbb{E}(|T_{nn'}|^2) = N_t \sum_{m\in\mathcal{M}} \beta_{mn}\xi_{mn'} + \tau_{\mathbf{u}} N_t^2 |\varphi_{n'n}|^2 \left(\sum_{m\in\mathcal{M}} \beta_{mn}\alpha_{mn'}\right)^2, \tag{18}$$

where β_{mn} and ξ_{mn} are deterministic quantities that are determined beforehand in (3) and (7).

Hence, (15) becomes

$$g_n(\mathbf{p}) = \frac{\lambda_n p_n^2}{\varphi_n(\mathbf{p}) + \sigma^2}, \tag{19}$$

with

$$\lambda_n \triangleq N_t^2 \left(\sum_{m\in\mathcal{M}} \xi_{mn}\right)^2, \tag{20}$$

and

$$\varphi_n(\mathbf{p}) \triangleq N_t \sum_{n'\in\mathcal{N}} p_{n'}^2 \sum_{m\in\mathcal{M}} \beta_{mn}\xi_{mn'} + \tau_{\mathbf{u}} N_t^2 \sum_{n'\in\mathcal{N}\setminus\{n\}} p_{n'}^2 |\varphi_{n'n}|^2 \left(\sum_{m\in\mathcal{M}} \beta_{mn}\alpha_{mn'}\right)^2, \tag{21}$$

which is a positive convex quadratic function.

The achievable rate in nats/sec/Hz for decoding the signal $p_n S_n s_n$ in (13) is given by

$$r_n(\mathbf{p}) = \ln\left(1 + g_n(\mathbf{p})\right), \tag{22}$$

where the UE n needs to only know the mean of the effective channel gain, S_n to detect its desired signal s_n. Let $r(\mathbf{p}) \triangleq (r_1(\mathbf{p}), \ldots, r_N(\mathbf{p}))$. Define the geometric mean (GM):

$$f(r_1, \ldots, r_N) \triangleq \left(\prod_{n=1}^{N} r_n\right)^{1/N}, \tag{23}$$

which is a concave function, and then $f(r(\mathbf{p}))$ as the composition of f and the vector-valued function $r(\mathbf{p})$, i.e. $f(\mathbf{r}(\mathbf{p}))$ is the GM of users rates. Motivated by [27], we consider the following problem for GM-rate maximization subject to the total power constraint:

$$\max_{\mathbf{p}\in\mathbb{R}^N} F(\mathbf{p}) \triangleq f(r(\mathbf{p})) \tag{24a}$$

$$\text{s.t.} \quad \sum_{m=1}^{M} \pi_m(\mathbf{p}) \le MP, \tag{24b}$$

where P is a given power budget. As shown by our simulations, (24) not only facilitates scalable computation but it also achieves multiple targets such as

- Inherently fair user rate allocations with high sum rate. There is no need to consider the problem of either sum rate maximization or GM maximization subject to users rate constraints [28], which only cause artificially tense computation.
- Inherently fair power allocations to the APs due to the fair rate allocations. There is no need to impose the individual power constraints $\pi_m(\mathbf{p}) \le P$, $m \in \mathcal{M}$, for the APs' physical transmission, which only cause artificially tense computation.

We develop an iterative process for computing (24). Let $\mathbf{p}^{(\kappa)} \triangleq \{p_n^{(\kappa)}, n = 1, \ldots, N\}$ be the feasible point for (24) that is found from the $(\kappa-1)$-st iteration. We note that the linearized function of $f(r_1, \ldots, r_N)$ at $r(\mathbf{p}^{(\kappa)})$ is

$$\frac{F(\mathbf{p}^{(\kappa)}))}{N} \sum_{n=1}^{N} \frac{r_n}{r_n(\mathbf{p}^{(\kappa)})}. \tag{25}$$

Since $F(\mathbf{p}^{(\kappa)}) > 0$, we generate the next feasible point $\mathbf{p}^{(\kappa+1)}$ by considering the following problem of steep descent optimization for the concave function for the concave function $f(r_1, \ldots, r_N)$:

$$\max_{\mathbf{p}} \frac{F(\mathbf{p}^{(\kappa)})}{N} \sum_{n=1}^{N} \frac{r_n(\mathbf{p})}{r_n(\mathbf{p}^{(\kappa)})} \quad \text{s.t.} \quad (24b), \tag{26}$$

which is equivalent to the following problem:

$$\max_{\mathbf{p}} f^{(\kappa)}(\mathbf{p}) \triangleq \sum_{n=1}^{N} \gamma_n^{(\kappa)} r_n(\mathbf{p}) \quad \text{s.t.} \quad (24b), \tag{27}$$

for

$$\gamma_n^{(\kappa)} \triangleq \frac{\max_{n'\in\mathcal{N}} r_{n'}(\mathbf{p}^{(\kappa)})}{r_n(\mathbf{p}^{(\kappa)})}, \quad n = 1, \ldots, N. \tag{28}$$

Though the objective function in (27) looks like conventional weighted sum-rate optimization, however, quite different from the conventional weighted sum-rate optimization, which is based on constant user-specific weights, our proposed implementation optimizes those weights analytically and updates them at each iteration, with an implicit objective of achieving fairer rate distribution.

By applying the inequality (1) for $(x, y) = (\sqrt{\lambda_n}p_n, \varphi_n(\mathbf{p}) + \sigma^2)$ and $(\bar{x}, \bar{y}) = (\sqrt{\lambda_n}p_n^{(\kappa)}, \varphi_n(\mathbf{p}^{(\kappa)}) + \sigma^2)$ we obtain

$$r_n(\mathbf{p}) \ge a_n^{(\kappa)} + 2b_n^{(\kappa)}p_n - c_n^{(\kappa)}(\lambda_n p_n^2 + \varphi_n(\mathbf{p})) \triangleq r_n^{(\kappa)}(\mathbf{p}), \tag{29}$$

for $a_n^{(\kappa)} \triangleq r_n(\mathbf{p}^{(\kappa)}) - g_n(\mathbf{p}^{(\kappa)}) - c_n^{(\kappa)}\sigma^2$, $b_n^{(\kappa)} \triangleq \frac{\lambda_n p_n^{(\kappa)}}{\varphi_n(\mathbf{p}^{(\kappa)}) + \sigma^2}$, and $c_n^{(\kappa)} \triangleq \frac{g_n(\mathbf{p}^{(\kappa)})}{\lambda_n \left(p_n^{(\kappa)}\right)^2 + \varphi_n(\mathbf{p}^{(\kappa)}) + \sigma^2}$. The function $r_n^{(\kappa)}(\mathbf{p})$ is concave, which matches with $r_n(\mathbf{p})$ at $\mathbf{p}^{(\kappa)}$,

i.e., $r_n(\mathbf{p}^{(\kappa)}) = r_n^{(\kappa)}(\mathbf{p}^{(\kappa)})$. We solve the following problem at the κ-th iteration to generate $\mathbf{p}^{(\kappa+1)}$:

$$\max_{\mathbf{p}} \bar{f}^{(\kappa)}(\mathbf{p}) \triangleq \sum_{n\in\mathcal{N}} \gamma_n^{(\kappa)} r_n^{(\kappa)}(\mathbf{p}) \quad \text{s.t.} \quad (24b), \tag{30}$$

for

$$\begin{aligned}
\bar{f}^{(\kappa)}(\mathbf{p}) &\triangleq \sum_{n\in\mathcal{N}} \gamma_n^{(\kappa)} r_n^{(\kappa)}(\mathbf{p}) \\
&= \sum_{n\in\mathcal{N}} \gamma_n^{(\kappa)} \left[a_n^{(\kappa)} + 2b_n^{(\kappa)} p_n - c_n^{(\kappa)}(\lambda_n p_n^2 + \varphi_n(\mathbf{p}))\right] \\
&= \sum_{n\in\mathcal{N}} \gamma_n^{(\kappa)} \Bigg[a_n^{(\kappa)} + 2b_n^{(\kappa)} p_n - c_n^{(\kappa)} \Bigg(\lambda_n p_n^2 \\
&\quad + \sum_{n'\in\mathcal{N}} p_{n'}^2 z_{nn'} + \tau_{\mathrm{u}} \sum_{n'\in\mathcal{N}\setminus\{n\}} p_{n'}^2 |\varphi_{n'n}|^2 \bar{z}_{nn'}\Bigg)\Bigg] \\
&= \sum_{n\in\mathcal{N}} \gamma_n^{(\kappa)} \left(a_n^{(\kappa)} + 2b_n^{(\kappa)} p_n - \frac{d_n^{(\kappa)} p_n^2}{\gamma_n^{(\kappa)}}\right),
\end{aligned} \tag{31}$$

where

$$\begin{aligned}
z_{nn'} &\triangleq N_t \sum_{m\in\mathcal{M}} \beta_{mn}\xi_{mn'}, \\
\bar{z}_{nn'} &\triangleq N_t^2 \left(\sum_{m\in\mathcal{M}} \beta_{mn}\alpha_{mn'}\right)^2, \\
d_n^{(\kappa)} &\triangleq \gamma_n^{(\kappa)} c_n^{(\kappa)} \lambda_n + \sum_{n'\in\mathcal{N}} \gamma_{n'}^{(\kappa)} c_{n'}^{(\kappa)} z_{n'n} \\
&\quad + \tau_{\mathrm{u}} \sum_{n'\in\mathcal{N}\setminus\{n\}} \gamma_{n'}^{(\kappa)} c_{n'}^{(\kappa)} |\varphi_{nn'}|^2 \bar{z}_{n'n}.
\end{aligned}$$

The problem (30) admits the closed-form solution

$$p_n^{(\kappa+1)} = \begin{cases} \dfrac{\gamma_n^{(\kappa)} b_n^{(\kappa)}}{d_n^{(\kappa)}} & \text{if } \upsilon^{(\kappa)} \le MP \\ \dfrac{\gamma_n^{(\kappa)} b_n^{(\kappa)}}{d_n^{(\kappa)} + \mu e_n^{(\kappa)}} & \text{otherwise,} \end{cases} \tag{32}$$

where $\upsilon^{(\kappa)} \triangleq N_t \sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \xi_{mn} \left(\gamma_n^{(\kappa)} b_n^{(\kappa)} / d_n^{(\kappa)}\right)^2$, $e_n^{(\kappa)} \triangleq N_t \sum_{m\in\mathcal{M}} \xi_{mn}$ and $\mu > 0$ can be found by bisection search such that

$$N_t \sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \xi_{mn} \left(\frac{\gamma_n^{(\kappa)} b_n^{(\kappa)}}{d_n^{(\kappa)} + \mu e_n^{(\kappa)}}\right)^2 = MP.$$

Algorithm 1 provides the pseudo-code for the proposed computational procedure. It should be noted that one still needs to form a line search for finding the step size $\theta^{(\kappa+1)} \in [0,1]$ such that $F\big(\mathbf{p}^{(\kappa)} + \theta^{(\kappa+1)}(\mathbf{p}^{(\kappa+1)} - \mathbf{p}^{(\kappa)})\big) > F(\mathbf{p}^{(\kappa)})$ [29] to update $\mathbf{p}^{(\kappa+1)} \to \mathbf{p}^{(\kappa)} + \theta^{(\kappa+1)}(\mathbf{p}^{(\kappa+1)} - \mathbf{p}^{(\kappa)})$. Fortunately, we always observe in our simulations that

$$F(\mathbf{p}^{(\kappa+1)})) > F(\mathbf{p}^{(\kappa)}), \tag{33}$$

i.e. the full step size of length one $\theta^{(\kappa+1)} = 1$ is achieved, bypassing the line search. This can be explained as follows.

Algorithm 1 Scalable Algorithm for PGS-Based GM-Rate Optimization Problem (24)

1: **Initialization**: Initialize a feasible point $p^{(0)}$. Set $\kappa = 0$.
2: **Repeat until convergence of** $p^{(\kappa)}$: Update $\gamma_n^{(\kappa)}$ by (28). Generate $\mathbf{p}^{(\kappa+1)}$ by (32). Reset $\kappa \leftarrow \kappa + 1$.
3: **Output** $p^{(\kappa)}$ as the optimal solution of (24).

By noting that

$$F(\mathbf{p}) = \min_{\boldsymbol{\gamma}\in\Gamma} \left[\sum_{n=1}^{N} \gamma_n r_n(\mathbf{p})\right],$$

for $\Gamma \triangleq \{\boldsymbol{\gamma} \triangleq (\gamma_1, \ldots, \gamma_N) : \prod_{n=1}^{N} \gamma_n = 1/N^N, \gamma_n > 0, n = 1, \ldots, N\}$, we can express the problem (24) as the following maximin problem

$$\max_{\mathbf{p}\in\mathcal{P}} \min_{\boldsymbol{\gamma}\in\Gamma} \left[\sum_{n=1}^{N} \gamma_n r_n(\mathbf{p})\right], \tag{34}$$

where $\mathcal{P}$ is the set of feasible points for the power constraint (24b). Now, for $\mathbf{p}^{(\kappa)} \in \mathcal{P}$ and $\gamma^{(\kappa)}$ defined from (28),

$$\gamma^{(\kappa)} = \arg\min_{\boldsymbol{\gamma}\in\Gamma} \left[\sum_{n=1}^{N} \gamma_n r_n(\mathbf{p}^{(\kappa)})\right],$$

so Algorithm 1 provides a procedure of alternating optimization between $\mathbf{p}$ and $\boldsymbol{\gamma}$. A general-purpose projective gradient algorithm for GM-rate maximization [30] is much higher computation consuming due to the subtle fact that the GM-rate is a very high-order function of the decision variables so it is not easy to determine the Lipschitz constant of its gradient that plays a crucial role for the algorithm convergence and then it is challenging to implement Armijo rule of a line search to update the step size. Note that the computational complexity of (32) is linear that in contrast with that in [12], [15] of third order polynomial complexity for iterating their convex problems.

Before closing this section, it should be noted that Algorithm 1 can be easily adjusted to address a similar problem of normalized conjugate beamforming [8].

C. Energy Efficient Payload Data Transmission

Recalling that the function f is defined from (23), we now provide a new model of EE maximization, which also promotes the quality-of-service (QoS) as

$$\max_{\mathbf{p}\in\mathbb{R}^N} \frac{f(r(\mathbf{p}))}{\pi_{te}(\mathbf{p})} \quad \text{s.t.} \quad (24b), \tag{35}$$

where

$$\begin{aligned}
\pi_{te}(\mathbf{p}) = \alpha \sum_{m\in\mathcal{M}} \pi_m(\mathbf{p}) + N_t M P_{sc} \\
+ M\big(P_0 + P_{bt} B N f(r(\mathbf{p}))/\ln 2\big),
\end{aligned} \tag{36}$$

which is the total power consumption [12], [31]. Here, α is the reciprocal of drain efficiency of the the power amplifier at APs, P_{sc} is the internal power required for running the circuit components at each antenna of the AP, P_0 is the power

consumption of each backhaul, and P_{bt} is the traffic-dependent power (in Watt per bits/s). The first and second terms in (36) recap the transmission power consumption, while the third term recaps the power consumption of the backhauls [32]. Note that unlike [33], [34], we set $N(\prod_{n\in\mathcal{N}} r_n(\mathbf{p}))^{1/N}$ instead of $\sum_{n\in\mathcal{N}} r_n(\mathbf{p})$. By maximizing the objective in (35) one achieves both the QoS in terms of users' minimum rates because the GM-rate in its numerator must be maximized and thus the QoS of all users are promoted, while the power consumption in its denominator must be minimized and thus the EE is promoted as well.

Observe that

$$\frac{f(r(\mathbf{p}))}{\pi_{te}(\mathbf{p})} = \left(\frac{\pi(\mathbf{p})}{f(r(\mathbf{p}))} + \frac{BNMP_{bt}}{\ln 2}\right)^{-1}, \tag{37}$$

for

$$\pi(\mathbf{p}) \triangleq \alpha \sum_{m\in\mathcal{M}} \pi_m(\mathbf{p}) + \nu, \tag{38}$$

where $\nu \triangleq N_t MP_{sc} + MP_0$ is a positive constant. Therefore, the problem (35) is equivalent to the following problem:

$$\max_{\mathbf{p}\in\mathbb{R}^N} f_E(r(\mathbf{p}), \pi(\mathbf{p})) \triangleq \frac{f(r(\mathbf{p}))}{\pi(\mathbf{p})} \quad \text{s.t.} \quad (24b). \tag{39}$$

Let $\mathbf{p}^{(\kappa)} \triangleq (p_1^{(\kappa)}, p_2^{(\kappa)}, \ldots, p_N^{(\kappa)})^T$ be the feasible point for (39) that is found from the $(\kappa-1)$-st iteration. The linearized function of $f_E(r(\mathbf{p}), \pi(\mathbf{p}))$ at $(r(\mathbf{p}^{(\kappa)}), \pi(\mathbf{p}^{(\kappa)}))$ is

$$f_E(r(\mathbf{p}^{(\kappa)}), \pi(\mathbf{p}^{(\kappa)})) + \frac{1}{N} f_E(r(\mathbf{p}^{(\kappa)}), \pi(\mathbf{p}^{(\kappa)})) \times \left(\sum_{n=1}^{N} \frac{r_n(\mathbf{p})}{r_n(\mathbf{p}^{(\kappa)})} - N\frac{\pi(\mathbf{p})}{\pi(\mathbf{p}^{(\kappa)})}\right). \tag{40}$$

The steepest ascents aim to solve the problem

$$\max_{\mathbf{p}} \left[\sum_{n=1}^{N} \gamma_n^{(\kappa)} r_n(\mathbf{p}) - \eta^{(\kappa)} \pi(\mathbf{p})\right] \quad \text{s.t.} \quad (24b), \tag{41}$$

for

$$\eta^{(\kappa)} \triangleq N \frac{\max_{n'\in\mathcal{N}} r_{n'}(\mathbf{p}^{(\kappa)})}{\pi(\mathbf{p}^{(\kappa)})}, \tag{42}$$

and

$$\gamma_n^{(\kappa)} \triangleq \frac{\max_{n'\in\mathcal{N}} r_{n'}(\mathbf{p}^{(\kappa)})}{r_n(\mathbf{p}^{(\kappa)})}. \tag{43}$$

Recalling that the function $r_n^{(\kappa)}(\mathbf{p})$ from (29), at the κth iteration we solve the following convex problem to generate the next iterative point $\mathbf{p}^{(\kappa+1)}$:

$$\max_{\mathbf{p}} \left[\sum_{n=1}^{N} \gamma_n^{(\kappa)} r_n^{(\kappa)}(\mathbf{p}) - \eta^{(\kappa)} \pi(\mathbf{p})\right] \quad \text{s.t.} \quad (24b), \tag{44}$$

which admits the closed-form solution

$$p_n^{(\kappa+1)} = \begin{cases} \dfrac{\gamma_n^{(\kappa)} b_n^{(\kappa)}}{d_n^{(\kappa)} + \alpha\eta^{(\kappa)} e_n^{(\kappa)}} & \text{if } \bar{v}^{(\kappa)} \le MP \\ \dfrac{\gamma_n^{(\kappa)} b_n^{(\kappa)}}{d_n^{(\kappa)} + (\alpha\eta^{(\kappa)} + \bar{\mu}) e_n^{(\kappa)}} & \text{otherwise,} \end{cases} \tag{45}$$

Algorithm 2 Scalable Algorithm for PGS-Based GM-EE Problem (35)

1: **Initialization**: Initialize a feasible point $p^{(0)}$. Set $\kappa = 0$.
2: **Repeat until convergence of** $p^{(\kappa)}$: Update $\eta^{(\kappa)}$ and $\gamma_n^{(\kappa)}$ by (42) and (43). Generate $\mathbf{p}^{(\kappa+1)}$ by (45). Reset $\kappa \leftarrow \kappa+1$.
3: **Output** $p^{(\kappa)}$ as the optimal solution of (35).

where $\bar{v}^{(\kappa)} \triangleq N_t \sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \xi_{mn} \left(\frac{\gamma_n^{(\kappa)} b_n^{(\kappa)}}{d_n^{(\kappa)} + \alpha\eta^{(\kappa)} e_n^{(\kappa)}}\right)^2$, $e_n^{(\kappa)} \triangleq N_t \sum_{m\in\mathcal{M}} \xi_{mn}$ and $\bar{\mu} > 0$ can be found by bisection search such that

$$N_t \sum_{m\in\mathcal{M}} \sum_{n\in\mathcal{N}} \xi_{mn} \left(\frac{\gamma_n^{(\kappa)} b_n^{(\kappa)}}{d_n^{(\kappa)} + (\alpha\eta^{(\kappa)} + \bar{\mu}) e_n^{(\kappa)}}\right)^2 = MP.$$

Algorithm 2 provides the pseudo-code for the proposed steep ascent. The computational complexity of (45) is linear.

D. Instantaneous Performance

This subsection will address the GM-rate and energy-efficiency performances when the optimization formulations are based on the online effective channel estimate available at UE n, which are referred as *instantaneous performances*.

Instead of (11), the received signal at the UE n can be expressed as:

$$\begin{aligned} y_n &= \sum_{m\in\mathcal{M}} \hat{\mathbf{h}}_{mn}^T \mathbf{x}_m + \sum_{m\in\mathcal{M}} \left(\mathbf{h}_{mn}^T - \hat{\mathbf{h}}_{mn}^T\right) \mathbf{x}_m + w_n \qquad (46) \\ &= \sum_{m\in\mathcal{M}} \hat{\mathbf{h}}_{mn}^T \left(\sum_{n'\in\mathcal{N}} p_{n'} \hat{\mathbf{h}}_{mn'}^* s_{n'}\right) + \sum_{m\in\mathcal{M}} \left(\mathbf{h}_{mn}^T - \hat{\mathbf{h}}_{mn}^T\right) \mathbf{x}_m + w_n \\ &= p_n \tilde{\lambda}_{nn} s_n + \sum_{n'\in\mathcal{N}\setminus\{n\}} p_{n'} \tilde{\lambda}_{nn'} s_{n'} + \tilde{\nu}_n + w_n, \qquad (47) \end{aligned}$$

for

$$\tilde{\lambda}_{nn'} \triangleq \begin{cases} \sum_{m\in\mathcal{M}} \|\hat{\mathbf{h}}_{mn}\|^2 & \text{for } n' = n \\ \sum_{m\in\mathcal{M}} \hat{\mathbf{h}}_{mn}^T \hat{\mathbf{h}}_{mn'}^* & \text{otherwise,} \end{cases} \tag{48a}$$

$$\tilde{\nu}_n \triangleq \sum_{m\in\mathcal{M}} \left(\mathbf{h}_{mn}^T - \hat{\mathbf{h}}_{mn}^T\right) \mathbf{x}_m. \tag{48b}$$

Using (8) and (9)

$$\begin{aligned} \mathbb{E}(|\tilde{\nu}_n|^2) &= \sum_{m\in\mathcal{M}} \varepsilon_{mn} \sum_{n'\in\mathcal{N}} p_{n'}^2 \|\hat{\mathbf{h}}_{mn'}^*\|^2 \\ &\triangleq e_n(\mathbf{p}), \end{aligned} \tag{49}$$

so the instantaneous SINR at UE n is

$$\tilde{g}_n(\mathbf{p}) = \frac{\tilde{\lambda}_{nn}^2 p_n^2}{\tilde{\varphi}_n(\mathbf{p}) + \sigma^2}, \tag{50}$$

with

$$\tilde{\varphi}_n(\mathbf{p}) \triangleq \sum_{n'\in\mathcal{N}\setminus\{n\}} p_{n'}^2 |\tilde{\lambda}_{nn'}|^2 + e_n(\mathbf{p}). \tag{51}$$

The achievable rate of decoding $p_n\tilde{\lambda}_{nn}s_n$ in (47) is

$$\chi_n(\mathbf{p}) \triangleq \ln(1+\tilde{g}_n(\mathbf{p})), \tag{52}$$

while the transmit power at AP m is defined by

$$\tilde{\pi}_m(\mathbf{p}) \triangleq \sum_{n\in\mathcal{N}} \|\hat{\mathbf{h}}_{mn}\|^2 p_n^2. \tag{53}$$

Algorithm 1 and Algorithm 2 can be transparently adjusted for computing the corresponding problems of GM-rate and GM-EE optimization, with modification of the definitions:, i.e., $a_n^{(\kappa)} \triangleq \chi_n(\mathbf{p}^{(\kappa)}) - \tilde{g}_n(\mathbf{p}^{(\kappa)}) - c_n^{(\kappa)}\sigma^2$, $b_n^{(\kappa)} \triangleq \frac{\tilde{\lambda}_{nn}^2 p_n^{(\kappa)}}{\tilde{\varphi}_n(\mathbf{p}^{(\kappa)})+\sigma^2}$, $c_n^{(\kappa)} \triangleq \frac{\tilde{g}_n(\mathbf{p}^{(\kappa)})}{\tilde{\lambda}_{nn}^2(p_n^{(\kappa)})^2+\tilde{\varphi}_n(\mathbf{p}^{(\kappa)})+\sigma^2}$, $d_n^{(\kappa)} \triangleq \sum_{n'\in\mathcal{N}} \gamma_{n'}^{(\kappa)} c_{n'}^{(\kappa)} |\tilde{\lambda}_{n'n}|^2 + \sum_{n'\in\mathcal{N}} \gamma_{n'}^{(\kappa)} c_{n'}^{(\kappa)} z_{n'n}$, $e_n^{(\kappa)} \triangleq \sum_{m\in\mathcal{M}} \|\hat{\mathbf{h}}_{mn}\|^2$, and $z_{nn'} \triangleq \sum_{m\in\mathcal{M}} \varepsilon_{mn} \|\hat{\mathbf{h}}_{mn'}\|^2$.

While the rate defined by (22) is achievable when the mean value $S_n \triangleq \mathbb{E}\left\{\sum_{m\in\mathcal{M}} \mathbf{h}_{mn}^T \hat{\mathbf{h}}_{mn}^*\right\}$ is available at the UE n, the rate defined by (52) is achievable when the instantaneous value of the effective channel gain, i.e., $\tilde{\lambda}_{nn} \triangleq \sum_{m\in\mathcal{M}} \|\hat{\mathbf{h}}_{mn}\|^2$ is available at the UE n, the information of which can be provided at the UE n by employing downlink channel estimation [35].[2]

III. Improper Gaussian Signaling to Improve the Downlink Payload

A. Rates' GM Efficient Payload Data Transmission

Until now, the transmit signal $\mathbf{x}_m$ by (9) is proper Gaussian ($\mathbb{E}(\mathbf{x}_m(\mathbf{x}_m)^T) = 0$) because each $\mathbb{E}(s_{n'}^2) = 0$ as $s_{n'} \in \mathcal{CN}(0,1)$. It has been recently shown that (see e.g. [18]–[22] and references therein) the release of Gaussian properness improves the signal degree of freedom, which helps to improve the rate performances of interfering networks. This section aims to use improper Gaussian signaling (IGS) to improve the users's GM rate as well as the energy efficiency of the considered cfm-MIMO system, where the signal transmitted by the AP m is given by

$$\mathbf{x}_m = \sum_{n'\in\mathcal{N}} \hat{\mathbf{h}}_{mn'}^* (v_{1,n'}s_{n'} + v_{2,n'}s_{n'}^*), \quad v_{n'} \triangleq (v_{1,n'}, v_{2,n'}) \in \mathbb{C}\times\mathbb{C}, \tag{54}$$

i.e. the symbol s_n is widely linearly beamformed by a pair of two beamformers $\hat{\mathbf{h}}_{mn'}^* v_{1,n'} \in \mathbb{C}^{N_t\times 1}$ and $\hat{\mathbf{h}}_{mn'}^* v_{2,n'} \in \mathbb{C}^{N_t\times 1}$. It is immediate to check that $\mathbb{E}(\mathbf{x}_m(\mathbf{x}_m)^T) \neq 0$ because $\mathbb{E}((v_{1,n'}s_{n'} + v_{2,n'}s_{n'}^*)^2) = 2v_{1,n'}v_{2,n'} \neq 0$, i.e. $\mathbf{x}_m$ is an improper Gussian signal.

Based on the channel statistical data, we can observe from (11) that pseudo-covariance $\mathbb{E}((\sum_{n'\in\mathcal{N}\setminus\{n\}} p_{n'}T_{nn's_{n'}})(\sum_{n'\in\mathcal{N}\setminus\{n\}} p_{n'}T_{nn's_{n'}})^T$ of the pre-last term, $\sum_{n'\in\mathcal{N}\setminus\{n\}} p_{n'}T_{nn's_{n'}}$, is still zero because $\mathbf{h}_{mn}$ and $\hat{\mathbf{h}}_{mn}$ are circular Gaussian. Thus, being sum of many RVs, this interference term can be approximated by proper Gaussian RV, so employing IGS (54) cannot be advantageous over PGS (9) because the latter is optimal for proper Gaussian interference networks [24]. Therefore, in this section, we shall consider only the case of instantaneous performance, i.e., the effective channel $\tilde{\lambda}_{nn} \triangleq \sum_{m\in\mathcal{M}} \|\hat{\mathbf{h}}_{mn}\|^2$ is available at UE n. Using (54), the received signal at the UE n under IGS, instead of (47), becomes

$$\begin{aligned} y_n = {} & \tilde{\lambda}_{nn}(v_{1,n}s_n + v_{2,n}s_n^*) + \sum_{n'\in\mathcal{N}\setminus\{n\}} \tilde{\lambda}_{nn'}(v_{1,n'}s_{n'} \\ & + v_{2,n'}s_{n'}^*) + \sum_{m\in\mathcal{M}} \left(\mathbf{h}_{mn}^T - \hat{\mathbf{h}}_{mn}^T\right) \\ & \times \sum_{n'\in\mathcal{N}} \hat{\mathbf{h}}_{mn'}^* (v_{1,n'}s_{n'} + v_{2,n'}s_{n'}^*) + w_n, \end{aligned} \tag{55}$$

with $\tilde{\lambda}_{nn'}$ defined from (48). In what follows, we use the following notations:

$$\bar{\mathbf{y}}_n \triangleq \begin{bmatrix} \Re\{y_n\} \\ \Im\{y_n\} \end{bmatrix}, \quad \bar{\mathbf{s}}_n \triangleq \begin{bmatrix} \Re\{s_n\} \\ \Im\{s_n\} \end{bmatrix}, \quad \bar{\mathbf{w}}_n \triangleq \begin{bmatrix} \Re\{w_n\} \\ \Im\{w_n\} \end{bmatrix},$$

$$\mathbf{V}_n \triangleq \begin{bmatrix} \Re\{v_{1,n}+v_{2,n}\} & -\Im\{v_{1,n}-v_{2,n}\} \\ \Im\{v_{1,n}+v_{2,n}\} & \Re\{v_{1,n}-v_{2,n}\} \end{bmatrix},$$

$$\mathbf{\Lambda}_{nn'} \triangleq \begin{bmatrix} \Re\{\tilde{\lambda}_{nn'}\} & -\Im\{\tilde{\lambda}_{nn'}\} \\ \Im\{\tilde{\lambda}_{nn'}\} & \Re\{\tilde{\lambda}_{nn'}\} \end{bmatrix}.$$

$$\mathbf{G}_{m,n} = \begin{bmatrix} \mathbf{e}_{mn,1}^T & -\mathbf{e}_{mn,2}^T \\ \mathbf{e}_{mn,2}^T & \mathbf{e}_{mn,1}^T \end{bmatrix} \triangleq \begin{bmatrix} \Re\left\{\mathbf{h}_{mn}^T - \hat{\mathbf{h}}_{mn}^T\right\} & -\Im\left\{\mathbf{h}_{mn}^T - \hat{\mathbf{h}}_{mn}^T\right\} \\ \Im\left\{\mathbf{h}_{mn}^T - \hat{\mathbf{h}}_{mn}^T\right\} & \Re\left\{\mathbf{h}_{mn}^T - \hat{\mathbf{h}}_{mn}^T\right\} \end{bmatrix},$$

and

$$\hat{\mathbf{H}}_{m,n} \triangleq \begin{bmatrix} \Re\left\{\hat{\mathbf{h}}_{mn}\right\} & \Im\left\{\hat{\mathbf{h}}_{mn}\right\} \\ -\Im\left\{\hat{\mathbf{h}}_{mn}\right\} & \Re\left\{\hat{\mathbf{h}}_{mn}\right\} \end{bmatrix},$$

and it can be readily shown that $\mathbb{E}\{[\bar{\mathbf{s}}_n]^2\} = \frac{1}{2}\mathbf{I}_2$ and $\mathbb{E}\{[\bar{\mathbf{w}}_n]^2\} = \frac{1}{2}\sigma^2\mathbf{I}_2$. Also, each element of a channel estimation error vector, $\mathbf{h}_{mn} - \hat{\mathbf{h}}_{mn}$, is a proper Gaussian with zero mean and variance ε_{mn}, so $\mathbb{E}(\|\mathbf{e}_{mn,\ell}\|^2) = N_t\varepsilon_{mn}/2$, $\ell = 1,2$, and

$$\mathbb{E}\left(\mathbf{e}_{mn,1}^T\mathbf{e}_{mn,2}\right) = 0.$$

As such for a deterministic and real matrix of size $2N_t \times 2$

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_{11} & \mathbf{a}_{12} \\ \mathbf{a}_{21} & \mathbf{a}_{22} \end{bmatrix},$$

where $\mathbf{a}_{\ell,\bar{\ell}} \in \mathbb{R}^{N_t\times 1}$, $\ell,\bar{\ell} = 1,2$, we can easily calculate

$$\begin{aligned} & \mathbb{E}\left([\mathbf{G}_{m,n}\mathbf{A}]^2\right) \\ & = \mathbb{E}\left(\begin{bmatrix} \mathbf{e}_{mn,1}^T\mathbf{a}_{11} - \mathbf{e}_{mn,2}^T\mathbf{a}_{21} & \mathbf{e}_{mn,1}^T\mathbf{a}_{12} - \mathbf{e}_{mn,2}^T\mathbf{a}_{22} \\ \mathbf{e}_{mn,2}^T\mathbf{a}_{11} + \mathbf{e}_{mn,1}^T\mathbf{a}_{21} & \mathbf{e}_{mn,2}^T\mathbf{a}_{12} + \mathbf{e}_{mn,1}^T\mathbf{a}_{22} \end{bmatrix}^2\right) \\ & = \frac{N_t\varepsilon_{mn}}{2}\langle[\mathbf{A}]^2\rangle\mathbf{I}_2. \end{aligned} \tag{56}$$

The equivalent real composite form of (55) is

$$\begin{aligned} \bar{\mathbf{y}}_n = {} & \mathbf{\Lambda}_{nn}\mathbf{V}_n\bar{\mathbf{s}}_n + \sum_{n'\in\mathcal{N}\setminus\{n\}} \mathbf{\Lambda}_{nn'}\mathbf{V}_{n'}\bar{\mathbf{s}}_{n'} \\ & + \sum_{m\in\mathcal{M}} \mathbf{G}_{m,n} \sum_{n'\in\mathcal{N}} \hat{\mathbf{H}}_{m,n'}\mathbf{V}_{n'}\bar{\mathbf{s}}_{n'} + \bar{\mathbf{w}}_n. \end{aligned} \tag{57}$$

[2]We assume perfect estimation of the effective channel gain in the downlink and under this assumption, our results can serve as an upper bound on the achievable rate and EE. Integration of downlink channel estimation using beamforming [35] can be the subject of future research.

Under the variable change

$$\mathbf{P}_n \triangleq \begin{bmatrix} p_n^{11} & p_n^{12} \\ p_n^{21} & p_n^{22} \end{bmatrix} = \begin{bmatrix} \Re\{v_{1,n}+v_{2,n}\} & -\Im\{v_{1,n}-v_{2,n}\} \\ \Im\{v_{1,n}+v_{2,n}\} & \Re\{v_{1,n}-v_{2,n}\} \end{bmatrix}, \quad (58)$$

which is invertible, i.e.,

$$\boldsymbol{\mathcal{V}}_n \triangleq \begin{bmatrix} \Re\{v_{1,n}\} & \Im\{v_{1,n}\} \\ \Re\{v_{2,n}\} & \Im\{v_{2,n}\} \end{bmatrix} = \frac{1}{2}\begin{bmatrix} p_n^{11}+p_n^{22} & p_n^{21}-p_n^{12} \\ p_n^{11}-p_n^{22} & p_n^{21}+p_n^{12}, \end{bmatrix}, \quad (59)$$

we have

$$\bar{\mathbf{y}}_n = \boldsymbol{\Lambda}_{nn}\mathbf{P}_n\bar{\mathbf{s}}_n + \sum_{n'\in\mathcal{N}\setminus\{n\}} \boldsymbol{\Lambda}_{nn'}\mathbf{P}_{n'}\bar{\mathbf{s}}_{n'} + \sum_{m\in\mathcal{M}} \mathbf{G}_{m,n} \sum_{n'\in\mathcal{N}} \hat{\mathbf{H}}_{m,n'}\mathbf{P}_{n'}\bar{\mathbf{s}}_{n'} + \bar{\mathbf{w}}_n. \quad (60)$$

Furthermore, we have

$$\|\boldsymbol{\mathcal{V}}_n\|^2 = \langle[\boldsymbol{\mathcal{V}}_n]^2\rangle = \frac{1}{2}\|\mathbf{P}_n\|^2 = \frac{1}{2}\langle[\mathbf{P}_n]^2\rangle,$$

and by using (56)

$$\begin{aligned} &\mathbb{E}\left\{\left[\sum_{m\in\mathcal{M}} \mathbf{G}_{m,n}\sum_{n'\in\mathcal{N}}\hat{\mathbf{H}}_{m,n'}\mathbf{P}_{n'}\bar{\mathbf{s}}_{n'}\right]^2\right\} \\ &= \sum_{m\in\mathcal{M}}\sum_{n'\in\mathcal{N}} \mathbb{E}\left\{\left[\mathbf{G}_{m,n}\hat{\mathbf{H}}_{m,n'}\mathbf{P}_{n'}\bar{\mathbf{s}}_n\right]^2\right\} \\ &= \frac{1}{2}\sum_{m\in\mathcal{M}}\sum_{n'\in\mathcal{N}} \mathbb{E}\left\{\left[\mathbf{G}_{m,n}\hat{\mathbf{H}}_{m,n'}\mathbf{P}_{n'}\right]^2\right\} \\ &= \frac{N_t}{4}\left(\sum_{m\in\mathcal{M}}\varepsilon_{mn}\sum_{n'\in\mathcal{N}}\left\langle\left[\hat{\mathbf{H}}_{m,n'}\mathbf{P}_{n'}\right]^2\right\rangle\right)\mathbf{I}_2 \\ &\triangleq \boldsymbol{\Phi}_n(\mathbf{P}), \end{aligned} \quad (61)$$

for $\mathbf{P} \triangleq \{\mathbf{P}_n, n = 1,\ldots,N\}$. Thus, the rate of s_n is given by $(1/2)\rho_n(\mathbf{P})$ [36] with

$$\rho_n(\mathbf{P}) = \ln\left|\mathbf{I}_2 + [\boldsymbol{\Lambda}_{nn}\mathbf{P}_n]^2 \times\left(\sum_{n'\in\mathcal{N}\setminus\{n\}}[\boldsymbol{\Lambda}_{nn'}\mathbf{P}_{n'}]^2 + 2\boldsymbol{\Phi}_n(\mathbf{P}) + \sigma^2\mathbf{I}_2\right)^{-1}\right|. \quad (62)$$

The problem of GM-rate maximization is

$$\max_{\mathbf{P}} \left[\prod_{n\in\mathcal{N}}\frac{1}{2}\rho_n(\mathbf{P})\right]^{1/N} \quad (63a)$$

$$\text{s.t.} \quad \sum_{m\in\mathcal{M}}\tilde{\pi}_m(\mathbf{P}) \le MP, \quad (63b)$$

where $\tilde{\pi}_m(\mathbf{P}) \triangleq \frac{1}{2}\sum_{n\in\mathcal{N}}\left\langle\left[\hat{\mathbf{H}}_{m,n}^T\right]^2[\mathbf{P}_n]^2\right\rangle$. With the function f defined from (23), and $\boldsymbol{\rho}(\mathbf{P}) \triangleq (\rho_1(\mathbf{P}),\ldots,\rho_N(\mathbf{P}))$, the problem (63) is equivalent to the following problem:

$$\max_{\mathbf{P}} f\left(\rho_1(\mathbf{P}),\ldots,\rho_N(\mathbf{P})\right) \quad \text{s.t.} \quad (63b). \quad (64)$$

Let $\mathbf{P}^{(\kappa)} \triangleq \{\mathbf{P}_n^{(\kappa)}, n = 1,\ldots,N\}$ be the feasible point for (64) that is found from the $(\kappa-1)$-st iteration. Like (27), we solve the following problem at the κ-th iteration to seek the steepest ascent:

$$\max_{\mathbf{P}} f^{(\kappa)}(\mathbf{P}) \triangleq \sum_{n=1}^{N}\gamma_n^{(\kappa)}\rho_n(\mathbf{P}) \quad \text{s.t.} \quad (63b), \quad (65)$$

for

$$\gamma_n^{(\kappa)} \triangleq \frac{\max_{n'\in\mathcal{N}}\rho_{n'}(\mathbf{P}^{(\kappa)})}{\rho_n(\mathbf{P}^{(\kappa)})}, \quad n = 1,\ldots,N. \quad (66)$$

By applying the inequality (2) for $(\mathbf{V},\mathbf{Y}) = (\boldsymbol{\Lambda}_{nn}\mathbf{P}_n, \sum_{n'\in\mathcal{N}\setminus\{n\}}[\boldsymbol{\Lambda}_{nn'}\mathbf{P}_{n'}]^2 + 2\boldsymbol{\Phi}_n(\mathbf{P}) + \sigma^2\mathbf{I}_2)$ and $(\bar{\mathbf{V}},\bar{\mathbf{Y}}) = (\boldsymbol{\Lambda}_{nn}\mathbf{P}_n^{(\kappa)}, \sum_{n'\in\mathcal{N}\setminus\{n\}}[\boldsymbol{\Lambda}_{nn'}\mathbf{P}_{n'}^{(\kappa)}]^2 + 2\boldsymbol{\Phi}_n(\mathbf{P}^{(\kappa)}) + \sigma^2\mathbf{I}_2)$, we obtain

$$\begin{aligned} \rho_n(\mathbf{P}) &\ge \bar{a}_n^{(\kappa)} + 2\langle\mathbf{B}_n^{(\kappa)},\mathbf{P}_n\rangle - \sum_{n'\in\mathcal{N}}\left\langle\mathbf{C}_{nn'}^{(\kappa)},[\mathbf{P}_{n'}]^2\right\rangle \\ &\quad - \sum_{m\in\mathcal{M}}\sum_{n'\in\mathcal{N}}\left\langle\mathbf{D}_{mnn'}^{(\kappa)},[\mathbf{P}_{n'}]^2\right\rangle \\ &\triangleq \rho_n^{(\kappa)}(\mathbf{P}), \end{aligned} \quad (67)$$

where $\bar{a}_n^{(\kappa)} \triangleq \rho_n(\mathbf{P}^{(\kappa)}) - \left\langle(\boldsymbol{\Psi}_n^{(\kappa)})^{-1},[\boldsymbol{\Lambda}_{nn}\mathbf{P}_n^{(\kappa)}]^2\right\rangle - \sigma^2\langle\boldsymbol{\Gamma}_n^{(\kappa)}\rangle$, $\mathbf{B}_n^{(\kappa)} \triangleq \boldsymbol{\Lambda}_{nn}^T(\boldsymbol{\Psi}_n^{(\kappa)})^{-1}\boldsymbol{\Lambda}_{nn}\mathbf{P}_n^{(\kappa)}$, $\mathbf{C}_{nn'}^{(\kappa)} \triangleq \boldsymbol{\Lambda}_{nn'}^T\boldsymbol{\Gamma}_n^{(\kappa)}\boldsymbol{\Lambda}_{nn'}$, $\mathbf{D}_{mnn'}^{(\kappa)} \triangleq \frac{1}{2}\varepsilon_{mn}\langle\boldsymbol{\Gamma}_n^{(\kappa)}\rangle[\hat{\mathbf{H}}_{m,n'}^T]^2$, $\boldsymbol{\Psi}_n^{(\kappa)} \triangleq \sum_{n'\in\mathcal{N}\setminus\{n\}}[\boldsymbol{\Lambda}_{nn'}\mathbf{P}_{n'}^{(\kappa)}]^2 + 2\boldsymbol{\Phi}_n(\mathbf{P}^{(\kappa)}) + \sigma^2\mathbf{I}_2$, and $\boldsymbol{\Gamma}_n^{(\kappa)} \triangleq \left(\boldsymbol{\Psi}_n^{(\kappa)}\right)^{-1} - \left(\boldsymbol{\Psi}_n^{(\kappa)} + [\boldsymbol{\Lambda}_{nn}\mathbf{P}_n^{(\kappa)}]^2\right)^{-1}$. The function $\rho_n^{(\kappa)}(\mathbf{P})$ is concave, which matches with $\rho_n(\mathbf{P})$ at $\mathbf{P}^{(\kappa)}$: $\rho_n(\mathbf{P}^{(\kappa)}) = \rho_n^{(\kappa)}(\mathbf{P}^{(\kappa)})$. We solve the following problem at the κ-th iteration to generate $\mathbf{P}^{(\kappa+1)}$:

$$\max_{\mathbf{P}} \bar{f}^{(\kappa)}(\mathbf{P}) \triangleq \sum_{n\in\mathcal{N}}\gamma_n^{(\kappa)}\rho_n^{(\kappa)}(\mathbf{P}) \quad \text{s.t.} \quad (63b), \quad (68)$$

for

$$\begin{aligned} &\bar{f}^{(\kappa)}(\mathbf{P}) \\ &\triangleq \sum_{n\in\mathcal{N}}\gamma_n^{(\kappa)}\rho_n^{(\kappa)}(\mathbf{P}) \\ &= \sum_{n\in\mathcal{N}}\gamma_n^{(\kappa)}\Big(\bar{a}_n^{(\kappa)} + 2\langle\mathbf{B}_n^{(\kappa)},\mathbf{P}_n\rangle - \sum_{n'\in\mathcal{N}}\left\langle\mathbf{C}_{nn'}^{(\kappa)},[\mathbf{P}_{n'}]^2\right\rangle \\ &\quad - \sum_{m\in\mathcal{M}}\sum_{n'\in\mathcal{N}}\left\langle\mathbf{D}_{mnn'}^{(\kappa)},[\mathbf{P}_{n'}]^2\right\rangle\Big) \\ &= \sum_{n\in\mathcal{N}}\gamma_n^{(\kappa)}\left(\bar{a}_n^{(\kappa)} + 2\langle\mathbf{B}_n^{(\kappa)},\mathbf{P}_n\rangle - \frac{1}{\gamma_n^{(\kappa)}}\langle\mathbf{E}_n^{(\kappa)},[\mathbf{P}_n]^2\rangle\right), \end{aligned} \quad (69)$$

where

$$\mathbf{E}_n^{(\kappa)} \triangleq \sum_{n'\in\mathcal{N}}\gamma_{n'}^{(\kappa)}\mathbf{C}_{n'n}^{(\kappa)} + \sum_{n'\in\mathcal{N}}\sum_{m\in\mathcal{M}}\gamma_{n'}^{(\kappa)}\mathbf{D}_{mn'n}^{(\kappa)},$$

The problem (68) admits the following closed-form solution of linear computational complexity

$$\mathbf{P}_n^{(\kappa+1)} = \begin{cases} (\mathbf{E}_n^{(\kappa)})^{-1}\gamma_n^{(\kappa)}\mathbf{B}_n^{(\kappa)} & \text{if } \omega^{(\kappa)} \le MP \\ \left(2\mathbf{E}_n^{(\kappa)} + \mu\mathbf{G}_n\right)^{-1}2\gamma_n^{(\kappa)}\mathbf{B}_n^{(\kappa)} & \text{otherwise,} \end{cases} \quad (70)$$

Algorithm 3 Scalable Algorithm for IGS-Based GM-Rate Optimization Problem (63)

1: **Initialization**: Initialize a feasible point $P^{(0)}$. Set $\kappa = 0$.
2: **Repeat until convergence of** $\mathbf{P}^{(\kappa)}$: Update $\gamma_n^{(\kappa)}$ by (66). Generate $\mathbf{P}^{(\kappa+1)}$ by (70). Reset $\kappa \leftarrow \kappa + 1$.
3: **Output** $\mathbf{P}^{(\kappa)}$ as the optimal solution of (63).

where $\omega^{(\kappa)} \triangleq \frac{1}{2}\sum_{n\in\mathcal{N}} \left\langle \mathbf{G}_n, \left[(\mathbf{E}_n^{(\kappa)})^{-1}\gamma_n^{(\kappa)}\mathbf{B}_n^{(\kappa)}\right]^2 \right\rangle$, $\mathbf{G}_n \triangleq \sum_{m=1}^{M}\left[\hat{\mathbf{H}}_{m,n}^T\right]^2$ and $\mu > 0$ can be found by bisection search such that

$$\frac{1}{2}\sum_{n\in\mathcal{N}} \left\langle \mathbf{G}_n, \left[\left(2\mathbf{E}_n^{(\kappa)} + \mu\mathbf{G}_n\right)^{-1} 2\gamma_n^{(\kappa)}\mathbf{B}_n^{(\kappa)}\right]^2 \right\rangle = MP.$$

Algorithm 3 provides the pseudo-code for the proposed scalable procedure for computing (63), which like Algorithm 1 converges at least to a locally optimal solution.

B. Energy Efficient Payload Data Transmission

Like (35), the GE-EE maximization under IGS is formulated as

$$\max_{\mathbf{P}} \frac{[\prod_{n\in\mathcal{N}} \frac{1}{2}\rho_n(\mathbf{P})]^{1/N}}{\tilde{\pi}_{te}(\mathbf{P})} \quad \text{s.t.} \quad (63b), \tag{71}$$

where $\tilde{\pi}_{te}(\mathbf{P})$ is the total power consumption [12], [31], given by

$$\tilde{\pi}_{te}(\mathbf{P}) = \alpha \sum_{m\in\mathcal{M}} \tilde{\pi}_m(\mathbf{P}) + N_t M P_{sc} + M\left(P_0 + \frac{P_{bt}BN}{\ln 2}\left(\prod_{n\in\mathcal{N}} \frac{1}{2}\rho_n(\mathbf{P})\right)^{1/N}\right). \tag{72}$$

By using a similar argument for obtaining (39) and by defining $\pi_{te}(\mathbf{P}) \triangleq \alpha\sum_{m\in\mathcal{M}}\tilde{\pi}_m(\mathbf{P}) + \nu$ and $\nu \triangleq N_t MP_{sc} + MP_0$, the equivalent problem for problem (71) is:

$$\max_{\mathbf{P}} f_E(\rho(\mathbf{P}), \tilde{\pi}(\mathbf{P})) \triangleq \frac{f(\rho(\mathbf{P}))}{\tilde{\pi}(\mathbf{P})} \quad \text{s.t.} \quad (63b). \tag{73}$$

Let $\mathbf{P}^{(\kappa)} \triangleq \{\mathbf{P}_n^{(\kappa)}, n = 1, \ldots, N\}$ be the feasible point for (39) that is found from the $(\kappa-1)$-st iteration. Like (41), we seek the steepest ascent by solving the problem

$$\max_{\mathbf{P}} \left[\sum_{n=1}^{N} \gamma_n^{(\kappa)}\rho_n(\mathbf{P}) - \eta^{(\kappa)}\tilde{\pi}(\mathbf{P})\right] \quad \text{s.t.} \quad (63b), \tag{74}$$

for

$$\eta^{(\kappa)} \triangleq N \max_{n'\in\mathcal{N}} \rho_{n'}(\mathbf{P}^{(\kappa)})/\tilde{\pi}(\mathbf{P}^{(\kappa)}), \tag{75}$$

and

$$\gamma_n^{(\kappa)} \triangleq \max_{n'\in\mathcal{N}} \rho_{n'}(\mathbf{P}^{(\kappa)})/\rho_n(\mathbf{P}^{(\kappa)}). \tag{76}$$

Recalling the function $\rho_n^{(\kappa)}(\mathbf{P})$ from (67), at the κth iteration, we solve the following convex problem to generate the next iterative point $\mathbf{P}^{(\kappa+1)}$:

$$\max_{\mathbf{P}} \left[\sum_{n=1}^{N} \gamma_n^{(\kappa)}\rho_n^{(\kappa)}(\mathbf{P}) - \eta^{(\kappa)}\tilde{\pi}(\mathbf{P})\right] \quad \text{s.t.} \quad (63b), \tag{77}$$

which admits the following closed form solution of linear computational complexity:

$$\mathbf{P}_n^{(\kappa+1)} = \begin{cases} \left(2\mathbf{E}_n^{(\kappa)} + \bar{\eta}^{(\kappa)}\mathbf{G}_n\right)^{-1} 2\gamma_n^{(\kappa)}\mathbf{B}_n^{(\kappa)} & \text{if } \bar{\omega}^{(\kappa)} \le MP \\ \left(2\mathbf{E}_n^{(\kappa)} + (\bar{\eta}^{(\kappa)} + \bar{\mu})\mathbf{G}_n\right)^{-1} 2\mathbf{B}_n^{(\kappa)} & \text{otherwise,} \end{cases} \tag{78}$$

where $\bar{\eta}^{(\kappa)} = \eta^{(\kappa)}\alpha$, $\mathbf{G}_n \triangleq \sum_{m=1}^{M}\left[\hat{\mathbf{H}}_{m,n}^T\right]^2$, $\bar{\omega}^{(\kappa)} \triangleq \frac{1}{2}\sum_{n\in\mathcal{N}} \left\langle \mathbf{G}_n, \left[\left(2\mathbf{E}_n^{(\kappa)} + \eta^{(\kappa)}\alpha\mathbf{G}_n\right)^{-1} 2\gamma_n^{(\kappa)}\mathbf{B}_n^{(\kappa)}\right]^2 \right\rangle$ and $\bar{\mu} > 0$ can be found by bisection search such that

$$\frac{1}{2}\sum_{n\in\mathcal{N}} \left\langle \mathbf{G}_n, \left[\left(2\mathbf{E}_n^{(\kappa)} + (\bar{\eta}^{(\kappa)} + \bar{\mu})\mathbf{G}_n\right)^{-1} \times 2\gamma_n^{(\kappa)}\mathbf{B}_n^{(\kappa)}\right]^2 \right\rangle = MP.$$

Algorithm 4 Scalable Algorithm for IGS-Based GM-EE Problem (71)

1: **Initialization**: Initialize a feasible point $P^{(0)}$. Set $\kappa = 0$.
2: **Repeat until convergence of** $\mathbf{P}^{(\kappa)}$: Update $\eta^{(\kappa)}$ and $\gamma_n^{(\kappa)}$ by (75) and (76). Generate $\mathbf{P}^{(\kappa+1)}$ by (78). Reset $\kappa \leftarrow \kappa + 1$.
3: **Output** $\mathbf{P}^{(\kappa)}$ as the optimal solution of (71).

Algorithm 4 provides the pseudo-code for the proposed scalable procedure for computing (71), which like Algorithm 2 converges at least to a locally optimal solution.

IV. Simulations

In this section, we analyze the performance of our proposed algorithms. We start by first describing our simulation parameters. The large-scale fading coefficient in (3) is given by

$$\beta_{mn} = 10^{(-\text{PL}_{mn} + X_{mn})/10}, \tag{79}$$

where PL_{mn} is the path-loss (in dB) and $10^{X_{mn}/10}$ represents the shadowing effect with $X_{mn} \sim \mathcal{N}(0, \sigma_{\text{sh}}^2)$. We use a three-slope model for path-loss [37], which is given by (in dB)

$$\text{PL}_{mn} = \begin{cases} L + 35\log_{10}(d_{mn}), & d_{mn} > d_1 \\ L + 15\log_{10}(d_1) + 20\log_{10}(d_{mn}), & d_o < d_{mn} \le d_1 \\ L + 15\log_{10}(d_1) + 20\log_{10}(d_o), & d_{mn} \le d_o, \end{cases} \tag{80}$$

where the constant factor L depends on the carrier frequency and the heights of the users and APs and d_{mn} is the distance between the AP m and UE n. In our simulations, we set

$\sigma_{\text{sh}} = 8$ dB, $d_o = 10$ m, $d_1 = 50$ m, and $L = 140.7$ dB. These parameters resemble those in [1], [12].

We have randomly deployed the APs and the UEs within a square of 1×1 km^2. The square is wrapped around at the edges to avoid boundary effects. Unless stated otherwise, the pilot sequences for N UEs are obtained from N orthogonal pilot sequences of length $t_u = N$, where the rows of fast-Fourier transform (FFT) matrix are used to generate pilot sequences. However, to study the effect of pilot contamination, we also analyze the performance of the proposed algorithms under smaller pilot sequences of length $t_u < N$, under which the pilot sequences would not remain orthogonal and will be based on the truncated rows of FFT matrix. The length of coherence interval is $t = 320$ samples, corresponding to a coherence bandwidth of $B = 320$ KHz and coherence time of 1 ms. In addition, we set the system bandwidth to 10 MHz, noise power density is set to -174 dBm/Hz and noise figure is equal to 9 dB [1], [2].

To simulate the energy efficiency performance of a network, the drain efficiency of power amplifier is set to 0.4, i.e., $\alpha = 1/0.4$ in (36), internal power for running the circuit components is set to $P_{sc} = 0.2$ W, fixed power consumption for each backhaul is $P_0 = 0.825$ W, and traffic-dependent backhaul power is $P_{bt} = 0.25$ W/(Gbits/s). These values are taken from [12].

Unless otherwise specified, we use $M = 128$ APs, $N = 32$ UE, training length $t_u = N$, $N_t = 2$ antennas per AP, and power budget $P = 20$ dBm. The power allocated to the pilots, p_{u}, is set to P. We use MATLAB for simulating our proposed algorithms.

In what follows, we refer to our proposed algorithms as follows for convenience:

- *PGS stat. perf. (Alg. 1)* refers to the performance (perf.) of PGS based Alg. 1 (for the rate defined by (22).
- *PGS instant. perf. (Alg. 1)* refers to the performance of PGS based modified Alg. 1 for the rate defined by (52).
- *IGS instant. perf. (Alg. 3)* refers to the performance of IGS based Alg. 3.

Alg. 1 and Alg. 3 refer to the GM-rate optimization algorithms. On the other hand, if we replace them by Alg. 2 and Alg. 4, respectively, the above description will refer to the GM-EE algorithms. Note that in all the figures, the rates are calculated in bps/Hz, which are obtained by dividing the rates (defined in nats/sec/Hz) by $\ln 2$. The results are averaged over random realizations of the AP/user locations, shadow fading, and small-scale fading.

In the following subsections, we separately analyze the GM-rate performance and the GM-EE performance of our proposed algorithms.

A. GM-Rate Performance

Fig. 2 plots the distribution of individual user-rate χ_n, for $n \in \mathcal{N}$, achieved by solving the PGS based instantaneous GM-rate optimization Alg. 1. Fig. 2 also plots the GM-rate, $\left(\prod_{n\in\mathcal{N}} \chi_n\right)^{1/N}$ and the max-min rate. The latter solves the max-min rate optimization problem using the approach in [15] and the individual user-rate achieved by solving that problem

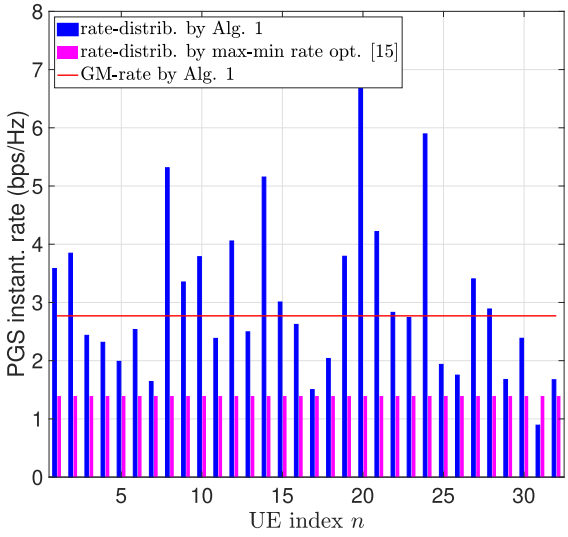

Fig. 2. Individual user-rate distribution by PGS instant. Alg. 1 with $t_u = N/2$ and $N_t = 1$.

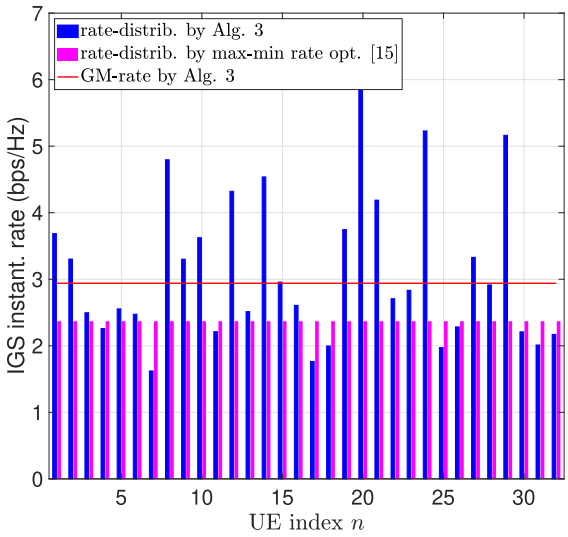

Fig. 3. Individual user-rate distribution by IGS instant. Alg. 3 with $t_u = N/2$ and $N_t = 1$.

is plotted for comparison. Since max-min rate optimization ensures maximization of the worst-case user-rate to achieve fair rate-distribution, comparing this approach with the proposed GM-rate optimization algorithms will indicate whether the latter are successful in achieving fair rate-distribution among the users. Similarly, Fig. 3 plots the distribution of individual user-rate $\frac{1}{2}\rho_n$, for $n \in \mathcal{N}$, achieved by solving the IGS based instantaneous GM-rate optimization Alg. 3 along with the GM-rate, $\left(\prod_{n\in\mathcal{N}} \frac{1}{2}\rho_n\right)^{1/N}$ and the max-min rate. Fig. 2 and Fig. 3 assume $t_u = N/2$ to provide comparison in the presence of pilot contamination. Moreover, N_t is set to 1 because the max-min rate algorithms in [15] assume single-antenna APs.

Remark 1: The following interesting insights can be observed from Figs. 2 and 3, which appreciate the importance of the proposed GM-rate optimization Algs. 1 and 3.

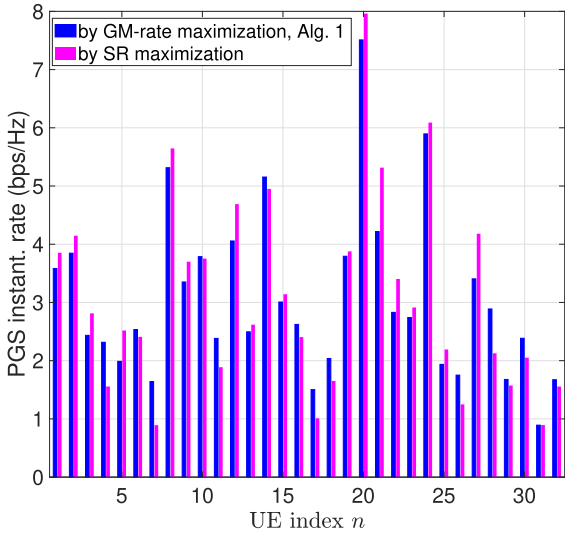

Fig. 4. Individual user-rate distribution by Alg. 1 (SR 97.65 bps/Hz; norm. rate-var. 0.21) and by QoS constrained SR maximization [15]. (SR 99.02 bps/Hz; normalized rate-var. 0.29.)

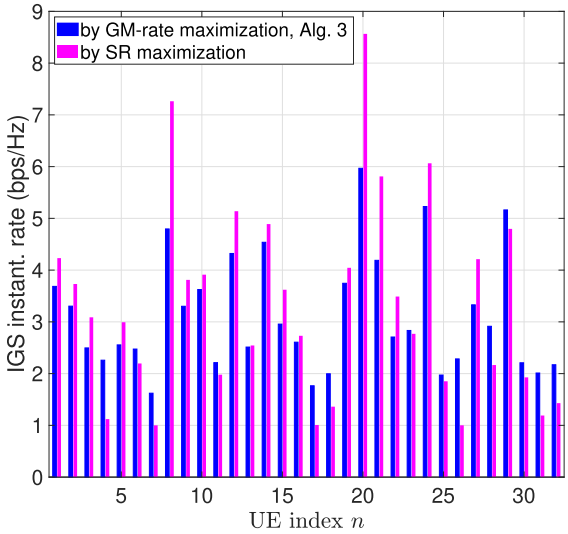

Fig. 5. Individual user-rate distribution by Alg. 3 (SR 99.76 bps/Hz; norm. rate-var. 0.13) and by QoS constrained SR maximization [15]. (SR 105.93 bps/Hz; normalized rate-var. 0.32.)

- The individual user rate by the GM-rate optimization is minimally compromised compared to that by the max-min rate optimization. Only for very few users, the user-rate by GM-rate optimization is smaller than that by the conventional max-min rate optimization. On the other hand, the computational complexity of the proposed GM-rate optimization Algorithms 1 and 3 is $\mathcal{O}(N)$, which is very small, thanks to the proposed closed-form based solutions, compared to that offered by the conventional max-min rate optimization algorithms. Particularly, due to the dependency on convex solvers, the computational complexities of PGS and IGS based conventional max-min rate algorithms are $\mathcal{O}(N^3)$ and $\mathcal{O}(N^4)$, respectively [38].
- In addition to the better rates provided by the IGS than those by the PGS, the former also promises fairer rate distribution (with smaller variance) among the user-rates. For the example in Figs. 2 and 3, the standard-deviation in the user-rates is 1.42 and 1.12 for the PGS and IGS based implementations, respectively.

Fig. 4 compares the individual user-rate distribution achieved by solving the PGS based instantaneous GM-rate optimization Alg. 1 with that achieved by solving conventional SR maximization with QoS user-rate constraints. The solution of the latter is based on the approach in [15], which relies on the use of convex-solvers. It can be observed from Fig. 4 that compared to the conventional SR maximization, the proposed GM-rate maximization leads to a fairer rate allocation to all users (smaller variance (var) among the users' rates) with a marginal compromise in the achievable sum-rate. In addition, the proposed GM-rate maximization is computationally very efficient compared to the conventional SR maximization (the complexity of former versus latter is $\mathcal{O}(N)$ versus $\mathcal{O}(N^4)$), thanks to our proposed solution based on closed-form expressions. Similar results are observed in Fig. 5 for the IGS based implementation.

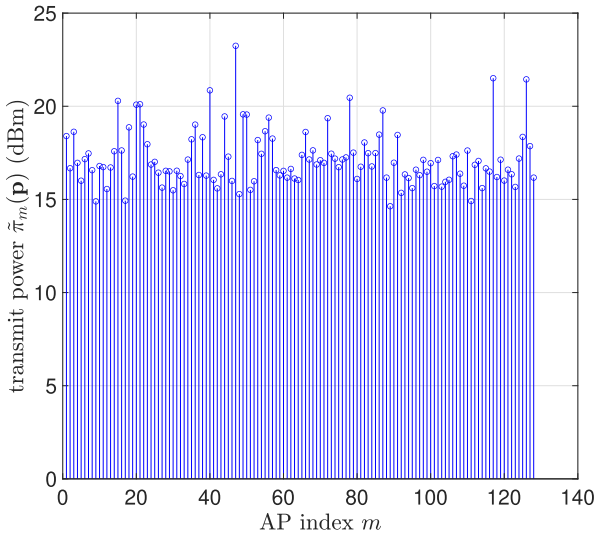

Fig. 6. Average transmit power $\pi_m(\mathbf{p})$ versus the AP index m for PGS instant. Alg. 1.

Figs. 6 and 7 plot the average transmit power $\tilde{\pi}_m(\mathbf{p})$ and $\tilde{\pi}_m(\mathbf{P})$, respectively, versus the AP index m, for PGS instant. Alg. 1 and IGS instant. Alg. 3, respectively. Figs. 6 and 7 confirm that though we are using a single sum-power constraint in (24b) and (63b) instead of separate M power constraint for all the APs, the average transmit power requirement is still balanced over all the APs. In addition, we have verified through simulations that we can keep the average transmit power at *each* AP within some specific power budget, say $\bar{P}$, by decreasing the total power budget P. This obviously results in a slight performance loss owing to the decrease in total power budget P. For example, if the average transmit power at each AP has to be kept smaller than $\bar{P} = 23$ dBm (the average transmit power of AP 47 in Fig. 6 exceeds this limit), we can achieve this by decreasing the total power budget P from 20 dBm to 19.5 dBm, which slightly decreases the achievable GM-rate from 3.48 bps/Hz to 3.43 bps/Hz.

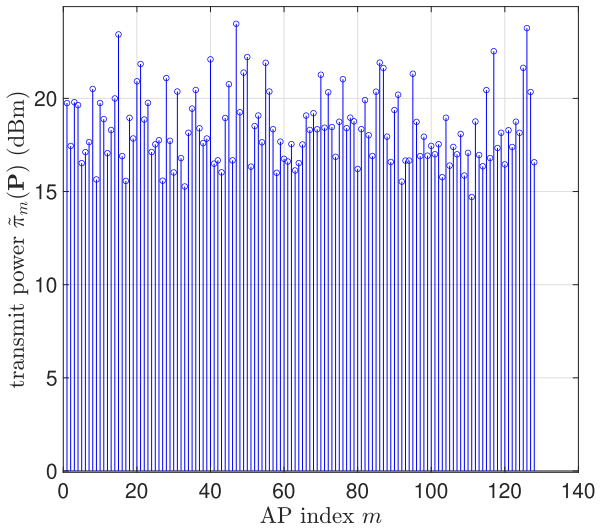

Fig. 7. Average transmit power $\tilde{\pi}_m(\mathbf{P})$ versus the AP index m for IGS instant. Alg. 3.

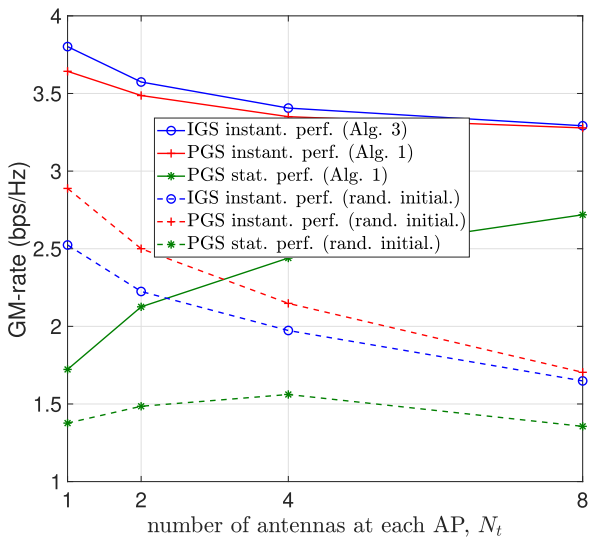

Fig. 9. GM-rate versus the number of antennas per AP N_t for $MN_t = 256$.

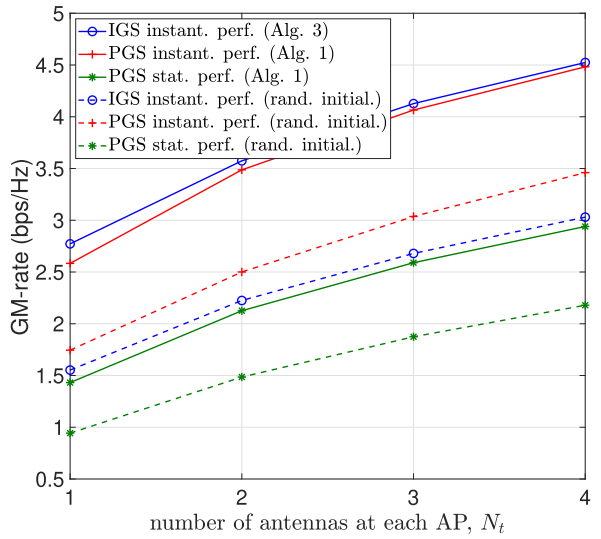

Fig. 8. GM-rate versus the number of antennas per AP N_t for $M = 128$.

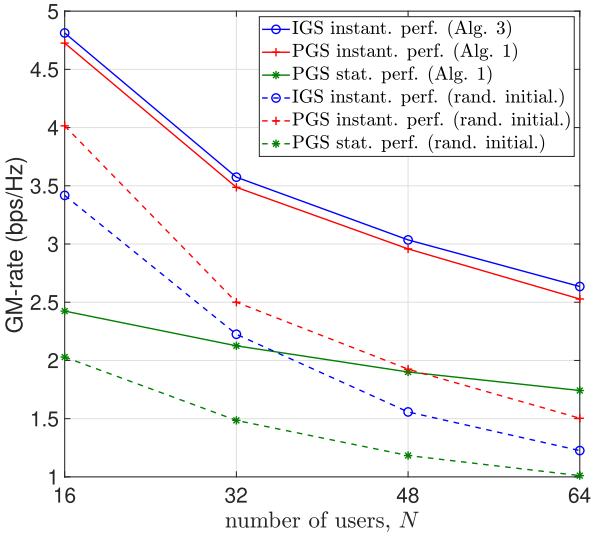

Fig. 10. GM-rate vs. the number of users N.

Fig. 8 and Fig. 9 plot the GM-rate versus the number of antennas at each AP N_t, such that Fig. 8 assumes fixed number of APs $M = 128$, while Fig. 9 assumes that the product $MN_t = 256$ is kept fixed. Fig. 8 and Fig. 9 also plot the GM-rate achieved by using random (rand.) power allocation, which is referred to as "rand. initialization". We can observe from these figures that compared to such random power allocation, the optimized power allocation by the proposed Algorithms 1 and 3 is substantially helpful in uplifting the GM-rate performance. Fig. 8 shows that the GM-rate increases with the increase in N_t due to increase in the number of resources. On the other hand, the GM-rate is decreasing upon increasing N_t in Fig. 9 because total number of distributed antennas MN_t is kept fixed. Fig. 9 shows that under fixed product MN_t, $N_t = 1$ yields the best performance.

Fig. 10 plots the GM-rate versus the number of users N. As expected, the GM-rate decreases with the increase in the the number of users N. This is because when more users will be competing for the same fixed resources (fixed power budget), their performance will be affected. Fig. 10 shows that as N increases, the instantaneous performance gets closer to the performance based on the statistical data.

Fig. 11 plots the optimized GM-rate versus the number of APs M. Fig. 11 also shows the performance of the proposed Algorithms 1 and 3 with user-centric (UC) approach, where each UE is served by a set of selected APs [12, Sec. V]. The APs selection in UC approach can save power consumption and reduce the backhaul/fronthaul requirements. Particularly, we employ largest large-scale fading based approach for APs selection, where the UE n is associated with only $M_n \leq M$ APs, corresponding to the M_n largest large-scale fading coefficients. Particularly, following [12, Alg. 3], the proposed algorithms can be easily modified to implement UC approach. It can be observed from Fig. 11 that by employing UC

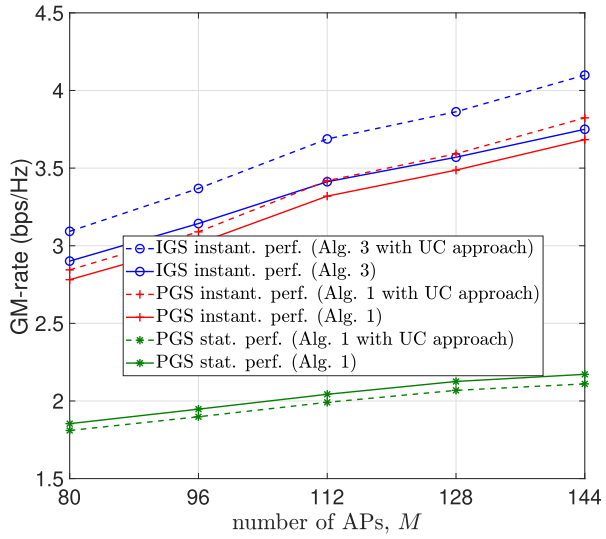

Fig. 11. GM-rate by Algorithms 1 and 3 adjusted for UC approach.

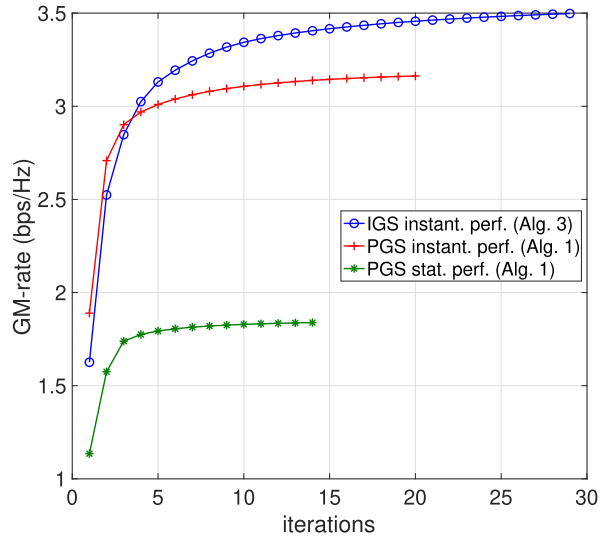

Fig. 12. Convergence of Algorithms 1 and 3.

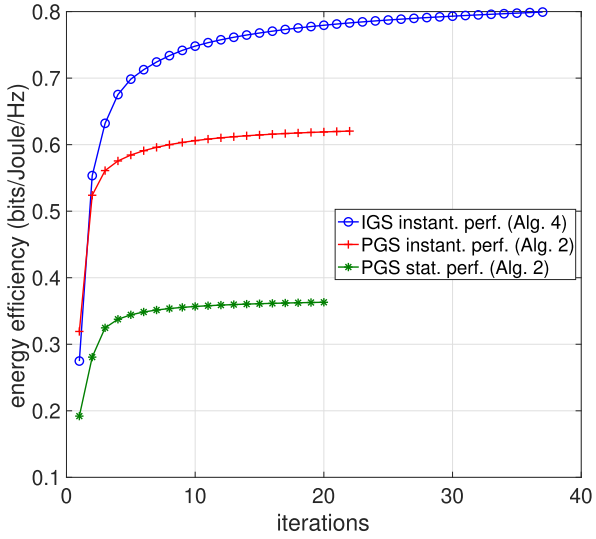

Fig. 13. Convergence of Algorithms 2 and 4.

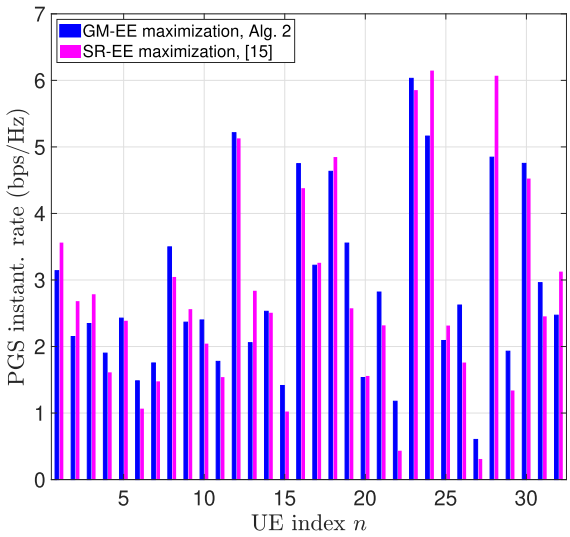

Fig. 14. Individual UE-rate distributions by GM-EE and SR-EE optimizations with $t_u = N/2$ and $N_t = 1$.

approach, the achievable GM-rate gets better for PGS instant. perf. (Alg. 1) and IGS instant. perf. (Alg. 3), however, it gets slightly worse for PGS stat. perf. (Alg. 1). The improvement in the achievable rate under UC approach is expected as this approach saves power consumption at the APs and thus reduces undue interference at the UEs. However, the improvement is not observed for PGS stat. perf. (Alg. 1), which can be explained owing to our simpler implementation, which is based on the same power allocation across all the APs for a given UE.

Fig. 12 plots the convergence performance of the proposed Algorithms 1 and 3. It can be observed from Fig. 12 that the proposed PGS based Alg. 1 converges fairly promptly within 15-20 iterations, whereas the proposed IGS based Alg. 3 requires more iterations (around 25-30) for convergence. This is because the latter has to optimize $3N$ additional optimization variables than the former. The same trend of the convergence performance has been observed for the proposed GM-EE maximization Algorithms 2 and 4, which is plotted in Fig. 13.

TABLE I

COMPARISON OF PGS BASED GM-EE (ALG. 2) WITH THE SR-EE ALGORITHM (APPROACH IN [15]) FOR $N_t = 1$, $t_u = N/2$

	GM-EE Alg. 2 (PGS instant. perf)	SR-EE (PGS) (using [15])
sum-EE	0.67 bits/J/Hz	0.69 bits/J/Hz
GM-EE	0.59 bits/J/Hz	0.57 bits/J/Hz
sum-rate	93.9 bps/Hz	92.9
rate-dev.	0.52	0.60
# iter.	33.0	14.33
complexity	$\mathcal{O}(N)$	$\mathcal{O}(N^4)$ [38]

B. Energy Efficiency Performance

In this section, we analyze the GM-rate based EE (GM-EE) performance by plotting $\frac{N[\prod_{n\in\mathcal{N}} r_n(\mathbf{p})]^{1/N}}{\pi_{te}(\mathbf{p})}$ for PGS

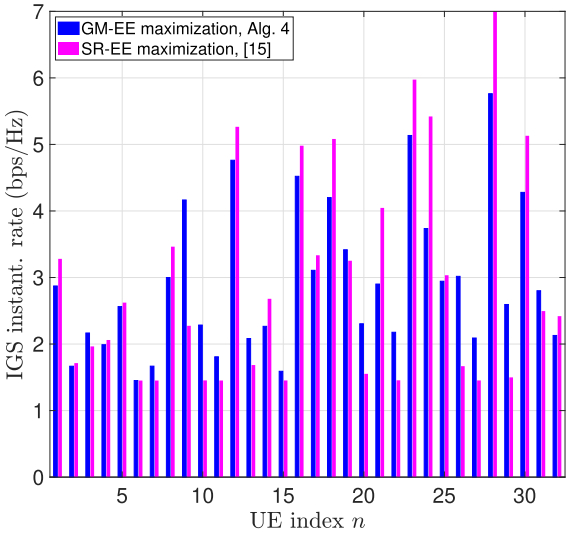

Fig. 15. Individual IGS UE-rate distributions by GM-EE and SR-EE optimizations with $t_u = N/2$ and $N_t = 1$.

stat. perf. (Alg. 2), $\frac{N[\prod_{n\in\mathcal{N}}\chi_n(\mathbf{p})]^{1/N}}{\pi_{te}(\mathbf{p})}$ for PGS instant. perf. (Alg. 2), and $\frac{N[\prod_{n\in\mathcal{N}}\frac{1}{2}\rho_n(\mathbf{P})]^{1/N}}{\tilde{\pi}_{te}(\mathbf{P})}$ for IGS instant. perf. (Alg 4). Note that we have multiplied the numerator of EEs by a factor N to have comparable numbers as that are obtained by the conventional sum-rate (SR) based EE (SR-EE) metrics, $\frac{\sum_{n\in\mathcal{N}} r_n(\mathbf{p})}{\pi_{te}(\mathbf{p})}$, $\frac{\sum_{n\in\mathcal{N}}\chi_n(\mathbf{p})}{\pi_{te}(\mathbf{p})}$, and $\frac{\sum_{n\in\mathcal{N}}\frac{1}{2}\rho_n(\mathbf{P})}{\tilde{\pi}_{te}(\mathbf{P})}$, respectively.

Fig. 14 plots the distribution of individual user-rate χ_n, for $n \in \mathcal{N}$, achieved by solving the proposed PGS based instantaneous GM-EE optimization Alg. 2. Similarly, Fig. 15 plots the individual user-rate $\frac{1}{2}\rho_n$, achieved by solving IGS based instantaneous GM-EE optimization Alg. 4. In both figures, we compare the achievable rates by the proposed GM-EE maximization (solution by closed-form iterations) with that by the conventional SR-EE maximization (solution based on the off-the-shelf convex solvers [12], [15]). To plot SR-EE maximization results, we use the solution approach of [15][3] and set the QoS user-rate constraint to be equal to the minimum user-rate achieved by the proposed GM-EE optimization. Fig. 14 and Fig. 15 assume $t_u = N/2$ to provide comparison in the presence of pilot contamination. Moreover, N_t is set to 1 because SR-EE algorithms in [15] assume single-antenna APs. The detailed comparison of the proposed GM-EE algorithms and the conventional SR-EE algorithms is provided in Tables I and II, which show the achievable average sum-EE (the conventional EE objective), GM-EE, sum-rate, standard-deviation among the user rates (rate-dev.), number of iterations required for convergence and computational complexity. It can be observed from Fig. 14 and Fig. 15 and Tables I and II that though the individual user rates, sum-rates, and sum-EE are minimally compromised by proposed GM-EE algorithms compared to the conventional SR-EE algorithms, the former promise fairer rate distribution and much smaller computational complexity.

[3]The path-following algorithms of [15] are used by employing convex solver, such as CVX, however, for fair comparison, we use single sum-power constraint and imperfect channel estimation, as adopted in this work.

TABLE II

COMPARISON OF IGS BASED GM-EE (ALG. 4) WITH THE SR-EE ALGORITHM (APPROACH IN [15]) FOR $N_t = 1, t_u = N/2$

	GM-EE Alg. 4 (IGS instant. perf)	SR-EE (IGS) (using [15])
sum-EE	0.68 bits/J/Hz	0.74 bits/J/Hz
GM-EE	0.61 bits/J/Hz	0.59 bits/J/Hz
sum-rate	96.9 bps/Hz	98.8
rate-dev.	0.43	0.57
# iter.	48.7	36.0
complexity	$\mathcal{O}(N)$	$\mathcal{O}(N^4)$ [38]

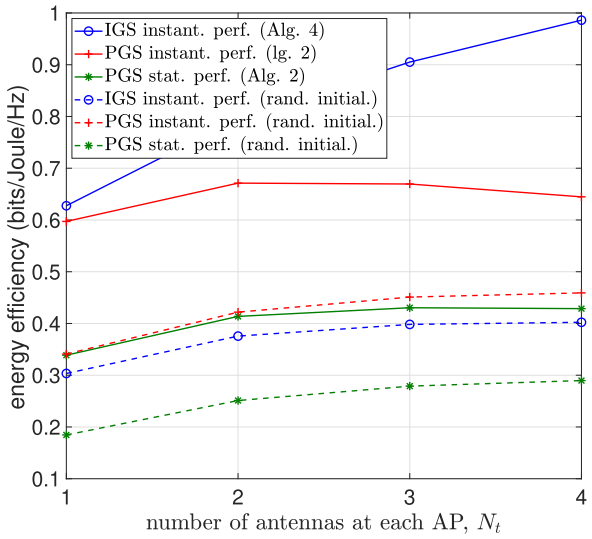

Fig. 16. GM-EE versus the number of antennas per AP N_t for $M = 128$.

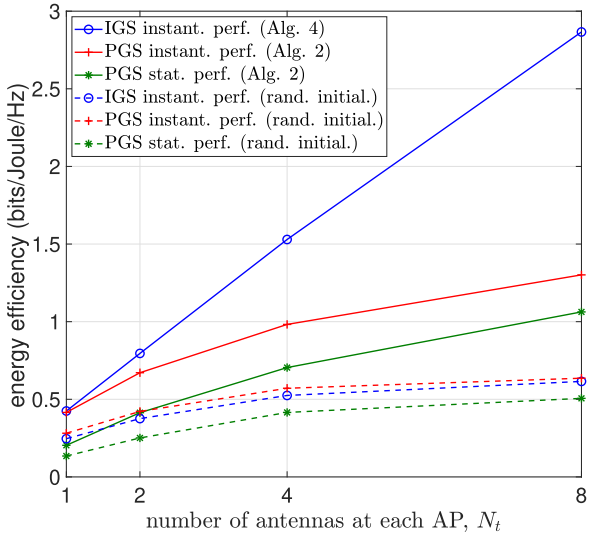

Fig. 17. GM-EE vs. the number of antennas per AP N_t for $MN_t = 256$.

Fig. 16 and 17 plot the GM-EE versus the number of antennas at each AP N_t, such that Fig. 16 assumes fixed number of APs $M = 128$, while Fig. 17 assumes that the product $MN_t = 256$ is kept fixed. Fig. 16 shows that the GM-EE increases upon increasing N_t due to increase in the number of resources, however, the improvement starts

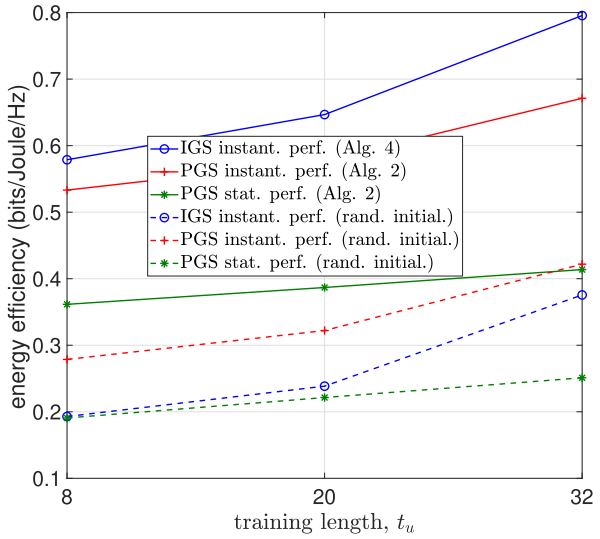

Fig. 18. GM-EE versus the training length t_u.

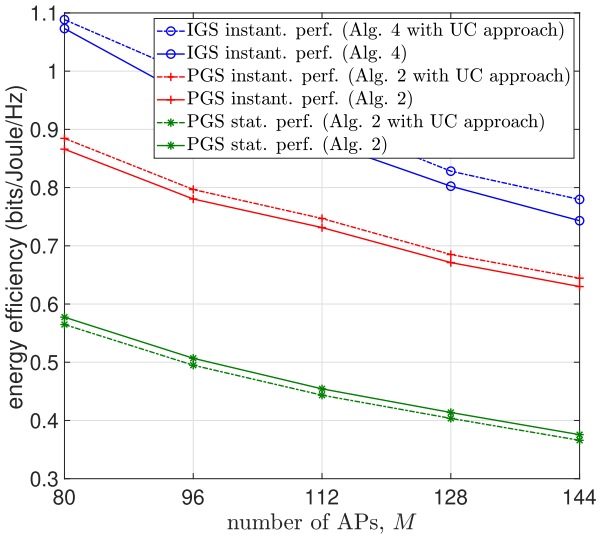

Fig. 19. GM-EE with UC approach.

diminishing for the larger values of N_t because of the increase in power consumption (the factor in the denominator of EE). However, the improvement in GM-EE is not affected in that way in Fig. 17 because the total number of distributed antennas MN_t is kept fixed. Fig. 16 and Fig. 17 show the supremacy of IGS over PGS as the performance gain by the former gets substantial for larger values of N_t.

Fig. 18 plots the GM-EE versus the length of the pilot sequences, t_u. Since the number of users is kept fixed to $N = 32$, Fig. 18 shows the effect of pilot contamination because when t_u is smaller than N, the pilots sequences would not remain orthogonal. The GM-EE performance degrades for smaller values of t_u due to pilot contamination. Fig. 19 plots the optimized GM-EE versus the number of APs M. Fig. 19 also plots the performance of the proposed Algorithms 2 and 4 with APs selection based UC approach and shows the advantage of employing this approach for PGS instant. perf. (Alg. 2) and IGS instant. perf. (Alg. 4).

V. Conclusion

The paper has proposed a particular power allocation for conjugate beamforming to deliver information to users of a cell-free massive MIMO network. It has introduced the maximization of the GM of users' rates as a new way of achieving balanced users' rates while keeping their sum high. New algorithms of very low computational complexity, which are based on closed-form expressions, have been developed for solutions under both proper Gaussian signaling and improper Gaussian signaling. The paper has also exploited the ratio of the user rates' GM to the total transmit power as a new way to gauge the network energy efficiency with users' rates balanced. Its maximization is also achieved by iterating closed-form expressions. The provided simulations have clearly demonstrated the advantages of the proposed approach in bringing about multiple performance advantages in terms of achievable individual user rates and their sum, and the energy efficiency with fair power allocations to the access points of the cfm-MIMO. A possible future research direction is to investigate the performance of the proposed algorithms under a Rician fading channel assumption.

References

[1] 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. 1834–1849, Mar. 2017.

[2] T. M. Hoang, H. Q. Ngo, T. Q. Duong, H. D. Tuan, and A. Marshall, "Cell-free massive MIMO networks: Optimal power control against active eavesdropping," *IEEE Trans. Commun.*, vol. 66, no. 10, pp. 4724–4737, Oct. 2018.

[3] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "On the uplink maxmin SINR of cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 18, no. 4, pp. 2021–2036, Apr. 2019.

[4] S.-N. Jin, D.-W. Yue, and H. H. Nguyen, "Spectral and energy efficiency in cell-free massive MIMO systems over correlated rician fading," *IEEE Syst. J.*, vol. 15, no. 2, pp. 2822–2833, Jun. 2021.

[5] A. Papazafeiropoulos, H. Q. Ngo, P. Kourtessis, S. Chatzinotas, and M. J. Senior, "Optimal energy efficiency in cell-free massive MIMO systems: A stochastic geometry approach," in *Proc. IEEE PIMRC*, Oct. 2020, pp. 1–7.

[6] 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.

[7] J. Zheng, J. Zhang, E. Bjornson, and B. Ai, "Impact of channel aging on cell-free massive MIMO over spatially correlated channels," *IEEE Trans. Wireless Commun.*, vol. 20, no. 10, pp. 6451–6466, Oct. 2021.

[8] G. Interdonato, H. Q. Ngo, and E. G. Larsson, "Enhanced normalized conjugate beamforming for cell-free massive MIMO," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 2863–2877, May 2021.

[9] J. Denis and M. Assaad, "Improving cell-free massive MIMO networks performance: A user scheduling approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 11, pp. 7360–7374, Nov. 2021.

[10] S. Chen, J. Zhang, E. Björnson, S. Wang, C. Xing, and B. Ai, "Wireless caching: Cell-free versus small cells," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2021, pp. 1–6.

[11] T. V. Chien, E. Bjornson, and E. G. Larsson, "Optimal design of energy-efficient cell-free massive MIMO: Joint power allocation and load balancing," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2020, pp. 5145–5149.

[12] H. Q. Ngo, L.-N. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 25–39, Jan. 2018.

[13] L. D. Nguyen, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Multi-user regularized zero-forcing beamforming," *IEEE Trans. Signal Process.*, vol. 67, no. 11, pp. 2839–2853, Jun. 2019.

[14] L. D. Nguyen, H. D. Tuan, T. Q. Duong, H. V. Poor, and L. Hanzo, "Energy-efficient multi-cell massive MIMO subject to minimum user-rate constraints," *IEEE Trans. Commun.*, vol. 69, no. 2, pp. 914–928, Feb. 2021.

[15] A. A. Nasir, H. D. Tuan, H. Q. Ngo, T. Q. Duong, and H. V. Poor, "Cell-free massive MIMO in the short blocklength regime for URLLC," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 5861–5871, Sep. 2021.

[16] A. A. Nasir, H. D. Tuan, T. Q. Duong, H. V. Poor, and L. Hanzo, "Hybrid beamforming for multi-user millimeter-wave networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 3, pp. 2943–2956, Mar. 2020.

[17] H. A. Ammar, R. Adve, S. Shahbazpanahi, G. Boudreau, and K. V. Srinivas, "Distributed resource allocation optimization for user-centric cell-free MIMO networks," *IEEE Trans. Wireless Commun.*, vol. 21, no. 5, pp. 3099–3115, May 2022.

[18] C. Hellings, M. Joham, and W. Utschick, "QoS feasibility in MIMO broadcast channels with widely linear transceivers," *IEEE Signal Process. Lett.*, vol. 20, no. 11, pp. 1134–1137, Nov. 2013.

[19] Y. Zeng, C. M. Yetis, E. Gunawan, Y. L. Guan, and R. Zhang, "Transmit optimization with improper Gaussian signaling for interference channels," *IEEE Trans. Signal Process.*, vol. 61, no. 11, pp. 2899–2913, Jun. 2013.

[20] S. Lagen, A. Agustin, and J. Vidal, "On the superiority of improper Gaussian signaling in wireless interference MIMO scenarios," *IEEE Trans. Commun.*, vol. 64, no. 8, pp. 3350–3368, Aug. 2016.

[21] A. A. Nasir, H. D. Tuan, T. Q. Duong, and H. V. Poor, "Improper Gaussian signaling for broadcast interference networks," *IEEE Signal Process. Lett.*, vol. 26, no. 6, pp. 808–812, Jun. 2019.

[22] H. D. Tuan, A. A. Nasir, H. H. Nguyen, T. Q. Duong, and H. V. Poor, "Non-orthogonal multiple access with improper Gaussian signaling," *IEEE J. Sel. Topics Signal Process.*, vol. 13, no. 3, pp. 496–507, Mar. 2019.

[23] M. Soleymani, C. Lameiro, I. Santamaria, and P. J. Schreier, "Energy-efficient improper signaling for K-user interference channels," in *Proc. IEEE EUSIPCO*, Sep. 2019, pp. 1–5.

[24] P. J. Schreier and L. L. Scharf, *Statistical Signal Processing of Complex-Valued Data: The Theory of Improper and Noncircular Signals*. Cambridge, U.K.: Cambridge Univ. Press, 2010.

[25] H. H. M. Tam, H. D. Tuan, and D. T. Ngo, "Successive convex quadratic programming for quality-of-service management in full-duplex MU-MIMO multicell networks," *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2340–2353, Jun. 2016.

[26] Z. Chen and E. Björnson, "Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5205–5219, Nov. 2018.

[27] H. Yu, H. D. Tuan, E. Dutkiewicz, H. V. Poor, and L. Hanzo, "Maximizing the geometric mean of user-rates to improve rate-fairness: Proper vs. improper Gaussian signaling," *IEEE Trans. Wireless Commun.*, vol. 21, no. 1, pp. 295–309, Jan. 2022.

[28] M. Bashar *et al.*, "Exploiting deep learning in limited-fronthaul cell-free massive MIMO uplink," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 8, pp. 1678–1697, Aug. 2020.

[29] R. Fletcher, *Practical Methods of Optimization*, 2nd ed. Hoboken, NJ, USA: Wiley, 1997.

[30] M. Farooq, H. Q. Ngo, E.-K. Hong, and L.-N. Tran, "Utility maximization for large-scale cell-free massive MIMO downlink," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 7050–7062, Oct. 2021.

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

[32] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, vol. 14, no. 6, pp. 3059–3075, Jun. 2015.

[33] S. Buzzi, I. Chih-Lin, T. E. Klein, H. V. Poor, C. Yang, and A. Zappone, "A survey of energy-efficient techniques for 5G networks and challenges ahead," *IEEE J. Sel. Areas Commun.*, vol. 34, no. 4, pp. 697–709, Apr. 2016.

[34] A. Zappone, L. Sanguinetti, G. Bacci, E. Jorswieck, and M. Debbah, "Energy-efficient power control: A look at 5G wireless technologies," *IEEE Trans. Signal Process.*, vol. 64, no. 7, pp. 1668–1683, Apr. 2016.

[35] G. Interdonato, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Downlink training in cell-free massive MIMO: A blessing in disguise," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5153–5169, Nov. 2019.

[36] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Hoboken, NJ, USA: Wiley, 2006.

[37] A. Tang, J. Sun, and K. Gong, "Mobile propagation loss with a low base station antenna for NLOS street microcells in urban area," in *Proc. IEEE VTS 53rd Vehicular Technol. Conf.* , vol. 1, May 2001, pp. 333–336.

[38] D. Peaucelle, D. Henrion, and Y. Labit. (2002). *Users Guide for SeDuMi Interface 1.03*. [Online]. Available: http://homepages.laas.fr/peaucell/software/sdmguide.pdf