Energy Efficiency Optimization In Cell-free Massive MIMO With Normalized Conjugate Beamforming

Bin Yan, Ningxin Zhou, Zheng Wang
School of Information Science and Engineering
Southeast University, Nanjing, China
Email: bin_yan@seu.edu.cn, ningxin_zhou@seu.edu.cn, wznuaa@gmail.com

***Abstract*—In this paper, we propose a power allocation scheme for optimizing the energy efficiency (EE) in cell-free multiple-input and multiple-output (MIMO) downlink systems using normalized conjugate beamforming (NCB) precoding. Specifically, the NCB precoding-based EE optimization problem under short-term power constraints is firstly formulated. To address its non-convex nature, a power allocation scheme is proposed, which employs the sequential convex approximation (SCA) to transform it into a series of accessible second-order cone programming (SOCP) problems, thereby enabling an iterative solution process. Moreover, an access point (AP) selection scheme based on K-means++ is also designed, which further improves EE by eliminating the pilot contamination. Simulation result confirms that the proposed power allocation scheme converges rapidly, and attains an improved EE compared to the existing NCB precoding-based power allocation schemes.**

***Index Terms*—Cell-free massive MIMO, energy efficiency, normalized conjugate beamforming, power allocation, AP selection.**

I. Introduction

The cell-free massive multiple-input and multiple-output (MIMO) has become a promising network architecture in beyond fifth-generation (B5G) and sixth-generation (6G), which embraces the user-centric idea to eliminate the concept of cell boundaries [1]. As one of the key technologies in cell free massive MIMO, power allocation allocates the downlink transmit power between users and access points (APs) [2], given the fact that the AP is subject to long-term power constraints (LTPCs) or short-term power constraints (STPCs) [3]. With the network speed rapidly rising, hardware resource processing has brought the challenging issues about energy consumption and environmental pollution, thereby rendering energy efficiency (EE) become an important criterion in power allocation [4]. As a result, this leads to the EE optimization problem in cell-free massive MIMO systems.

Specifically, in [4], under conjugate beamforming (CB) precoding, the EE optimization problem is solved by sequential convex approximation (SCA), where two AP selection schemes based on received-power-based (RPB) and largest-large-scale-fading-based (LLSFB) are also proposed to reduce the power consumption caused by the backhaul links. Although CB precoding has low complexity and backhaul requirements, it suffers from the high inter-user interference.

Corresponding author: Zheng Wang (e-mail: wznuaa@gmail.com)

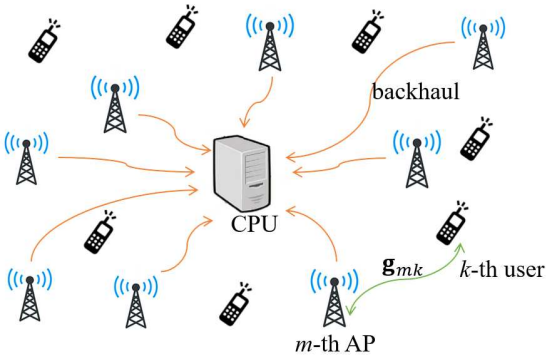

Fig. 1. Illustration of the cell free massive MIMO downlink systems.

To further improve the performance, zero-forcing (ZF) precoding is employed in addressing the EE optimization problem [5]. Unfortunately, ZF precoding includes complex matrix inversion operations that grow cubically with the increasing user numbers, rendering it infeasible in practice.

Different from CB and ZF, normalized conjugate beamforming (NCB) precoding is subject to STPCs rather than LTPCs, making it very appealing in practical implementation [3]. On one hand, compared to CB, NCB exploits the *channel hardening* property and reduces the beamforming uncertainty, consequently leading to a higher achievable rate [3]. On the other hand, NCB avoids the complex matrix inversion operation with guaranteed minimum rate user service quality as the network load increases, which results in lower complexity and stronger robustness than ZF [6]. Therefore, in this paper, the NCB precoding-based EE optimization problem is considered under constraints of per-AP STPCs, followed by a novel AP selection scheme for the further improvement of EE.

II. System model

As shown in Fig. 1, consider a cell-free massive MIMO downlink system with M multi-antenna APs providing service to K single-antenna user in the same time-frequency resources, each AP is equipped with N antennas. With the Rayleigh fading scenario, $\mathbf{g}_{mk} = \sqrt{\beta_{mk}}\mathbf{h}_{mk} \in \mathbb{C}^{N\times 1}$ reflects the channel coefficient between the k-th user and the m-th AP, β_{mk} is the large-scale fading coefficient, and $\mathbf{h}_{mk} \in \mathbb{C}^{N\times 1}$ represents the small-scale fading vector. All APs are connected to the central processing unit (CPU) through backhaul links, and the system operates in time division duplex (TDD).

A. Uplink Training

During the uplink training phase, all K users simultaneously transmit pilot sequences to APs. Let τ_u be the length of coherence interval slot for the uplink training (in samples). The pilot signal received at the m-th AP is

$$\mathbf{y}_m = \sqrt{\rho_u \tau_u} \sum_{k=1}^{K} \mathbf{g}_{mk} \boldsymbol{\psi}_k^H + \mathbf{n}_m. \tag{1}$$

Here, ρ_u is the normalized uplink transmit signal-to-noise ratio (SNR), $\boldsymbol{\psi}_k \in \mathbb{C}^{\tau_u \times 1}$ is the pilot sequence transmitted from the k-th user, $\mathbf{n}_m \in \mathbb{C}^{N \times \tau_u}$ is a Gaussian noise matrix whose elements are $\mathcal{CN}(0,1)$. Under the minimum mean square error (MMSE) criterion, the variance of the channel estimate α_{mk} between the k-th user and the m-th AP is given by [1]

$$\alpha_{mk} \triangleq \frac{\rho_u \tau_u \beta_{mk}^2}{1 + \rho_u \tau_u \sum_{k'=1}^{K} \beta_{mk'} |\boldsymbol{\psi}_{k'}^H \boldsymbol{\psi}_k|^2}. \tag{2}$$

When the coherence interval is shorter than the number of users, i.e $\tau_u < K$, some users will use the same pilot so that APs are not able to spatially separate the linearly dependent channels. This is known as pilot contamination phenomenon.

B. Downlink Payload Data Transmission

After obtaining the estimated channel state information (CSI), the m-th AP transmits the signal

$$\mathbf{x}_m = \sqrt{\rho_d} \sum_{k=1}^{K} \sqrt{\eta_{mk}} \frac{\widehat{\mathbf{g}}_{mk}^*}{\|\widehat{\mathbf{g}}_{mk}\|_2} s_k. \tag{3}$$

Here s_k is the symbol intended for the k-th user with $\mathbb{E}\{\|s_k\|^2\} = 1$, $\widehat{\mathbf{g}}_{mk} \sim \mathcal{CN}(0, \alpha_{mk})$ is the channel estimation of $\mathbf{g}_{mk}$, ρ_d represents the maximum downlink normalized transmit power at each AP, η_{mk} is the power allocation coefficient assigned to the link between the k-th user and the m-th AP. When using NCB precoding, each AP needs to satisfy the following STPCs [3]

$$\mathbb{E}\{\|\mathbf{x}_m\|_2^2\} = \rho_d \sum_{k=1}^{K} \eta_{mk} \leq \rho_d, \forall m. \tag{4}$$

The signal-to-interference-plus-noise ratio (SINR) of the k-th user using NCB precoding is shown in (5) with

$$\Gamma_N \triangleq \frac{\Gamma(N + 1/2)}{\Gamma(N)} \tag{6}$$

and

$$\gamma_{kk'} \triangleq (N-1) \sum_{m=1}^{M} \eta_{mk'} \alpha_{mk'} \frac{\beta_{mk}^2}{\beta_{mk'}^2} + \Gamma_N^2 \sum_{m=1}^{M} \sum_{n \neq m}^{M} \sqrt{\eta_{mk'} \eta_{nk'} \alpha_{mk'} \alpha_{nk'}} \frac{\beta_{mk} \beta_{nk}}{\beta_{mk'} \beta_{nk'}}, \tag{7}$$

where $\Gamma(\cdot)$ denotes the *Gamma* function [7]. Then, based on $SINR_k(\{\eta_{mk}\})$ in (5), the spectral efficiency (SE) of the k-th user using NCB precoding can be obtained based on *Shannon's theorem*, i.e.,

$$S_{ek}(\{\eta_{mk}\}) = \log_2 \left(1 + SINR_k \left(\{\eta_{mk}\}\right)\right). \tag{8}$$

C. Downlink Energy Efficiency

In a cell-free massive MIMO downlink systems, the total EE (in bit/Joule) is defined as [4]

$$E_e(\{\eta_{mk}\}) = \frac{BS_e(\{\eta_{mk}\})}{P_{\text{total}}} = \frac{B \sum_{k=1}^{K} S_{ek}(\{\eta_{mk}\})}{P_{\text{total}}} \tag{9}$$

with the power consumption of the downlink transmission

$$P_{\text{total}} = \rho_d N_0 \sum_{m=1}^{M} \frac{1}{\mu_m} \left(\sum_{k=1}^{K} \eta_{mk} \right) + P_{fix} + BS_e(\{\eta_{mk}\}) \sum_{m=1}^{M} P_{\text{bt},m}. \tag{10}$$

Here, B represents the system bandwidth, $S_e(\{\eta_{mk}\})$ is the sum SE, $0 < \mu_m \leq 1$ represents the power amplifier efficiency, N_0 denotes the noise power, P_{fix} is an element independent of $\{\eta_{mk}\}$ and is determined by the number of APs and the circuit, $P_{\text{bt},m}$ denotes the traffic-dependent power consumption.

To ensure the reliable communication for all users, the SE threshold S_{ok} is set, and each AP should satisfy STPCs in (4). In particular, the NCB precoding-based EE optimization problem aims at finding the optimal power allocation coefficients $\{\eta_{mk}\}$ to maximize the system EE in (9) [4]. Mathematically, it can be formulated as

$$\max_{\boldsymbol{\eta}} \quad E_e(\{\eta_{mk}\}), \tag{11a}$$

$$\text{s.t.} \quad S_{ek}(\{\eta_{mk}\}) \geq S_{ok}, \forall k, \tag{11b}$$

$$\rho_d \sum_{k=1}^{K} \eta_{mk} \leq \rho_d, \forall m, \tag{11c}$$

$$\eta_{mk} \geq 0, \forall m, \forall k, \tag{11d}$$

which is further equivalent to [4]

$$\max_{\boldsymbol{\eta}} \quad \frac{BS_e(\{\eta_{mk}\})}{\rho_d N_0 \sum_{m=1}^{M} \frac{1}{\mu_m} \left(\sum_{k=1}^{K} \eta_{mk} \right) + P_{fix}}, \tag{12a}$$

$$\text{s.t.} \quad \text{(11b), (11c), (11d).} \tag{12b}$$

III. Power allocation scheme for NCB precoding-based EE optimization problem

Since the NCB precoding-based EE optimization problem (12) is neither convex nor concave, we try to propose a power allocation scheme that employs SCA for iterative problem-solving in this section.

Firstly, note that the dual summation in (7) renders the objective function complex and intractable. Therefore, the

$$SINR_k(\{\eta_{mk}\}) = \frac{\rho_d \Gamma_N^2 \left(\sum_{m=1}^{M} \sqrt{\eta_{mk} \alpha_{mk}} \right)^2}{1 + \rho_d \sum_{k'=1}^{K} \sum_{m=1}^{M} \eta_{mk'} \beta_{mk} + (N - 1 - \Gamma_N^2) \rho_d \sum_{m=1}^{M} \eta_{mk} \alpha_{mk} + \rho_d \sum_{k' \neq k}^{K} \gamma_{kk'} |\boldsymbol{\psi}_k^H \boldsymbol{\psi}_{k'}|^2}, \tag{5}$$

$$SINR_k(\{c_{mk}\}) = \frac{\Gamma_N^2(\boldsymbol{\alpha}_k^T\mathbf{c}_k)^2}{\frac{1}{\rho_d}+\sum_{k'=1}^K\|\boldsymbol{\beta}_k\mathbf{c}_{k'}\|_2^2+(N-1-\Gamma_N^2)\|\boldsymbol{\alpha}_k\cdot\mathbf{c}_k\|_2^2+\sum_{k'\neq k}^K\left((N-1)\|\boldsymbol{\xi}_{kk'}\cdot\mathbf{c}_{k'}\|_2^2+\Gamma_N^2\left(\boldsymbol{\xi}_{kk'}^T\mathbf{c}_{k'}\right)^2\right)\zeta_{kk'}}. \quad (14)$$

following first-order approximation is taken to deal with the user interference caused by the pilot contamination in the NCB precoding scheme [8]

$$\gamma_{kk'}\approx(N-1)\sum_{m=1}^M\eta_{mk'}\alpha_{mk'}\frac{\beta_{mk}^2}{\beta_{mk'}^2}+\Gamma_N^2\left(\sum_{m=1}^M\sqrt{\eta_{mk'}\alpha_{mk'}}\frac{\beta_{mk}}{\beta_{mk'}}\right)^2. \quad (13)$$

Here, in order to promote the quadratic convex transformation, we denote $\{\eta_{mk}\}$ in (12) as $\{c_{mk}^2\}$, i.e. $c_{mk}\triangleq\sqrt{\eta_{mk}}$. Then, based on the following definitions:

$$\begin{aligned}\mathbf{c}_k&\triangleq[c_{1k},\cdots,c_{mk}]^T\in\mathbb{R}^M,\zeta_{kk'}\triangleq|\boldsymbol{\psi}_{k'}^H\boldsymbol{\psi}_k|^2\in\mathbb{R},\\ \boldsymbol{\alpha}_k&\triangleq[\sqrt{\alpha_{1k}},\cdots,\sqrt{\alpha_{Mk}}]^T\in\mathbb{R}^M,\\ \boldsymbol{\beta}_k&\triangleq diag(\sqrt{\beta_{1k}},\cdots,\sqrt{\beta_{Mk}})\in\mathbb{R}^{M\times M},\\ \boldsymbol{\xi}_{kk'}&\triangleq[\sqrt{\alpha_{1k'}}\frac{\beta_{1k}}{\beta_{1k'}},\ldots,\sqrt{\alpha_{Mk'}}\frac{\beta_{Mk}}{\beta_{Mk'}}]^T\in\mathbb{R}^M,\\ P_{abs}&\triangleq\rho_dN_0\sum_{m=1}^M\frac{1}{\mu_m}\left(\sum_{k=1}^Kc_{mk}^2\right)+P_{fix}\in\mathbb{R},\end{aligned}$$

we can characterize the user $SINR_k(\{\eta_{mk}\})$ in (5) as a function of $\{c_{mk}\}$, resulting in a more refined formulation shown in (14). For notational simplicity, here we short the expression $SINR_k(\{c_{mk}\})$ as $SINR_k$. Then, by introducing the auxiliary variables $\{t_k\}$, we can equivalently express the original problem (12) by a more tractable one [4], which gives rise to

$$\begin{aligned}\max_{\mathbf{c},\mathbf{t}}\quad&\frac{B\sum_{k=1}^Kt_k}{\ln 2},&(15a)\\ \text{s.t.}\quad&\frac{\ln(1+SINR_k)}{P_{abs}}\geq t_k,\forall k,&(15b)\\ &S_{ek}(\{c_{mk}\})\geq S_{ok},\forall k,&(15c)\\ &\sum_{k=1}^Kc_{mk}^2\leq 1,\forall m,&(15d)\\ &c_{mk}\geq 0,\forall m,\forall k.&(15e)\end{aligned}$$

On one hand, with respect to the constraint in (15c), it can be expressed by a second-order cone (SOC) constraint for a given SE threshold S_{ok}, which is shown in (16).

On the other hand, as for non-convex constraint in (15b), we invoke the following inequality [9]

$$\frac{\ln(1+x)}{t}\geq a-\frac{b}{x}-d\cdot t,\quad\forall x>0,t>0, \quad (17)$$

where $a\triangleq 2\frac{\ln(1+\hat{x})}{\hat{t}}+\frac{\hat{x}}{\hat{t}(\hat{x}+1)}>0, b\triangleq\frac{\hat{x}^2}{\hat{t}(\hat{x}+1)}>0, d\triangleq\frac{\ln(1+\hat{x})}{\hat{t}^2}>0,\forall\hat{x}>0,\hat{t}>0$. Then, based on it, by letting $\{c_{mk}^n\}$ be the feasible solution obtained from the n-th iteration, we can obtain the corresponding $SINR_k^n$ and P_{abs}^n. Subsequently, the logarithmic numerator term $\ln(1+SINR_k)$ with variables $\{c_{mk}\}$ can be seamlessly separated from the quadratic denominator term P_{abs} through (17), i.e.,

$$\begin{aligned}&\frac{\ln(1+SINR_k)}{P_{abs}}\geq a_k^n-d_k^nP_{abs}-\frac{b_k^n}{\rho_d\Gamma_N^2({\boldsymbol{\alpha}_k}^T\mathbf{c}_k)^2}-\\ &\frac{\sum_{k'=1}^K\|\boldsymbol{\beta}_k\mathbf{c}_{k'}\|_2^2+(N-1-\Gamma_N^2)\|\boldsymbol{\alpha}_k\cdot\mathbf{c}_k\|_2^2}{\Gamma_N^2({\boldsymbol{\alpha}_k}^T\mathbf{c}_k)^2(b_k^n)^{-1}}-\\ &\frac{\sum_{k'\neq k}^K\left((N-1)\|\boldsymbol{\xi}_{kk'}\cdot\mathbf{c}_{k'}\|_2^2+\Gamma_N^2\left(\boldsymbol{\xi}_{kk'}^T\mathbf{c}_{k'}\right)^2\right)\zeta_{kk'}}{\Gamma_N^2({\boldsymbol{\alpha}_k}^T\mathbf{c}_k)^2(b_k^n)^{-1}}\end{aligned} \quad (18)$$

with

$$a_k^n=2\frac{\ln(1+SINR_k^n)}{P_{abs}^n}+\frac{SINR_k^n}{P_{abs}^n(SINR_k^n+1)}, \quad (19a)$$

$$b_k^n=\frac{(SINR_k^n)^2}{P_{abs}^n(SINR_k^n+1)}, \quad (19b)$$

$$d_k^n=\frac{\ln(1+SINR_k^n)}{(P_{abs}^n)^2}. \quad (19c)$$

Obviously, the right-hand-side (RHS) of (18) is still a non-convex function with respect to the independent variable $\{c_{mk}\}$. For this reason, another convex inequality in the following is introduced [5]

$$\frac{x^2}{t}\geq 2\frac{\hat{x}x}{\hat{t}}-\frac{\hat{x}^2}{\hat{t}^2}t,\quad\forall x>0,\hat{x}>0,t>0,\hat{t}>0. \quad (20)$$

By (20), the quadratic term of $\|\boldsymbol{\beta}_k\mathbf{c}_{k'}\|_2^2$, $\|\boldsymbol{\alpha}_k\cdot\mathbf{c}_k\|_2^2$, $\|\boldsymbol{\xi}_{kk'}\cdot\mathbf{c}_{k'}\|_2^2$ and $\left(\boldsymbol{\xi}_{kk'}^T\mathbf{c}_{k'}\right)^2$ will become the first-order term of $(\|\boldsymbol{\beta}_k\mathbf{c}_{k'}^n\|_2\|\boldsymbol{\beta}_k\mathbf{c}_{k'}\|_2)$, $(\|\boldsymbol{\alpha}_k\cdot\mathbf{c}_k^n\|_2\|\boldsymbol{\alpha}_k\cdot\mathbf{c}_k\|_2)$, $(\|\boldsymbol{\xi}_{kk'}\cdot\mathbf{c}_{k'}^n\|_2\|\boldsymbol{\xi}_{kk'}\cdot\mathbf{c}_{k'}\|_2)$ and $\left(\left(\boldsymbol{\xi}_{kk'}^T\mathbf{c}_{k'}^n\right)\left(\boldsymbol{\xi}_{kk'}^T\mathbf{c}_{k'}\right)\right)$ respectively. Meanwhile, for the quadratic terms $(\boldsymbol{\alpha}_k^T\mathbf{c}_k)^2$, we then recall another convex inequality [9]

$$x^2\geq 2\hat{x}x-\hat{x}^2,\quad\forall x\geq 0,\hat{x}\geq 0,2x\geq\hat{x}, \quad (21)$$

and we can obtain

$$\begin{aligned}(\boldsymbol{\alpha}_k^T\mathbf{c}_k)^2&\geq 2(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)(\boldsymbol{\alpha}_k^T\mathbf{c}_k)-(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^2,\forall k,&(22a)\\ 2c_{mk}&\geq c_{mk}^n,\quad\forall m,\forall k.&(22b)\end{aligned}$$

$$(\boldsymbol{\alpha}_k^T\mathbf{c}_k)^2\geq(2^{S_{ok}}-1)\left(\frac{1}{\rho_d\Gamma_N^2}+\frac{\sum_{k'=1}^K\|\boldsymbol{\beta}_k\mathbf{c}_{k'}\|_2^2}{\Gamma_N^2}+\frac{(N-1-\Gamma_N^2)}{\Gamma_N^2}\|\boldsymbol{\alpha}_k\cdot\mathbf{c}_k\|_2^2+\sum_{k'\neq k}^K\left(\boldsymbol{\xi}_{kk'}^T\mathbf{c}_{k'}\right)^2\zeta_{kk'}+\frac{(N-1)}{\Gamma_N^2}\sum_{k'\neq k}^K\|\boldsymbol{\xi}_{kk'}\cdot\mathbf{c}_{k'}\|_2^2\,\zeta_{kk'}\right),\forall k. \quad (16)$$

Furthermore, as shown in (22a), the terms containing the quadratic independent variables $(\boldsymbol{\alpha}_k^T\mathbf{c}_k)^2$ can also be replaced with the first-order terms $\varsigma_k \triangleq 2(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)(\boldsymbol{\alpha}_k^T\mathbf{c}_k)-(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^2 \in \mathbb{R}, \forall k$, which leads to a more computationally efficient formulation [10].

Based on the two convex inequalities (20) and (21), we have the following results to approximate the non-convex terms in the RHS of (18)

$$
\begin{aligned}
t_k \le\ & a_k^n - d_k^n P_{\text{abs}} - \frac{b_k^n}{\rho_d \Gamma_N^2 (\boldsymbol{\alpha}_k{}^T \mathbf{c}_k)^2} - \\
& \frac{b_k^n}{\Gamma_N^2}\left\{\frac{\sum_{k'=1}^K(\|\boldsymbol{\beta}_k\mathbf{c}_{k'}^n\|_2\|\boldsymbol{\beta}_k\mathbf{c}_{k'}\|_2)}{0.5(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^2} - \frac{\sum_{k'=1}^K\|\boldsymbol{\beta}_k\mathbf{c}_{k'}^n\|_2^2}{(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^4\varsigma_k^{-1}}\right\} - \\
& \frac{b_k^n(N-1-\Gamma_N^2)}{\Gamma_N^2}\left\{\frac{\|\boldsymbol{\alpha}_k\cdot\mathbf{c}_k^n\|_2\|\boldsymbol{\alpha}_k\cdot\mathbf{c}_k\|_2}{0.5(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^2} - \frac{\|\boldsymbol{\alpha}_k\cdot\mathbf{c}_k^n\|_2^2}{(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^4\varsigma_k^{-1}}\right\} \\
& - \frac{b_k^n(N-1)}{\Gamma_N^2}\left\{\frac{\sum_{k'\neq k}^K(\|\boldsymbol{\xi}_{kk'}\cdot\mathbf{c}_{k'}^n\|_2\|\boldsymbol{\xi}_{kk'}\cdot\mathbf{c}_{k'}\|_2)\zeta_{kk'}}{0.5(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^2} - \right. \\
& \left.\frac{\sum_{k'\neq k}^K\|\boldsymbol{\xi}_{kk'}\cdot\mathbf{c}_{k'}^n\|_2^2\zeta_{kk'}}{(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^4\varsigma_k^{-1}}\right\} + b_k^n\left\{\frac{\sum_{k'\neq k}^K\left(\boldsymbol{\xi}_{kk'}^T\mathbf{c}_{k'}\right)^2\zeta_{kk'}}{(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^4\varsigma_k^{-1}}\right. \\
& \left. - \frac{\sum_{k'\neq k}^K\left(\left(\boldsymbol{\xi}_{kk'}^T\mathbf{c}_{k'}^n\right)\left(\boldsymbol{\xi}_{kk'}^T\mathbf{c}_{k'}\right)\right)\zeta_{kk'}}{0.5(\boldsymbol{\alpha}_k^T\mathbf{c}_k^n)^2}\right\}, \forall k,
\end{aligned} \tag{23}
$$

which means that the non-convex constraint (15b) can be substituted by the convex approximation (23).

To sum up, the constraints (15d), (16) and (23) are SOC, and the constraint (15e) is linear matrix inequality (LMI). Based on them, since the objective function in (15) is a linear fractional function, the NCB precoding-based EE optimization problem can be reformulated as a series of accessible SOCP problems. Hence, the $(n+1)$-th iteration of the proposed power allocation scheme to solve this problem can be designed as follows

$$
\max_{\mathbf{c},\mathbf{t}} \quad \frac{B\sum_{k=1}^K t_k}{\ln 2}, \tag{24a}
$$

$$
\text{s.t.} \quad (15\text{d}), (15\text{e}), (16), (22\text{b}), (23), \tag{24b}
$$

where the computational complexity order of each iteration is $\mathcal{O}(M^{3.5}K^{3.5} + 2M^{3.5}K^{2.5} + 2M^{2.5}K^{3.5})$ [10]. On the other hand, at the beginning of the iterations, a suitable initial feasible solution can be obtained by simply solving the constrained optimization problem

$$
\max_{\mathbf{c^0}} \quad 0, \tag{25a}
$$

$$
\text{s.t.} \quad (15\text{d}), (15\text{e}), (16). \tag{25b}
$$

Finally, we follow the standard framework of SCA to analyze the convergence of the proposed power allocation scheme in (24). Suppose $\{c_{mk}^*\}$ is the optimal solution for the $(n+1)$-th iteration to (24), by relaxing convex inequalities (17), (20) and (21), solution $\{c_{mk}^*\}$ is also feasible to (15). When $c_{mk}^* = c_{mk}^n$, the equalities in (17), (20) and (21) hold, indicating that the optimal solution $\{c_{mk}^n\}$ for the n-th iteration is still valid for the $(n+1)$-th iteration. Therefore, the objective function consistently exhibits non-decreasing behavior, which ensures the convergence during the iterations.

IV. Further Improved by AP selection

In the previous section, we assumed that the connections are established between all users and APs. As the network terminal density experiences rapid growth, in the scenario where the number of network users is larger than coherent interval, i.e. $K > \tau_u$, the quality of cell-free network will significantly decline due to the pilot contamination, especially when the adjacent users being assigned the same pilot sequence [1].

Therefore, we now propose an AP selection scheme based on the K-means++ in this section. On one hand, this scheme considers clustering to enable users within each cluster to use orthogonal pilot sequences. On the other hand, it combines power allocation to eliminate the inter user interference caused by multiplexing pilots between different clusters. Specifically, the AP selection scheme consists of the following 5 steps.

- **Initialization**: Consider dividing the user set $\mathcal{K} = \{1, \cdots, K\}$ into L disjoint clusters $\mathcal{K}_1, \cdots, \mathcal{K}_L$. To ensure that the users within each cluster can be assigned with orthogonal pilot sequences, the size of each cluster should satisfy $|\mathcal{K}_l| \le \tau_u, \forall l$. Hence, we set the number of clusters as $L = \lceil \frac{K}{\tau_u} \rceil$ and invoke the K-means++ to initialize the cluster centroid position $\boldsymbol{\kappa}_1^0, \ldots, \boldsymbol{\kappa}_L^0 \in \mathbb{R}^M$.
- **User clustering**: Adopt large-scale fading coefficients $\boldsymbol{\beta}_k \triangleq [\beta_{1k}, \ldots, \beta_{Mk}]^T \in \mathbb{R}^M$ as the k-th user feature vectors. The *Euclidean distance* between each user and the clusters centroid is taken as the measurement

$$
d_{lk} = \|\boldsymbol{\beta}_k - \boldsymbol{\kappa}_l^i\|_2^2,\ \forall l, \forall k, \tag{26}
$$

  based on the obtained distance at the i-th iteration, where the following criteria is used for user clustering [11]

$$
\mathcal{K}_l = \{k\text{-th user belongs to cluster } l \,|\text{if } d_{lk} < d_{jk},\ \forall j \neq l|\}. \tag{27}
$$

- **Centroid position update**: Update the centroid position of each cluster according to

$$
\boldsymbol{\kappa}_l^{i+1} = \frac{\sum_{k\in\mathcal{K}_l^i}\boldsymbol{\beta}_k}{|\mathcal{K}_l^i|},\ \forall l. \tag{28}
$$

  Once all the centroid positions stop updating, i.e. $\boldsymbol{\kappa}_l^{i+1} = \boldsymbol{\kappa}_l^i, \forall l$, move to the next step. Otherwise, continue iterating the user clustering process.
- **Modify cluster size**: If all cluster sizes obey the requirement $|\mathcal{K}_l| \le \tau_u, \forall l$, proceed to the next step. Otherwise, we need to take some modified measures. Specifically, for clusters that exceed the size limit, i.e. $|\mathcal{K}_{l'}| > \tau_u$, we record the number of clusters that exceed the limit as L' and calculate the variance between all the users and the centroid

$$
\sigma_{l'k} = \frac{\|\boldsymbol{\beta}_k - \boldsymbol{\kappa}_{l'}^i\|_2^2}{M},\ \forall k, \tag{29}
$$

  then order the variance in descending way and remove the corresponding users from cluster $\mathcal{K}_{l'}$, until $|\mathcal{K}_{l'}| \le \tau_u$. Meanwhile, we fix the clusters $\mathcal{K}_{l'}$ after the modified completion. After that, the removed users and the clusters that remain unmodified will be regrouped for another

Algorithm 1 The Proposed Power Allocation Scheme For EE Optimization With NCB

Input: $S_{ok}, \rho_d, N, \{\alpha_{mk}\}, \{\beta_{mk}\}, N_I$
Output: power allocation coefficients $\{\eta_{mk}\} = \{c_{mk}^2\}$
Step 1: perform AP selection, go to Step 2; without AP selection, go to Step 4
Step 2: perform AP selection scheme based on the K-means++ to obtain the connectivity matrix $\mathbf{X}$
Step 3: if $\mathbf{X}_{mk} = 1$, let $\widehat{\alpha}_{mk} = \alpha_{mk}$; else $\widehat{\alpha}_{mk} = 0$, $\forall m, \forall k$. Replace $\{\alpha_{mk}\}$ with $\{\widehat{\alpha}_{mk}\}$ as Step 4 input, proceed to the next step
Step 4: obtain an initial feasible solution $\mathbf{c}^0$ by solving (25), set $n = 1$
Step 5: perform the n-th iteration: solving problem (24) by using SOCP solver, obtain optimal solution $\mathbf{c}^*$
Step 6: when $n = N_I$, terminate the algorithm; else go to Step 7
Step 7: update $\mathbf{c}^n = \mathbf{c}^*, n = n + 1$, go to Step 5

iteration of user clustering. Significantly, not only the user participation in clustering is reduced, but there is also a reduction in the number of clusters, i.e. $L = (L - L')$.

- **AP selection**: Select the serving APs for each cluster based on the final location of the cluster centroids matrix $\boldsymbol{\kappa}^f \in \mathbb{R}^{M \times L}$

$$\mathcal{A}_l = \{m\text{-th AP serves cluster } l \ | \text{if } \kappa_{ml}^f > \kappa_{mj}^f, \ \forall j \neq l|\}, \tag{30}$$

based on the obtained $\mathcal{A}_l$ and $\mathcal{K}_l$, we can establish the unique connectivity between APs and users. Define connectivity matrix $\mathbf{X} \in \mathbb{Z}^{M \times K}$, where $\mathbf{X}_{mk} \in \{0, 1\}$. $\mathbf{X}_{mk} = 0$ means that the m-th AP does not serve the k-th user, and $\eta_{mk} = 0$. If and only if m-th AP $\in \mathcal{A}_l$ and k-th user $\in \mathcal{K}_l$, it has $\mathbf{X}_{mk} = 1$.

Then, we analyze the impact of the proposed AP selection scheme with respect to EE in the downlink cell-free massive MIMO systems. Looking back at (5), when the k-th user and the k'-th user belong to the same cluster, the orthogonal pilot sequences will be assigned, i.e $|\boldsymbol{\psi}_k^H \boldsymbol{\psi}_{k'}|^2 = 0$. Alternatively, if the k-th user and the k'-th user belong to different clusters, they might potentially reuse pilot sequences. Nevertheless, our AP selection scheme guarantees that the m-th AP cannot simultaneously serve two clusters. For example, if $\eta_{mk} \neq 0$, then there must have $\eta_{mk'} = 0$, which leads to $\gamma_{kk'} = 0$. In conclusion, the interference caused by pilot contamination will be completely eliminated, i.e. $\sum_{k' \neq k}^{K} \gamma_{kk'} |\boldsymbol{\psi}_k^H \boldsymbol{\psi}_{k'}|^2 = 0$. Additionally, AP selection also reduces the backhaul power consumption, which helps to further improve the EE [4].

In a nutshell, after obtaining the connectivity matrix $\mathbf{X}$, channel estimation are reassigned as the input to EE optimization, based on the duality of $\{\eta_{mk}\}$ and $\{\alpha_{mk}\}$ [4]. To summarize, the proposed power allocation scheme for solving the NCB precoding-based EE optimization problem in cell-free massive MIMO downlink systems is outlined in Algorithm 1.

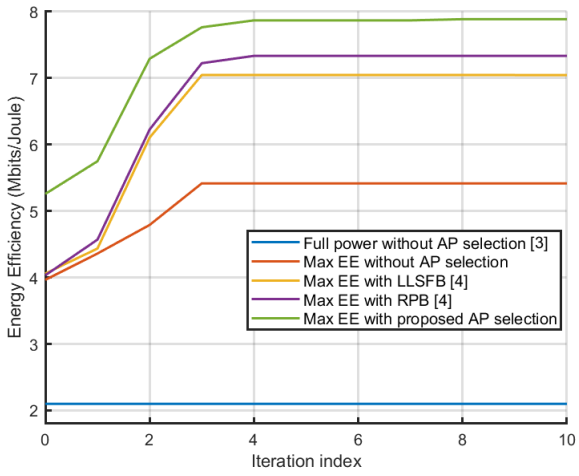

Fig. 2. Illustration of the energy efficiency versus the number of iterations ($M = 100, N = 1, K = 20, \tau_u = 5, D = 1$).

V. Simulation Results

In this section, we consider a $D \times D$ Km^2 cell-free area, where wrapped-around technique is used to imitate an infinite network. In particular, the large-scale fading coefficient follows the model in [1]

$$\beta_{mk} = \text{PL}_{mk} \cdot 10^{\frac{\sigma_{sh} z_{mk}}{10}}, \tag{31}$$

where PL_{mk} represents the path loss following the three-slope model, $10^{\frac{\sigma_{sh} z_{mk}}{10}}$ represents the shadow fading coefficient with the standard deviation σ_{sh}, and $z_{mk} \sim \mathcal{N}(0, 1)$. Other parameters in our system follow the same settings as in [4].

First of all, we verify the convergence of the proposed power allocation scheme, where MOSEK is employed as the solver for large-scale convex optimization problems. Meanwhile, we also apply the full power transmission scheme as a comparative benchmark. As shown in Fig. 2, the proposed power allocation scheme facilitates the rapid iterative growth in EE and achieve stability after a few iterations. Furthermore, as expected, in combination with the AP selection scheme, significant enhancements are achieved.

Fig. 3 compares the proposed scheme with other power allocation schemes in terms of the EE performance. Obviously, the proposed scheme significantly improves EE compared to other power allocation schemes. Specifically, compared to full power transmission, the EE performance is increased by 2 to 3 times. Meanwhile, NCB precoding benefits from STPCs and has some improvements in EE compared to CB precoding. In addition, the EE performance tends to decrease as the number of APs increases, which is caused by the energy consumed due to the increase in backhaul links.

In Fig. 4, we investigate the impact of the proposed AP selection scheme on EE under short coherence interval. The proposed AP selection scheme not only mitigates the reduced throughput caused by pilot contamination but also minimizes backhaul links power consumption [4], thus

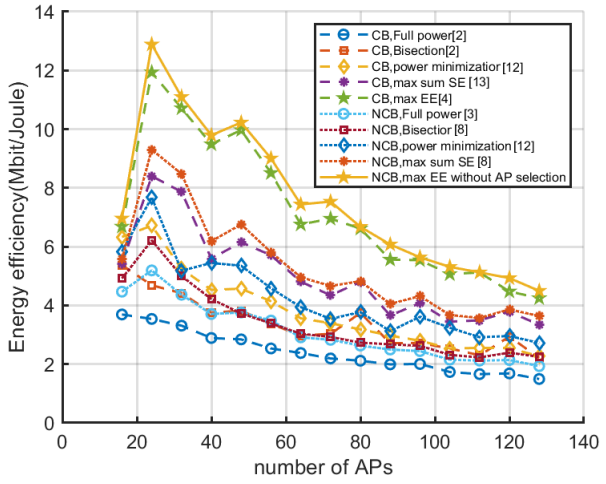

Fig. 3. Illustration of downlink energy efficiency versus the number of APs, comparison of proposed scheme with other power allocation schemes ($N = 2, K = 16, \tau_u = 16, D = 1$).

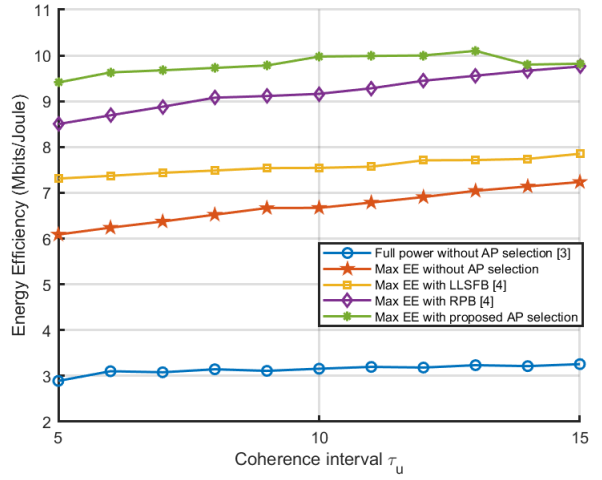

Fig. 4. Illustration of the downlink energy efficiency versus the coherence interval ($M = 100, N = 1, K = 40, D = 1$).

resulting in an enhanced EE at the cost of the increased complexity. Specifically, LLSFB, RPB, and the proposed AP selection scheme have computational complexities $\mathcal{O}(KM(1+\log_2 M), \mathcal{O}(KM(1+N+\log_2 M), \mathcal{O}((KM\lceil\frac{K}{\tau_u}\rceil)I+M\lceil\frac{K}{\tau_u}\rceil+M\log_2 M+(K-1)\lceil\frac{K}{\tau_u}\rceil)$ respectively, where I is the number of iterations in which cluster size modifications are performed. It is worth noting that, the RPB AP selection needs to firstly execute the optimization algorithm, which requires extra $\mathcal{O}(M^{3.5}K^{3.5}+2M^{3.5}K^{2.5}+2M^{2.5}K^{3.5})$ computational complexity. Significantly, as the number of coherence interval increases, the performance of proposed AP selection scheme may be degraded due to the reduction in the number of clusters, which results in users being served by redundant APs. When $K \leq \tau_u$, the proposed AP selection scheme will degenerate into the fully connected model.

VI. Conclusion

In this paper, a novel power allocation scheme for NCB precoding-based EE optimization is proposed. It converts the non-convex problem into a series of SOCP problems by SCA, so that it can be iteratively solved with polynomial complexity. Furthermore, an AP selection scheme based on K-means++ is designed to eliminate the pilot contamination and further improve EE. The simulation results show that the proposed power allocation scheme significantly improves the system EE compared to the other traditional power allocation schemes.

Acknowledgment

This work was supported in part by National Natural Science Foundation of China under Grants No. 62371124, in part by ZTE Corporation under Research Program 2023ZTE01-04.

References

[1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Ma rzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834C1850, Mar. 2017.

[2] 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. 4445C4459, Jul. 2017.

[3] G. Interdonato, H. Q. Ngo, E. G. Larsson and P. Frenger, "On the performance of cell-free massive MIMO with short-term power constraints," *2016 IEEE 21st International Workshop on Computer Aided Modelling and Design of Communication Links and Networks (CAMAD)*, Toronto, ON, Canada, 2016, pp. 225-230.

[4] 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. Netw.*, vol. 2, no. 1, pp. 25-39, Mar. 2018.

[5] 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 Communications Letters*, vol. 21, no. 8, pp. 1871-1874, Aug. 2017.

[6] F. Riera-Palou and G. Femenias, "Trade-offs in cell-free massive MIMO networks: Precoding power allocation and scheduling," *14th IEEE TELSIKS conference*, pp. 158-165, 2019.

[7] A. Polegre, F. Riera-Palou, G. Femenias, and A. G. Armada, "New insights on channel hardening in cell-free massive MIMO networks," *IEEE International Conference on Communications Work shops (ICC Workshops)*, 2020, pp. 1C7

[8] Y. Zhang, H. Cao, Y. Guo and L. Yang, "SCA Power optimization in Cell-Free Massive MIMO with Short-Term Power Constraints," *2018 10th International Conference on Wireless Communications and Signal Processing (WCSP)*, Hangzhou, China, 2018, pp. 1-6.

[9] H. Tuy, *Convex Analysis and Global Optimization*, 2nd ed. New York, NY, USA: Springer, 2016.

[10] K. -Y. Wang, A. M. -C. So, T. -H. Chang, W. -K. Ma and C. -Y. Chi, "Outage Constrained Robust Transmit Optimization for Multiuser MISO Downlinks: Tractable Approximations by Conic Optimization," *IEEE Transactions on Signal Processing*, vol. 62, no. 21, pp. 5690-5705, Nov.1, 2014.

[11] G. Femenias, N. Lassoued and F. Riera-Palou, "Access Point Switch ON/OFF Strategies for Green Cell-Free Massive MIMO Networking," *IEEE Access*, vol. 8, pp. 21788-21803, 2020.

[12] Y. Zhang, H. Cao, M. Zhou, Y. Li and L. Yang, "Power Minimization for Cell-Free Massive MIMO," *IEEE International Conference on Consumer Electronics - Taiwan (ICCE-TW)*, Yilan, Taiwan, 2019, pp. 1-2

[13] S. Chakraborty, Ö. T. Demir, E. Björnson, and P. Giselsson, "Efficient downlink power allocation algorithms for cell-free massive MIMO systems," *IEEE Open J. Commun. Soc.*, vol. 2, pp. 168C186, 2021.