

NOMA/OMA Mode Selection-Based Cell-Free Massive MIMO

Manijeh Bashar^{*§}, Kanapathippillai Cumanan^{*}, Alister G. Burr^{*}, Hien Quoc Ngo[†], Lajos Hanzo[‡], and Pei Xiao[§]

^{*}University of York, UK, [†]Queen's University Belfast, UK, [‡]University of Southampton, UK,

[§]Home of the 5G Innovation Centre, Institute for Communication Systems, University of Surrey, UK.

Email: {mb1465, kanapathippillai.cumanan, alister.burr}@york.ac.uk, hien.ngo@qub.ac.uk, lh@ecs.soton.ac.uk, p.xiao@surrey.ac.uk

Abstract—In this paper, non-orthogonal-multiple-access (NOMA)-based cell-free massive multiple-input multiple-output (MIMO) is investigated, where the users are grouped into multiple clusters. Exploiting conjugate beamforming, the bandwidth efficiency (BE) of the system is derived while the assumption that the users performing realistic successive interference cancellation (SIC) based on only the knowledge of channel statistics. The max-min fairness problem of maximizing the lowest user BE is investigated and an iterative bisection method is developed to determine the optimal solution to the max-min BE problem. Numerical results are presented for validating the proposed design's performance, and a mode switching scheme is conceived for selecting a specific Mode = {OMA, NOMA} that maximizes the system's BE.

Keywords: Cell-free massive MIMO, convex optimization, max-min bandwidth efficiency, NOMA.

I. INTRODUCTION

In recent years, sophisticated disruptive technologies have been conceived to meet the unprecedented bandwidth efficiency (BE) versus energy efficiency requirements of next generation wireless networks. Cell-free massive multiple-input multiple-output (MIMO) [1]–[4] and non-orthogonal-multiple-access (NOMA) [5]–[15] have been identified as two of the key elements of future wireless networks. In conventional orthogonal multiple access (OMA), orthogonal time, frequency, spreading code and pilot resources are assigned to the users to eliminate inter-user interference [1]. By contrast, the popular power domain-based NOMA concept invokes multiplexing and successive interference cancellation (SIC) at the receivers to separate multiple users having different powers and mapped to the same resources [9], [16], [17]. On the other hand, the next generation cell-free massive MIMO concept relies on a large number of distributed access points (APs) connected to a central processing unit (CPU) and invoking distributed signal processing to avoid assigning excessive bandwidth to unique cell-specific pilots, non-orthogonal pilot sequences must be assigned to the users which results in pilot contamination. In [18]–[22], the authors show that exploiting optimal uniform quantization and wireless microwave links with capacity 100 Mbits/s, the performance of limited-backhaul cell-free Massive MIMO system closely approaches the performance of cell-free Massive MIMO with perfect backhaul links. Recently,

Li et al. [23], proposed to combine the cell-free massive MIMO concept with NOMA by employing normalized conjugate beamforming lead to cognizance was given to power allocation. Against this backdrop, we study a NOMA-based cell-free massive MIMO system using conjugate beamforming and derive its closed-form signal-to-interference-plus-noise ratio (SINR) expression under the assumption of no downlink (DL) training. Additionally, we consider the max-min fairness based BE problem and develop a bisection search method for overcoming the non-convexity of the original max-min fairness problem. Note that Cheng et al. [1] do not investigate any power allocation scheme. In a nutshell, our contributions are summarized as follows:

1. A closed-form expression is derived for the DL BE efficiency under the assumption that only the channel statistics are known by the users, whilst considering the effects of both pilot contamination and imperfect SIC.
2. A max-min BE optimization problem is formulated under per-AP power constraints. A bisection search method is employed for optimally solving the max-min fairness problem, where second order cone programming (SOCP) is chosen for solving the power minimization problem in each iteration of the classic bisection search.
3. A mode selection scheme is conceived where the set Mode = {OMA, NOMA} is defined for maximizing the BE of the system. Numerical results are presented to demonstrate the superiority of NOMA over OMA in terms of the max-min BE for dispersive scenarios associated with short coherence time and a large number of users.

II. SYSTEM MODEL

In this paper, we consider DL transmission in a NOMA-based cell-free massive MIMO system having M APs and K_{tot} single-antenna users, who are uniformly distributed in the coverage area. It is assumed that the users are grouped into L clusters with K users in each cluster, resulting in $K_{\text{tot}} = KL$. Moreover, each AP has N DL transmit antennas. We model the channel vector, $\mathbf{g}_{mlk} \in \mathbb{C}^{N \times 1}$, between the m th AP and the k th user in the l th cluster as $\mathbf{g}_{mlk} = \sqrt{\beta_{mlk}} \mathbf{h}_{mlk}$, where β_{mlk} and \mathbf{h}_{mlk} represent the large-scale fading and the small-scale fading, respectively. The elements of \mathbf{h}_{mlk} are independent and identically distributed (i.i.d.) $\mathcal{CN}(0, 1)$ random variables [24].

The work of K. Cumanan and A. G. Burr was supported by H2020- MSCA-RISE-2015 under grant number 690750. L. Hanzo would like to acknowledge the ERC's financial support of his Advanced Fellow Grant QuantCom. This work was also supported in part by the U.K. Engineering and Physical Sciences Research Council under Grant EP/P03456X/1.

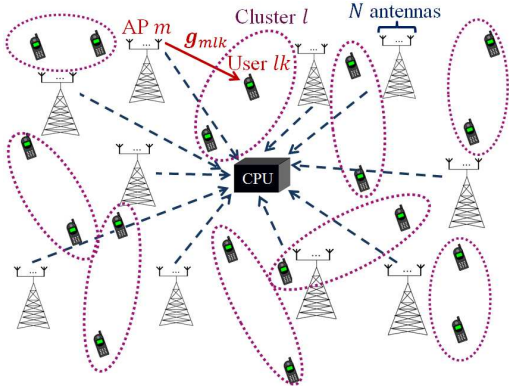


Figure 1. A cell-free massive MIMO system with K_{tot} single-antenna users and M APs. Each AP is equipped with N antennas. The solid lines denote the DL channels and the dashed lines present the backhaul links from the APs to the CPU. The users are grouped into L clusters which are shown by dotted lines. Each cluster includes K users, and it is assumed $K = 2$ here.

A. Uplink Channel Estimation

It is assumed that all pilot sequences transmitted by all the K_{tot} users in the channel estimation phase are collected in a matrix $\Phi = [\phi_{11} \cdots \phi_{K1} \cdots \phi_{1L} \cdots \phi_{KL}] \in \mathbb{C}^{\tau_p \times K_{\text{tot}}}$, where $\|\phi_{lk}\|^2 = 1$, τ_p refers to the length of the pilot sequence for each user, while the kl th column, ϕ_{kl} , represents the pilot sequence used for the k th user in the l th cluster. Moreover, the same pilot sequences ($\phi_{kl} = \hat{\phi}_l, \forall k$) are assigned to the users within the same cluster, whereas orthogonal pilots are assigned to different clusters. Based on the analysis in [1], the alternative approach to exploit the uplink pilots in NOMA-based massive MIMO system is a linear combination of the users' channels expressed as follows:

$$\mathbf{f}_{ml} = \sum_{k=1}^K \mathbf{g}_{mlk}, \forall l. \quad (1)$$

After performing a de-spreading operation, the MMSE estimate of \mathbf{f}_{ml} is given by

$$\hat{\mathbf{f}}_{ml} = c_{ml} \left(\sqrt{\tau_p \rho_p} \sum_{k=1}^K \mathbf{g}_{mlk} + \mathbf{W}_{p,m} \hat{\phi}_l \right), \quad (2)$$

where $\mathbf{W}_{p,m} \in \mathbb{C}^{M \times K}$ represents the noise sequence at the m th AP whose elements are i.i.d. $\mathcal{CN}(0, 1)$ and ρ_p introduces the normalized signal-to-noise ratio (SNR) of each pilot sequence (defined in Section V). Moreover, c_{ml} is given by $c_{ml} = \frac{\sqrt{\tau_p \rho_p} \sum_{k'=1}^K \beta_{mlk'}}{\tau_p \rho_p \sum_{k'=1}^K \beta_{mlk'} + 1}$. Note that as in [24], we assume that the large-scale fading, β_{mlk} , is known at both the CPU and the users.

B. Downlink Transmission with Conjugate Beamforming

Upon employing conjugate beamforming at the APs, the signal transmitted from the m th AP is given by

$$\mathbf{x}_m = \sqrt{\rho_d} \sum_{l=1}^L \sum_{k=1}^K \sqrt{\eta_{mlk}} \hat{\mathbf{f}}_{ml}^* s_{lk}, \quad (3)$$

where s_{lk} ($\mathbb{E}\{|s_{lk}|^2\} = 1$) and η_{mlk} are the DL transmitted symbol and the DL transmit power at the m th AP, respectively. Additionally, ρ_d refers to the maximum transmit power

normalized by the noise power p_n at the APs, which is given by

$$\mathbb{E}\{\|\mathbf{x}_m\|^2\} = \rho_d N \sum_{l=1}^L \sum_{k=1}^K \eta_{mlk} \gamma_{ml}, \quad (4)$$

where $\gamma_{ml} = \mathbb{E}\{\|\hat{\mathbf{g}}_{mlk}\|^2\} = \frac{\tau_p \rho_p (\sum_{k'=1}^K \beta_{mlk'})^2}{\tau_p \rho_p \sum_{k'=1}^K \beta_{mlk'} + 1}$.

Moreover, note that $\gamma_{ml} = \sqrt{\tau_p \rho_p} (\sum_{k'=1}^K \beta_{mlk'}) c_{ml}$. The power elements η_{mlk} are designed for satisfying that the following per-AP power constraints are satisfied:

$$\sum_{l=1}^L \sum_{k=1}^K \eta_{mlk} \gamma_{ml} \leq \frac{1}{N}, \forall m. \quad (5)$$

III. PERFORMANCE ANALYSIS

The DL BE of NOMA-based cell-free massive MIMO is derived in this section without any DL training. Hence, the users rely on the channel statistics to perform SIC. First, the basic concept of NOMA combined with cell-free massive MIMO is presented, followed by the derivations of the achievable DL rate and BE using conjugate beamforming.

A. NOMA without Downlink Training

Note that NOMA is employed only within each cluster but not between the clusters. It is assumed that within the l th cluster, “user-1” is the least-contaminated user whose signal is detected first. The signals of other users are then detected by exploiting SIC, while “user- L ” is the weakest user whose signal becomes automatically available after SIC. Again, similar to the terminology in [1], since there is no DL training, users rely on the channel statistics. Furthermore, the users are sorted based on their channel statistics, which is studied in Sections III-B. The following condition is necessary in order to successfully perform SIC and to decode the weaker user signals:

$$\mathbb{E}\{\log_2(1 + \text{SINR}_{lj}^{lk})\} \geq \mathbb{E}\{\log_2(1 + \text{SINR}_{lk}^{lk})\}, \forall j < k, \forall l, \quad (6)$$

where SINR_{lj}^{lk} is the effective SINR of user j in cluster l , when user j in cluster l decodes the signal intended for user k in the same cluster l . Based on this necessary condition, the achievable rate of the k th user in the l th cluster is obtained as:

$$\mathbf{R}_{lk}^{lk, \text{final}} = \min(\mathbb{E}\{\log_2(1 + \text{SINR}_{lj}^{lk})\}, \mathbb{E}\{\log_2(1 + \text{SINR}_{lk}^{lk})\}), \forall l, k, \quad (7)$$

where $\mathbf{R}_{lk}^{lk, \text{final}}$ is the achievable rate of user k in cluster l . Several clustering schemes will be investigated in Section VI-B.

B. Received Signal

NOMA is applied when the users with stronger channels decode data from the users with weaker channels and subtract it from the received signal. Given that only the statistics of the channels are available at the users' ends, and exploiting

the analysis in [1], the signal received by the k th user in the l th cluster is given by

$$\begin{aligned}
r_{lk}^{lk} &= \sum_{m=1}^M \sum_{l'=1}^L \sum_{k'=1}^K \sqrt{\eta_{ml'k'}} \mathbf{g}_{mlk'}^T \hat{\mathbf{f}}_{ml'}^* s_{l'k'} + n_{lk} \\
&- \sqrt{\rho_d} \sum_{k''=k+1}^K \mathbb{E} \left\{ \sum_{m=1}^M \sqrt{\eta_{mlk}} \mathbf{g}_{mlk}^T \hat{\mathbf{f}}_{ml}^* \right\} s_{lk''} \\
&= \underbrace{\sqrt{\rho_d} \mathbb{E} \left\{ \sum_{m=1}^M \sqrt{\eta_{mlk}} \mathbf{g}_{mlk}^T \hat{\mathbf{f}}_{ml}^* \right\}}_{\text{DS}_{lk}} s_{lk} \\
&+ \underbrace{\sqrt{\rho_d} \left(\sum_{m=1}^M \sqrt{\eta_{mlk}} \mathbf{g}_{mlk}^T \hat{\mathbf{f}}_{ml}^* - \mathbb{E} \left\{ \sum_{m=1}^M \sqrt{\eta_{mlk}} \mathbf{g}_{mlk}^T \hat{\mathbf{f}}_{ml}^* \right\} \right)}_{\text{BU}_{lk}} s_{lk} \\
&+ \underbrace{\sum_{k' \neq k}^{k-1} \sqrt{\rho_d} \sum_{m=1}^M \sqrt{\eta_{mlk'}} \mathbf{g}_{mlk'}^T \hat{\mathbf{f}}_{ml'}^* s_{lk'}}_{\text{IUI}_{lk'}} + \sum_{k''=k+1}^K \\
&\underbrace{\sqrt{\rho_d} \left(\sum_{m=1}^M \sqrt{\eta_{mlk''}} \mathbf{g}_{mlk''}^T \hat{\mathbf{f}}_{ml''}^* - \mathbb{E} \left\{ \sum_{m=1}^M \sqrt{\eta_{mlk''}} \mathbf{g}_{mlk''}^T \hat{\mathbf{f}}_{ml''}^* \right\} \right)}_{\text{ISIC}_{lk''}} s_{lk''} \\
&+ \underbrace{\sum_{l' \neq l}^L \sum_{k'=1}^K \sqrt{\rho_d} \sum_{m=1}^M \sqrt{\eta_{ml'k'}} \mathbf{g}_{ml'k'}^T \hat{\mathbf{f}}_{ml'}^* s_{l'k'}}_{\text{ICI}_{l'l'k'}} + n_{lk}, \quad (8)
\end{aligned}$$

where DS_{lk} and BU_{lk} are the desired signal (DS) and beamforming uncertainty (BU) for the k th user in the l th cluster, respectively, and $\text{IUI}_{lk'}$ denotes the inter-user-interference (IUI) caused by the k' th user in the l th cluster. In addition, $\text{ISIC}_{lk''}$ accounts for the interference imposed by the k'' th user on the k th user in the l th cluster due to imperfect SIC (ISIC), and $\text{ICI}_{l'l'k'}$ is the inter-cluster-interference (ICI) imposed by the users in clusters $l' \neq l$. Moreover, the superscript CB in (8) refers to conjugate beamforming. Upon exploiting the fact that the terms DS_{lk} , BU_{lk} , $\text{IUI}_{lk'}$, $\text{ISIC}_{lk''}$, and $\text{ICI}_{l'l'k'}$ are mutually uncorrelated, the achievable SINR of the system is given by (9) (defined at the top of the next page). The closed-form expression for the achievable DL rate of the k th user in the l th cluster is given in the following theorems:

Theorem 1. Having the channel statistics at the users and employing conjugate beamforming at the APs, the closed-form expression for the achievable DL rate of the signal intended for the k th user in the l th cluster is given by $R_{lk}^{lk} = \log_2(1 + \text{SINR}_{lk}^{lk})$, where the SINR_{lk}^{lk} is formulated as

$$\begin{aligned}
\text{SINR}_{lk}^{lk} &= \frac{N^2 \left(\sum_{m=1}^M \sqrt{\eta_{mlk}} \frac{\gamma_{ml}}{\sum_{i=1}^K \beta_{mli}} \beta_{mli} \right)^2}{N^2 \sum_{k'=1}^{k-1} \left(\sum_{m=1}^M \sqrt{\eta_{mlk'}} \gamma_{ml} \beta_{mlk} \right)^2 + N \sum_{l'=1}^L \sum_{k'=1}^K \sum_{m=1}^M \eta_{ml'k'} \beta_{mlk} \gamma_{ml'} + \frac{1}{\rho_d}}. \quad (10)
\end{aligned}$$

Proof: Please refer to the Appendix. ■

Theorem 2. The closed-form expression for the achievable DL rate of the signal intended for the k th user (weaker user) at the j th user (stronger user) in the l th cluster is given by $R_{lj}^{lk} = \log_2(1 + \text{SINR}_{lj}^{lk})$, where the SINR_{lj}^{lk} is given by (11) (defined at the top of the next page).

Proof: The proof follows the similar steps the one in the Appendix and it is omitted due to the space limit. ■

Assuming that the users are sorted based on their path loss and exploiting the power domain, NOMA employs SIC at the receivers to detect the desired signals [17]. Considering the expression in (10), we take the term $\mathbf{h}_{lk}^{\text{vir}} = \left[\frac{\gamma_{1l}}{\sum_{i=1}^K \beta_{1li}} \beta_{1lk}, \frac{\gamma_{2l}}{\sum_{i=1}^K \beta_{2li}} \beta_{2lk}, \dots, \frac{\gamma_{Ml}}{\sum_{i=1}^K \beta_{Mli}} \beta_{Mlk} \right]^T$, $\forall l, k$, as the virtual channel of the k th user at the l th cluster. Next we sort the users based on the quality of this virtual channel, i.e., $\|\mathbf{h}_{l1}^{\text{vir}}\|_2 \geq \|\mathbf{h}_{l2}^{\text{vir}}\|_2 \geq \dots \geq \|\mathbf{h}_{lK}^{\text{vir}}\|_2, \forall l$.

C. Bandwidth Efficiency

The BE (in bit/s/Hz) of the k th user in the l th cluster is defined as follows:

$$S_{lk}^{lk, \text{final}} = \left(1 - \frac{\tau_p}{\tau_c} \right) \log_2 \left(1 + \text{SINR}_{lk}^{lk, \text{final}} \right), \quad (12)$$

where τ_c denotes the number of samples for each coherence interval and

$$\text{SINR}_{lk}^{lk, \text{final}} = \min \left(\text{SINR}_{lj}^{lk}, \text{SINR}_{lk}^{lk} \right), \forall l, k. \quad (13)$$

IV. MAX-MIN BANDWIDTH EFFICIENCY

In this section, the max-min BE problem is presented, where the minimum user BE is maximized under per-AP power constraints, while can be formulated as follows:

$$P_1 : \max_{\eta_{mlk}} \min_{k=1 \dots K, l=1 \dots L} S_{lk}^{lk, \text{final}} \quad (14a)$$

$$\text{s.t.} \quad \sum_{l=1}^L \sum_{k=1}^K \eta_{mlk} \gamma_{ml} \leq \frac{1}{N}, \forall m, \quad (14b)$$

$$\eta_{mlk} \geq 0, \quad \forall m, \forall l, \forall k. \quad (14c)$$

Problem P_1 can be re-written as Problem P_2 (defined at the top of next page), where we have $s_{mlk} = \sqrt{\eta_{mlk}}$. By defining new slack variables, Problem P_2 can be re-formulated as:

$$P_3 : \max_t \{s_{mlk}, \varrho_{lk'j}, \nu_m\} \quad (16a)$$

$$\text{s.t.} \quad \frac{N^2 \left(\sum_{m=1}^M s_{mlk} \frac{\gamma_{ml} \beta_{mlj}}{\sum_{i=1}^K \beta_{mli}} \right)^2}{N^2 \sum_{k'=1}^{k-1} \varrho_{lk'j}^2 + N \sum_{m=1}^M \beta_{mlj} \nu_m^2 + \frac{1}{\rho_d}} > t, \forall j < k \quad (16b)$$

$$\left(\sum_{l' \neq l, k'=1}^L \sum_{m=1}^K \gamma_{ml'} s_{ml'k'}^2 + \sum_{k'=1}^k \gamma_{ml} s_{mlk}^2 \right) \leq \nu_m^2, \forall m, \forall j < k \quad (16c)$$

$$0 \leq \nu_m \leq \frac{1}{\sqrt{N}}, \quad \forall m, \quad (16d)$$

$$\sum_{m=1}^M s_{mlk'} \frac{\gamma_{ml} \beta_{mlj}}{\sum_{i=1}^K \beta_{mli}} \leq \varrho_{lk'j}, 1 \leq k' \leq k-1, \forall j < k \quad (16e)$$

$$s_{mlk} \geq 0, \quad \forall m, \forall l, \forall k. \quad (16f)$$

$$\text{SINR}_{lk} = \frac{|\text{DS}_{lk}|^2}{\mathbb{E}\left\{\left|\text{BU}_{lk}\right|^2\right\} + \sum_{k'=1}^{k-1} \mathbb{E}\left\{\left|\text{IU}_{lk'}\right|^2\right\} + \sum_{k''=k+1}^K \mathbb{E}\left\{\left|\text{ISIC}_{lk''}\right|^2\right\} + \sum_{l' \neq l}^L \sum_{k'=1}^K \mathbb{E}\left\{\left|\text{ICI}_{l'k'}\right|^2\right\} + 1}. \quad (9)$$

$$\text{SINR}_{lj}^{lk} = \frac{N^2 \left(\sum_{m=1}^M \sqrt{\eta_{mlk}} \frac{\gamma_{ml}}{\sum_{i=1}^K \beta_{mli}} \beta_{mlj} \right)^2}{N^2 \sum_{k'=1}^{k-1} \left(\sum_{m=1}^M \sqrt{\eta_{mlk'}} \frac{\gamma_{ml}}{\sum_{i=1}^K \beta_{mli}} \beta_{mlj} \right)^2 + N \sum_{l' \neq l}^L \sum_{k'=1}^K \sum_{m=1}^M \eta_{ml'k'} \beta_{mlj} \gamma_{ml'} + N \sum_{k'=1}^k \sum_{m=1}^M \eta_{mlk'} \beta_{mlj} \gamma_{ml} + \frac{1}{\rho_d}}. \quad (11)$$

$$\begin{aligned} P_2 : \max_{\varsigma_{mlk}} \min_{\{k=1 \dots K, l=1 \dots L\}} \text{SINR}_{lk}^{lk, \text{final}} &= \min \left(\text{SINR}_{lk}^{lk}, \text{SINR}_{lj}^{lk} \right), \forall j < k \\ &= \min \left(\frac{N^2 \left(\sum_{m=1}^M \sqrt{\eta_{mlk}} \frac{\gamma_{ml}}{\sum_{i=1}^K \beta_{mli}} \beta_{mlj} \right)^2}{N^2 \sum_{k'=1}^{k-1} \left(\sum_{m=1}^M \sqrt{\eta_{mlk'}} \frac{\gamma_{ml}}{\sum_{i=1}^K \beta_{mli}} \beta_{mlj} \right)^2 + N \sum_{l' \neq l}^L \sum_{k'=1}^K \sum_{m=1}^M \eta_{ml'k'} \beta_{mlj} \gamma_{ml'} + N \sum_{k'=1}^k \sum_{m=1}^M \eta_{mlk'} \beta_{mlj} \gamma_{ml} + \frac{1}{\rho_d}}, \right. \\ &\quad \left. \frac{N^2 \left(\sum_{m=1}^M \sqrt{\eta_{mlk}} \frac{\gamma_{ml}}{\sum_{i=1}^K \beta_{mli}} \beta_{mlk} \right)^2}{N^2 \sum_{k'=1}^{k-1} \left(\sum_{m=1}^M \sqrt{\eta_{mlk'}} \frac{\gamma_{ml}}{\sum_{i=1}^K \beta_{mli}} \beta_{mlk} \right)^2 + N \sum_{l'=1}^L \sum_{k'=1}^K \sum_{m=1}^M \eta_{ml'k'} \beta_{mlk} \gamma_{ml'} + \frac{1}{\rho_d}} \right), \forall j < k \\ \text{s.t. } \sum_{l=1}^L \sum_{k=1}^K \varsigma_{mlk}^2 \gamma_{ml} &\leq \frac{1}{N}, \quad \forall m, \quad \varsigma_{mlk} \geq 0, \quad \forall m, \forall l, \forall k, \end{aligned} \quad (15a)$$

$$(15b)$$

Problem P_3 is quasi-convex, therefore a bisection search maybe invoked for determining the optimal solution of the original Problem P_3 by solving the following power minimization problem, given a specific SINR t in each iteration:

$$P_4 : \min_{\{\varsigma_{mlk}, \varrho_{lk'j}, \nu_m\}} \sum_{m=1}^M \sum_{l'=1}^L \sum_{k'=1}^K \gamma_{ml'} \varsigma_{ml'k'}^2 \quad (17a)$$

$$\text{s.t. } \frac{N^2 \left(\sum_{m=1}^M \varsigma_{mlk} \frac{\gamma_{ml} \beta_{mlj}}{\sum_{i=1}^K \beta_{mli}} \right)^2}{N^2 \sum_{k'=1}^{k-1} \varrho_{lk'j}^2 + N \sum_{m=1}^M \beta_{mlj} \nu_m^2 + \frac{1}{\rho_d}} \geq t, \forall j < k, \quad (17b)$$

$$(16b) - (16f).$$

Problem P_4 can be formulated as a standard SOCP. More precisely, for a given t , Problem P_4 is reformulated as follows:

$$P_5 : \min_{\{\varsigma_{mlk}, \varrho_{lk'j}, \nu_m\}} \sum_{m=1}^M \sum_{l'=1}^L \sum_{k'=1}^K \gamma_{ml'} \varsigma_{ml'k'}^2 \quad (18a)$$

$$\text{s.t. } \|\mathbf{z}_{lj}\| \leq \frac{N \sum_{m=1}^M \varsigma_{mlk} \frac{\gamma_{ml} \beta_{mlj}}{\sum_{i=1}^K \beta_{mli}}}{\sqrt{t}}, \forall j < k \quad (18b)$$

$$(16b) - (16e),$$

where we have $\mathbf{z}_{lj} \triangleq \left[N \mathbf{v}_{lj,1}^T \quad \sqrt{N} \mathbf{v}_{lj,2}^T \quad \frac{1}{\sqrt{\rho_d}} \right]^T$, and $\mathbf{v}_{lj,1} = [\varrho_{l1j} \dots \varrho_{lk-1j}]^T$, $\mathbf{v}_{lj,2} = [\sqrt{\beta_{l1j}} \nu_1 \dots \sqrt{\beta_{lMj}} \nu_M]^T$. It can be seen that (18) represents a second order cone (SOC) [25], [26]. Hence, Problem P_5 is a standard SOCP, which is a convex problem.

V. NUMERICAL RESULTS AND DISCUSSIONS

We consider a NOMA-based cell-free massive MIMO system of M APs and K_{tot} single-antenna users spreading over a $D \times D$ coverage area. Both the APs and users are uniformly distributed at random points and the APs are equipped with N antennas. The simulation parameters and the numerical results are discussed in the following subsections.

A. Simulation Parameters

The channel coefficients between users and APs are modelled in Section II, where the coefficient β_{mlk} is defined by $\beta_{mlk} = \text{PL}_{mlk} 10^{\frac{\sigma_{sh} z_{mlk}}{10}}$, and PL_{mlk} is the path loss between the k th user in the l th cluster to the m th AP, while the second term $10^{\frac{\sigma_{sh} z_{mlk}}{10}}$ denotes the shadow fading having a standard deviation of $\sigma_{sh} = 8$ dB, and $z_{mlk} \sim \mathcal{N}(0, 1)$ [24]. In the simulations, an uncorrelated shadowing and a three-slope path loss models similar to [24] are considered. The noise power p_n is calculated as in [24, Section VI]. Furthermore, it is assumed that \bar{p}_p and \bar{p}_d denote the power of the pilot sequence and of the DL data, respectively, where $\rho_p = \frac{\bar{p}_p}{p_n}$ and $\rho_d = \frac{\bar{p}_d}{p_n}$ [24], [27], [28]. In our simulations, we set $\bar{p}_p = 100$ mW and $\bar{p}_d = 200$ mW. Similar to [24], we assume that the coverage area is wrapped around at the edges, which can simulate an area without boundaries. Hence, the square simulation area has eight neighbours. Finally, note that the solution of the max-min problem will ensure achieving the same per-user DL BE at all users.

Throughout this section, we assume that there are only two users in each cluster, known which is termed as user pairing

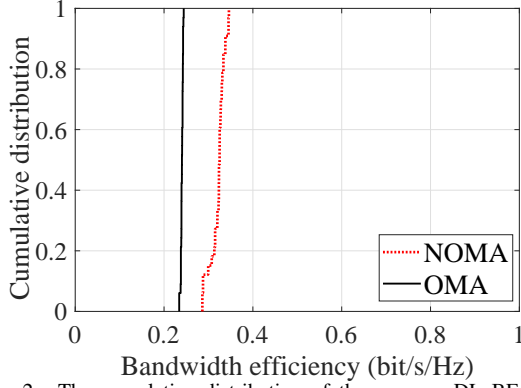


Figure 2. The cumulative distribution of the per-user DL BE of cell-free massive MIMO with per-AP power constraints. We set $M = 20$, $N = 15$, $K = 2$, $K_{\text{tot}} = 100$, $\tau_c = 110$, $D = 1$ km, $\bar{\rho}_p = 100$ mW and $\bar{\rho}_d = 200$ mW.

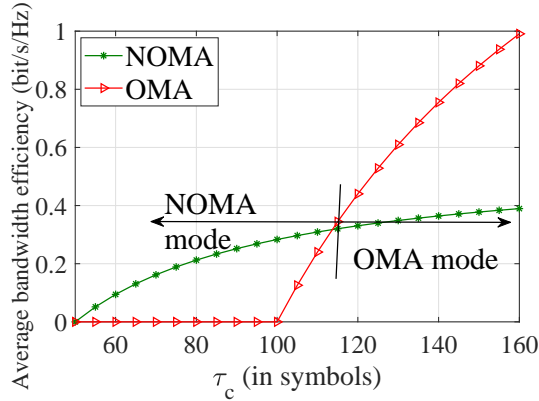


Figure 3. The average DL BE of cell-free massive MIMO versus coherence time (in symbols). We solve the max-min spectral efficiency problems with per-AP power constraints. The simulation parameters are the same as Fig. 2.

in the literature [29]. However, in contrast to [29], the CSI is not available at the transmitter and the receiver. In [1], [23], the authors propose to pair the users based on their positions. However, the clustering schemes proposed in this paper are different from those in [1], [23]. In this work, we randomly pair users.

Since the length of pilot sequences has to be at least equal to the total number of users ($\tau_c \geq K_{\text{tot}}$) in OMA, less time is left for data transmission leading to effective throughput degradation. By contrast, NOMA requires the length of pilot sequences to be at least equal to the number of clusters ($\tau_c \geq K_{\text{tot}}/2$) which leaves more time for payload data transmission compared to OMA. Note that within each cluster, the user having stronger received signal power is assumed to have perfect detection whereas other users experience some residual IUI. As a result, we are capable of enhancing the BE by selecting the optimal mode from the set $\text{Mode} = \{\text{OMA}, \text{NOMA}\}$, which depends both on the total number of users and on the length of channel coherence time.

B. Simulation Results

First, we consider a NOMA-based cell-free massive MIMO system with 20 APs ($M = 20$) and supporting 100 users

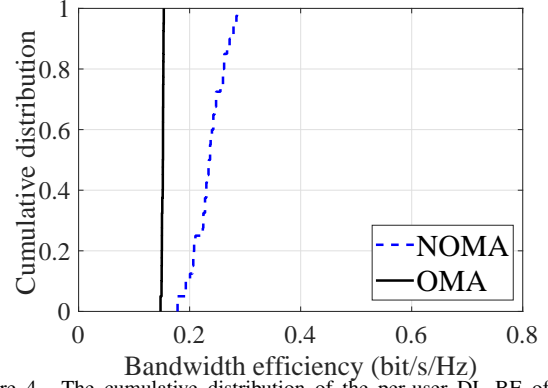


Figure 4. The cumulative distribution of the per-user DL BE of cell-free massive MIMO with per-AP power constraints for conjugate beamforming technique. We set $M = 20$, $N = 15$, $K = 2$, $K_{\text{tot}} = 140$, $\tau_c = 150$, $D = 1$ km, $\bar{\rho}_p = 100$ mW and $\bar{\rho}_d = 200$ mW.

($K_{\text{tot}} = 100$), who are randomly distributed over the area of size $D \times D$ km² and $\tau_c = 110$. Fig. 2 represents the cumulative distribution of the achievable DL BE, where we use the scheme proposed in Section IV to obtain the max-min BE. Moreover, the figure shows that NOMA outperforms OMA. In order to investigate the role of the channel's coherence time with the above-mentioned parameters, the average BE of NOMA-based cell-free massive MIMO is evaluated against τ_c in Fig. 3. These results demonstrate that there is an optimal mode switching point depending on the length of the channel's coherence time for maximizing the minimum max-min DL BE.

VI. CONCLUSIONS

We have investigated the performance of NOMA-based cell-free massive MIMO relying on conjugate beamforming and on clustering the users. A closed-form expression was derived for the DL BE of the system by exploiting SIC with no DL training and optimally solved the max-min DL BE problem. Numerical results were presented for several clustering schemes. Moreover, the effect of the system's coherence time has been investigated and a switching point from the mode set $\text{mode} = \{\text{OMA}/\text{NOMA}\}$ has been proposed to maximize the minimum DL BE of the system.

APPENDIX: PROOF OF THEOREM 1

The desired signal of user k in cluster l is given by

$$\text{DS}_{lk} = \sqrt{\rho_d} \mathbb{E} \left\{ \sum_{m=1}^M \eta_{mlk} \mathbf{g}_{mlk}^T \hat{\mathbf{f}}_{ml}^* \right\} = N \sum_{m=1}^M \frac{\sqrt{\rho_d \eta_{mlk}} \gamma_{ml} \beta_{mlk}}{\sum_{i=1}^K \beta_{mli}}. \quad (19)$$

Next, we have

$$\begin{aligned} \mathbb{E} \left\{ |\text{BU}_{lk}|^2 \right\} &= \rho_d \mathbb{E} \left\{ \left| \sum_{m=1}^M \sqrt{\eta_{mlk}} \mathbf{g}_{mlk}^T \hat{\mathbf{f}}_{ml}^* \right|^2 \right\} \\ &= \rho_d \mathbb{E} \left\{ \left| \sum_{m=1}^M \sqrt{\eta_{mlk}} \mathbf{g}_{mlk}^T \hat{\mathbf{f}}_{ml}^* \right|^2 \right\} \\ &= \rho_d N \sum_{m=1}^M \eta_{mlk} \beta_{mlk} \gamma_{ml} \end{aligned} \quad (20)$$

Next, $\mathbb{E}\{|I_{l'k'}|^2\}$ is calculated as follows:

$$\begin{aligned}\mathbb{E}\{|I_{l'k'}|^2\} &= \rho_d \mathbb{E}\left\{\left|\sum_{m=1}^M \sqrt{\eta_{mlk'}} \mathbf{g}_{mlk'}^T \hat{\mathbf{f}}_{ml}^*\right|^2\right\} = \\ &= \rho_d \mathbb{E}\left\{\left|\sum_{m=1}^M \sqrt{\eta_{mlk'}} \mathbf{g}_{mlk'}^T c_{ml} \left(\sqrt{\tau_p \rho_p} \sum_{i=1}^K \mathbf{g}_{mli} + \underbrace{\mathbf{w}_{p,m} \phi_{lk}}_{\tilde{\mathbf{w}}_{ml}}\right)^*\right|^2\right\} \\ &= N \rho_d \sum_{m=1}^M \eta_{mlk'} \beta_{mlk} \gamma_{ml} \\ &+ N^2 \rho_d \left(\sum_{m=1}^M \sqrt{\eta_{mlk'}} \gamma_{ml} \frac{\beta_{mlk}}{\sum_{i=1}^K \beta_{mli}}\right)^2.\end{aligned}\quad (21)$$

In the next step, we calculate the term $\mathbb{E}\{|I_{l'k'}|^2\}$ as follows:

$$\begin{aligned}\mathbb{E}\{|I_{l'k'}|^2\} &= \mathbb{E}\left\{\left|\sum_{m=1}^M \eta_{ml'k'} \mathbf{g}_{mlk'}^T \hat{\mathbf{f}}_{ml'k'}^*\right|^2\right\} \\ &= \rho_d \mathbb{E}\left\{\underbrace{\left|\sum_{m=1}^M \sqrt{\eta_{ml'k'}} c_{ml'} \mathbf{g}_{mlk'}^T \tilde{\mathbf{w}}_{ml'}^*\right|^2}_E\right\} \\ &+ \underbrace{\rho_d \tau_p \rho_p \mathbb{E}\left\{\left|\sum_{m=1}^M \sqrt{\eta_{ml'k'}} c_{ml'} \mathbf{g}_{mlk'}^T \left(\sum_{i=1}^K \mathbf{g}_{ml'i}^*\right)\right|^2\right\}}_F,\end{aligned}\quad (22)$$

where the term E is obtained by $E = N \sum_{m=1}^M \eta_{ml'k'} c_{ml'}^2 \beta_{mlk}$, and the term F can be calculated as $F = N \sum_{m=1}^M \eta_{ml'k'} \beta_{mlk} \gamma_{ml'} - N \sum_{m=1}^M \eta_{ml'k'} c_{ml'}^2 \beta_{mlk}$.

The term $\mathbb{E}\{|I_{l'k'}|^2\}$ is obtained as follows:

$$\begin{aligned}E\{|I_{l'k'}|^2\} &= \rho_d \mathbb{E}\left\{\underbrace{\left|\sum_{m=1}^M \sqrt{\eta_{mlk'}} \mathbf{g}_{mlk'}^T \hat{\mathbf{f}}_{ml'}^*\right|^2}_{I_1}\right\} - \rho_d \mathbb{E}\left\{\underbrace{\left|\sum_{m=1}^M \sqrt{\eta_{mlk'}} \mathbf{g}_{mlk'}^T \mathbf{f}_{ml}^*\right|^2}_{I_2}\right\},\end{aligned}\quad (23)$$

where the term I_1 is obtained by

$$\begin{aligned}I_1 &= N^2 \left(\sum_{m=1}^M \frac{\sqrt{\eta_{mlk'}} \gamma_{ml} \beta_{mlk}}{\sum_{i=1}^K \beta_{mli}}\right)^2 \\ &+ N \sum_{m=1}^M \eta_{mlk'} \beta_{mlk} \gamma_{ml},\end{aligned}\quad (24)$$

and $I_2 = N \sum_{m=1}^M \sqrt{\eta_{mlk'}} \gamma_{ml} \frac{\beta_{mlk}}{\sum_{i=1}^K \beta_{mli}}$. Finally, using (23) and (24), we have

$$\mathbb{E}\{|I_{l'k'}|^2\} = N \rho_d \sum_{m=1}^M \eta_{mlk'} \beta_{mlk} \gamma_{ml}.\quad (25)$$

REFERENCES

- [1] H. V. Cheng, E. Björnson, and E. G. Larsson, "Performance analysis of NOMA in training-based multiuser MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 372–385, Jan. 2018.
- [2] G. Interdonato, E. Björnson, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Ubiquitous cell-free massive MIMO communications," *IEEE Commun. Mag.*, pp. 1–19, submitted.
- [3] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "Enhanced max-min SINR for uplink cell-free massive MIMO systems," in *Proc. IEEE ICC*, May 2018, pp. 1–6.
- [4] A. Burr, M. Bashar, and D. Maryopi, "Ultra-dense radio access networks for smart cities: Cloud-RAN, Fog-RAN and cell-free Massive MIMO," in *Proc. IEEE PIMRC*, Sep. 2018.
- [5] Y. Liu, Z. Qin, M. El Kashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal multiple access for 5G and beyond," *IEEE Trans. Wireless Commun.*, vol. 105, no. 12, pp. 2347–2381, Dec. 2017.
- [6] Y. Liu, H. Xing, C. Pan, A. Nallanathan, M. El Kashlan, and L. Hanzo, "Multiple-antenna-assisted non-orthogonal multiple access," *IEEE Trans. Wireless Commun.*, vol. 25, no. 2, pp. 17–23, Apr. 2018.
- [7] Z. Ding, P. Fan, and H. V. Poor, "Impact of user pairing on 5G non-orthogonal multiple-access downlink transmissions," *IEEE Trans. Veh. Technol.*, vol. 85, no. 5, pp. 6010–6023, Aug. 2016.
- [8] P. Xu and K. Cumanan, "Optimal power allocation scheme for non-orthogonal multiple access with α -fairness," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 10, pp. 2357–2369, Oct. 2017.
- [9] P. Xu, K. Cumanan, and Z. Yang, "Optimal power allocation scheme for NOMA with adaptive rates and alpha-fairness," in *Proc. IEEE Globecom*, Dec. 2017.
- [10] H. Alobiedollah, K. Cumanan, J. Thiyagalingam, A. G. Burr, Z. Ding, and O. A. Dobre, "Energy efficiency fairness beamforming design for MISO NOMA systems," in *Proc. IEEE WCNC*, Apr. 2019.
- [11] —, "Sum rate fairness trade-off-based resource allocation technique for MISO NOMA systems," in *Proc. IEEE WCNC*, Apr. 2019.
- [12] —, "Energy efficient beamforming design for miso non-orthogonal multiple access systems," *IEEE Trans. Commun.*, Accepted.
- [13] F. Alavi, K. Cumanan, A. Burr, and Z. Ding, "Robust beamforming techniques for non-orthogonal multiple access systems with bounded channel uncertainties," *IEEE Commun. Lett.*, vol. 21, no. 9, pp. 2033–2036, Sep. 2017.
- [14] —, "Beamforming techniques for non-orthogonal multiple access in 5G cellular networks," *IEEE Trans. Veh. Technol.*, vol. 5, pp. 565–577, 2018.
- [15] —, "Outage constraint based robust beamforming design for non-orthogonal multiple access in 5G cellular networks," in *Proc. IEEE PIMRC*, Nov. 2017.
- [16] G. Wunder, P. Jung, M. Kasparick, T. Wild, F. Schaich, Y. Chen, S. T. Brink, I. Gaspar, N. Michailow, A. Festag, L. Mendes, N. Cassiau, D. Ktenas, M. Dryjanski, S. Pietrzyk, B. Eged, P. Vago, and F. Wiedmann, "5GNow: non-orthogonal, asynchronous waveforms for future mobile applications," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 97–105, Feb. 2014.
- [17] R. Zhang and L. Hanzo, "A unified treatment of superposition coding aided communications: Theory and practice," *IEEE Commun. Surveys Tutorials*, vol. 13, no. 3, pp. 503–520, Mar. 2011.
- [18] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, and M. Debbah, "Cell-free massive MIMO with limited backhaul," in *Proc. IEEE ICC*, May 2018, pp. 1–7.
- [19] —, "Max-min SINR of cell-free massive MIMO uplink with optimal uniform quantization," *IEEE Trans. Commun.*, Submitted.
- [20] A. G. Burr, M. Bashar, and D. Maryopi, "Cooperative access networks: Optimum fronthaul quantization in distributed massive MIMO and cloud RAN," in *Proc. IEEE VTC*, Jun. 2018, pp. 1–7.
- [21] M. Bashar, H. Q. Ngo, A. Burr, D. Maryopi, K. Cumanan, and E. G. Larsson, "On the performance of backhaul constrained cell-free Massive MIMO with linear receivers," in *Proc. IEEE Asilomar*, Nov. 2018, pp. 1–7.
- [22] D. Maryopi, M. Bashar, and A. Burr, "On the uplink throughput of zero-forcing in cell-free Massive MIMO with coarse quantization," *IEEE Trans. Veh. Technol.*, pp. 1–5, Accepted.
- [23] Y. Li and G. A. A. Baduge, "NOMA-Aided cell-free massive MIMO systems," *IEEE Wireless Commun. Lett.*, pp. 1–4, 2018.

- [24] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, Mar. 2017.
- [25] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, UK: Cambridge University Press, 2004.
- [26] M. Bashar, K. Cumanan, A. G. Burr, M. Debbah, and H. Q. Ngo, "On the uplink max-min SINR of cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, pp. 1–17, Jan. 2019.
- [27] M. Bashar, K. Cumanan, A. G. Burr, H. Q. Ngo, and H. V. Poor, "Mixed quality of service in cell-free Massive MIMO," *IEEE Commun. Lett.*, vol. 22, no. 7, pp. 706–709, Jul. 2018.
- [28] M. Bashar, K. Cumanan, A. Burr, H. Q. Ngo, E. Larsson, and P. Xiao, "On the energy efficiency of limited-backhaul cell-free Massive MIMO," in *Proc. IEEE ICC*, May 2019.
- [29] S. Ali, E. Hossain, and D. I. Kim, "Non-orthogonal multiple access (NOMA) for downlink multiuser MIMO systems: User clustering, beamforming, and power allocation," *IEEE Access*, vol. 5, pp. 565–577, Mar. 2017.