Weighted MMSE Optimization of Conjugate Beamforming for Cell-Free Massive MIMO

Daesung Yu[1], Hoon Lee[2], Seung-Eun Hong[3] and Seok-Hwan Park[1]

[1]Division of Electronic Engineering, Jeonbuk National University, Jeonju 54896, South Korea

[2]Department of Information and Communications Engineering, Pukyong National University, Busan 48513, South Korea

[3]Electronics and Telecommunications Research Institute (ETRI), Daejeon 34129, South Korea

Email: imcreative93@jbnu.ac.kr, hlee@pknu.ac.kr, iptvguru@etri.re.kr, seokhwan@jbnu.ac.kr

***Abstract*—Cell-free massive multiple-input multiple-output (MIMO) systems are envisioned to achieve the improved spectral efficiency by supporting users via nearby access points (APs). This work addresses the optimization of beamforming weights for cell-free massive MIMO systems. The connectivity level constraints are taken into account to accommodate finite-rate fronthaul links. The minimum rate maximization problem is tackled by the weighted minimum mean squared error (WMMSE) algorithm. Numerical results show that the proposed scheme achieves significantly improved performance than a baseline scheme in overall signal-to-noise ratio (SNR) regime.**

***Index Terms*—Cell-free massive MIMO, conjugate beamforming, weighted MMSE algorithm.**

I. Introduction

In cell-free massive multiple-input multiple-output (MIMO) systems, users are jointly served by multiple nearby access points (APs) [1], [2]. The system capacity can be improved by means of cooperative transmission and reception strategies across APs, which are centrally managed by a cloud processor (CP). The system can be seen as an instance of coordinated multi-point joint transmission [3], network MIMO [4], and cloud radio access network (C-RAN) [5]. The optimization of cooperative downlink beamforming for cell-free systems was studied in [6], [7]. Computation structures of these algorithms are not scalable to the network size, especially to the numbers of APs and antennas, thereby leading to the prohibitive calculation cost in massive connectivity scenarios.

A practical but efficient strategy is to fix radiation patterns of APs and focus on designing their beam steering directions. Possible candidates for predefined beam patterns are conjugate and zero-forcing beamforming schemes [1], [2]. The transmission from massive APs is simply controlled by optimizing scalar beam weights. Consequently, the corresponding optimization algorithms become scalable to the size of the cell-free MIMO systems. These conventional studies have been confined to single-antenna APs. Furthermore, the limited coordination capability among a CP and APs have not been adequately examined.

This work was supported by Basic Science Research Program through the National Research Foundation (NRF) of Korea grants funded by the Ministry of Education [NRF-2019R1A6A1A09031717, 2021R1C1C1006557, 2021R1I1A3054575]. This work was also supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (No. 2018-0-01659, 5G Open Intelligence-Defined RAN (ID-RAN) Technique based on 5G New Radio).

This work investigates the optimization of cell-free MIMO systems by adjusting scalar beam weights at APs. We consider a practical scenario where the capacity of fronthaul links connecting a CP and APs is limited. In this configuration, bitstreams of intended users would not be available at all APs due to the finite-rate fronthaul links. Such a limited fronthaul overhead imposes constraints on the connectivity level of wireless access links. We tackle the problem of maximizing minimum-rate based on the weighted minimum mean squared error (WMMSE) framework. The resulting algorithm achieves a locally optimal solution. Numerical results demonstrate the performance gain of the developed WMMSE optimization algorithm over a baseline scheme.

II. System Model

We consider a downlink cell-free massive MIMO system where K single-antenna users are served by M multi-antenna APs each equipped with N antennas. Let $\mathcal{K} = \{1, 2, \ldots, K\}$ and $\mathcal{M} = \{1, 2, \ldots, M\}$ be the sets of user and AP indices, respectively. The APs are controlled by a CP connected by fronthaul links. The received signal y_k of user k is written as

$$y_k = \sum\nolimits_{i\in\mathcal{M}} \mathbf{h}_{k,i}^H \mathbf{x}_i + z_k, \tag{1}$$

where $\mathbf{h}_{k,i} \in \mathbb{C}^{N\times 1}$ denotes the channel vector from AP i to user k, $\mathbf{x}_i \in \mathbb{C}^{N\times 1}$ is the transmitted signal vector from AP i which satisfies the power contraint $\mathbb{E}[||\mathbf{x}_i||^2] \leq P$, and $z_k \sim \mathcal{CN}(0, \sigma^2)$ represents the additive noise signal. The SNR of the system is defined as P/σ^2.

For the simple implementation, each AP adopts conjugate beamforming [2]. Then, the transmitted signal vector $\mathbf{x}_i$ of AP i is given as

$$\mathbf{x}_i = \sum\nolimits_{k\in\mathcal{K}} \frac{\mathbf{h}_{k,i}}{||\mathbf{h}_{k,i}||} \alpha_{k,i} s_k, \tag{2}$$

where s_k stands for the data symbol for user k, and $\alpha_{k,i} \in \mathbb{C}$ indicates a scalar beam weight at AP i intended for user k. To convey $\mathbf{x}_i$, AP i downloads the set of the data symbols $\{s_k\}_{k\in\mathcal{K}}$ from the CP through the fronthaul link. The transmit power constraint at AP i with the beamforming model (2) can be represented as

$$\mathbb{E}[||\mathbf{x}_i||^2] = \sum\nolimits_{k\in\mathcal{K}} |\alpha_{k,i}|^2 \leq P. \tag{3}$$

Consequently, the received signal y_k in (1) can be rewritten as

$$y_k = \sum_{i\in\mathcal{M}} g_{k,k,i}\alpha_{k,i}s_k + \sum_{l\in\mathcal{K}\setminus\{k\}}\sum_{i\in\mathcal{M}} g_{k,l,i}\alpha_{l,i}s_l + z_k, \quad (4)$$

where the effective channel coefficient $g_{k,l,i}$ is defined as $g_{k,l,i} \triangleq (\mathbf{h}_{k,i}^H\mathbf{h}_{l,i})/||\mathbf{h}_{l,i}|| \in \mathbb{C}$. It is noted that $g_{k,l,i} = ||\mathbf{h}_{k,i}||$ for $k = l$.

Denoting $\boldsymbol{\alpha} = \{\alpha_{k,i}\}_{k\in\mathcal{K},i\in\mathcal{M}}$, the signal-to-interference-plus-noise ratio (SINR) for user k is written by

$$\gamma_k(\boldsymbol{\alpha}) = \frac{|\mathbf{g}_{k,k}^H\boldsymbol{\alpha}_k|^2}{\sigma^2 + \sum_{l\in\mathcal{K}\setminus\{k\}}|\mathbf{g}_{k,l}^H\boldsymbol{\alpha}_l|^2}, \quad (5)$$

where

$$\boldsymbol{\alpha}_k = \begin{bmatrix} \alpha_{k,1} \\ \alpha_{k,2} \\ \vdots \\ \alpha_{k,M} \end{bmatrix} \text{ and } \mathbf{g}_{k,l} = \begin{bmatrix} g_{k,l,1}^* \\ g_{k,l,2}^* \\ \vdots \\ g_{k,l,M}^* \end{bmatrix}. \quad (6)$$

Then, the achievable data rate R_k of user k is bounded as

$$R_k \le \log_2\left(1 + \gamma_k(\boldsymbol{\alpha})\right). \quad (7)$$

We aim at optimizing the scalar beam weights $\boldsymbol{\alpha}$ for maximizing the minimum rate $R_{\min} \triangleq \min_{k\in\mathcal{K}} R_k$ [2], [7], [8] subject to the power constraints (3). The corresponding problem is formulated as

$$\underset{\boldsymbol{\alpha},\mathbf{R}}{\text{maximize}} \ \min_{k\in\mathcal{K}} R_k \quad (8a)$$

$$\text{s.t. } R_k \le \log_2(1+\gamma_k(\boldsymbol{\alpha})), \ k\in\mathcal{K}, \quad (8b)$$

$$\sum\nolimits_{k\in\mathcal{K}} |\alpha_{k,i}|^2 \le P, \ i\in\mathcal{M}, \quad (8c)$$

where $\mathbf{R} = \{R_k\}_{k\in\mathcal{K}}$ accounts for the collection of the achievable rate variables. It is not straightforward to identify the globally optimum solution of (8) due to the non-convexity of the rate constraint (8b).

III. Optimization of Conjugate Beamforming

A. WMMSE Optimization Algorithm

We tackle the non-convex optimization problem (8) based on the WMMSE approach [9], [10]. To this end, we first transform (8b) into the weighted mean square error (MSE) formula. Let $e_k(\boldsymbol{\alpha}, u_k) = \mathbb{E}[|s_k - \hat{s}_k|^2]$ be the MSE between the data signal s_k and a linear estimate $\hat{s}_k = u_k^* y_k$. It is written as

$$e_k(\boldsymbol{\alpha}, u_k) = \left|1 - u_k^*\mathbf{g}_{k,k}^H\boldsymbol{\alpha}_k\right|^2 + \sum\nolimits_{l\in\mathcal{K}\setminus\{k\}} \left|u_k^*\mathbf{g}_{k,l}^H\boldsymbol{\alpha}_l\right|^2 + |u_k|^2\sigma^2. \quad (9)$$

From (9), the achievable rate R_k in (7) can be expressed as

$$R_k \le \log_2\left(1+\gamma_k(\boldsymbol{\alpha})\right) = \max_{u_k\in\mathbb{C}} \log_2\left(\frac{1}{e_k(\boldsymbol{\alpha}, u_k)}\right), \quad (10)$$

where the optimal u_k that maximizes the right-hand side (RHS) of (10), or, equivalently, minimizes the MSE $e_k(\boldsymbol{\alpha}, u_k)$, is obtained as

$$u_k = \frac{1}{\sigma^2 + \sum_{l\in\mathcal{K}}\left|\mathbf{g}_{k,l}^H\boldsymbol{\alpha}_l\right|^2}\mathbf{g}_{k,k}^H\boldsymbol{\alpha}_k. \quad (11)$$

Algorithm 1 WMMSE algorithm for solving (14)

1. Initialize $\boldsymbol{\alpha}$ as arbitrary values that satisfy the power constraints (3), and set $t \leftarrow 1$.
2. Compute the $R_{\min}$ with the initialized $\boldsymbol{\alpha}$ and store to $R_{\min}^{(0)}$.
3. Update $\mathbf{u}$ and $\mathbf{w}$ according to (11) and (13).
4. Update $\boldsymbol{\alpha}$ as a solution of the convex problem obtained by fixing $\mathbf{u}$ and $\mathbf{w}$ in problem (14).
5. Compute the $R_{\min}$ with the updated $\boldsymbol{\alpha}$ and store to $R_{\min}^{(t)}$.
6. Stop if $|R_{\min}^{(t)} - R_{\min}^{(t-1)}| \le \delta$ or $t > t_{\max}$. Otherwise, set $t \leftarrow t+1$ and go back to Step 3.

According to [11, Lem. 1], (11) is further refined as

$$R_k \le \max_{u_k\in\mathbb{C}, w_k>0} \log_2 w_k + \frac{1}{\ln 2}(1 - w_k e_k(\boldsymbol{\alpha}, u_k)), \quad (12)$$

the optimal weight w_k given as

$$w_k = 1/e_k(\boldsymbol{\alpha}, u_k). \quad (13)$$

As a result, we can equivalently restate problem (8) as

$$\underset{\boldsymbol{\alpha},\mathbf{R},\mathbf{u},\mathbf{w}}{\text{maximize}} \ \min_{k\in\mathcal{K}} R_k \quad (14)$$

$$\text{s.t. } R_k \le \log_2 w_k + \frac{1}{\ln 2}(1 - w_k e_k(\boldsymbol{\alpha}, u_k)), \ k\in\mathcal{K},$$

$$\sum\nolimits_{k\in\mathcal{K}} |\alpha_{k,i}|^2 \le P, \ i\in\mathcal{M},$$

$$w_k > 0, \ k\in\mathcal{K},$$

with $\mathbf{u} = \{u_k\}_{k\in\mathcal{K}}$ and $\mathbf{w} = \{w_k\}_{k\in\mathcal{K}}$. Problem (14) becomes convex by fixing $\mathbf{u}$ and $\mathbf{w}$. With the optimum structure in (11) and (13), the minimum rate $R_{\min}$ can be gradually improved via an alternating optimization between $\boldsymbol{\alpha}$ and $\{\mathbf{u}, \mathbf{w}\}$. The corresponding procedure is summarized in Algorithm 1.

B. Limiting Connectivity Level

Practical fronthaul links suffer from the limited capacity. In this configuration, the data symbol s_k is available only at a subset of the APs denoted by $\mathcal{M}_k \subset \mathcal{M}$. Thus, user k can be supported by APs $i \in \mathcal{M}_k$, degrading the access link performance. Such a concept is formalized by the *connectivity level* [12], [13] denoted by $\tilde{M} = |\mathcal{M}_k|$. It quantifies the number of the APs that can get the data symbols via the capacity-constrained fronthaul links.

It is assumed that data fetching strategies, i.e., the sets $\mathcal{M}_1, \cdots, \mathcal{M}_K$, are predefined by the network. These are equivalently represented by a group of binary variables $c_{k,i} \in \{0,1\}, \forall k\in\mathcal{K}, \forall i\in\mathcal{M}$, defined as

$$c_{k,i} = \begin{cases} 1, & i\in\mathcal{M}_k \\ 0, & i\notin\mathcal{M}_k \end{cases}. \quad (15)$$

The connectivity level constraints are readily included into the minimum rate maximization task (8). Since user k can only be served by APs $i\in\mathcal{M}_k$, the associated beam weight $\alpha_{k,i}$ becomes inactive if $c_{k,i} = 0$. This is written by

$$(1 - c_{k,i})\alpha_{k,i} = 0, \ k\in\mathcal{K}, i\in\mathcal{M}. \quad (16)$$

The proposed WMMSE algorithm can straightforwardly handle the linear constraint (16).

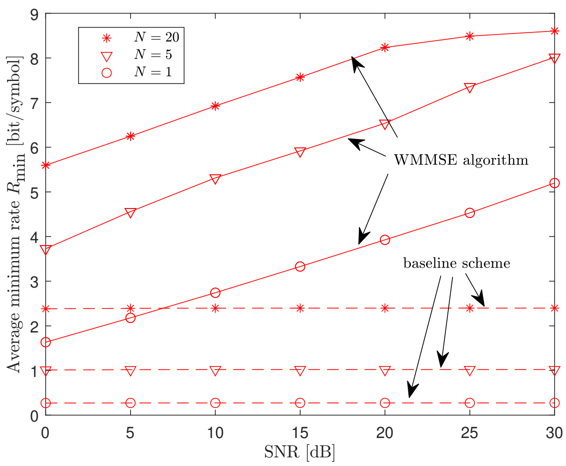

Fig. 1. Average minimum rate $R_{\min}$ versus SNR [dB] ($M = \tilde{M} = K = 10$ and $N \in \{1, 5, 20\}$).

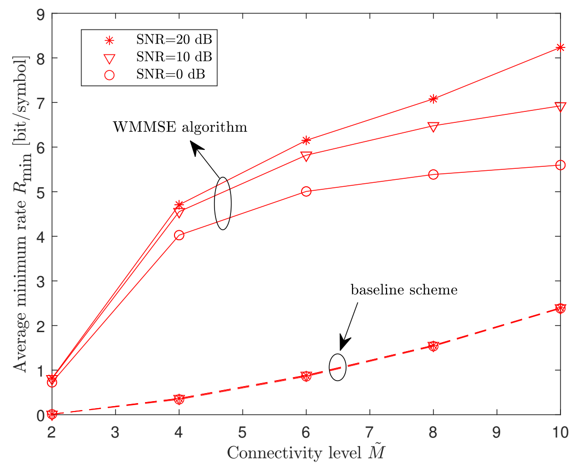

Fig. 2. Average minimum rate $R_{\min}$ versus connectivity level $\tilde{M}$ ($M = K = 10$, $N = 20$ and SNR $\in \{0, 10, 20\}$ dB).

IV. Numerical Results

We validate the effectiveness of the WMMSE algorithm via numerical simulations. The positions of the APs and users are uniformly distributed within a circular area of radius 100 m. We model each channel vector $\mathbf{h}_{k,i}$ as $\mathbf{h}_{k,i} = (\rho_0(d_{k,i}/d_0)^{-\eta})^{1/2}\tilde{\mathbf{h}}_{k,i}$, where $d_{k,i}$ denotes the distance between user k and AP i, d_0 is the reference distance set to $d_0 = 30$ m, ρ_0 indicates the path-loss at the reference distance d_0 set to $\rho_0 = 10$ dB, η is the path-loss exponent set to $\eta = 3$, and $\tilde{\mathbf{h}}_{k,i}$ represents the small-scale Rayleigh fading vector, i.e., $\tilde{\mathbf{h}}_{k,i} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$. We assume without claim of optimality that the set $\mathcal{M}_k$ of APs serving user k consists of the $\tilde{M}$ closest APs to user k.

Fig. 1 plots the average minimum rate $R_{\min}$ versus the SNR for $M = \tilde{M} = K = 10$ and $N \in \{1, 5, 20\}$. A baseline scheme adopts a simple equal power transmission strategy with $\alpha_{k,i} = \sqrt{P/K_i}$, $\forall k \in \mathcal{K}$, $\forall i \in \mathcal{M}$, where K_i stands for the number of the users served by AP $i \in \mathcal{M}_k$. The performance of the WMMSE gradually increases as the SNR grows, whereas that of the baseline method is not enhanced. The impact of WMMSE optimization becomes more significant with a larger number of AP antennas N. This verifies the effectiveness of the proposed beam weight optimiztion policy.

Fig. 2 presents the average minimum rate $R_{\min}$ versus the connectivity level $\tilde{M}$ for $M = K = 10$, $N = 20$, and SNR $\in \{0, 10, 20\}$ dB. The figure shows that the performance gap between the baseline scheme and WMMSE is considerable in all connectivity levels $\tilde{M}$. The impact of SNR on the minimum rate performance of WMMSE is more clearly observed when the system is equipped with fronthaul links of higher capacity, i.e., larger connectivity level $\tilde{M}$.

V. Conclusion

We have studied the design of conjugate beamforming for a cell-free massive MIMO system under constraints on the connectivity level. The minimum rate maximization problem has been tackled by non-convex optimization techniques. Numerical results have demonstrated the effectiveness of the optimized conjugate beamforming and the impacts of various system parameters such as SNR, the number of AP antennas, and the connectivity level.

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. 1834–1850, 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. 4445–4459, Jul. 2017.

[3] D. Lee *et al.*, "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 148–155, Feb. 2012.

[4] D. Gesbert, S. Hanly, H. Huang, S. Shamai (Shitz), O. Simeone, and W. Yu, "Multi-Cell MIMO Cooperative Networks: A New Look at Interference," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 9, pp. 1380–1408, Dec. 2010.

[5] O. Simeone, A. Maeder, M. Peng, O. Sahin, and W. Yu, "Cloud radio access network: Virtualizing wireless access for dense heterogeneous systems," *J. Commun. Netw.*, vol. 18, no. 2, pp. 135–149, Apr. 2016.

[6] S.-H. Park, O. Simeone, O. Sahin, and S. Shamai (Shitz), "Joint precoding and multivariate backhaul compression for the downlink of cloud radio access networks," *IEEE Trans. Signal Process.*, vol. 61, no. 22, pp. 5646–5658, Nov. 2013.

[7] A. Zhou, J. Wu, E. G. Larsson, and P. Fan, "Max-min optimal beamforming for cell-free massive MIMO," *IEEE Commun. Lett.*, vol. 24, no. 10, pp. 2344–2348, Oct. 2020.

[8] J. Park, H. Lee, S. Eom, and I. Lee, "UAV-aided wireless powered communication networks: Trajectory optimization and resource allocation for minimum throughput maximization," *IEEE Access*, vol. 7, no. 1, pp. 134978–134991, Dec. 2019.

[9] S. S. Christensen, R. Agarwal, E. De Carvalho, and J. M. Cioffi, "Weighted sum-rate maximization using weighted MMSE for MIMO-BC beamforming design," *IEEE Trans. Wireless Commun.*, vol. 7, no. 12, pp. 4792–4799, Dec. 2008.

[10] D. Yu, J. Kim, and S.-H. Park, "An efficient rate-splitting multiple access scheme for the downlink of C-RAN systems," *IEEE Wireless Commun. Lett.*, vol. 8, no. 6, pp. 1555–1558, Dec. 2019.

[11] Y. Zhou and W. Yu, "Fronthaul compression and transmit beamforming optimization for multi-antenna uplink C-RAN," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 4138–4151, Aug. 2016.

[12] A. Maeder, and *et al*, "A scalable flexible radio access network architecture for fifth generation mobile networks," *IEEE Commun. Mag.*, vol. 54, no. 11, pp. 16–23, Nov. 2016.

[13] J. Kim, H. Lee, S.-E. Hong, and S.-H. Park, "Flexible connectivity level control via rate splitting in fog RAN downlink with finite-capacity backhaul," *2020 International Conf. Inf. Commun. Technol. Conv. (ICTC)*, Jeju, Korea, 2020, pp. 984–988, Oct. 2020.