SCIENCE CHINA
Information Sciences

• LETTER •

November 2023, Vol. 66 219301:1–219301:2
https://doi.org/10.1007/s11432-023-3858-8

Manifold optimization assisted centralized hybrid precoding for cell-free massive MIMO systems

Jie LUO[1], Jiancun FAN[1,2*], Yawen WANG[1] & Jinbo ZHANG[2]

[1] *School of Information and Communications Engineering, Xi'an Jiaotong University, Xi'an* 710049*, China;*
[2] *Science and Technology on Communication Networks Laboratory, the 54th Research Institute of CETC, Shijiazhuang* 050081*, China*

Received 8 March 2023/Revised 28 June 2023/Accepted 31 August 2023/Published online 22 September 2023

Citation Luo J, Fan J C, Wang Y W, et al. Manifold optimization assisted centralized hybrid precoding for cell-free massive MIMO systems. Sci China Inf Sci, 2023, 66(11): 219301, https://doi.org/10.1007/s11432-023-3858-8

In the sixth generation (6G) wireless communication, cell-free massive multiple-input multiple-output (MIMO) has emerged as a promising technology due to it can improve macro diversity gain, enhance edge user service quality, and reduce path loss [1]. To suppress inter-user interference and improve spectral efficiency, the design of the precoding scheme is important.

However, the fully digital precoding scheme requires a dedicated radio frequency (RF) link for each antenna, which leads to high energy consumption and hardware deployment cost [2]. Compared with fully digital precoding, hybrid precoding composed of analog and digital precoders is very necessary in user-centric and cell-free massive MIMO systems because it can balance the achievable rate and energy consumption [3]. To deal with the non-convex constraint such as block diagonalization introduced by cell-free structures during hybrid precoding design, we propose a manifold optimization assisted centralized hybrid precoding (MO-CHP) scheme, which is implemented by alternate optimization and Riemann gradient descent methods.

System model and problem formulation. We consider a cell-free massive MIMO system with CHP architecture, where M access point (APs) serve K single-antenna users at the same frequency, and each AP is equipped with N antennas and N_{RF} RF links. The uplink channel matrix between user k and AP m is $\boldsymbol{H}_{m,k} \in \mathbb{C}^{N\times 1}$. At AP m, the digital precoder and analog precoder are $\boldsymbol{F}_m^{\mathrm{DD}} = [\boldsymbol{F}_{m,1}^{\mathrm{DD}}, \ldots, \boldsymbol{F}_{m,K}^{\mathrm{DD}}] \in \mathbb{C}^{N_{\mathrm{RF}}\times K}$ and $\boldsymbol{F}_m^{\mathrm{RF}} \in \mathbb{C}^{N\times N_{\mathrm{RF}}}$, respectively. The received baseband signal of the kth user can be expressed as

$$y_k = \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{F}_k s_k + \sum_{i=1,i\neq k}^{K} \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{F}_i s_i + n_k, \tag{1}$$

where $\boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{F}_i = \sum_{m=1}^{M} \boldsymbol{H}_{m,k}^{\mathrm{H}}\boldsymbol{F}_m^{\mathrm{RF}}\boldsymbol{F}_{m,i}^{\mathrm{DD}}$, $n_k \sim (0, \sigma_k^2)$ is the additive white Gaussian noise, and $s_k \sim \mathcal{CN}(0,1)$ is the transmit data symbol.

In practical applications, not all APs need to serve all users, and the connection between them depends on channel conditions. In Appendix A, we define a series of diagonal matrices $\boldsymbol{D}_k = \mathrm{blkdiag}\left(\boldsymbol{D}_{1,k}, \ldots, \boldsymbol{D}_{M,k}\right)$ to construct the equivalent channel $\hat{\boldsymbol{H}}_k^{\mathrm{H}} = \boldsymbol{H}_k^{\mathrm{H}}\boldsymbol{D}_k$ for $k = 1, \ldots, K$, where $\boldsymbol{D}_{m,k} = \boldsymbol{I}_N$ means that user k is served by the mth AP. We assume that the transmitted signal follows the Gaussian distribution, the achievable rate of kth user is

$$R_k = \log\left(1 + \frac{|\hat{\boldsymbol{H}}_k^{\mathrm{H}}\boldsymbol{F}^{\mathrm{RF}}\boldsymbol{F}_k^{\mathrm{DD}}|^2}{\sum_{i=1,i\neq k}^{K}|\hat{\boldsymbol{H}}_k^{\mathrm{H}}\boldsymbol{F}^{\mathrm{RF}}\boldsymbol{F}_i^{\mathrm{DD}}|^2 + \sigma_k^2}\right). \tag{2}$$

In order to maximize the sum achievable rate, the overall optimization problem can be expressed as

$$\begin{aligned}
(\mathrm{P1}):\ \max_{\boldsymbol{F}^{\mathrm{RF}},\boldsymbol{F}_k^{\mathrm{DD}}}\ & \sum_{k=1}^{K} R_k & \text{(3a)}\\
\text{s.t.}\ & \left\|\boldsymbol{F}^{\mathrm{RF}}\boldsymbol{F}_k^{\mathrm{DD}}\right\|_{\mathrm{F}}^2 \leqslant 1,\ \forall k, & \text{(3b)}\\
& \sum_{k=1}^{K}\left\|\boldsymbol{F}_m^{\mathrm{RF}}\boldsymbol{F}_{m,k}^{\mathrm{DD}}\right\|^2 \leqslant \rho_{\mathrm{AP}},\ \forall m, & \text{(3c)}\\
& \boldsymbol{F}^{\mathrm{RF}} = \mathrm{blkdiag}\left(\boldsymbol{F}_1^{\mathrm{RF}}, \ldots, \boldsymbol{F}_M^{\mathrm{RF}}\right), & \text{(3d)}\\
& \left|\boldsymbol{F}_m^{\mathrm{RF}}(p,q)\right| = 1,\ \forall m,\ \forall p,\ \forall q, & \text{(3e)}
\end{aligned}$$

where Eq. (3b) represents the transmit power constraints, Eq. (3c) means per-AP power constraints, ρ_{AP} is the maximum transmit power for each AP, Eq. (3d) represents the block-diagonal constraint, and Eq. (3e) is the unit modulus constraint of analog precoder, respectively.

It can be seen that the optimization problem P1 needs to maximize the sum achievable rate under consideration of multiple non-convex constraints, which is not easy to solve. In this regard, we can adopt a two-step approach. In the first step, the hybrid precoding matrix $\boldsymbol{F}_k = \boldsymbol{F}^{\mathrm{RF}}\boldsymbol{F}_k^{\mathrm{DD}}$ is regarded as a whole, so that the internal non-convex constraints are not considered, and the process of solving $\boldsymbol{F}_k$ is simplified. In Appendix B, we solve the fully digital precoder $\boldsymbol{F}_k$ from two different perspectives: interference cancellation

* Corresponding author (email: fanjc0114@gmail.com)

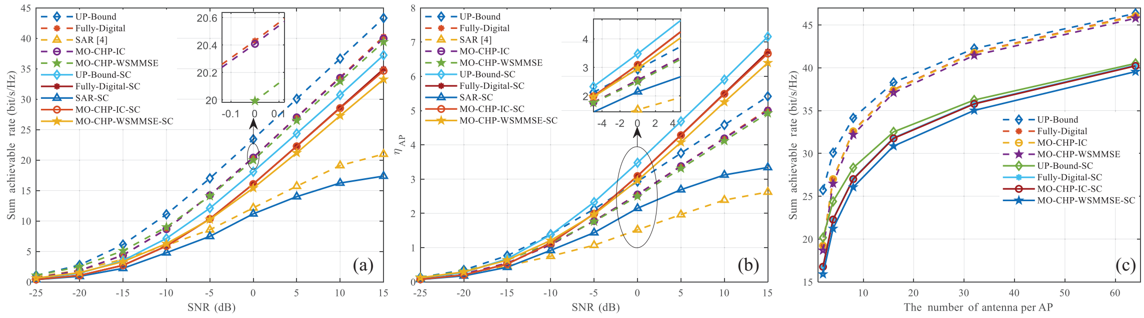

Figure 1 (Color online) Simulation results. (a) Comparison of the sum achievable rate versus SNR; (b) comparison of η_{AP} versus SNR; (c) comparison of the sum achievable rate versus the number of antenna per AP.

(IC) and weighted sum minimum mean square error (WS-MMSE) criteria. As indicated by P2, the second step is to decompose $\boldsymbol{F}_k$ into two parts, $\boldsymbol{F}^{\mathrm{RF}}$ and $\boldsymbol{F}_k^{\mathrm{DD}}$ based on the first step, considering the non-convex constraints. From [4], we have known that minimizing the objective function in P2 is approximately equivalent to maximizing spectral efficiency in P1.

MO-CHP. To remove non-convex constraints of analog precoder such as block diagonalization constraint and unit modulus constraint, we transform the optimization space from Euclidean space to Riemann space and define the Riemann gradient (see Appendix C for details).

Once we obtain the fully-digital precoders $\boldsymbol{F}^{\mathrm{IC}}$ or $\boldsymbol{F}^{\mathrm{WS\text{-}MMSE}}$ according to different criteria, the next task is to decompose them into analog and baseband digital precoders while satisfying the non-convex constraint, which is expressed as

$$(\mathrm{P2}): \min_{\boldsymbol{F}^{\mathrm{RF}}, \boldsymbol{F}_k^{\mathrm{DD}}} \left\| \boldsymbol{F}^{\mathrm{IC}} - \boldsymbol{F}^{\mathrm{RF}} \boldsymbol{F}^{\mathrm{DD}} \right\|_{\mathrm{F}}^2 \quad (4a)$$

$$\text{s.t.} \quad (3b)\text{–}(3e). \quad (4b)$$

In the process of solving problem P2, since the two variables $\boldsymbol{F}^{\mathrm{RF}}$ and $\boldsymbol{F}^{\mathrm{DD}}$ will affect each other, the core idea of our solution is alternate optimization, that is, solving problem P3 and P4 alternately until the convergence condition is satisfied.

$$(\mathrm{P3}): \min_{\boldsymbol{F}_k^{\mathrm{DD}}} \left\| \boldsymbol{F}^{\mathrm{IC}} - \boldsymbol{F}^{\mathrm{RF}} \boldsymbol{F}^{\mathrm{DD}} \right\|_{\mathrm{F}}^2 \quad (5a)$$

$$\text{s.t.} \quad (3b) \text{ and } (3c). \quad (5b)$$

In P3, due to $\boldsymbol{F}^{\mathrm{RF}}$ is fixed and there is no constraint of non-convex conditions such as constant modulus constraint. The problem of solving $\boldsymbol{F}^{\mathrm{DD}}$ is a classic convex optimization problem, whose closed-form least squares solution can be solved by Lagrangian multiplier method [5].

$$(\mathrm{P4}): \min_{\boldsymbol{F}^{\mathrm{RF}}} \left\| \boldsymbol{F}^{\mathrm{IC}} - \boldsymbol{F}^{\mathrm{RF}} \boldsymbol{F}^{\mathrm{DD}} \right\|_{\mathrm{F}}^2 \quad (6a)$$

$$\text{s.t.} \quad (3b)\text{–}(3e). \quad (6b)$$

In P4, although $\boldsymbol{F}^{\mathrm{DD}}$ is fixed, we still cannot solve $\boldsymbol{F}^{\mathrm{RF}}$ directly because the analog precoder $\boldsymbol{F}^{\mathrm{RF}}$ has non-convex constraint. In order to remove the non-convex constraints of analog precoder such as block diagonalization constraint and unit modulus constraint, we first vectorize the non-zero diagonal elements in the analog precoding matrix, then define the optimization space as the Riemann space. In the Riemann manifold, we determine the entire optimization search space as the product of V circles on the complex plane, that is, the product geometry of the Riemann manifold $\mathbb{C}^V$. Based on the Riemann gradient defined in Appendix D, we can use the conjugate gradient descent method to solve the analog precoder $\boldsymbol{F}^{\mathrm{RF}}$ with fixed $\boldsymbol{F}^{\mathrm{DD}}$. Besides, we analyze the computational complexity of the proposed MO-CHP scheme in Appendix D.

Conclusion and simulation results. To evaluate the performance of the proposed MO-CHP scheme (see Appendix E for details), Figure 1 shows the curve of the sum achievable rate versus signal-to-noise ratio (SNR), $\eta_{\mathrm{AP}} = \frac{R_{\mathrm{total}}}{M_{\mathrm{active}}}$, and the number of antenna per AP. It can be seen that the proposed MO-CHP scheme is very close to the fully-digital precoder, which shows the effectiveness of the manifold optimization algorithm in cell-free massive MIMO hybrid precoding.

To maximize the sum achievable rate, we propose a novel MO-CHP scheme for the cell-free mmWave massive MIMO system. To eliminate the non-convex constraint in hybrid precoding design, the analog precoder is vectorized and then defined as a complex circular manifold. Thus, the Riemann gradient descent method can be used for manifold optimization.

Acknowledgements This work was supported by National Natural Science Foundation of China (Grant No. 61671367) and Research Foundation of Science and Technology on Communication Networks Laboratory.

Supporting information Appendixes A–E. The supporting information is available online at info.scichina.com and link.springer.com. The supporting materials are published as submitted, without typesetting or editing. The responsibility for scientific accuracy and content remains entirely with the authors.

References

1 Chen S, Zhang J, Bjornson E, et al. Structured massive access for scalable cell-free massive MIMO systems. IEEE J Sel Areas Commun, 2021, 39: 1086–1100
2 Guo J C, Yu Q Y, Sun W B, et al. Robust efficient hybrid pre-coding scheme for mmWave cell-free and user-centric massive MIMO communications. IEEE Trans Wirel Commun, 2021, 20: 8006–8022
3 Yu X, Shen J C, Zhang J, et al. Alternating minimization algorithms for hybrid precoding in millimeter wave MIMO systems. IEEE J Sel Top Signal Process, 2016, 10: 485–500
4 Ayach O E, Rajagopal S, Abu-Surra S, et al. Spatially sparse precoding in millimeter wave MIMO systems. IEEE Trans Wirel Commun, 2014, 13: 1499–1513
5 Luo J, Fan J, Zhang J. MDL-AltMin: a hybrid precoding scheme for mmWave systems with deep learning and alternate optimization. IEEE Wirel Commun Lett, 2022, 11: 1925–1929