Local Uplink Processing in Cell-Free Networks: A New Approach

Vida Ranjbar, Marc Moonen, Sofie Pollin
Department of Electrical Engineering, KU Leuven, Belgium
Email: {vida.ranjbar, marc.moonen, sofie.pollin}@kuleuven.be

***Abstract*—Cell-free massive MIMO is a distributed version of cellular massive MIMO in which a user can be served by all or a subset of distributed access points (APs). In other words, there are no cell boundaries between APs as opposed to conventional cellular networks. Such a network can enhance user experience in terms of spectral efficiency by suppressing interference from other users using different local or centralized combining techniques. However, increasing the number of users and APs make some of these techniques unscalable. In this paper, we consider the uplink performance of the cell-free massive MIMO using an estimation-based scalable distributed combining method and compare its performance with existing distributed methods that rely on full or partial channel state information. Simulation results show performance improvements in terms of computational complexity and spectral efficiency under dense user distribution.**

***Index Terms*—Scalable cell-free Massive MIMO implementation, local combining vectors**

I. Introduction

Massive MIMO is a promising technology for future wireless networks as it can increase the spectral efficiency of the users by coherently combining the signals received by or transmitted from a massive number of antennas at the APs. In cell-free massive MIMO, which is a new version of network MIMO, the antennas of each APs are distributed, exploiting the spatial diversity to mitigate large-scale fading. In this paper, we consider the uplink of a cell-free massive MIMO network.

In the original cell-free massive MIMO [1], each AP serves all the users and has network-wide channel state information (CSI). These assumptions are not practical, as they require huge fronthaul signaling for both CSI and data sharing, in addition to a large computational cost when it comes to more sophisticated reception techniques than maximum ratio combining (MRC) such as centralized MMSE (CMMSE) [2], [3].

To reduce the computational cost, local MMSE (LMMSE) was proposed, computing local combining weights that can be applied in each AP locally, but it still requires full network-wide CSI from all AP to all users [3]. To enhance scalability as a function of user density, [2] proposed a dynamic cooperation clustering (DCC) method that selects a subset of AP for serving each user. This method is also referred to as user-centric cell-free massive MIMO. The user selection can be applied to the CMMSE, or LMMSE, resulting in a PMMSE or LPMMSE variant of cell-free massive MIMO. In this partial processing, the reception of a particular user is defined by only a subset of other users. However, ignoring the effect of the users outside this subset, makes partial combining methods susceptible to interference in scenarios with a large number of users. In this paper, we propose an alternative method to determine the local combining vectors, not relying on full network-wide CSI, but still achieving interference suppression for all users.

In this paper, we take a new approach in computing MMSE-based combining vectors. We estimate the LMMSE with less computational complexity and even without the knowledge of all channels to all users. Our results show that under full CSI and high user density, the complexity and performance of our method are similar to the ideal LMMSE. Compared to LPMMSE, our method has the advantage that it can eliminate interference to all users in the network while serving and having CSI of only a subset of users. Especially in dense scenarios, our method hence performs better than LPMMSE, while assuming the same level of CSI.

The rest of this paper is organized as follows. In section II, we introduce the system model used in this paper and review channel estimation and uplink data reception. In section III, we review state-of-art local combining vectors and then state the problem associated with these methods. Then, we come up with an estimation-based combining vector which is supposed to solve the problem of high interference inter-AP users and high complexity associated with LPMMSE and LMMSE respectively.

Notaion: In this paper, a boldface lowercase letter denotes a vector and a boldface uppercase letter denotes a matrix. The superscripts $()^H$, $()^T$ and $()^*$ denote conjugate transpose, transpose, and conjugate operation respectively. The cardinality of set $\mathcal{S}$ and expectation operators are indicated by $|\mathcal{S}|$ and $\mathbb{E}\{.\}$, respectively. Identity matrix with size $N \times N$ is shown as $\mathbf{I}_N$. $\mathcal{N}_{\mathbb{C}}(0, \mathbf{R})$ indicates a multivariate circulary symetric complex Gaussian distribution with $\mathbf{R}$ as correlation matrix.

II. System Model

We consider the uplink of a cell-free massive MIMO network with K single-antenna users, L APs and N antennas per AP distributed, over a coverage area. The channel between a particular user k and AP l is denoted by $\mathbf{h}_{kl} \in \mathbb{C}^N$ and

$$\mathbf{h}_{kl} \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{R}_{kl}) \tag{1}$$

where $\mathbf{R}_{kl}$ is the spatial correlation matrix.

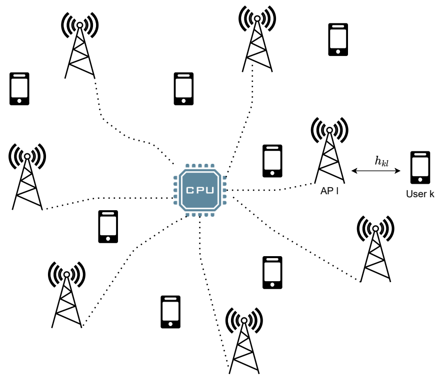

Fig. 1. Cell-Free Network

We assume that users are operating based on a Time Division Duplex (TDD) protocol in which every coherence block is divided into pilot and data transmission phases. Suppose τ_c samples can be transmitted in each coherence block. Then $\tau_c = \tau_p + \tau_d$ with τ_p number of pilot samples and τ_d number of data samples. During pilot transmission which occurs at the beginning of each coherence block, users send their pilot sequence to the APs and the APs will estimate the channel from each user using the received pilot sequence. As there may not be one pilot per user, some users may have to share pilot which results in pilot contamination [4].

A. Pilot assignment and User-AP association

For methods such as PMMSE and LPMMSE mentioned earlier, a user-AP association algorithm is used, which forms clusters of APs, each cluster serving one particular user. To decide which APs serve a particular user, we use the pilot assignment and cluster formation algorithm proposed in [2] in which eventually a user is served by a master AP and a subset of nearby APs, together forming a cluster. For a detailed description of the algorithm, the reader is referred to [2].

B. Pilot transmission and channel estimation

For channel estimation, τ_p mutually orthogonal τ_p-length pilots $\{\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_{\tau_p}\}$ are used. Suppose user k is assigned pilot $\boldsymbol{\phi}_{t_k}$ with $t_k \in \{1, 2, \ldots, \tau_p\}$. The set of users sharing pilot $\boldsymbol{\phi}_{t_k}$ with user k is $\mathcal{S}_{t_k}$. When users send their pilot to the APs, the received pilot matrix at AP l will be:

$$\mathbf{Y}_l^p = \sum_{i=1}^{K} \sqrt{p_i}\mathbf{h}_{il}\boldsymbol{\phi}_{t_i}^T + \mathbf{N}_l \tag{2}$$

in which $\mathbf{Y}_l^p \in \mathbb{C}^{N\times\tau_p}$ is the received pilot matrix at AP l, p_i is the transmit power of user i , $\boldsymbol{\phi}_{t_i} \in \mathbb{C}^{\tau_p}$ is the pilot sequence transmitted by user i and $\mathbf{N}_l \in \mathbb{C}^{N\times\tau_p}$ is the receiver noise at AP l with elements independently distributed as $\mathcal{N}_{\mathbb{C}}(0, \sigma^2)$.

For the estimation of user k's channel, AP l first correlates the received pilot signal matrix with the normalized pilot sequence $\boldsymbol{\phi}_{t_k}/\sqrt{\tau_p}$ sent by user k as follows (with $\boldsymbol{\phi}_{t_k} * \boldsymbol{\phi}_{t_k}^* = \tau_p$):

$$\mathbf{y}_{t_k l} = \mathbf{Y}_l^p \boldsymbol{\phi}_{t_k}^*/\sqrt{\tau_p} = \sum_{i\in\mathcal{S}_{t_k}} \sqrt{p_i\tau_p}\mathbf{h}_{il} + \mathbf{N}_l\boldsymbol{\phi}_{t_k}^*/\sqrt{\tau_p}. \tag{3}$$

We define $\mathbf{n}_{t_k l} \triangleq \mathbf{N}_l\boldsymbol{\phi}_{t_k}^*/\sqrt{\tau_p} \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2\mathbf{I}_N)$, as the entries of $\mathbf{N}_l$ are independent normal random variables and the entries of $\mathbf{n}_{t_k l}$ are a weighted sum of the entries of $\mathbf{N}_l$.

From $\mathbf{y}_{t_k l}$, the MMSE estimate of channel $\mathbf{h}_{kl}$ will be:

$$\hat{\mathbf{h}}_{kl} = \sqrt{p_k}\mathbf{R}_{kl}\boldsymbol{\psi}_{t_k l}^{-1}\mathbf{y}_{t_k l}. \tag{4}$$

With $\boldsymbol{\psi}_{t_k l}$ defined as below :

$$\boldsymbol{\psi}_{t_k l} = \mathbb{E}\{\mathbf{y}_{t_k l}(\mathbf{y}_{t_k l})^H\} = \sum_{i\in\mathcal{S}_{t_k}} \tau_p p_i \mathbf{R}_{il} + \sigma^2\mathbf{I}_N. \tag{5}$$

The channel estimation error $\tilde{\mathbf{h}}_{kl} = \mathbf{h}_{kl} - \hat{\mathbf{h}}_{kl} \sim \mathcal{N}_{\mathbb{C}}(0, \mathbf{C}_{kl})$ is independent of $\hat{\mathbf{h}}_{kl}$ and

$$\mathbf{C}_{kl} = \mathbb{E}\{\tilde{\mathbf{h}}_{kl}\tilde{\mathbf{h}}_{kl}^H\} = \mathbf{R}_{kl} - p_k\tau_p\mathbf{R}_{kl}\boldsymbol{\psi}_{t_k l}^{-1}\mathbf{R}_{kl}. \tag{6}$$

C. Uplink data reception

During uplink reception, the received signal vector $\mathbf{y}_l \in \mathbb{C}^N$ at AP l is

$$\mathbf{y}_l = \sum_{i=1}^{K} \mathbf{h}_{il}s_i + \mathbf{n}_l, \tag{7}$$

where $s_i \sim \mathcal{N}_{\mathbb{C}}(0, p_i)$ is the signal transmitted by user i and $\mathbf{n}_l \sim \mathcal{N}_{\mathbb{C}}(0, \sigma^2\mathbf{I}_N)$ is the receiver noise at AP l.

Suppose after user-AP association, the subset of APs serving user k is $\mathcal{M}_k$ and the subset of users served by AP l is $\mathcal{D}_l$. Then matrix $\mathbf{D}_{kl}$ is an $N \times N$ matrix defined as follows:

$$\mathbf{D}_{kl} = \begin{cases} \mathbf{I}_N & k \in \mathcal{D}_l \\ 0 & k \notin \mathcal{D}_l \end{cases} \tag{8}$$

In distributed combining, assuming AP l selects a combining vector v_{kl} for a particular user k, then it obtains a local estimate of user k's signal as follows:

$$\hat{s}_{kl} = \mathbf{v}_{kl}^H\mathbf{D}_{kl}\mathbf{y}_l. \tag{9}$$

After local estimation, all the APs send their local estimate for the users they are serving to the central processing unit (CPU) which takes care of collective signal detection for each user. So for every user k we have

$$\hat{s}_k = \sum_{l=1}^{L} \hat{s}_{kl}. \tag{10}$$

Note that in a partial combining method, the local estimates of the users that are not served by AP l will be zero, based on (8) and (9).

By defining a collective combining vector for user k as $\mathbf{v}_k = [\mathbf{v}_{k1}^T\mathbf{v}_{k2}^T \ldots \mathbf{v}_{kL}^T]^T \in \mathbb{C}^{NL}$, a collective channel of user k as $\mathbf{h}_k = [\mathbf{h}_{k1}^T\mathbf{h}_{k2}^T \ldots \mathbf{h}_{kL}^T]^T \in \mathbb{C}^{NL}$ and

$\mathbf{D}_k$ =diag$(\mathbf{D}_{k1}, \mathbf{D}_{k2}, \dots, \mathbf{D}_{kL})$, a block diagonal $NL \times NL$ matrix, (10) is rewritten as

$$\hat{s}_k = \sum_{l=1}^{L} \mathbf{v}_{kl}^H \mathbf{D}_{kl} \mathbf{y}_l = \mathbf{v}_k^H \mathbf{D}_k \mathbf{h}_k s_k + \sum_{i \neq k} \mathbf{v}_k^H \mathbf{D}_k \mathbf{h}_i s_i + \mathbf{v}_k^H \mathbf{D}_k \mathbf{n} \tag{11}$$

where $\mathbf{n} = [\mathbf{n}_1^T \mathbf{n}_2^T \dots \mathbf{n}_L^T]^T$ is the collective noise.

D. Spectral efficiency

A commonly used bound for spectral efficiency in distributed combining methods is the "use and then forget bound" [2]- [4],

$$\text{SE}_k = \frac{\tau_d}{\tau_c} \log_2(1 + \text{SINR}_k) \tag{12}$$

where SINR_k is defined as below:

$\text{SINR}_k =$

$$\frac{p_k |\mathbb{E}\{\mathbf{v}_k^H \mathbf{D}_k \mathbf{h}_k\}|^2}{\sum_{i=1}^{K} p_i \mathbb{E}\{|\mathbf{v}_k^H \mathbf{D}_k \mathbf{h}_i|^2\} - p_k |\mathbb{E}\{\mathbf{v}_k^H \mathbf{D}_k \mathbf{h}_k\}|^2 + \sigma^2 \mathbb{E}\{\|\mathbf{D}_k \mathbf{v}_k\|^2\}} \tag{13}$$

The rationale behind the bound is that CSI is only used in the APs when designing the combining vectors, but not for the collective signal detection at the CPU.

III. Combining Vectors

In this section, we define some commonly-used local combining vectors and propose an estimation-based combining vector as an alternative.

We assume that after user-AP association, user k is served by AP l.

A. Maximum ratio combining

The simplest form of combining is MRC in which

$$\mathbf{v}_{kl} = \hat{\mathbf{h}}_{kl}. \tag{14}$$

With MRC, the power of the desired signal $\mathbf{v}_k^H \mathbf{h}_k$ is maximized but it does not necessarily suppress the interference from other users. It is a sub-optimal yet low-complexity method.

B. Local MMSE combining

In LMMSE, each AP uses local channel estimates to calculate a local MMSE combining vector for user k. The LMMSE minimizes the squared error between the transmitted signal and the locally estimated signal, i.e.:

$$\mathbf{v}_{kl}^{\text{LMMSE}} = \arg\min_{\mathbf{v}_{kl}} \mathbb{E}\{|s_k - \mathbf{v}_{kl}^H \mathbf{y}_l|^2\}, \tag{15}$$

which has the following solution:

$$\mathbf{v}_{kl}^{\text{LMMSE}} = \mathbf{R}_{yy}^{-1} \mathbf{R}_{sy}, \tag{16}$$

where $\mathbf{R}_{yy}$ is the correlation matrix of the received signal vector and $\mathbf{R}_{sy}$ is the correlation vector between transmitted signal s_k and received signal vector $\mathbf{y}_l$ given the channel estimates. In LMMSE, each AP uses the channel estimates for all the users so that $\mathbf{R}_{yy}$ is defined as

$$\mathbf{R}_{yy} = \mathbb{E}\{\mathbf{y}_l \mathbf{y}_l^H | \hat{\mathbf{h}}_{il}, \forall i\} = \sum_{i=1}^{K} p_i (\hat{\mathbf{h}}_{il} \hat{\mathbf{h}}_{il}^H + \mathbf{C}_{il}) + \sigma^2 \mathbf{I}_N \tag{17}$$

and $\mathbf{R}_{sy}$ is defined as

$$\mathbf{R}_{sy} = \mathbb{E}\{\mathbf{y}_l s_k^* | \hat{\mathbf{h}}_{il}, \forall i\} = p_k \hat{\mathbf{h}}_{kl}. \tag{18}$$

Then LMMSE combining vector of user k is then given as:

$$\mathbf{v}_{kl} = p_k \left(\sum_{i=1}^{K} p_i (\hat{\mathbf{h}}_{il} \hat{\mathbf{h}}_{il}^H + \mathbf{C}_{il}) + \sigma^2 \mathbf{I}_N \right)^{-1} \hat{\mathbf{h}}_{kl}. \tag{19}$$

C. Local partial MMSE combining

An alternative sub-optimal method to reduce complexity is LPMMSE in which AP l uses only the channel estimates of the users it is serving to compute the combining vector of user k. In this case, the combining vector is given as:

$$\mathbf{v}_{kl} = p_k \left(\sum_{i \in \mathcal{D}_l} p_i (\hat{\mathbf{h}}_{il} \hat{\mathbf{h}}_{il}^H + \mathbf{C}_{il}) + \sigma^2 \mathbf{I}_N \right)^{-1} \hat{\mathbf{h}}_{kl}. \tag{20}$$

Although LPMMSE may be a low-complexity yet powerful method, the fact that it ignores the users that are not served by AP l makes it susceptible in high user density scenarios. In such scenarios, LPMMSE performance may deviate significantly from the LMMSE performance.

D. Local MMSE combining vector based on estimation of R_{yy}

As an alternative to (17), $\mathbf{R}_{yy}$ can be estimated from the received signal vectors $\mathbf{y}_l$ at AP l as follows

$$\mathbf{R}_{yy}^e = \frac{\sum_{n=1}^{\tau_s} \mathbf{y}_l^n \mathbf{y}_l^{nH}}{\tau_s}. \tag{21}$$

Where $\mathbf{R}_{yy}^e$ denotes the estimation of $\mathbf{R}_{yy}$ and $\mathbf{y}_l^n$ is the received signal vector at AP l for the n_{th} of τ_d data samples in one coherence block. τ_s is the number of data samples that are used in the estimation of $\mathbf{R}_{yy}$ and $\tau_s \leq \tau_d$. Then, we replace $\mathbf{R}_{yy}$ with $\mathbf{R}_{yy}^e$ in (16) and end up with a combining vector as follows:

$$\mathbf{v}_{kl}^e = p_k (\mathbf{R}_{yy}^e)^{-1} \hat{\mathbf{h}}_{kl}. \tag{22}$$

Note that (21) implies that the AP starts decoding data after the reception of τ_s samples, which may impose a delay. However, the fact that in cell-free networks, the antennas are distributed and the number of antennas in each AP is relatively small, makes it possible to estimate $\mathbf{R}_{yy}$ with a small number of received signal vectors (small value of τ_s) at any AP, therefore the delay will be small.

IV. Computational Complexity of Computing Different Combining Vectors

In table I, we will see the number of complex multiplications per user in one coherence block is given for different methods. As we can see, the computational complexity of the LMMSE scales with the number of users K. Also, we can see that the computational complexity of the proposed method is proportional to τ_s and independent of K. As in cell-free networks, the number of antenna per AP is small, τ_s can also be small.

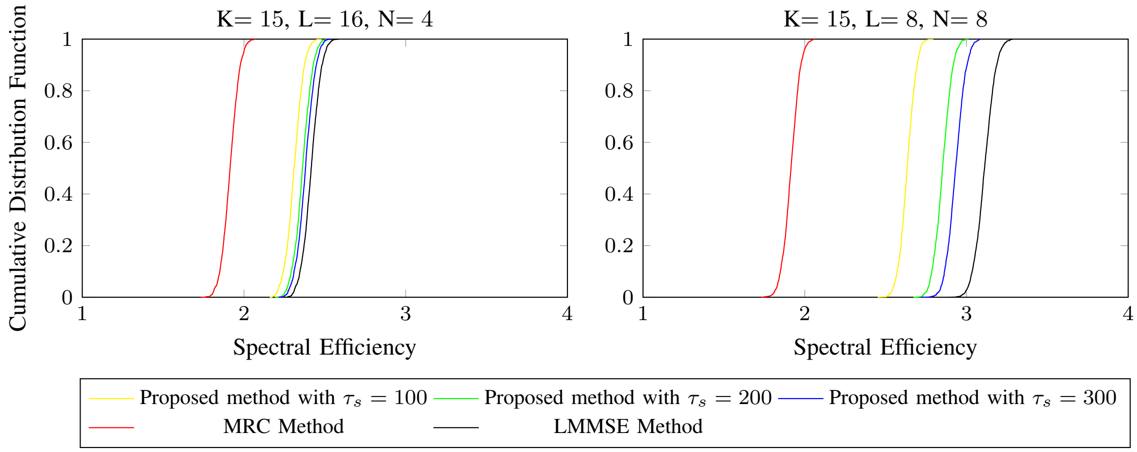

Fig. 2. Two scenarios with different number of antenna per AP using measurement data

TABLE I
COMPUTATIONAL COMPLEXITY OF COMBINING VECTOR CALCULATION PER USER IN ONE COHERENCE BLOCK

Proposed Method with User-AP association
$\frac{N^2+N}{2} * \lvert\mathcal{M}_k\rvert * \tau_s + (\frac{N^3-N}{3} + N^2) * \lvert\mathcal{M}_k\rvert$
Proposed Method with All APs serveing all users
$\frac{N^2+N}{2} * L * \tau_s + (\frac{N^3-N}{3} + N^2) * L$
LPMMSE
$\frac{N^2+N}{2} \sum_{l \in \mathcal{M}_k} \lvert\mathcal{D}_l\rvert + (\frac{N^3-N}{3} + N^2)\lvert\mathcal{M}_k\rvert$
LMMSE
$\frac{N^2+N}{2} * L * K + (\frac{N^3-N}{3} + N^2) * L$

V. NUMERICAL RESULTS

In this section, to analyze and compare different methods discussed in the previous section, we provide numerical results using measured data from the KU Leuven Massive MIMO testbed as well as simulated data.

A. KU Leuven testbed data

We use channel measurements from an indoor distributed massive MIMO testbed. In the testbed, we have 64 antennas. We divide them into 16 APs each having 4 antennas and 8 APs each having 8 antennas. Considering 15 users in the room, we plot the cumulative distribution function (CDF) of the spectral efficiency bound of different methods in Fig. 2. We assume perfect CSI for this part and that a particular user is served by all APs. As expected, Fig. 2 shows that, by increasing the number of samples τ_s used in (21), the proposed method approaches the LMMSE . Also, we observe that, by having a smaller number of antennas per APs, the convergence of the proposed method's performance is faster towards the performance of LMMSE.

B. Simulation data

To be able to evaluate the different methods in networks with more antennas and users, we used Matlab to simulate such scenarios. In our simulation, we consider $L = 100$ and $N = 4$ in a $100 \times 100 m^2$ area. We assume imperfect CSI at the APs. The computational complexity of the different methods is shown in figure 3. We can observe that:

- The LMMSE computational complexity is increasing with K which makes it unscalable.
- The proposed method with all APs serving all users has a constant computational complexity with K increasing.
- The LPMMSE and the proposed method with the user-AP association have much smaller computational complexity than the two other schemes. Furthermore, increasing K decreases the computational complexity of these two methods. This is because by increasing K, the number of AP serving a particular user k, $\lvert\mathcal{M}_k\rvert$ decreases, as every APs prefers to serve up to τ_p users, based on user-AP association that we used [2].

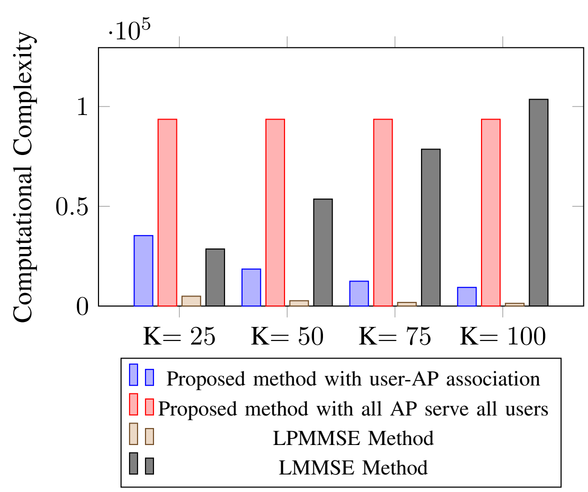

Fig. 3. Computational complexity of different methods. $L = 100, N = 4, K = \{25, 50, 75, 100\}$ and $\tau_s = \tau_d = 90$

The CDF function of spectral efficiency is shown in Fig. 4. We observe that with increasing K, the performance of

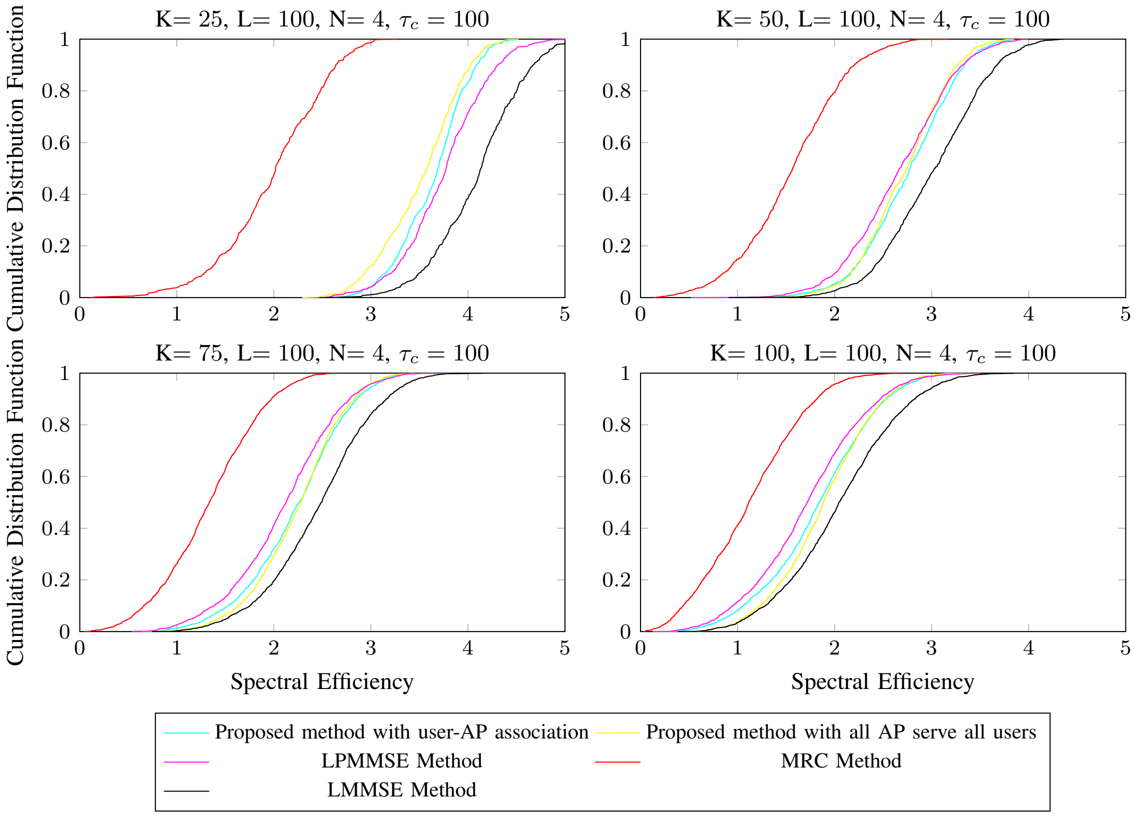

Fig. 4. SE CDF, $L = 100, N = 4$, with increasing number of users from $K = 25$ to $K = 100$ and $\tau_s = \tau_d = 90$

the LPMMSE will degrade and the proposed method will approach the performance of the LMMSE. This behavior is expected from the LPMMSE, as it ignores the effect of non-served users in the combining vector calculation. On the other hand, with an increasing K, the proposed method's performance approaches the performance of the LMMSE and it outperforms the performance of the LPMMSE. Also, we can enhance the performance of the proposed method further by increasing the parameter τ_s in a scenario with a larger τ_c, but at the cost of more computational complexity and delay.

VI. Conclusion

In this paper, we have focused on the processing in uplink cell-free Massive MIMO systems and have proposed an estimation-based combining vector in which we use received signal vectors to estimate the correlation matrix part of the LMMSE. The advantage of the proposed method is that, unlike the LMMSE, no channel state information is needed for the other users when calculating the combining vector of a particular user. Our method has been shown to outperform the LPMMSE in dense user scenarios. Also, the computational complexity of the proposed method is independent of the number of users in the system, which makes it scalable. Furthermore, by assuming a small number of antennas in each AP which is the case in cell-free scenarios, a good estimation of the correlation matrix can be obtained using only a small number of received signal vectors.

References

[1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-Free Massive MIMO Versus Small Cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[2] E. Björnson and L. Sanguinetti, "Scalable Cell-Free Massive MIMO Systems," *IEEE Transactions on Communications*, vol. 68, no. 7, pp. 4247–4261, 2020.

[3] E. Björnson and L. Sanguinetti, "Making Cell-Free Massive MIMO Competitive With MMSE Processing and Centralized Implementation," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 77–90, 2020.

[4] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO networks: Spectral, energy, and hardware efficiency," *Foundations and Trends® in Signal Processing*, vol. 11, no. 3-4, pp. 154–655, 2017.

[5] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

[6] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-Free Massive MIMO: Uniformly great service for everyone," in *IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pp. 201–205, 2015.

[7] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: ten myths and one critical question," *IEEE Communications Magazine*, vol. 54, no. 2, pp. 114–123, 2016.

[8] J. Zhang, S. Chen, Y. Lin, J. Zheng, B. Ai, and L. Hanzo, "Cell-Free Massive MIMO: A New Next-Generation Paradigm," *IEEE Access*, vol. 7, pp. 99878–99888, 2019.