Generalized Partial Full-Pilot Zero-Forcing Combining for Distributed Massive MIMO

Mohd Saif Ali Khan*, Karthik R.M.† and Samar Agnihotri*
*School of Computing & EE, Indian Institute of Technology Mandi, HP, India
†Ericsson India Pvt. Ltd., Chennai, TN, India
Email: d21013@students.iitmandi.ac.in, karthik.r.m@ericsson.com, samar@iitmandi.ac.in

***Abstract*—Efficient uplink processing in distributed massive multiple-input multiple-output (D-mMIMO) systems requires effective local combining to significantly mitigate inter-user interference. Recent zero-forcing (ZF) based combining schemes, such as partial full-pilot ZF (PFZF) and protected weak PFZF (PWPFZF), rely on heuristic threshold-based user grouping that may lead to inefficient utilization of spatial degrees of freedom across access points. To address this limitation, we propose an adaptive pilot-aware local combining scheme, generalized PFZF (G-PFZF), that dynamically allocates spatial degrees of freedom based on local channel conditions and replaces heuristic grouping with a decentralized pilot-level optimization framework. Numerical results demonstrate that the proposed G-PFZF scheme achieves significantly higher sum spectral efficiency compared to PFZF and PWPFZF.**

I. Introduction

Distributed massive MIMO (D-mMIMO) systems have emerged as a promising architecture for next-generation wireless systems, as they may provide uniform service quality across the network by enabling numerous distributed access points (APs) to coherently serve user equipments (UEs) [1], [2]. However, achieving scalable signal processing in uplink D-mMIMO systems remains challenging because centralized processing architectures require significant fronthaul signaling and computational overhead [3], [4].

In the recent years, several (partial) centralized combining architectures have been proposed to mitigate inter-user interference in D-mMIMO systems by jointly processing signals or channel statistics across multiple APs. [3], [5]–[8]. While these approaches may achieve strong interference suppression, they require the central processing unit (CPU) to collect signals or channel statistics from multiple APs, resulting in substantial fronthaul and computational costs. To reduce such costs, a two-layer decoding architecture has been proposed in [3], [6], where each AP performs local combining while the CPU performs the second-stage decoding. In this work, we discuss the design of a new local combining scheme and take up the design of a new decoding scheme in a separate work.

Local maximum ratio (MR) combining [1], offers low computational complexity, but suffers from limited interference suppression capability. The local partial minimum mean square error (LP-MMSE) combining scheme [3], [6], achieves better performance. However, it does not provide tractable closed-form spectral efficiency (SE) expressions, making system-level analysis and optimization challenging.

To address these limitations of the MR and LP-MMSE combining, local combining schemes such as full-pilot ZF (FZF), partial FZF (PFZF), and protected weak PFZF (PW-PFZF) [9], [10], exploit the structure of pilot reuse to construct combining vectors that partially suppress interference while retaining analytical tractability. In these schemes, each AP classifies the UEs into strong and weak groups using large-scale fading coefficients (LSFCs), allowing some spatial degrees of freedom to be allocated for interference suppression while preserving array gain for weaker UEs. Although these schemes provide better interference suppression compared to the MR combining, their user grouping decisions rely on heuristic threshold-based rules. As a result, the number of spatial DoF allocated for interference mitigation is fixed by predetermined thresholds rather than being adapted to the local channel and interference conditions at each AP.

To address such shortcomings of these zero-forcing schemes, we propose an adaptive pilot-level combining scheme, generalized PFZF (G-PFZF), that dynamically allocates spatial degrees of freedom based on locally observed channel conditions. Since users sharing the same pilot cannot be fully distinguished by an AP, grouping decisions are performed at the pilot-level rather than at the user level. Here, each AP independently partitions its set of pilots into strong and weak groups by solving a local sum SE optimization problem. Assigning a pilot to the strong group dedicates a spatial degree of freedom (DoF) to mitigate interference via ZF combining, whereas assigning a pilot to the weak group preserves that DoF for signal enhancement via MR combining. This novel pilot-level optimization allows each AP to adaptively operate across a range of combining schemes from pure MR to FZF. We provide a novel solution approach to this optimization problem and also, a closed-form SE expression for the G-PFZF combining. Performance evaluation of the proposed G-PFZF scheme establishes that it significantly outperforms PFZF and PWPFZF schemes.

Organization: The remainder of the paper is organized as follows. Section II describes the system model. Section III presents the proposed combining scheme and derives its corresponding closed-form SE expression. Section IV provides the performance evaluation. Finally, Section V concludes the paper and outlines potential directions for future work.

II. System Model

The system consists of T single-antenna UEs; and M APs, each with A antennas. Both APs and UEs are distributed uniformly within the coverage area, and APs serve the UEs using the same time-frequency resources. The channel follows a

block fading model with a coherence block of L_c symbols, of which L_p symbols are used for pilot training. The small-scale fading is represented by a Rayleigh fading vector $\mathbf{h}_{mt}\in\mathbb{C}^{A\times 1}$, and the LSFC, encompassing both path-loss and shadowing effects, is denoted by β_{mt}. The channel vector between AP m and UE t is defined as $\mathbf{g}_{mt}=\beta_{mt}^{1/2}\mathbf{h}_{mt}$. We assume that $\mathbf{h}_{mt}\sim\mathcal{CN}(0,\mathbf{I}_A)$, for all m and t, is an independent and identically distributed (i.i.d) complex Gaussian random vector, thus $\mathbf{g}_{mt}\sim\mathcal{CN}(0,\beta_{mt}\mathbf{I}_A)$. We consider user-centric design [6], where $\mathcal{T}_m$ denotes the set of UEs served by AP m and $\mathcal{M}_t$ denotes the set of APs that are serving UE t.

A. Channel Estimation

A pilot $\sqrt{L_p}\boldsymbol{\psi}_{i_t}\in\mathbb{C}^{L_p\times 1}$ is transmitted by each UE t during the uplink training stage, with $|\boldsymbol{\psi}_{i_t}|^2=1$, where i_t is the pilot index of UE t. At the m-th AP, the received signal is:

$$\mathbf{y}_m^{pilot}=\sum\nolimits_{t=1}^{T}\sqrt{p_t^pL_p}\mathbf{g}_{mt}\boldsymbol{\psi}_{i_t}^H+\mathbf{N}_m,$$

where $\mathbf{N}_m\in\mathbb{C}^{A\times L_p}$ represents the additive white Gaussian noise (AWGN) matrix characterized by i.i.d. complex Gaussian entries. Additionally, p_t^p represents the normalized pilot power for UE t during pilot transmission. The MMSE estimate of the true channel vector $\mathbf{g}_{mt}$ is provided by [1]:

$$\hat{\mathbf{g}}_{mt}=\sqrt{p_t^pL_p}\beta_{mt}\theta_{mi_t}^{-1}\mathbf{y}_m^{pilot}\boldsymbol{\psi}_{i_t},\tag{1}$$

where $\theta_{mi_t}=\big(\sum_{k=1}^{T}p_k^pL_p\beta_{mk}\left|\boldsymbol{\psi}_{i_t}^H\boldsymbol{\psi}_{i_k}\right|^2+1\big)$. The estimate $\hat{\mathbf{g}}_{mt}$ and estimated error $\tilde{\mathbf{g}}_{mt}=\mathbf{g}_{mt}-\hat{\mathbf{g}}_{mt}$ are independent Gaussian with distributions $\hat{\mathbf{g}}_{mt}\sim\mathcal{CN}(0,\gamma_{mt}\mathbf{I}_A)$ and $\tilde{\mathbf{g}}_{mt}\sim\mathcal{CN}(0,(\beta_{mt}-\gamma_{mt})\mathbf{I}_A)$, where

$$\gamma_{mt}=\mathbb{E}\big\{|[\hat{\mathbf{g}}_{mt}]_l|^2\big\}=p_t^pL_p\beta_{mt}^2\theta_{mi_t}^{-1}.\tag{2}$$

Remark 1: The channel estimates of UEs sharing the same pilot sequence are linearly dependent. As a result, an AP cannot fully distinguish among these UEs, leading to residual pilot contamination.

The channel estimate matrix at AP m is denoted by $\hat{\mathbf{G}}_m=[\hat{\mathbf{g}}_{m1},\hat{\mathbf{g}}_{m2},\ldots,\hat{\mathbf{g}}_{mT}]\in\mathbb{C}^{A\times T}$. Owing to pilot reuse, the columns of $\hat{\mathbf{G}}_m$ exhibit linear dependence, rendering the matrix $\hat{\mathbf{G}}_m$ rank-deficient. The full rank matrix is designed by removing the linearly dependent columns and by including only those columns that contain linearly independent channel estimates, representing one vector for each pilot sequence. Thus, the matrix $\bar{\mathbf{G}}_m$ is designed as:

$$\bar{\mathbf{G}}_m=\mathbf{y}_m^{pilot}\boldsymbol{\Psi},\quad\in\mathbb{C}^{A\times L_p},$$

where $\boldsymbol{\Psi}=[\boldsymbol{\psi}_1,\boldsymbol{\psi}_2,\ldots,\boldsymbol{\psi}_{L_p}]$. Then $\hat{\mathbf{g}}_{mt}=c_{mt}\bar{\mathbf{G}}_m\mathbf{e}_{i_t}$, where $\mathbf{e}_{i_t}$ is the i_t-th column of identity matrix $\mathbf{I}_{L_p}$.

B. Uplink Transmission and Spectral Efficiency

Let x_t denote the unit power uplink data signal transmitted by UE t. The uplink normalised data power corresponding to UE t and AWGN vector are denoted by p_t^u and $\mathbf{n}_m\sim\mathcal{N}_{\mathbb{C}}(0,\mathbf{I}_A)$, respectively. The signal received at AP m is

$$\mathbf{y}_m=\sum\nolimits_{t=1}^{T}\mathbf{g}_{mt}\sqrt{p_t^u}x_t+\mathbf{n}_m.$$

The locally estimated data symbol for UE t at AP m is

$$\hat{x}_{mt}=\mathbf{v}_{mt}^H\mathbf{y}_m,\tag{3}$$

where $\mathbf{v}_{mt}\in\mathbb{C}^{A\times 1}$ denotes the local combining vector for UE t at AP m. These estimates are forwarded to the CPU, which aggregates them using large-scale fading decoding (LSFD) to obtain the final estimate of x_t as

$$\begin{aligned}\hat{x}_t=&\sum_{m\in\mathcal{M}_t}\sqrt{p_t^u}a_{mt}\mathbf{v}_{mt}^H\hat{\mathbf{g}}_{mt}x_t+\sum_{m\in\mathcal{M}_t}\sqrt{p_t^u}a_{mt}\mathbf{v}_{mt}^H\tilde{\mathbf{g}}_{mt}x_t\\&+\sum_{\substack{k=1\\k\neq t}}^{T}\sum_{m\in\mathcal{M}_t}\sqrt{p_k^u}a_{mt}\mathbf{v}_{mt}^H\mathbf{g}_{mk}x_k+\sum_{m\in\mathcal{M}_t}a_{mt}\mathbf{v}_{mt}^H\mathbf{n}_m.\end{aligned}\tag{4}$$

Based on (4), the achievable uplink SE is obtained using the bounding technique and the corresponding Shannon capacity lower bound [11] is given by Theorem 1.

Theorem 1 *[9], [11] : A lower bound on the uplink ergodic SE for UE t is given by*

$$\mathrm{SE}_t^u=L_u\log_2\big(1+\mathrm{SINR}_t\big),\tag{5}$$

where $L_u=\left(\frac{1-\frac{L_p}{L_c}}{2}\right)$ *and* SINR_t *is*

$$\mathrm{SINR}_t=\frac{p_t^u|\mathrm{DS}_t|^2}{\sum\limits_{k=1}^{T}p_k^u\mathbb{E}\left\{|\mathrm{UI}_{tk}|^2\right\}-p_t^u|\mathrm{DS}_t|^2+\mathbb{E}\left\{|\mathrm{GN}_t|^2\right\}},\tag{6}$$

$$\begin{aligned}\textit{where,}\quad\mathrm{DS}_t&=\sum\nolimits_{m\in\mathcal{M}_t}a_{mt}\mathbb{E}\left\{\mathbf{v}_{mt}^H\hat{\mathbf{g}}_{mt}\right\},\\\mathrm{UI}_{tk}&=\sum\nolimits_{m\in\mathcal{M}_t}a_{mt}\mathbf{v}_{mt}^H\boldsymbol{g}_{mk},\\\mathrm{GN}_t&=\sum\nolimits_{m\in\mathcal{M}_t}a_{mt}\mathbf{v}_{mt}^H\boldsymbol{n}_m.\end{aligned}$$

The SE lower bound, in (5), is valid regardless of the combining scheme used. Here a_{mt} can be calculated using the optimal LSFD method as in [6], [9], [10].

III. Adaptive Combining Analysis

In this section, we discuss the proposed G-PFZF combining scheme to overcome the limitations of threshold-based UE partitioning and derive closed-form expression for the corresponding uplink SE.

A. Generalized Partial Full-Pilot Zero-Forcing Scheme

We propose the G-PFZF scheme, where each AP independently designs combining vectors at the pilot level. The pilot-level combining design is governed by a local sum SE optimization that determines whether each pilot is treated as strong or weak, thereby explicitly controlling the allocation of spatial degrees of freedom between interference suppression and desired signal enhancement. Unlike PFZF, the G-PFZF scheme does not require at least one strong pilot, thereby allowing each AP to flexibly operate in MR, PFZF, or full ZF mode depending on local channel and interference conditions.

For AP m, let $\mathcal{S}_m$ and $\mathcal{W}_m$ denote the sets of strong and weak pilots, respectively, and let $L_{\mathcal{S}_m}$ be the number of strong pilots. Define the set of strong pilot in-

$$\text{SINR}_t^{\text{G-PFZF}}=\frac{p_t^u\Big|\sum\limits_{m\in\mathcal{M}_t}a_{mt}\sqrt{(A-\delta_{mi_t}L_{\mathcal{S}_m})\gamma_{mt}}\Big|^2}{\sum\limits_{k\in\mathcal{P}_{i_t}\setminus\{t\}}p_k^u\Big|\sum\limits_{m\in\mathcal{M}_t}a_{mt}\sqrt{(A-\delta_{mi_t}L_{\mathcal{S}_m})\gamma_{mk}}\Big|^2+\sum\limits_{k=1}^{T}p_k^u\sum\limits_{m\in\mathcal{M}_t}|a_{mt}|^2(\beta_{mk}-\delta_{mi_t}\delta_{mi_k}\gamma_{mk})+\sum\limits_{m\in\mathcal{M}_t}|a_{mt}|^2}. \quad (9)$$

$$\text{SINR}_{mt}^{\text{G-PFZF}}=\frac{S_{mt}}{I_{mt}}=\frac{p_t^u\Big(A-\delta_{mi_t}-\delta_{mi_t}\sum\limits_{i=1,i\neq i_t}^{L_p}\delta_{mi}\Big)\gamma_{mt}}{\sum\limits_{k\in\mathcal{P}_{i_t}\setminus\{t\}}p_k^u\Big(A-\delta_{mi_t}-\delta_{mi_t}\sum\limits_{i=1,i\neq i_t}^{L_p}\delta_{mi}\Big)\gamma_{mk}+\sum\limits_{k\in\mathcal{P}_{i_t}}p_k^u(\beta_{mk}-\delta_{mi_t}\gamma_{mk})+\sum\limits_{k\notin\mathcal{P}_{i_t}}p_k^u(\beta_{mk}-\delta_{mi_t}\delta_{mi_k}\gamma_{mk})+1}. \quad (10)$$

dices as $\mathcal{R}_{\mathcal{S}_m}=\{r_{m,1},\ldots,r_{m,L_{\mathcal{S}_m}}\}$. Let $\bar{\mathbf{G}}_m\in\mathbb{C}^{A\times L_p}$ collect the estimated channel vectors corresponding to all L_p orthogonal pilots, and define the selection matrix $\mathbf{E}_{\mathcal{S}_m}=[\mathbf{e}_{r_{m,1}},\ldots,\mathbf{e}_{r_{m,L_{\mathcal{S}_m}}}]$. For a given pilot $i\in\mathcal{S}_m$, let $j_{mi}\in\{1,2,\ldots,L_{\mathcal{S}_m}\}$ denote the index of pilot i within $\mathcal{R}_{\mathcal{S}_m}$, and define the vector $\boldsymbol{\varepsilon}_{j_{mi}}\in\mathbb{C}^{L_{\mathcal{S}_m}}$ as the j_{mi}-th column of the identity matrix $\mathbf{I}_{L_{\mathcal{S}_m}}$, such that $\mathbf{E}_{\mathcal{S}_m}\boldsymbol{\varepsilon}_{j_{mi}}=\mathbf{e}_i$. The pseudo-inverse of matrix for ZF structure can be expressed as $(\mathbf{E}_{\mathcal{S}_m}^H\bar{\mathbf{G}}_m^H\bar{\mathbf{G}}_m\mathbf{E}_{\mathcal{S}_m})^{-1}$. To obtain the combining vector associated with a specific pilot, this pseudo-inverse term is multiplied by $\boldsymbol{\varepsilon}_{j_{mi}}$ and subsequently normalized. Thus, the local ZF vector for pilot $i\in\mathcal{S}_m$ at AP m is given by:

$$\mathbf{v}_{mi}^{\text{LZF}}=\sqrt{(A-L_{\mathcal{S}_m})\theta_{mi}}\bar{\mathbf{G}}_m\mathbf{E}_{\mathcal{S}_m}(\mathbf{E}_{\mathcal{S}_m}^H\bar{\mathbf{G}}_m^H\bar{\mathbf{G}}_m\mathbf{E}_{\mathcal{S}_m})^{-1}\boldsymbol{\varepsilon}_{j_{mi}}. \quad (7)$$

Similarly, the MR combining vector for pilot $i\in\mathcal{W}_m$ at AP m corresponds directly to the channel vector associated with that pilot and is expressed as:

$$\mathbf{v}_{mi}^{\text{MR}}=\frac{\bar{\mathbf{G}}_m\mathbf{e}_i}{\sqrt{A\theta_{mi}}}. \quad (8)$$

By using the combining vectors in (7) and (8), the closed-form expression of achievable SE can be computed using Theorem 1 with the corresponding SINR given by (9), where $\mathcal{P}_{i_t}$ represents the subset of UEs sharing pilot i_t and $\delta_{mi_t}=1$, if $i_t\in\mathcal{S}_m$, otherwise $\delta_{mi_t}=0$. The derivation of the closed-form SE expression is given in Appendix.

Remark 2: Pilot-level grouping establishes an explicit mapping, allowing the number of required spatial degrees of freedom to be written explicitly as $L_{\mathcal{S}_m}=\sum_{i=1}^{L_p}\delta_{mi}$, which is essential for formulating a tractable optimization problem.

Adaptive Pilot Grouping for G-PFZF Combining: For local optimization framework, the local SINR of UE t at AP m is obtained by assuming that only AP m contributes to the reception of UE t. Moreover, by substituting $L_{\mathcal{S}_m}=\sum_{i=1}^{L_p}\delta_{mi}$, the dependence of the SINR on the pilot-level grouping decision at AP m becomes explicit. The resulting local SINR expression is given in (10).

Allocating a pilot to the strong group ($\delta_{mi}=1$) assigns a ZF combining vector to that pilot, thereby suppressing coherent interference from UEs sharing the same pilot as well as from UEs associated with other strong pilots. This interference suppression is achieved by consuming one spatial DoF per strong pilot, which reduces the available array gain to $(A-L_{\mathcal{S}_m})$ for all UEs served by AP m. Conversely, classifying a pilot as weak preserves the full array gain but provides no interference suppression for the corresponding UEs. This fundamental trade-off between array gain and interference mitigation makes the pilot grouping problem non-trivial and renders fixed, network-wide threshold-based grouping highly suboptimal. To navigate this trade-off, AP m independently solves a local optimization problem, which is formulated as

$$\max_{\delta_{m1},\delta_{m2},\cdots,\delta_{mL_p}}\ \sum\nolimits_{t=1}^{T}\log_2\Big(1+\frac{S_{mt}}{I_{mt}}\Big), \quad (11a)$$

$$\text{subject to: } \delta_{mi}\in\{0,1\},\ \forall i, \quad (11b)$$

$$A-1\geq\sum\nolimits_{i=1}^{L_p}\delta_{mi}, \quad (11c)$$

The constraint (11c) represents the necessary condition for interference suppression. The problem in (11) is a Mixed-Integer Non-Linear Program (MINLP), a class of problems known for its exponential computational cost. To develop a more tractable formulation, we begin by relaxing the binary constraint (11b) into a continuous one as

$$0\leq\delta_{mi}\leq 1,\ \forall i, \quad (12)$$

$$\delta_{mi}-\delta_{mi}^2\leq 0,\ \forall i. \quad (13)$$

The combined constraints (12) and (13) are equivalent to the original binary constraint (11b), as the only values between 0 and 1 that satisfy $\delta_{mi}(1-\delta_{mi})\leq 0$ are precisely 0 and 1. Thus, the problem can be written as

$$\max_{\boldsymbol{\delta}_m}\ \sum\nolimits_{t=1}^{T}\log_2\Big(1+\frac{S_{mt}}{I_{mt}}\Big), \quad (14a)$$

$$\text{subject to: (11c), (12), (13).} \quad (14b)$$

To solve this problem with low computation cost, we employ gradient based proximal gradient ascent (PGA) method [12]. We first reformulate the problem as:

$$\max_{\boldsymbol{\delta}_m}\ f(\boldsymbol{\delta}_m), \quad (15a)$$

$$\text{subject to: (12), where,} \quad (15b)$$

$$f(\boldsymbol{\delta}_m) = \sum_{t=1}^{T} \log_2(1 + \frac{S_{mt}}{I_{mt}}) - \chi\Big(\lambda_1 \sum_{i=1}^{L_p} \max\Big(0, \delta_{mi} - \delta_{mi}^2\Big)^2 + \lambda_2 \max\Big(0, \sum_{i=1}^{L_p} \delta_{mi} - A + 1\Big)^2\Big),$$

where χ is a penalty parameter, and λ_1 and λ_2 are weights. To solve Problem (15) using the PGA method at each AP independently, we follow the steps outlined in Algorithm 1, where $\nabla f(\boldsymbol{\delta}_m)$ represents the gradient of $f(\boldsymbol{\delta}_m)$.

Algorithm 1 Proximal Gradient Ascent (PGA) Method

Initialize: $j = 1$, $k = 1$, $\boldsymbol{\delta}_m^{(1)}$, $s^{(k)} = f(\boldsymbol{\delta}_m^{(1)})$, $\chi = 1$, $\Delta > 1$, $\epsilon = 5e^{-3}$
repeat
  repeat
    Compute $\boldsymbol{\delta}_m^{(j+1)} = \min\big(1, \max\big(0, \boldsymbol{\delta}_m^{(j)} + \alpha \nabla f(\boldsymbol{\delta}_m^{(j)})\big)\big)$
    Update $j = j + 1$
  until $\Big|\frac{f(\boldsymbol{\delta}_m^{(j)}) - f(\boldsymbol{\delta}_m^{(j-1)})}{f(\boldsymbol{\delta}_m^{(j-1)})} \leq \epsilon\Big|$
  Update $\chi = \chi \times \Delta$, $k = k + 1$ and $s^{(k)} = f(\boldsymbol{\delta}_m^{(j)})$
until $\Big|\frac{s^{(k)} - s^{(k-1)}}{s^{(k-1)}} \leq \epsilon\Big|$

Note 1: In UE-based grouping, each AP would have T binary variables (one per UE), leading to TM total variables across the network. In the proposed pilot-based grouping, each AP has L_p binary variables, leading to ML_p total variables. Since $L_p \ll T$ in dense deployments, this significantly reduces the computational cost.

IV. Performance Evaluation

We consider a network with M=100 APs with A=8 antennas each and T=100 UEs, uniformly distributed in a square area of size 1×1 km^2. The LSFCs are generated according to the model in [1], incorporating shadow fading with standard deviation of 8 dB. Pilots are assigned using the scalable pilot assignment scheme in [3]. All UEs are assumed to operate at the full transmit power of 100 mW. For the baseline PFZF and PWPFZF schemes, UE grouping is performed using 90% of total LSFCs received at the APs, as identified via numerical simulation, to provide the best performance [9]. Unless otherwise specified, the simulation setup follows [1], with a carrier frequency of 1.9 GHz and a bandwidth of 20 MHz. To ensure statistical reliability, all results are averaged over 500 independent network realizations.

Fig. 1 shows that the baseline PFZF/PWPFZF scheme exhibits a pronounced peak at $L_{\mathcal{S}_m}$=6-7, revealing its rigid threshold-based design that sacrifices excessive degrees of freedom for interference suppression regardless of local conditions. In contrast, the proposed G-PFZF scheme shows fundamentally different behavior by peaking at 1-4 strong pilots. This efficient allocation preserves spatial resources for signal enhancement. Most notably, the significant non-zero count at $L_{\mathcal{S}_m} = 0$ demonstrates our scheme's ability to adaptively default to simple MR combining when it is the optimal scheme for a given AP's local conditions.

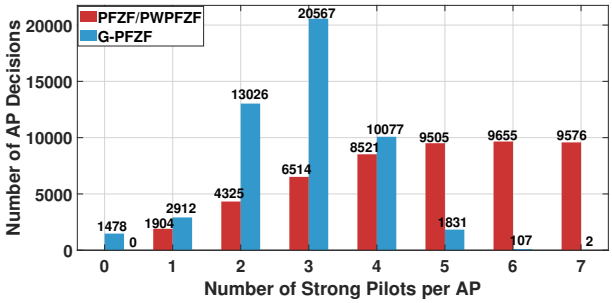

Fig. 1. Distribution of strong pilot decisions.

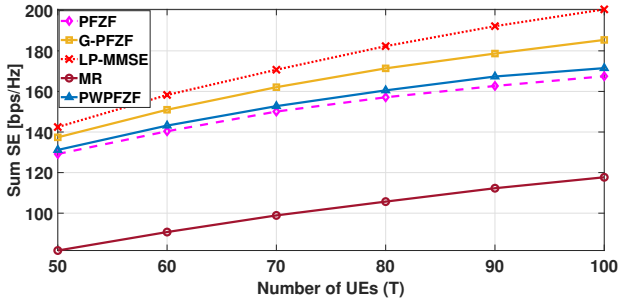

Fig. 2. The uplink sum SE comparison for various combining schemes.

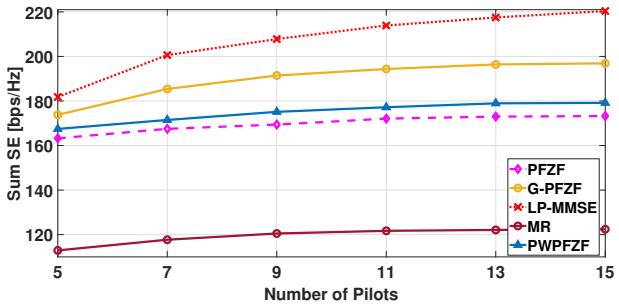

Fig. 3. The uplink sum SE comparison for various combining schemes.

Fig. 2 depicts the sum SE versus the number of UEs (T) for various combining schemes. Compared with the baseline PFZF/PWPFZF, the proposed G-PFZF achieves a gain of 5%-10.5%, with the gaps widening with increasing number of UEs. This performance improvement comes from the adaptive per-AP pilot classification, that enables a more efficient and dynamic trade-off between utilizing degrees of freedom for interference suppression and desired signal enhancement. Although LP-MMSE performs better, but as it relies on Monte Carlo realizations and does not provide tractable closed-form expressions for SE and LSFD weights, it is less desirable.

Fig. 3 illustrates the sum SE versus the number of orthogonal pilots L_p. The proposed G-PFZF scheme achieves significant performance improvements of 4-14% compared to PFZF/PWPFZF, with the performance gap widening as the number of pilots increases. This expanding gap highlights the advantage of the G-PFZF scheme over the rigid network-wide threshold approach. The PFZF/PWPFZF scheme must rigidly limit the number of strong pilots to avoid violating the fundamental condition for zero-forcing ($L_{\mathcal{S}_m} > A$), which leads to inefficient interference management. In contrast, our proposed optimization, ensures that $L_{\mathcal{S}_m} < A$, while maximizing array

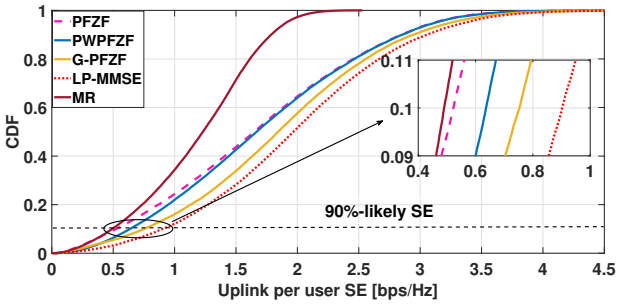

Fig. 4. The CDF of the SE per user.

gain and suppressing interference.

Fig. 4 represents the empirical cumulative distribution function (CDF) of the SE per user. In terms of 90%-likely SE, the proposed G-PFZF outperforms the PWPFZF, PFZF and MR schemes by a margin of 19%, 44% and 53% respectively. The LP-MMSE scheme, which relies on Monte Carlo simulations, though has the best performance.

V. Conclusion and Future Work

This paper establishes a framework for uplink processing in D-mMIMO systems by addressing local combining. We propose an adaptive pilot-aware local combining scheme, G-PFZF, where pilots are partitioned using local optimization. Numerical results confirmed that the G-PFZF combining outperforms the PFZF and PWPFZF combining by a significant margin for different network parameters. In future, we plan to investigate the robustness of generalized combining scheme under hardware impairments.

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. on Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, 2017.

[2] M. S. A. Khan, S. Agnihotri, and R. M. Karthik, "Joint AP-UE association and power factor optimization for distributed massive MIMO," in *Proc. IEEE PIMRC*, Valencia, Spain, Sept. 2024.

[3] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *IEEE Trans. on Commun.*, vol. 68, no. 7, pp. 4247–4261, 2020.

[4] M. S. A. Khan, S. Agnihotri, and R. M. Karthik, "Distributed pilot assignment for distributed massive-MIMO networks," in *Proc, IEEE WCNC*, Dubai, UAE, April 2024.

[5] D. Maryopi, M. Bashar, and A. Burr, "On the uplink throughput of zero forcing in cell-free massive MIMO with coarse quantization," *IEEE Trans. on Veh. Techno.*, vol. 68, no. 7, pp. 7220–7224, 2019.

[6] Ö. T. Demir, E. Björnson, L. Sanguinetti *et al.*, "Foundations of user-centric cell-free massive MIMO," *Foundations and Trends® in Signal Processing*, vol. 14, no. 3-4, pp. 162–472, 2021.

[7] L. Du, L. Li, H. Q. Ngo, T. C. Mai, and M. Matthaiou, "Cell-free massive MIMO: Joint maximum-ratio and zero-forcing precoder with power control," *IEEE Trans. on Commun.*, vol. 69, no. 6, pp. 3741–3756, 2021.

[8] X. Wang, J. Cheng, C. Zhai, and A. Ashikhmin, "Partial cooperative zero-forcing decoding for uplink cell-free massive MIMO," *IEEE Internet of Things Journal*, vol. 9, no. 12, pp. 10 327–10 339, 2021.

[9] J. Zhang, J. Zhang, E. Björnson, and B. Ai, "Local partial zero-forcing combining for cell-free massive MIMO systems," *IEEE Trans. on Commun.*, vol. 69, no. 12, pp. 8459–8473, 2021.

[10] M. S. A. Khan, S. Agnihotri, and R. M. Karthik, "Comments on "Local partial zero-forcing combining for cell-free massive MIMO systems"," *IEEE Trans. on Commun.*, vol. 73, no. 12, pp. 14 986–14 989, 2025.

[11] 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.

[12] C. Hao, T. T. Vu, H. Q. Ngo, M. N. Dao, X. Dang, C. Wang, and M. Matthaiou, "Joint user association and power control for cell-free massive MIMO," *IEEE Internet of Things Journal*, vol. 11, no. 9, pp. 15 823–15 841, 2024.

Appendix

The desired signal term in (6) can be written as

$$|\mathrm{DS}_t|^2 = \left|\mathbb{E}\Big\{\sum_{m\in\mathcal{Z}_{i_t}} a_{mt}(\mathbf{v}_{mi_t}^{\mathrm{LZF}})^H \hat{\mathbf{g}}_{mt} + \sum_{m\in\mathcal{Y}_{i_t}} a_{mt}(\mathbf{v}_{mi_t}^{\mathrm{MR}})^H \hat{\mathbf{g}}_{mt}\Big\}\right|^2,$$

where $\mathcal{Z}_{i_t}$ and $\mathcal{Y}_{i_t}$ denote the subsets of APs assigned to pilot i_t and classified as strong and weak APs, respectively. The $|\mathrm{DS}_t|^2$ can be written as:

$$|\mathrm{DS}_t|^2 = \Big|\sum_{m\in\mathcal{M}_t} a_{mt}\sqrt{(A-\delta_{mi_t}L_{\mathcal{S}_m})\gamma_{mt}}\Big|^2, \tag{A1}$$

The interference term $\mathbb{E}\{|\mathrm{UI}_{tk}|^2\}$ in (6) can be expanded as:

$$\begin{aligned}
&\mathbb{E}\Big\{\Big|\sum_{m\in\mathcal{Z}_{i_t}} a_{mt}(\mathbf{v}_{mi_t}^{\mathrm{LZF}})^H\mathbf{g}_{mk} + \sum_{m\in\mathcal{Y}_{i_t}} a_{mt}(\mathbf{v}_{mi_t}^{\mathrm{MR}})^H\mathbf{g}_{mk}\Big|^2\Big\}\\
&=\mathbb{E}\Big\{\Big|\sum_{m\in\mathcal{Z}_{i_t}} a_{mt}(\mathbf{v}_{mi_t}^{\mathrm{LZF}})^H\mathbf{g}_{mk}\Big|^2\Big\}+\mathbb{E}\Big\{\Big|\sum_{m\in\mathcal{Y}_{i_t}} a_{mt}(\mathbf{v}_{mi_t}^{\mathrm{MR}})^H\mathbf{g}_{mk}\Big|^2\Big\}\\
&+2\mathbb{E}\Big\{\sum_{m\in\mathcal{Z}_{i_t}}\sum_{n\in\mathcal{Y}_{i_t}} a_{mt}a_{nt}(\mathbf{v}_{mi_t}^{\mathrm{LZF}})^H\mathbf{g}_{mk}(\mathbf{v}_{ni_t}^{\mathrm{MR}})^H\mathbf{g}_{nk}\Big\}.
\end{aligned} \tag{A2}$$

The first term (A_1) of (A2) can be rewritten as:

$$\begin{aligned}
A_1 = &\Big|\sum_{m\in\mathcal{Z}_{i_t}} a_{mt}\sqrt{(A-L_{\mathcal{S}_m})\gamma_{mk}}\Big|^2\left|\boldsymbol{\psi}_{i_t}^H\boldsymbol{\psi}_{i_k}\right|^2\\
&+\sum_{m\in\mathcal{Z}_{i_t}}|a_{mt}|^2(\beta_{mk}-\delta_{mi_k}\gamma_{mk}).
\end{aligned} \tag{A3}$$

The second term (A_2) of (A2) can be rewritten as:

$$A_2 = \Big|\sum_{m\in\mathcal{Z}_{i_t}} a_{mt}\sqrt{A\gamma_{mk}}\Big|^2\left|\boldsymbol{\psi}_{i_t}^H\boldsymbol{\psi}_{i_k}\right|^2+\sum_{m\in\mathcal{Z}_{i_t}}|a_{mt}|^2\beta_{mk}. \tag{A4}$$

Finally, the third term (A_3) of (A2) can be rewritten as:

$$A_3 = 2\Big(\sum_{m\in\mathcal{Z}_{i_t}} a_{mt}\sqrt{(A-L_{\mathcal{S}_m})\gamma_{mt}}\Big)\Big(\sum_{m\in\mathcal{Y}_{i_t}} a_{mt}\sqrt{A\gamma_{mt}}\Big). \tag{A5}$$

Using (A3), (A4) and (A5)

$$\begin{aligned}
\mathbb{E}\Big\{|\mathrm{UI}_{tk}|^2\Big\} = &\Big|\sum_{m\in\mathcal{M}_t} a_{mt}\sqrt{(A-\delta_{mi_t}L_{\mathcal{S}_m})\gamma_{mk}}\Big|^2\left|\boldsymbol{\psi}_{i_t}^H\boldsymbol{\psi}_{i_k}\right|^2\\
&+\sum_{m\in\mathcal{M}_t}|a_{mt}|^2(\beta_{mk}-\delta_{mi_t}\delta_{mi_k}\gamma_{mk}).
\end{aligned} \tag{A6}$$

The noise term of interference in (6) can be written as:

$$\mathbb{E}\{|\mathrm{GN}_t|^2\} = \sum_{m\in\mathcal{M}_t}|a_{mt}|^2 \tag{A7}$$

Substituting (A1), (A6) and (A7) in (6) gives (9).