# A Probabilistic QoS-Aware User Association in Cell-Free Massive MIMO for Industry 4.0

Rachuri, Sanyasi Vishnu Vardhan; Manchon, Carles Navarro; Berardinelli, Gilberto; Amiri, Abolfazl

*Published in:*
IEEE Wireless Communications and Networking Conference, WCNC 2024



*Publication date:*
2024

*Document Version*
Accepted author manuscript, peer reviewed version

# A Probabilistic QoS-Aware User Association in Cell-Free Massive MIMO for Industry 4.0

Vishnu Rachuri*†, Carles Navarro Manchón*, Gilberto Berardinelli*, Abolfazl Amiri†
*Department of Electronic Systems, Aalborg University, 9220 Aalborg, Denmark
†Nokia Standards, 9220 Aalborg, Denmark
{svvr, cnm, gb}@es.aau.dk, {abolfazl.amiri}@nokia.com

***Abstract*—Industry 4.0 requires wireless solutions to meet the demanding reliability and latency requirements of cyber-physical systems serving many industrial sensors and actuators. The cell-free massive multiple input multiple output (MIMO) paradigm, exploiting both massive MIMO gains and cell-free properties, is expected to play a crucial role in addressing Industry 4.0 communication requirements. However, finding an optimal user association to satisfy the quality of service (QoS) requirements is critical. In this article, we propose a novel centralized probabilistic user association algorithm to identify and maximize the number of served users without relying on short-term channel state information to satisfy the QoS requirements. In larger factories, approximately $90\%$ of the total users are satisfied. However, as the system becomes interference-limited and is expected to meet challenging requirements, the number of served users decreases. Nevertheless, the proposed solution exhibits adaptability in maximizing the number of satisfied users. Simulation results further prove the effectiveness of the proposed solution in terms of achievable spectral efficiency and maximizing the number of satisfied users across varying requirements and factory sizes.**

***Index Terms*—User association, cell-free massive MIMO, Industry 4.0**

## I. Introduction

The Industry 4.0 vision relies on the advancements in fifth-generation (5G) and beyond 5G (B5G) or 6G technologies, eliminating the need for costly and high-maintenance wired solutions vital for critical communication needs and improving modularity and production management efficiency. Massive multiple input multiple output (mMIMO) is a crucial technology for 5G and beyond. It can serve multiple users simultaneously using the same time and frequency resources and can function as a co-located or distributed architecture. In industrial environments, with harsh propagation conditions, distributed mMIMO can offer better link stability, coverage, and signal-to-noise ratio (SNR) gains. In [1], such gains have been verified via measurement campaigns taken from two factory setups using an automated guided vehicle. The analysis in [2] shows that, in an indoor line-of-sight scenario, distributed mMIMO requires a lower transmit power in downlink than a co-located massive MIMO setup to provide high data rates. Building upon the benefits of distributed mMIMO, cell-free massive MIMO (cf-mMIMO) proposes an alternative architecture by removing the cell boundaries. In this architecture, a large set of access points (APs) jointly serve the users, making it a promising solution for networks with diverse user requirements. Authors in [4] have shown the superiority of cf-mMIMO with respect to other distributed mMIMO transmission schemes in terms of achievable data rates and computational complexity in industrial scenarios. Initial work on cf-mMIMO assumed all APs to be connected via fronthaul links to a central unit jointly serving all users, known as a full-scale system. This configuration entails high computational and fronthaul bandwidth requirements, resulting in poor scalability [5]. Fig. 1 depicts the formation of serving clusters (the set of APs assigned for a user) in an industrial scenario.

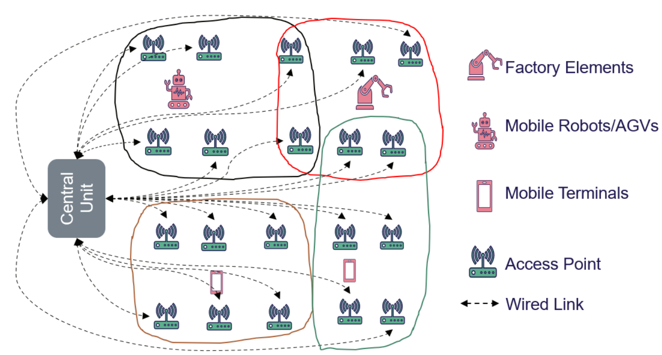


Fig. 1: Serving clusters of cf-mMIMO in an industrial scenario

For a scalable cf-mMIMO system, an intelligent user association strategy is needed. It reduces the number of serving APs per user, enhancing scalability by decreasing joint processing complexity. Choosing the right user association and forming a serving cluster is crucial. Authors in [5] proposed a user-centric dynamic cooperation clustering (DCC), a large-scale fading (LSF) based algorithm. Each user chooses a master AP with the most significant gain, which allocates a pilot sequence and instructs neighboring APs to form a serving cluster. Alternatively, in [6], the association problem is formulated as a matching problem, and a Hungarian algorithm is proposed to maximize the uplink (UL) sum rate. A sub-optimal joint user clustering and access point selection scheme proposed in

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement No 956670. The authors would like to extend their appreciation to Klaus Pedersen for the valuable inputs and support during the research stay at Nokia.

[7] divides the users into multiple clusters based on similar channel characteristics and performs the AP selection process only once for each user cluster. To reduce the joint processing complexity and signaling overhead, [8] formulated a bipartite graph partitioning problem and proposed a rate-constrained network decomposition algorithm that forms multiple weakly interfered serving clusters. The authors evaluated this algorithm at mmWave bands, assuming APs are equipped with multiple antennas and a well-defined beam space. In [9], a new performance indicator, user satisfaction rate (USR), is introduced to measure the ratio of scheduled users satisfying their performance requirements; a dynamic programming-based user scheduling solution is proposed to maximize the system throughput and USR. The authors considered a downlink transmission where all the APs jointly serve all the users and rely on full CSI availability.

Nonetheless, the presented state-of-the-art either disregards guaranteeing user-specific QoS requirements, as often demanded by industrial applications for closed-loop control, or relies on the assumption that instantaneous CSI from all users is available at the central unit; this might be impractical to obtain as small-scale fading may rapidly vary over time. The main contributions of this paper are the following:

- We propose a QoS-aware user association algorithm, which finds suitable user-AP associations to maximize the number of satisfied users. This article focuses on guaranteeing a requested spectral efficiency (SE) as the key QoS indicator.
- The proposed algorithm uses a method to predict the minimum number of APs needed to meet user requirements without relying on accurate instantaneous CSI. This method adopts closed-form expressions based on LSF coefficients to predict the achievable SE and a probability distribution model to measure the reliability of predictions.

The rest of the paper is organized as follows: Section II outlines the system model, Section III introduces the closed-form expressions for prediction based on LSF coefficients, and devises a distribution model to capture the reliability of the closed-form expressions. Section IV describes the proposed algorithm, and Section V presents the simulation results. Finally, Section VI provides the conclusion and outlines future work.

**Notation:** Boldface lowercase $\mathbf{x}$ and uppercase $\mathbf{X}$ denote column vectors and matrices respectively. $\mathcal{N}_{\mathbb{C}}$ represents the complex Gaussian distribution and $\mathbb{E}\{.\}$ denotes expected value. $^{\mathrm{T}}$, $^{\mathrm{H}}$, and $^{\dagger}$ denote matrix operations transpose, hermitian, and pseudo-inverse, respectively. $\mathrm{diag}\,(.)$ transforms square matrices into block-diagonal matrix. $\setminus$, $\bigcup$, represent set difference and union. Whereas $\subset$, $|\mathcal{X}|$ denote subset relation and cardinality of a set, respectively.

## II. System Model and Spectral Efficiency

### A. System Model

We consider a cf-mMIMO system, with a set $\mathcal{L}$ of distributed APs, with cardinality $|\mathcal{L}| = L$. Each AP has $N$ antennas connected to a central unit, jointly serving a group $\mathcal{K}$ of single-antenna users, where $|\mathcal{K}| = K$, over the same radio resources. The UL received signal at the $l$-th AP is

$$\mathbf{y}_l^{ul} = \sum_{k=1}^{K} \sqrt{\eta_k}\mathbf{h}_{k,l} s_k + \mathbf{n}_l \tag{1}$$

where the $k$-th user, with $k = 1, \cdots, K$, transmits complex symbol $s_k$ with transmit power $\eta_k$ on channel $\mathbf{h}_{k,l} \in \mathbb{C}^N$, where $\mathbf{n}_l \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \sigma_{ul}^2 \mathbf{I}_N\right)$ is thermal noise. The links $\mathbf{h}_{k,l}$ are mutually independent Rayleigh fading channels, with $\mathbf{h}_{kl} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \mathbf{R}_{kl}\right)$. The complex Gaussian distribution models small scale fading where $\mathbf{R}_{kl} = \beta_{kl}\mathbf{I}_N$ is a spatial correlation matrix formulated considering LSF values describing geometric path loss and shadow fading. The collective channel of the $k$-th user is denoted as $\mathbf{h}_k = \left[\mathbf{h}_{k1}^{\mathrm{T}} \quad .... \quad \mathbf{h}_{kL}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^M$, where $M = LN$.

This paper focuses on the UL user association. The channel is considered constant across a UL frame of length $\tau_f$, divided into UL pilot ($\tau_p$) and UL data ($\tau_{ul}$) samples. We assume that the CSI is acquired using UL pilot sequences of length $\tau_p$ and the available number of orthogonal pilot sequences $\tau_p < K$. Given that $\Phi_{tk} \in \mathbb{C}^{\tau_p}$ is a pilot sequence assigned to user $k$, $\mathcal{P}_k$ denote the set of users assigned with the same pilot sequence, including the $k$-th user, i.e.

$$\Phi_{tj}^{\mathrm{H}}\Phi_{tk} = \begin{cases} 0 & j \notin \mathcal{P}_k \\ \tau_p & j \in \mathcal{P}_k \end{cases} \tag{2}$$

The pilot assignment is a critical initial step to mitigate pilot contamination. The pilot assignment strategy proposed in [3] based on large-scale fading values is adopted in this article.

In user-centric cf-mMIMO, APs in the serving cluster participate in the UL receive processing for each user. Therefore, the channel estimates of the user are available only at those APs. The MMSE channel estimator proposed in [3] is adopted to obtain the channel estimate $\hat{\mathbf{h}}_k = \left[\hat{\mathbf{h}}_{k1}^{\mathrm{T}} \quad .... \quad \hat{\mathbf{h}}_{kL}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^M$.

We consider a centralized architecture and define an association matrix $\mathbf{D} \in \{0,1\}^{\mathrm{LN}\times\mathrm{KN}} = [\mathbf{D}_1 \dots \mathbf{D}_K]$, where $\mathbf{D}_k = [\mathbf{D}_{k1} \dots \mathbf{D}_{kL}]^{\mathrm{T}}$ represent the association of the $k$-th user. $\mathbf{D}_{kl} \in \{0,1\}^{N\times N}$ is either $\mathbf{I}_N$ or $\mathbf{0}_N$, depending on whether the $l$-th AP is in the $k$-th user's serving cluster or not. The central unit further estimates the received symbols using *combining vectors* $\mathbf{v}_k = \left[\mathbf{v}_{k1}^{\mathrm{T}} \dots \mathbf{v}_{kL}^{\mathrm{T}}\right]^{\mathrm{T}} \in \mathbb{C}^M$. Considering $\mathbf{O}_k = \mathrm{diag}\,(\mathbf{D}_k)$ as a block diagonal matrix, combining vectors are formulated according to the channel estimates of the serving clusters. We adopted the *partial*-MMSE (PMMSE) combiner (3) from [3], where the set of users $\mathcal{S}_k$ that are partially or completely served by the serving cluster of the $k$-th user are considered, i.e.

$$\begin{aligned} \mathbf{v}_k &= \eta_k \left(\sum_{i\in\mathcal{S}_k} \eta_i \mathbf{O}_k \hat{\mathbf{h}}_i \hat{\mathbf{h}}_i^{\mathrm{H}} \mathbf{O}_k + \mathbf{Z}_k\right)^{\dagger} \mathbf{O}_k \hat{\mathbf{h}}_k, \\ \mathbf{Z}_k &= \mathbf{O}_k \left(\sum_{i\in\mathcal{S}_k} \eta_i \mathbf{C}_i + \sigma_{ul}^2 \mathbf{I}_M\right) \mathbf{O}_k, \end{aligned} \tag{3}$$

where $\mathbf{C}_i \in \mathbb{C}^{\mathrm{LN}\times\mathrm{LN}}$ is the error correlation matrix of the collective channel estimate of the $i$-th user [3]. The effective signal-to-interference-and-noise ratio of $k$-th user $\mathrm{SINR}_k$ using P-MMSE can be represented as [10]

$$\mathrm{SINR}_k = \frac{\mathrm{DS}_k}{\mathrm{IUI}_k + \mathrm{n}_k} \tag{4}$$

where $\mathrm{DS}_k$, $\mathrm{IUI}_k$, and $\mathrm{n}_k$ denote desired signal, inter-user interference, and noise, respectively (5), given by:

$$\begin{aligned} \mathrm{DS}_k &= \eta_k \mathbb{E}\left|\left\{\mathbf{v}_k^{\mathrm{H}}\mathbf{O}_k\mathbf{h}_k\right\}\right|^2, \\ \mathrm{IUI}_k &= \sum_{i=1}^{K}\eta_i \mathbb{E}\left\{\left|\mathbf{v}_k^{\mathrm{H}}\mathbf{O}_k\mathbf{h}_i\right|^2\right\} - \eta_k \mathbb{E}\left|\left\{\mathbf{v}_k^{\mathrm{H}}\mathbf{O}_k\mathbf{h}_k\right\}\right|^2, \\ \mathrm{n}_k &= \sigma_{ul}^2 \mathbb{E}\left\{||\mathbf{O}_k\mathbf{v}_k||^2\right\}. \end{aligned} \tag{5}$$

This results in an achievable SE for the $k$-th user as:

$$\mathrm{SE}_k = \left(1 - \frac{\tau_p}{\tau_f}\right)\log_2\left(1 + \mathrm{SINR}_k\right). \tag{6}$$

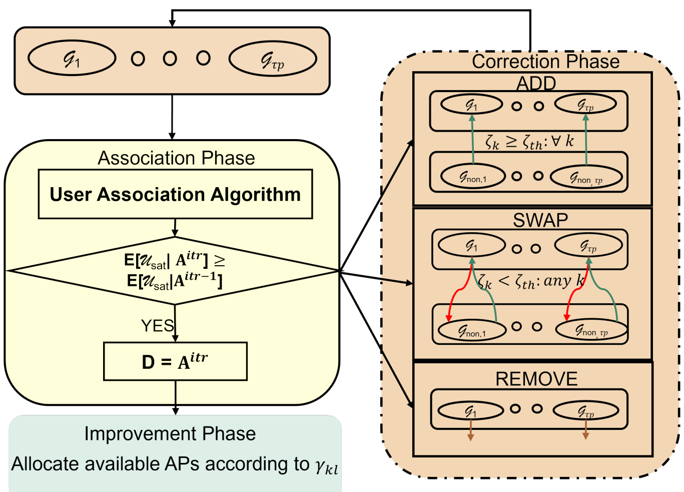


Fig. 2: Proposed user association algorithm

## III. Spectral Efficiency Prediction

Selecting an appropriate serving cluster (association matrix) is essential for developing an efficient cf-mMIMO system. Determining the minimum number of APs required per user is crucial to ensure fairness and equal opportunity for each user to meet their respective requirements. However, ensuring fairness might require accurate knowledge of the instantaneous CSI for each user. This can be computationally complex and incurs significant signaling overhead, making it infeasible.

On the other hand, LSF characteristics slowly vary, so it is feasible for the network to estimate them accurately over time. Closed-form expressions for the uplink SE have been derived in [10], using LSF coefficients with different modes of Zero-Forcing (ZF) combining schemes. Using such approximations can provide insight into the impact of interference suppression when interference cancellation combiners, such as P-MMSE (3), are applied.

Channel estimate $\hat{h}_{kl}$ and estimation error $\widetilde{h}_{kl} = h_{kl} - \hat{h}_{kl}$ are independent Gaussian vectors with distributions [3]:

$$\hat{h}_{kl} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \gamma_{kl}\mathbf{I}_N\right), \tag{8}$$

$$\tilde{h}_{kl} \sim \mathcal{N}_{\mathbb{C}}\left(\mathbf{0}, \left(\beta_{kl} - \gamma_{kl}\right)\mathbf{I}_N\right), \tag{9}$$

$$\gamma_{kl} = \mathbb{E}\left\{\left|\hat{h}_{kl}\right|^2\right\} = \frac{\eta_k\tau_p\beta_{kl}^2}{\tau_p\sum_{t\in\mathcal{P}_k}\eta_t\beta_{tl} + \sigma_{ul}^2}. \tag{10}$$

When users $k$ and $t$ use the same pilot, channel estimates $\hat{h}_{kl}$ and $\hat{h}_{tl}$ are linearly dependent [10], as follows:

$$\hat{h}_{kl} = \frac{\sqrt{\eta_k}\beta_{kl}}{\sqrt{\eta_t}\beta_{tl}}\hat{h}_{tl}. \tag{11}$$

The achievable $\mathrm{SINR}_k$ can be calculated as in (7) at the top of the next page, and using (6), we can obtain a prediction of the achievable SE ($\mathrm{SE}_{\mathrm{pred}}$). Evaluating the accuracy of $\mathrm{SE}_{\mathrm{pred}}$ compared to actual SE ($\mathrm{SE}_{\mathrm{act}}$) is crucial in finding the best association matrix. A distribution model is formulated to evaluate the deviation between $\mathrm{SE}_{\mathrm{pred}}$ and $\mathrm{SE}_{\mathrm{act}}$.

### *A. Distribution Model*

To model the deviation between predicted and actual SE values, we fit a distribution that captures the statistics of $\Delta\mathrm{SE} = \mathrm{SE}_{\mathrm{pred}} - \mathrm{SE}_{\mathrm{act}}$, where $\mathrm{SE}_{\mathrm{pred}}$ is obtained using the closed-form approximation (7) and $\mathrm{SE}_{\mathrm{act}}$ is obtained using instantaneous CSI (4),(6). As an initial step, we collect $\mathrm{SE}_{\mathrm{pred}}$ and $\mathrm{SE}_{\mathrm{act}}$ for a full-scale system and subsequently employ a distribution model that approximates the $\Delta\mathrm{SE}$. In this paper, we model $\Delta\mathrm{SE}$ as a Gaussian distribution, i.e.

$$\Delta\mathrm{SE} \sim \mathcal{N}\left(\mu, \sigma\right) \tag{12}$$

where the mean ($\mu$) and standard deviation ($\sigma$) are empirically calculated from the data obtained via simulations on a full-scale system. This can further guide the selection of a favorable association that maximizes the expected number of satisfied users (14) according to the guaranteed SE requirements ($\mathrm{SE}_{\mathrm{req}}$). Using the model (12), the probability of user $k$ satisfying its SE requirement, conditioned on its prediction, can be calculated as

$$\zeta_k = \Pr\left(\mathrm{SE}_k \geq \mathrm{SE}_{\mathrm{req}}|\mathrm{SE}_{\mathrm{pred}}\right). \tag{13}$$

and, with it, the expected number of satisfied users in the network becomes

$$\mathbb{E}\left\{\mathcal{U}_{\mathrm{sat}}\right\} = \sum_{k=1}^{|\mathcal{U}_{\mathrm{schd}}|}\zeta_k \tag{14}$$

where $\mathcal{U}_{\mathrm{schd}}$ and $\mathcal{U}_{\mathrm{sat}}$ correspond to a set of scheduled users and satisfied users, respectively. We can assess prediction credibility by using the *satisfaction rate* defined as

$$\rho_s = \frac{\mathbb{E}\left\{\mathcal{U}_{\mathrm{sat}}\right\}}{|\mathcal{U}_{\mathrm{schd}}|} \tag{15}$$

The central unit can leverage such a distribution model to find optimal associations to maximize the $\mathbb{E}\left\{\mathcal{U}_{\mathrm{sat}}\right\}$.

$$\text{SINR}_k = \frac{\eta_k \left| \sum_{l=1}^{L} \gamma_{kl} \right|^2}{\sum_{t=1}^{K} \eta_t \sum_{l=1}^{L} \gamma_{kl} \left(\beta_{tl} - \gamma_{tl}\right) + \sum_{t \in \mathcal{P}_k \setminus \{k\}} \eta_t \left( \sum_{l=1}^{L} \gamma_{tl} \right)^2 + \sigma_{ul}^2 \sum_{l=1}^{L} \gamma_{kl}} \tag{7}$$

## IV. QoS-aware User Association

This section introduces a QoS-based user association algorithm that uses the equations and distribution model from the previous section. The equations provide two key details: the expected number of satisfied users and the optimal association for meeting SE requirements. Considering single antenna ($N = 1$) per AP, we formulate our problem to maximize the expected number of satisfied users transmitting with power $\eta_{ul}$ and find an optimal association matrix $\mathbf{D} \in \{0,1\}^{L \times K}$, where $d_{lk} = 1$ if $l$-th AP is assigned to $k$-th user, otherwise 0, to satisfy guaranteed SE requirements, i.e.

$$\begin{aligned} \max_{\mathbf{D}} \quad & \mathbb{E}\left\{\mathcal{U}_{\text{sat}}\right\} \\ \text{s.t.} \quad & d_{l,k} \in \{0,1\} \\ & \eta_k = \eta_{ul}, \forall k \in \mathcal{U}_{\text{schd}} \end{aligned} \tag{16}$$

The proposed algorithm has three stages, shown in Fig. 2. The first stage is the *association phase* (Alg. 1), aiming at finding an initial association. The second stage is *correction phase* (Alg. 2) to update the designated user set to improve the satisfied users, and the third stage is *improvement phase* (Alg. 3) to enhance the reliability of formed associations.

To initialize the algorithm, a random subset of users ($\mathcal{U}_{\text{schd}} \subset \mathcal{K}$), where $|\mathcal{U}_{\text{schd}}| = K/2$ is selected. To limit the pilot contamination effect, we prevent two users sharing the same pilot sequence from being served by the same AP. To accomplish this, we divide the subset $\mathcal{U}_{\text{schd}}$ into groups $\mathcal{G} = \left\{\mathcal{G}_1, \dots, \mathcal{G}_{\tau_p}\right\}$ based on their assigned pilots, where $\mathcal{K}_{\mathcal{G}} = \left\{\mathcal{K}_{\mathcal{G}_1}, \dots, \mathcal{K}_{\mathcal{G}_{\tau_p}}\right\}$ represent set of selected users for each group and $\mathcal{U}_{\text{schd}} = \bigcup_{i=1,\dots,\tau_p} \mathcal{K}_{\mathcal{G}_i}$. Based on the same criterion, we further divide the non-selected users ($\mathcal{U}_{\text{non-select}} = \mathcal{K} \setminus \mathcal{U}_{\text{schd}}$) into $\mathcal{G}_{\text{non}} = \left\{\mathcal{G}_{\text{non},1}, \dots, \mathcal{G}_{\text{non},\tau_p}\right\}$. This is to ensure that users from the same group always choose different APs, thus forming disjoint AP clusters in each group.

### A. Association Phase

A sequential association approach is adopted to ensure the same AP does not serve two users from the same group. For each user APs are sorted according to strongest $\gamma_{kl}$ values, and one by one, each AP is added to $\mathcal{L}_{\text{selected}} \subset \mathcal{L}$ until $\text{SE}_{\text{pred}} \geq \text{SE}_{\text{req}}$. This ensures that the minimum number of APs are assigned per user to give a fair chance to other users. The same process is applied to all groups based on their respective requirements. Alg. 1 outlines the association phase, where $\Gamma \in \mathbb{R}^{K \times L}$ with each entry $\gamma_{kl}$ calculated using (10), and the $\text{SE}_{\text{req}}$ of each user is given as input. Additionally, the available APs for each group are initialized with $\mathcal{L}_{\mathcal{G}_i} = \mathcal{L}$.

**Algorithm 1** Association Phase

**Input:** $\Gamma$, $\mathcal{L}$, $\mathcal{G}$, $\mathcal{G}_{\text{non}}$, $\mathcal{K}_{\mathcal{G}}$, $\text{SE}_{\text{req}}$

Initialization : $\mathbf{T} \in \mathbf{0}^{L \times K} \leftarrow$ Intermediate Assn. matrix
$\mathcal{L}_{\mathcal{G}_i} \leftarrow \mathcal{L} : \forall \mathcal{G}_i \in \mathcal{G}$
**for** $\mathcal{G}_i$ in $\left\{\mathcal{G}_1, \mathcal{G}_2, \dots, \mathcal{G}_{\tau_p}\right\}$ **do**
  **for** $k = 1, 2, \dots |\mathcal{K}_{\mathcal{G}_i}|$ **do**
    Sort APs in descending order based on $\gamma_{kl}$.
    $\mathcal{L}_{\text{selected}} \leftarrow \{\}$
    **for** $l = 1, 2, \dots, |\mathcal{L}_{\mathcal{G}_i}|$ **do**
      $\mathcal{L}_{\text{selected}} \leftarrow \mathcal{L}_{\text{selected}} \cup \{l\}$
      Calculate $\text{SE}_{\text{pred}}$ for $k$-th user using (7) & (6)
      **if** $\text{SE}_{\text{pred}} \geq \text{SE}_{\text{req}}$ **then**
        $\mathbf{T}[\mathcal{L}_{\text{selected}}, k] = 1$
        $\mathcal{L}_{\mathcal{G}_i} \leftarrow \mathcal{L}_{\mathcal{G}_i} \setminus \mathcal{L}_{\text{selected}}$
        break;
      **end if**
    **end for**
  **end for**
**end for**

**Output:** $\mathbf{T}$

### B. Correction Phase

Once the association phase is finished, the distribution model in the previous section calculates the $\mathbb{E}\left\{\mathcal{U}_{\text{sat}}\right\}$ (14) and respective probabilities (13); if a user $k$ is not assigned to any APs, then the user is assigned with $\zeta_k = 0$. Once the association is performed and probabilities are obtained, we move on to the correction phase with three steps for each group: ADD, SWAP, and REMOVE.

- **ADD:** If every user in group $\mathcal{G}_i$ has achieved $\text{SE}_{\text{req}}$ with a probability higher than a threshold value $\zeta_{th}$, then more users can be added from the non-selected group ($\mathcal{G}_{\text{non}}$).
- **SWAP:** Users whose probability of achieving the $\text{SE}_{\text{req}}$ is lower than $\zeta_{th}$ can be swapped with users from the non-selected group ($\mathcal{G}_{\text{non}}$).
- **REMOVE:** If a user repeatedly fails to find a suitable association to meet the requirements, it is considered to be in '*worse state*' and can be removed from the user set (i.e., turned off or not scheduled). This step is necessary to prevent falling into the same loop repeatedly.

A detailed description of the correction phase is given in Alg. 2, which repeats for *nItr* iterations.

### C. Improvement Phase

After the correction phase, we can identify the users most likely to meet the requirements and maximize $\mathbb{E}\left\{\mathcal{U}_{\text{sat}}\right\}$. We can assign additional available APs based on their strongest channel gains to improve the probability of the selected users

**Algorithm 2** Correction Phase

**Input:** $\Gamma$, $\mathcal{L}$, $\mathcal{K}$, $\mathcal{G}$, $\mathcal{G}_{\text{non}}$, $\text{SE}_{\text{req}}$, *nItr*

```
Initialization : D ← 0^{L×K}
for itr = 1, 2, ... nItr do
    A^{itr} ∈ {0,1}^{L×K} ← Intermediate Assn. matrix
    A^{itr} ← Algorithm 1
    Calculate E{U_sat} using (14)
    if (E{U_sat} | A^{itr}) > (E{U_sat} | A^{itr-1}) then
        D ← A^{itr}
    end if
    for G_i in {G_1, 2, ..., G_{τ_p}} do
        Calculate ζ_t, ∀t ∈ G_i using (12) & (13).
        if ζ_t ≥ ζ_th : ∀t ∈ G_i then
            ADD: G_i ∪ {q} where q ∈ G_{non,i}
        else if ζ_t < ζ_th : ∃t ∈ G_i then
            SWAP: t ∈ G_i replaced with q ∈ G_{non,i}.
        else if ∃t ∈ G_i: worse state then
            REMOVE: Remove t from process.
        end if
    end for
end for
```

**Output:** $\mathbf{D}$

meeting the requirement and update $\mathbf{D}$. This enhances the initial association and improves SE as well as *satisfaction rate*. Details are captured in Alg. 3.

**Algorithm 3** Improvement Phase

**Input:** $\Gamma$, $\mathcal{L}$, $\mathcal{G}$, $\mathbf{D}$

```
for l = 1, 2, ..., L do
    for G_i in {G_1, G_2, ..., G_{τ_p}} do
        if no user from G_i is scheduled on l then
            Find the user k' with the strongest γ_{k'l} value
            D[l, k'] = 1
        end if
    end for
end for
```

**Output:** $\mathbf{D}$

TABLE I: Simulation Parameters

| **Parameter** | **Value** |
|---|---|
| Bandwidth (BW) | 20 MHz |
| $f_c$ | 3.5 GHz |
| UL tx. power $\eta_{ul}$ | 20 dBm |
| Noise Figure (nF) | 7 dB |
| Noise Power | -174 + 10 * $log_{10}(BW)$[MHz] + nF |
| Factory size | small = $50m \times 50m$ |
| | medium = $80m \times 80m$ |
| | large = $100m \times 100m$ |
| InF-DH Clutter Prop. | Density[%]=80, height & size[m]=6 & 2 |
| Height of AP & User | 10m & 1.5m |
| UL Frame Prop. | $\tau_p$ = 5 & $\tau_f$ = 200 |
| Combiner | P-MMSE (3) |

## V. Numerical Results

In this section, we evaluate the performance of the proposed algorithm under various factory sizes and different SE requirements. We considered three factory setups with a square geometry, small $= 50m \times 50m$, medium $= 80m \times 80m$ and large $= 100m \times 100m$. $L$ = 100 single antenna APs and $K$ = 40 single antenna users are deployed randomly and uniformly placed in the chosen geometry. We further adopt a dense clutter scenario with high AP (InF-DH) from the 3GPP Indoor Factory (InF) channel model [11] considering an industrial environment, where radio propagation is affected by the presence of a high clutter density. Table I captures the main simulation parameters. We first collect the $\text{SE}_{\text{pred}}$ and $\text{SE}_{\text{act}}$ values for a given geometry on a full-scale system and model a normal distribution to capture the statistics of $\Delta$SE. The proposed algorithm leverages this distribution model to find the optimal associations for maximizing $\mathbb{E}\{\mathcal{U}_{\text{sat}}\}$.

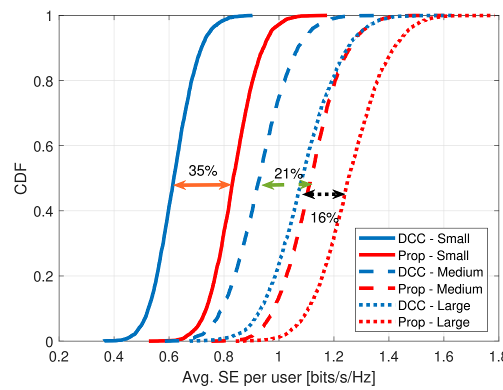


Fig. 3: Avg. per user SE analysis for all factory sizes $\text{SE}_{\text{req}} = 0.50$ bits/s/Hz

Results are generated using 4,000 random setups and evaluated the performance of the proposed algorithm by choosing DCC [5] as a benchmark. The proposed algorithm runs for 50 iterations considering a threshold probability of $\zeta_{th} = 0.3$. If $\zeta_k = 0$ over 10 times, that user is deemed to be in a worse state. Fig. 3 shows the Cumulative Distribution Function (CDF) of average SE per user. The proposed algorithm outperforms DCC in all scenarios, especially in smaller geometries where interference dominates. For instance, to guarantee $\text{SE}_{\text{req}} = 0.50$ bits/s/Hz in a smaller factory setup, the proposed algorithm shows a gain of 16-35% when compared to DCC depending on the environment. Sequentially allocating APs and finding a suitable schedulable user set improves resource utilization in interference-dominated scenarios. Giving up on a few highly interfering users can improve $\mathbb{E}\{\mathcal{U}_{\text{sat}}\}$.

The proposed algorithm aims to schedule a desirable set of users to maximize satisfaction rate; hence, it is crucial to evaluate the percentage of non-scheduled users and satisfaction rate to gauge decision credibility. Fig. 4 captures this, for the chosen factory setup for a minimum requirement of $\text{SE}_{\text{req}} = 0.25$, 0.50, and 0.75 bits/s/Hz. As expected, the percentage of not scheduled users increases as $\mathrm{SE_{req}}$ increases. To guarantee a satisfaction rate of $94\%$, approximately $9\%$ of total users are not scheduled for a smaller factory setup with $\mathrm{SE_{req}} = 0.25$. For a $\mathrm{SE_{req}} = 0.75$ setup, this increases to $37\%$ of not scheduled users and a satisfaction rate of $83\%$. From observations, it is evident that the satisfaction rate of scheduled users improves with an increase in the factory size. Indicating LSF approximations can replace instantaneous CSI in non-interference-limited scenarios and still make reliable decisions in interference-dominated situations.

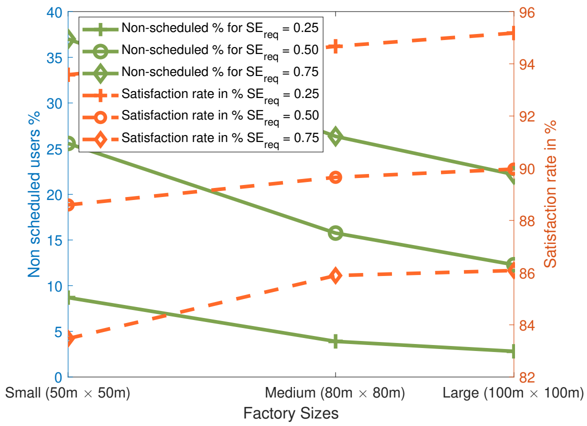


Fig. 4: Percentage of turned-off users and satisfaction rate for $\mathrm{SE_{req}} = 0.25$, $0.50$ and $0.75$ bits/s/Hz

Fig. 5 provides insights into the $10\%$-tile of satisfied users across various factory sizes and $\mathrm{SE_{req}}$ to evaluate challenging satisfaction scenarios. With bigger factories, more users can be satisfied, resulting in a higher number of satisfied users at any given $\mathrm{SE_{req}}$. To meet $\mathrm{SE_{req}} = 0.25$, approximately $90\%$ of total users are satisfied for a large factory, and this reduces to $80\%$ for a smaller factory. However, as $\mathrm{SE_{req}}$ increases, fewer users meet the requirements. For a smaller factory, as $\mathrm{SE_{req}}$ increases from $0.25$ to $0.75$, the total number of satisfied users reduces from $80\%$ to $45\%$. The proposed algorithm has demonstrated adaptability to different scenarios, resulting in significant gains in terms of the average per-user SE, satisfaction rate, and total number of satisfied users, particularly in interference-limited or challenging requirements.

## VI. Conclusion and Future Work

We have proposed a user association algorithm for cell-free mMIMO that aims to increase the number of satisfied users according to their QoS requirements. The developed method predicts the number of satisfied users based on large-scale fading knowledge, thus removing the need to acquire instantaneous channel knowledge of all users. The proposed algorithm identifies the users who can meet the requirements and find a suitable association to maximize the number of satisfied users based on an estimated probability to meet their QoS. Simulation results confirm the efficiency and adaptability of the proposed solution in terms of average achievable SE per user and number of satisfied users. Large factory deployments consistently achieve higher satisfaction rates, approximately $95\%$, at lower SE requirements; satisfaction rates decline to $83\%$ in interference-limited scenarios or in the case of higher SE requirements. Future work will further investigate integrating the probabilistic model into various radio resource allocation problems such as power control, dynamic channel assignment, etc., considering heterogeneous requirements and priorities according to realistic industrial scenarios.

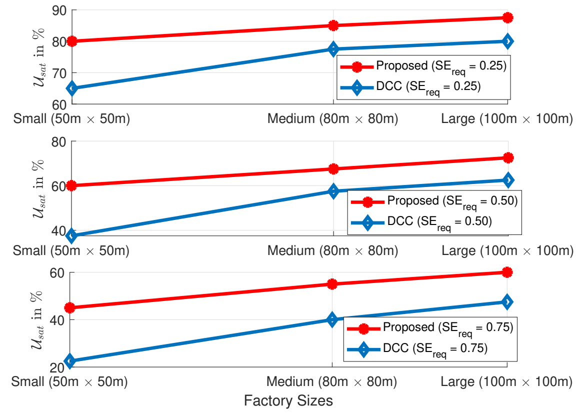


Fig. 5: 10 percentile of no.of satisfied users for varying $\mathrm{SE_{req}}$ and geometries in percentage

## References

[1] M. Arnold, P. Baracca, T. Wild, F. Schaich and a. S. t. Brink, "Measured Distributed vs Co-located Massive MIMO in Industry 4.0 Environments," Joint European Conference on Networks and Communications and 6G Summit (EuCNC/6G Summit), pp. 306-310, 2021.

[2] U. K. Ganesan, E. Björnson and a. E. G. Larsson, "RadioWeaves for Extreme Spatial Multiplexing in Indoor Environments," 54th Asilomar Conf. Signals Syst. and Comput, pp. 1007-1011, 2020.

[3] Özlem Tuğfe Demir, Emil Björnson, and Luca Sanguinetti (2021) "Foundations of User-Centric Cell-Free Massive MIMO", Foundations and Trends in Signal Processing: Vol. 14, No. 3-4, pp. 162-472.

[4] M. Alonzo, P. Baracca, S. R. Khosravirad and S. Buzzi, "Cell-Free and User-Centric Massive MIMO Architectures for Reliable Communications in Indoor Factory Environments," in IEEE Open Journal of the Communications Society, vol. 2, pp. 1390-1404, 2021.

[5] E. Björnson and a. L. Sanguinetti, "Scalable Cell-Free Massive MIMO Systems," IEEE Transactions on Communications, vol. 68, pp. 4247-4261, July 2020.

[6] C. D'Andrea and a. E. G. Larsson, "User Association in Scalable Cell-Free Massive MIMO Systems," 54th Asilomar Conference on Signals, Systems, and Computers, pp. 826-830, 2020.

[7] R. Wang, M. Shen, Y. He and a. X. Liu, "Performance of Cell-Free Massive MIMO With Joint User Clustering and Access Point Selection," IEEE Access, vol. 9, pp. 40860-40870, 2021.

[8] J. Wang, L. Dai, L. Yang and B. Bai, "Clustered Cell-Free Networking: A Graph Partitioning Approach," in IEEE Transactions on Wireless Communications, doi: 10.1109/TWC.2022.3233444.

[9] X. Gong and G. Wu, "Dynamic User Scheduling with User Satisfaction Rate in Cell-Free Massive MIMO," 2022 IEEE/CIC International Conference on Communications in China (ICCC Workshops), Sanshui, Foshan, China, 2022, pp. 100-105.

[10] J. Zhang, J. Zhang, E. Björnson and B. Ai, "Local Partial Zero-Forcing Combining for Cell-Free Massive MIMO Systems," IEEE Transactions on Communications, vol. 69, pp. 8459-8473, 2021.

[11] T. Jiang et al., "3GPP Standardized 5G Channel Model for IIoT Scenarios: A Survey," in IEEE Internet of Things Journal, vol. 8, no. 11, pp. 8799-8815, 1 June1, 2021.