Effective Channel DL Pilot-Based Estimation in User-Centric Cell-Free Massive MIMO Networks

Daynara D. Souza*, Marx M. M. Freitas*, Daniel B. da Costa†, Gilvan S. Borges*, André Mendes Cavalcante‡, Luca Valcarenghi⋆, and João C. Weyl Albuquerque Costa*
*Federal University of Pará, Belém, Brazil
†Technology Innovation Institute, Abu Dhabi, UAE
‡Ericsson Telecommunications S.A., Indaiatuba, Brazil
⋆Scuola Superiore Sant'Anna, Pisa, Italy
daynara@ufpa.br, marx@ufpa.br, daniel.costa@tii.ae, gilvan.borges@ifpa.edu.br,
andre.mendes.cavalcante@ericsson.com, luca.valcarenghi@santannapisa.it, jweyl@ufpa.br

***Abstract*—This paper investigates the performance of downlink (DL) pilot-based training to estimate the effective channel in user-centric cell-free massive multiple-input multiple-output (MIMO) networks. An algorithm for DL pilot assignment is proposed based on the level of interference between each user equipment (UE). It is proposed a refinement method for access point (AP) selection that controls the maximum AP cluster size of UEs. The strategy aims to control the maximum number of APs serving each UE to reduce the disparities among the AP cluster sizes. DL pilot-based training is compared with the blind, perfect and statistical channel state information (CSI) methods, assuming different precoding techniques, AP selection schemes, and the presence of pilot contamination. Our results demonstrate the following: (*i*) the proposed DL pilot assignment algorithm outperforms the baseline solutions; (*ii*) the proposed AP selection refinement method can improve the energy efficiency up to 86.6% without compromising the spectral efficiency; and (*iii*) DL pilot-based estimation reduces the normalized mean-square error significantly compared with blind and statistical CSI methods.**

***Index Terms*—AP selection refinement, cell-free massive MIMO networks, DL pilot-based estimation, DL pilot assignment, effective channel estimation, user-centric approach.**

I. Introduction

Cell-free (CF) massive multiple-input multiple-output (MIMO) networks consist of a large number of access points (APs) spread out in the coverage area, cooperating to serve the user equipments (UEs). Due to their distributed nature, these systems can provide increased macro-diversity and a more uniform spectral efficiency (SE) than the cell-based systems. The canonical version of CF considers that all APs serve all UEs. However, such a system demands enormous resource requirements (e.g., fronthaul signaling and processing) from the network, making it unscalable. [1], [2]. In this regard, the user-centric (UC) approach has emerged as an alternative to solve these drawbacks. By performing AP selection and limiting the number of UEs that each AP can serve, one can achieve scalability when the network resources (i.e., signal processing, signaling on fronthaul/backhaul, and total power) are independent of the number of UEs. However, strategies to control the size of AP clusters are still missing in the literature. The number of APs serving each UE can be very small or very large depending on the UE's position. In case of large AP clusters, the APs may become overloaded, impacting negatively the energy efficiency (EE) [3].

Another critical aspect of CF systems is that channel hardening may be less pronounced than in cellular systems due to the geographical distribution of the APs, with each one being equipped with a few antennas. The channel hardening phenomenon makes random channels behave almost deterministically when the number of antennas is large, enabling downlink (DL) data decoding based on statistical channel state information (CSI) of the effective channel [1], [2]. The low degree of hardening in CF systems may lead to the need for UEs to perform a more reliable channel estimation method. One of the main alternatives is the DL pilot-based estimation of the effective channel. This approach can be made more efficient by beamforming the DL pilots based on the uplink (UL) channel estimates at the APs, making the number of resource samples used on orthogonal pilot sequences a function of the number of UEs [4], [5]. However, a drawback of DL pilot-based estimation is that it has to use additional resource samples in training pilots, requiring small number of pilot sequences and, consequently, their reuse among the UEs.

This paper investigates the performance of the DL pilot-based estimation approach in UC CF massive MIMO networks, proposing novel algorithms for DL pilot assignment and cluster control for AP selection refinement. To the best of the authors' knowledge, such analysis has not been carried out in the literature yet. The analyzes are made considering different centralized and distributed precoding techniques. Numerical results are provided in terms of normalized mean square error (NMSE), SE, EE, and computational complexity (CC). The impact of different system configurations is investigated by varying some key parameters, such as coherence interval and pilot sequence length. Insightful discussions on the trade-off between estimation accuracy and overhead, and the performance of different estimation techniques are provided.

This work was supported by the Innovation Center, Ericsson Telecomunicações S.A, Brazil, by the National Council for Scientific and Technological Development (CNPq) and by the Coordination for the Improvement of Higher Education Personnel (CAPES).

The results show that the proposed DL pilot assignment can reduce the NMSE by 85%, while cluster control can improve the EE by 86.6%.

II. System Model

It is assumed a time-division duplex (TDD) CF massive MIMO system with M APs, equipped with N antennas each, and K single-antenna UEs, where $L = MN > K$. The TDD frame length is equal to the coherence interval. Hence, the channel is assumed to be static within a frame and but it varies independently for each frame. The channel between the AP m and UE k (and vice-versa), $\mathbf{h}_{mk} \in \mathbb{C}^{N\times 1}$, undergoes independent correlated Rayleigh fading, being defined as

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

where $\mathbf{R}_{mk} \in \mathbb{C}^{N\times N}$ represents the covariance matrix modeling the large-scale fading behavior, considering spatial channel correlation, path loss, and shadowing.

The UL channels are estimated by correlating a received UL pilot signal with a corresponding known pilot sequence and performing minimum mean square error (MMSE) estimation. For each coherence interval of length τ_c (in symbols), all UEs simultaneously send UL pilot sequences of length τ_{up} samples. The pilot sequences are assumed to be pair-wisely orthogonal, and different UEs can be assigned to the same pilot sequence, i.e., they are reused when $K > \tau_{up}$.

In the DL, it is performed an AP selection scheme to determine the subset $\mathcal{M}_k \subset \{1, \ldots, M\}$ of APs serving UE k, which can also be represented by the diagonal matrix $\mathbf{D}_{mk} \in \mathbb{N}^{N\times N}$, i.e.,

$$\mathbf{D}_{mk} = \begin{cases} \mathbf{I}_N & m \in \mathcal{M}_k \\ \mathbf{0}_N & m \notin \mathcal{M}_k. \end{cases} \tag{2}$$

Then, each AP serves the UEs by implementing power control and precoding based on the UL estimates. Each AP only serves a limited number of UEs to address scalability aspects [1], [2]. Thus, it follows that $|\mathcal{D}_m| \leq \tau_p$, where $\mathcal{D}_m$ is a subset containing the UEs served by the AP m. Let $q_k(n)$ be the n-th symbol intended for UE k. It is assumed that $\mathbb{E}\left\{\mathbf{q}(n)\mathbf{q}(n)^H\right\} = \mathbf{I}_K$, where $\mathbf{q}(n) \triangleq [q_1(n), \cdots, q_K(n)]^T$. The data signal sent by AP m can be written as

$$\mathbf{x}_m(n) = \sum_{k=1}^{K} \mathbf{D}_{mk}\mathbf{w}_{mk} q_k(n), \tag{3}$$

where the term $\mathbf{w}_{mk} \in \mathbb{C}^{N\times 1}$ represents the precoding vector, such that $\mathbb{E}\left\{\|\mathbf{w}_{mk}\|^2\right\} = \bar{\rho}_{mk} = \rho_{mk}/\sigma_{dl}^2$, with ρ_{mk} being the transmit power that AP m assigns to the UE k and σ_{dl}^2 is the noise power. UE k receives a linear combination of the signals transmitted by the APs, i.e.,

$$\begin{aligned} y_{d,k}(n) &= \sum_{m=1}^{M} \mathbf{h}_{mk}\mathbf{x}_m(n) + n_{d,k}(n) \\ &= \underbrace{\alpha_{kk} q_k(n)}_{\text{desired signal}} + \underbrace{\sum_{k'\neq k}^{K} \alpha_{kk'} q_{k'}(n)}_{\text{inter-user interference}} + \underbrace{n_{d,k}(n)}_{\text{noise}}, \end{aligned} \tag{4}$$

where

$$\alpha_{kk'} = \sum_{m=1}^{M} \mathbf{h}_{mk}^{\mathrm{H}} \mathbf{D}_{mk'}\mathbf{w}_{mk'}, \quad k' = 1, \cdots, K. \tag{5}$$

The noise at the receiver follows a complex Gaussian distribution with zero mean and unit variance, i.e., $n_{d,k}(n) \sim \mathcal{CN}(0,1)$, α_{kk} is the effective channel for UE k, and $\alpha_{kk'}, k' \neq k$ is the effective interfering channel. In order to coherently detect the transmitted data symbol q_k, UE k should have sufficient knowledge of the effective channel α_{kk}, which can be acquired by estimation or statistical CSI [4]–[7].

III. AP Selection Refinement Method

This paper proposes a strategy to control the maximum AP cluster size of the UEs, by restricting the cardinality ($|\mathcal{M}_k|$) of the AP clusters to a limit called C_{max}, $\forall k \in \{1, \cdots, K\}$. The latter represents the maximum number of APs that each UE can connect. Let A_c denote the number of connections that the M APs (each serving at most τ_{up} UEs) can provide to the network. C_{max} can be calculated as $\mathrm{C}_{max} = \max\left(1, \alpha A_c/K\right)$, where $A_c = \tau_{up}M$, and $0 < \alpha \leq 1$ is a refinement parameter that modifies the stringency of C_{max}.

Therefore, when a UE is connected to an excessive number of APs (i.e., $|\mathcal{M}_k| \geq \mathrm{C}_{max}$) after performing AP selection, a central processing unit (CPU) is activated to drop the UE's connection with the APs having the weakest channel gains so that $|\mathcal{M}_k| = \mathrm{C}_{max}$. To this end, the CPU performs a sort operation in ascending order to identify the APs with the weakest channel gain to the UE k, with $\bar{\beta}_{m1} \leq \bar{\beta}_{m2} \leq \cdots \leq \bar{\beta}_{mk}$, where $\bar{\beta}_{mk}$ denotes the sorted version of $\beta_{mk} = \mathrm{tr}\{\mathbf{R}_{mk}\}/N$, $\forall m \in \mathcal{M}_k$. Then, it calculates $\mathrm{E}_k = |\mathcal{M}_k| - \mathrm{C}_{max}$, where E_k is the number of excessive APs in the AP cluster of the UE k. Let $\mathcal{E}_k$ denote the subset containing indexes of the first E_k APs presenting the lowest values in $\bar{\beta}_{mk}$. Thus, the CPU imposes that $\mathbf{D}_{mk} = \mathbf{0}_N$, $\forall m \in \mathcal{E}_k$. This policy aims to reduce the disparity among the cluster sizes of UEs by using the single refinement parameter α.

IV. DL Pilot-Based Channel Estimation and Pilot Assignment

When using DL pilot-based estimation scheme, the APs send DL pilots to the UEs by precoding them based on the UL channel estimates [4], [6]. The AP $m \in \mathcal{M}_{k'}$ precodes the DL pilot sequences $\boldsymbol{\psi}_{z_{k'}} \in \mathbb{C}^{\tau_{dp}\times 1}$, where τ_{dp} is the DL pilot length, such that the received DL pilot at the UE k can be written as

$$\mathbf{y}_{dp,k} = \sqrt{\tau_{dp}} \sum_{k'=1}^{K} \alpha_{kk'} \boldsymbol{\psi}_{z_{k'}} + \mathbf{n}_{dp,k}, \tag{6}$$

where $\mathbf{n}_{dp,k}$ is the noise vector whose elements are independent and identically distributed (i.i.d.) $\mathcal{CN}(0,1)$ random variables (RVs). It is assumed that the DL pilot sequences are pair-wisely orthonormal, i.e.,

$$\boldsymbol{\psi}_{z_1}^{\mathrm{H}} \boldsymbol{\psi}_{z_2} = \begin{cases} 1, \text{ if } z_1 = z_2 \\ 0, \text{ if } z_1 \neq z_2, \end{cases} \tag{7}$$

and are also reused among the UEs when $K > \tau_{dp}$. Accordingly, UE k correlates the received signal with a known pilot sequence $\boldsymbol{\psi}_{z_k}$ in order to estimate its effective channel α_{kk}, such that

$$\hat{y}_{dp,k} = \boldsymbol{\psi}_{z_k}^{\mathrm{H}} \mathbf{y}_{dp,k} = \sqrt{\tau_{dp}}\alpha_{kk} + \sqrt{\tau_{dp}} \sum_{k' \neq k}^{K} \alpha_{kk'} \boldsymbol{\psi}_{z_k}^{\mathrm{H}} \boldsymbol{\psi}_{z_{k'}} + n_{p,k}, \tag{8}$$

where $n_{p,k} = \boldsymbol{\psi}_{z_{k'}}^{\mathrm{H}} \mathbf{n}_{dp,k}$. Then, the UE k performs the linear MMSE estimation of its channel α_{kk}, which is given by

$$\hat{\alpha}_{kk} = \mathbb{E}\left\{\alpha_{kk}\right\} + \frac{\mathrm{Cov}\left\{\alpha_{kk}, \hat{y}_{dp,k}\right\}}{\mathrm{Cov}\left\{\hat{y}_{dp,k}, \hat{y}_{dp,k}\right\}} \left(\hat{y}_{dp,k} - \mathbb{E}\left\{\hat{y}_{dp,k}\right\}\right). \tag{9}$$

The second term in (8) contains the pilot contamination effect generated by the pilot-sharing UEs. Algorithm 1 presents a DL pilot assignment method that minimizes the pilot contamination interference, where the index of the pilot assigned to the UE k is denoted as $z_k \in \{1, \ldots, \tau_{dp}\}$. In this one, the first τ_{dp} UEs are assigned to orthogonal sequences. The remaining UEs are assigned to the pilot that causes the lowest pilot contamination, given by the average power of the DL pilot contamination term $\mathbb{E}\{|\alpha_{ki}|^2\}$. This method is similar to the UL pilot assignment proposed by [2], with the difference being the pilot contamination term.

Algorithm 1: DL pilot assignment that aims to minimize pilot contamination

---

Input: DL pilot length τ_{dp}.
for $k = 1, \ldots, \tau_{dp}$ **do**
  $z_k \leftarrow k$
end
for $k = \tau_{dp} + 1, \ldots, K$ **do**
  $\zeta \leftarrow \underset{z \in \{1,\ldots,\tau_{dp}\}}{\arg\min} \sum_{i=1, z_i = z}^{k-1} \mathbb{E}\{|\alpha_{ki}|^2\}$
  $z_k \leftarrow \zeta$
end
Output: Pilot assignment indexes $z_1, \ldots, z_K$

V. Normalized Mean Square Error, Spectral and Energy Efficiency

The performance of the DL pilot-based estimation method can be computed by the NMSE between the channel estimate $\hat{\alpha}_{kk}$ and the effective channel α_{kk}, i.e.,

$$\mathrm{NMSE}_k = \frac{\mathbb{E}\left\{\left|\alpha_{kk} - \hat{\alpha}_{kk}\right|^2\right\}}{\mathbb{E}\left\{\left|\alpha_{kk}\right|^2\right\}}. \tag{10}$$

The achievable DL SE for UE k based on the use-and-then-forget (UatF) lower bound can be computed as [4], [7]

$$\mathrm{SE}_k = \frac{\tau_d}{\tau_c} \log_2\left(1 + \mathrm{SINR}_k\right), \tag{11}$$

with

$$\mathrm{SINR}_k = \frac{\left|\mathbb{E}\left\{\frac{\alpha_{kk}}{\hat{\alpha}_{kk}}\right\}\right|^2}{\mathrm{Var}\left\{\frac{\alpha_{kk}}{\hat{\alpha}_{kk}}\right\} + \sum\limits_{k' \neq k}^{K} \mathbb{E}\left\{\left|\frac{\alpha_{kk'}}{\hat{\alpha}_{kk}}\right|^2\right\} + \mathbb{E}\left\{\frac{1}{|\hat{\alpha}_{kk}|^2}\right\}}, \tag{12}$$

where $\hat{\alpha}_{kk}$ is the estimate of the effective channel α_{kk} and SINR_k is the signal-to-interference-plus-noise ratio (SINR) of UE k. To account for channel estimation overhead, the SE is multiplied by a pre-log factor $\tau_d/\tau_c = (\tau_c - \tau_p)/\tau_c = 1 - \tau_p/\tau_c$, i.e., the fraction of samples used for DL data transmission, where τ_p is the total number of samples used for UL and DL pilot estimation.

The total EE in bit/Joule is defined as the ratio between the total throughput $R_t = B\sum_{k=1}^{K} \mathrm{SE}_k$ in bit/s, where B is the bandwidth in Hz, and the total power consumed by all APs in Watts, including the consumption of amplifiers, circuits and backhaul links connecting them to the CPU [3], i.e.,

$$\mathrm{EE}_t = \frac{R_t}{\sum\limits_{m=1}^{M} \left\{\frac{\sigma_{dl}^2}{\gamma_m}\mathbb{E}\left\{\|\mathbf{x}_m\|^2\right\} + NP_{\mathrm{tc},m} + P_{\mathrm{bh},m}\right\}}, \tag{13}$$

where $0 < \gamma_m \leq 1$ denotes the efficiency of the power amplifier, and $P_{\mathrm{tc},m}$ is the power required of each antenna of the AP m to run internal components, such as converters and filters. Additionally, $P_{bh,m}$ is the power that the backhaul link connecting the CPU and AP m consumes, given by $P_{\mathrm{bh},m} = P_{0,m} + P_{\mathrm{bt},m}\, B \sum_{k \in \mathcal{D}_m} \mathrm{SE}_k$, where $P_{0,m}$ is a fixed power consumption of each backhaul and $P_{\mathrm{bt},m}$ is the traffic-dependent power in Watt per bit/s.

VI. Numerical Results

In order to evaluate the network performance of the proposed solutions, Monte-Carlo simulations are run. The simulation scenario consists of $M = 100$ APs, each equipped with $N = 4$ antennas, and $K = 20$ UEs covering a 1 km^2 rectangular area. It is assumed that the UEs are uniformly distributed into the coverage area and that the APs are placed following a hard core point process (HCPP). The propagation model adopted for the simulations is the 3GPP Urban Micro (UMi) path-loss model in which the line-of-sight (LOS) condition uses the probability functions defined in 3GPP TR 38.901 [8]. The correlation matrices $\mathbf{R}_{mk}$ are computed using the UMi path-loss model and the local scattering spatial correlation model presented in [2]. The simulations parameters are presented in Table I. The parameters for EE are set as $\gamma_m = 0.4$, $P_{\mathrm{tc},m} = 0.2$ W, $P_{0,m} = 0.825$ W, and $P_{\mathrm{bt},m} = 0.25$ W/(Gbit/s) [3].

It is considered two types of network implementations: (i) distributed configuration, where each AP performs the channel estimation, precoding and power allocation locally, and (ii) centralized configuration, where these tasks are performed by the CPU [1]. For the distributed implementation, the power coefficients at AP m are set as $\rho_{mk} = \rho_d\sqrt{\beta_{mk}}/\sum_{k' \in \mathcal{D}_m}\sqrt{\beta_{mk'}}$, where ρ_d is the maximum DL transmit power. For the centralized one, it is used the scalable fractional power control [2]. In order to compute the precoding vectors, it is employed the partial MMSE (P-MMSE) and partial regularized zero-forcing (P-RZF) for the centralized implementation. For the distributed, it is utilized the local partial MMSE (LP-MMSE) and maximum ratio (MR). These techniques were chosen due to their scalability features [1].

TABLE I: Parameters and models used in the simulations.

Parameter	Value
Effective environment height, h_E	1.0 m
Shadow fading standard deviation, σ_{SF}	4 dB
Antenna height AP, UE, h_{AP}, h_{UE}	11.65 m, 1.65 m
RX noise Figure, NF	8 dB
Coherence interval τ_c	200 samples
UL and DL pilot length	$\tau_{up} = \tau_{dp} = 10$ samples
Carrier Frequency f, Bandwidth	3.5 GHz, 100 MHz
AP's total DL power	23 dBm
UE's total UL power	20 dBm
Angular standard deviations (ASDs)	$\sigma_\varphi = \sigma_\theta = 15^\circ$
Antenna spacing	1/2 wavelength distance

Fig. 1 shows the cumulative distribution function (CDF) of the NMSE of different DL pilot assignment methods. One can note that the proposed DL pilot assignment in Algorithm 1, which is a non-joint method that aims to minimize pilot contamination, performs better than the other ones, reducing the NMSE by 85% compared with the joint method in the 50th percentile. The joint UL and DL strategy assigns orthogonal DL pilots to those UEs that use the same UL pilots [4]. This implies that non-joint methods can improve NMSE, even though they do not consider the number of pilot sequences as a function of the number of UEs. Therefore, providing higher performance with the advantage of being scalable, differently from the joint strategies which require $\tau_{up}\tau_{dp} \geq K$.

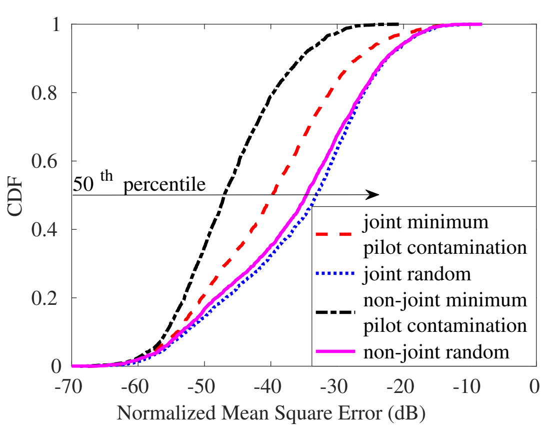


Fig. 1: CDF of the NMSE for different DL pilot assignment methods using MR precoding.

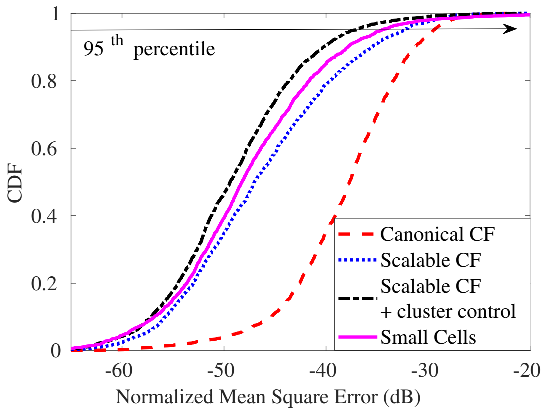
Fig. 2 shows the impact of AP selection in the DL pilot-based estimation. In Fig. 2a, one can note that AP selection techniques are crucial for reducing channel estimation errors since the NMSE is reduced in all AP clustering schemes compared to canonical CF. More specifically, the NMSE becomes even smaller in strategies that makes the UE connect to fewer APs, such as the proposed cluster control and small cells. The cluster control can reduce the NMSE up to 5 dB for the 95th percentile compared to the scalable CF strategy. Nonetheless, a solution like small cells (a single AP serving the UE, simulated by setting $C_{max} = 1$) can damage the SE as Fig. 2b depicts. One can note that even though it provides higher levels of EE in Fig. 2c, it does not provide a uniform coverage. This implies that the parameter α, employed in the cluster control strategy, has to be well fitted to avoid decreases in SE. For $\alpha = 25\%$, the cluster control can provide gains of up 19% to the 95th percentile of the SE for MR precoding and can improve the EE up to 86.6% for the P-MMSE. It happens because the cluster control provides the same SE while reducing the average number of APs serving each UE ($|\mathcal{M}_k|$) and the average number of UE served by an AP ($|\mathcal{D}_m|$), as Table II demonstrates, which also reduces the number of complex scalars exchanged with the fronthaul/backhaul and CC [1].

TABLE II: Average number of APs per UE ($|\mathcal{M}_k|$) and UEs per AP ($|\mathcal{D}_m|$). All standard deviations were around zero.

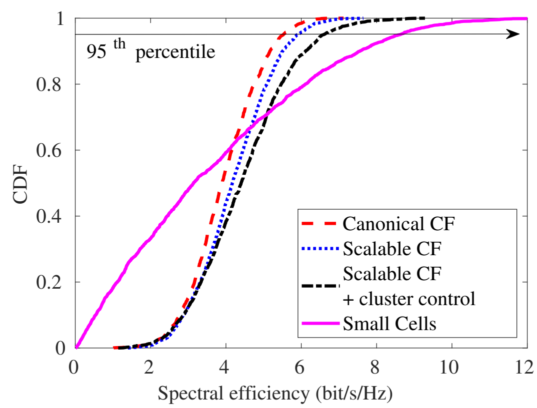
Method	mean $\lvert\mathcal{M}_k\rvert$	mean $\lvert\mathcal{D}_m\rvert$
Scalable CF	50	10
Scalable CF + cluster control	12	2.4

Fig. 3 evaluates the CDF of the NMSE and SE for MR and LP-MMSE precoding. The performance of DL pilot-based estimation is compared with blind estimation, statistical CSI, and perfect CSI. Blind estimation is a method that does not require time-frequency resource samples. It uses the average power of the received DL data signals to estimate the effective channel and is implemented following the steps presented in [9]. The conventional statistical CSI method uses average effective channel $\mathbb{E}\{\alpha_{kk}\}$ as the estimates, and an achievable DL SE can be computed using the *hardening bound* [1], [2]. The perfect CSI curves represent the SE when the UE has perfect knowledge of the effective channel, achieved in a genie-aided manner [2]. The estimation overhead is $\tau_p = \tau_{up} + \tau_{dp}$ for DL pilot-based estimation, whereas for the other estimation methods $\tau_p = \tau_{up}$, which is used in the pre-log factor of the SE. From Fig. 3, one can note that the DL pilot-based estimation decreases the NMSE compared with blind and conventional statistical CSI approaches for both precoding schemes. For instance, DL pilot-based estimation can reduce the NMSE by 95% and 71% compared with blind estimation for MR and LP-MMSE, respectively. From the SE results using MR precoding, it can be seen that blind and DL pilot-based estimation methods improve the system performance significantly compared with statistical CSI, up to 80.7% improvement for the 80% likely UEs. However, for LP-MMSE, DL pilot estimation degrades the SE due to the estimation overhead, whereas the improvement is slight for the blind approach. It is expected that the need to estimate the effective channel with LP-MMSE would be lower due to the higher degree of channel hardening.

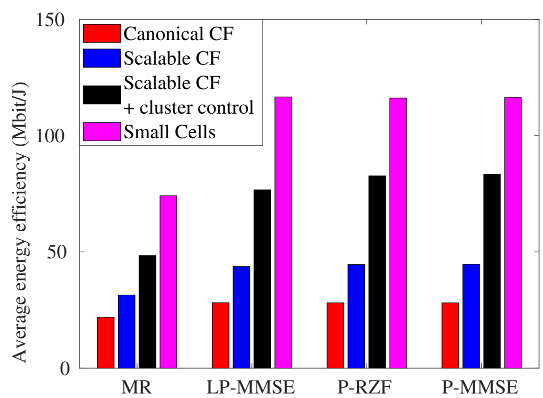
In Fig. 4, the coherence interval length, τ_c, is varied to evaluate the performance in terms of average SE for higher mobility and dispersion scenarios. One can note that blind



(a) NMSE of MR

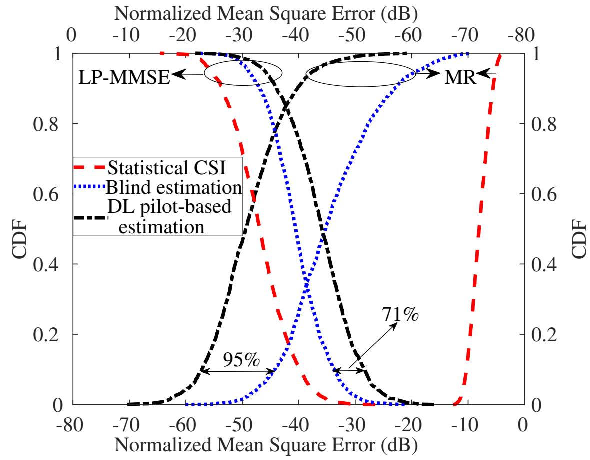


(b) SE of MR

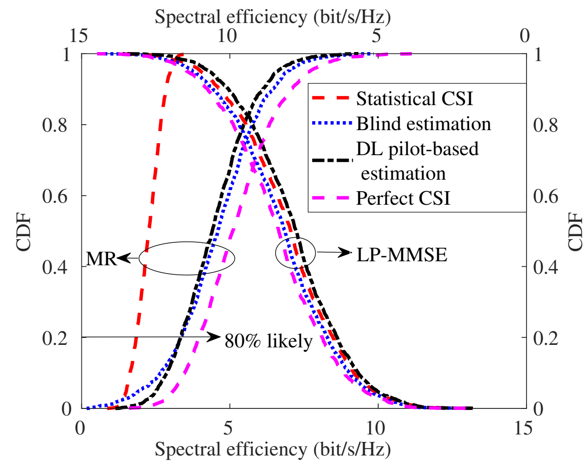


(c) EE

Fig. 2: CDF of the NMSE, SE and average total EE of DL pilot-based estimation for different AP selection schemes.



(a) NMSE



(b) SE

Fig. 3: CDF of the NMSE and SE for different estimation methods.

estimation is the best method for $\tau_c \leq 200$ samples, while DL pilot estimation has the best performance for $\tau_c \geq 300$ samples. The reason for that is the low pre-log factor value of the DL pilot estimation for the $\tau_c \leq 200$, although it has higher estimation accuracy. These results indicate that deciding which method is best for the UC CF scenario may vary depending on the system parameters.

To analyze the impact of the number of DL pilot sequences (τ_{dp}) in the performance of DL pilot-based method, Fig. 5 shows the average NMSE and SE versus τ_{dp}. As expected, the estimation accuracy improves as τ_{dp} increases since there is less pilot contamination interference. The SE for MR precoding can be maximized by setting $\tau_{dp} = 5$, which balances its estimation overhead and accuracy. On the other hand, the average SE decreases as the number of DL pilot sequences increases for the other precoding schemes, becoming best to set τ_{dp} at values as small as one, or performing another channel estimation method such as the blind one.

The CC of blind and DL pilot-based estimation methods can also be compared. Assuming that the statistical values are precomputed, known, and stay the same throughout communication, there is no need to evaluate its CC for each coherence interval. This also means that the CC for statistical CSI is zero. The CC of DL pilot-based estimation is a function of the pilot length, τ_{dp}, as each UE has to correlate its pilot sequence with the received pilot signal, an operation that requires $\tau_{dp} - 1$ additions and τ_{dp} multiplications. For blind estimation, each UE has to compute the average sample power of the received signal, its CC depends on the number of samples used for DL

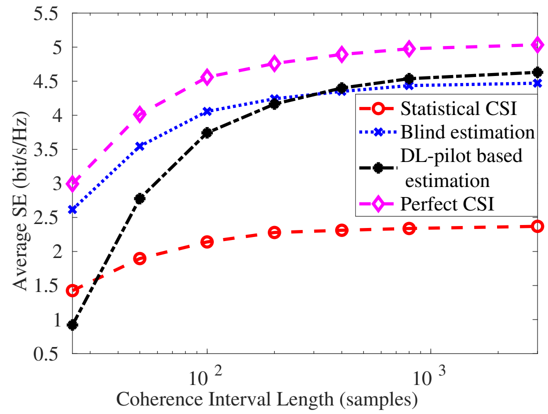
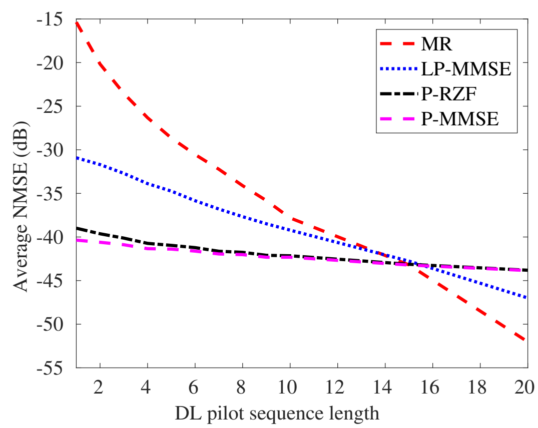
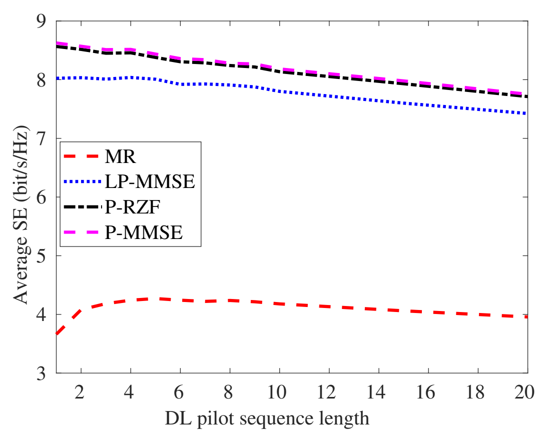


Fig. 4: Average SE versus coherence interval length.



(a) NMSE



(b) SE

Fig. 5: DL pilot-based estimation average NMSE and SE versus the DL pilot sequence length.

data transmission and requires $\tau_d - 1$ additions [9]. Table III summarizes these values, where it can be noted that their total CC is the same if $\tau_{dp} = \tau_d/2$. As typically $\tau_{dp} < \tau_d/2$, it is expected that the UEs using blind estimation will have higher CC than DL pilot-based estimation. Furthermore, there is also the CC of the proposed DL pilot assignment in Algorithm 1, which is $\mathcal{O}((K - \tau_{dp})\tau_{dp})$ for all UEs K, but the UEs do not perform this task.

TABLE III: CC for each UE by using different estimation methods, in every coherence interval.

Method	# additions	# multiplications	Total
DL pilot estimation	$\tau_{dp} - 1$	τ_{dp}	$2\tau_{dp} - 1$
Blind estimation	$\tau_d - 1$	0	$\tau_d - 1$

VII. Conclusions

This paper investigated the performance of DL pilot-based estimation in UC CF massive MIMO systems. It is proposed an algorithm for DL pilot assignment that aims to minimize pilot contamination. The paper also proposed a refinement method for AP selection that controls the maximum AP cluster size of UEs. The results demonstrate that the proposed DL pilot assignment algorithm can outperform the baseline solutions, reducing by 85% the NMSE, and has the advantage of being scalable. The proposed AP selection refinement method can improve the EE by up to 86.6% without compromising the SE. The results also demonstrated that for MR precoding, blind and DL pilot-based estimation methods can improve the system performance significantly compared to using only the statistical CSI. DL pilot-based estimation decreases NMSE and increases the SE of the 80% likely UEs by about 80.7% compared with statistical CSI. It is demonstrated that when the pilot overhead is small, DL pilot-based estimation is the best method since it has higher estimation accuracy. The analyzes indicate that it is possible to get the best performance with any one of the three estimation methods depending on the system parameters. This opens the way for future works to design self-regulated resource management strategies adapted for each scenario.

References

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

[2] Ö. Demir, E. Björnson, and L. Sanguinetti, *Foundations of User-Centric Cell-Free Massive MIMO*, ser. Foundations and Trends in Signal Processing. Hanover, MA, USA: Now Publ., 2021.

[3] H. Q. Ngo, L. Tran, T. Q. Duong, M. Matthaiou, and E. G. Larsson, "On the total energy efficiency of cell-free massive MIMO," *IEEE Trans. Green Commun. Netw*, vol. 2, no. 1, pp. 25–39, 2018.

[4] G. Interdonato, H. Q. Ngo, P. Frenger, and E. G. Larsson, "Downlink training in cell-free massive MIMO: A blessing in disguise," *IEEE Trans. Wireless Commun.*, vol. 18, no. 11, pp. 5153–5169, 2019.

[5] G. Interdonato, P. Frenger, and E. G. Larsson, "Utility-based downlink pilot assignment in cell-free massive MIMO," in *WSA 2018; 22nd International ITG Workshop on Smart Antennas*, 2018, pp. 1–8.

[6] G. Interdonato, H. Q. Ngo, E. G. Larsson, and P. Frenger, "How much do downlink pilots improve cell-free massive MIMO?" in *2016 IEEE Global Communications Conference (GLOBECOM)*, 2016, pp. 1–7.

[7] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, 2017.

[8] "Study on channel model for frequencies from 0.5 to 100 GHz, (release 16)," 3GPP, Sophia Antipolis, France, Tech. Rep. TR 38.901, 2019.

[9] D. D. Souza, M. M. M. Freitas, G. S. Borges, A. M. Cavalcante, D. B. da Costa, and J. C. W. A. Costa, "Effective channel blind estimation in cell-free massive MIMO networks," *IEEE Wireless Commun. Lett.*, vol. 11, no. 3, pp. 468–472, 2022.