Alvarez Polegre, A., Sanguinetti, L. & García Armada, A. (2021). Pilot Decontamination Processing in Cell-Free Massive MIMO. *IEEE Communications Letters*, 25(12), 3990-3994.

DOI: 10.1109/lcomm.2021.3118890

Pilot Decontamination Processing in Cell-Free Massive MIMO

Alberto Álvarez Polegre, *Student Member, IEEE*, Luca Sanguinetti, *Senior Member, IEEE*, and Ana García Armada, *Senior Member, IEEE*

***Abstract*—This letter focuses on the pilot contamination problem in the uplink and downlink of cell-free massive multiple-input multiple-output networks with different degrees of cooperation between access points. The optimum minimum mean square error processing can take advantage of large-scale fading coefficients for canceling the interference of pilot-sharing user-equipments and thus achieves asymptotically unbounded capacity. However, it is computationally demanding and can only be implemented in a fully centralized network. Here, sub-optimal schemes are derived that provide unbounded capacity with linear-growing complexity and using only local channel estimates but global channel statistics. This makes them suited for both centralized and distributed networks. In this latter case, the best performance is achieved with a generalized maximum ratio combiner that maximizes a capacity bound based on channel statistics only.**

***Index Terms*—Cell-free massive MIMO, pilot contamination, uplink combining, downlink precoding, reduced-complexity minimum mean square error, generalized maximum ratio.**

I. Introduction

The interference generated by user equipments (UEs) that transmit the same pilot sequence for channel acquisition is known as pilot contamination [1], [2]. This interference not only reduces the estimation quality but also makes the channel estimates statistically dependent. This latter effect makes it particularly hard to mitigate the interference (known as coherent interference) between these UEs. Although existing in most networks, this problem has a greater impact on those where the large number of UEs requires a high pilot-reuse factor. This was the case of cellular massive multiple-input multiple-output (mMIMO) networks [1], [2] and has been inherited by the new cell-free (CF) mMIMO, but to a larger extent due to the CF implementation [3], [4]. In the beyond 5G era, CF mMIMO is seen as a promising solution to provide uniform service for massive Machine Type Communicantions (mMTC) and Ultra-Reliable Low-Latency Communications (URLLC). These two use cases can turn pilot contamination into a totally unmanageable difficulty due to the amount of active UEs.

Signal processing strategies play a key role when dealing with pilot contamination. Particularly, the minimum mean square error (MMSE) scheme was shown to achieve optimal performance in CF mMIMO networks, in contrast to the classical maximum ratio (MR) strategy, which is severely limited by coherent interference [5]. The main drawback of MMSE processing is that it comes at the cost of requiring a tight synchronized and centralized implementation, and high computational resources when the number of access points (APs) grows large. Considering capacity-limited wired connections and scalability concerns, such an implementation may limit the benefits of the CF architecture. As far as we are concerned, there are not any low-complexity processing schemes that can deal with pilot contamination available in the CF mMIMO literature.

Building on [6], in this letter we introduce a reduced-complexity, though sub-optimal, MMSE-based combining scheme that achieves unbounded capacity in high pilot-contaminated CF mMIMO networks (rather than a distributed mMIMO network) under different degrees of cooperation. The scheme can be implemented either in a centralized or a distributed fashion since only local channel estimates are required, in addition to global channel statistics. An alternative, yet optimal, scheme is also proposed by using a more conservative capacity bound that depends solely on statistical channel knowledge and provides an achievable rate when channel state information (CSI) is not available for final data detection. This is the case of a decentralized network in which the channels are estimated at the APs while data detection is performed at the central processing unit (CPU). These combiners are built based on the linear independence experienced by the global covariance matrices and the asymptotic behavior for a large number of APs, though they are valid under a finite regime. Neither of these processing schemes are constrained by pilot contamination and both show linear-growing complexity. Thanks to the uplink (UL)-downlink (DL) duality [7], precoding schemes can be obtained for the DL segment. These schemes are also valid if a user-centric approach is taken (e.g., [7], [8]).

Notation: The superscripts ^{T}, * and ^{H} denote transpose, conjugate, and Hermitian transpose, respectively. We use $\mathcal{CN}(\mathbf{0}, \mathbf{R})$ to denote the circularly symmetric complex Gaussian distribution with zero mean and covariance matrix $\mathbf{R}$. The expected value of a random variable x is denoted by $\mathbb{E}\{x\}$, while diag denotes the main diagonal components of a matrix. We use $a_n \asymp b_n$ to denote $a_n - b_n \to_{n\to\infty} 0$ almost surely for two (random) sequences a_n, b_n.

Acknowledgements: This work has been supported by project IRENE-EARTH (PID2020-115323RB-C33 / AEI / 10.13039/501100011033).

II. Network Model

We consider a CF mMIMO network with M single-antenna APs. The APs serve jointly K single-antenna UEs, and are connected via fronthaul links to a CPU. The standard time division duplex (TDD) protocol of cellular mMIMO is used [7, Sec. 2.3.2], where the τ_c available channel uses are employed

A. Álvarez Polegre and A. García Armada are with the Department of Signal Theory and Communications, Universidad Carlos III de Madrid, Avda. Universidad 30, 28911, Leganés (Madrid), Spain (email: aalvarez@tsc.uc3m.es, agarcia@tsc.uc3m.es).

L. Sanguinetti is with the Dipartimento di Ingegneria dell'Informazione, University of Pisa, 65122 Pisa, Italy (email: luca.sanguinetti@unipi.it).

for: τ_p for UL training phase, τ_d for DL payload transmission and τ_u for UL payload transmission. Clearly, $\tau_c \geq \tau_p+\tau_d+\tau_u$.

The channel between AP m and UE k is $h_{mk} \sim \mathcal{CN}(0, \beta_{mk})$ where β_{mk} is the large-scale fading. We assume that h_{mk} and $h_{m'k}$ are uncorrelated for $m \neq m'$ due to the different location of the APs. We define $\boldsymbol{h}_k = [h_{1k}, \ldots, h_{Mk}]^T \in \mathbb{C}^M$, which follows a $\mathcal{CN}(\boldsymbol{0}_M, \boldsymbol{R}_k)$ distribution, where $\boldsymbol{R}_k = \text{diag}(\beta_{1k}, \ldots, \beta_{Mk})$ is diagonal. The matrices $\{\boldsymbol{R}_k : k = 1, \ldots, K\}$ are assumed to be known, but practical estimation methods can be found in the literature (e.g., [9, Sec. IV]).

The same pilot sequence $\boldsymbol{\varphi} \in \mathbb{C}^{\tau_p}$ with $\|\boldsymbol{\varphi}\|^2 = 1$ is used by all the UEs during the UL training phase. The channel coefficients $\{h_{mk} : k = 1 \ldots, K\}$ can be estimated either directly at AP m or at the the CPU. In the latter case, AP m sends the received pilot signal to the CPU via the fronthaul link. Regardless of where the channel estimation is carried out, the MMSE estimate of h_{mk} is [7, Sec. 4.1]

$$\widehat{h}_{mk} = \frac{\beta_{mk}}{\psi_m} \left(\frac{1}{\sqrt{\rho_p}} \boldsymbol{\varphi}^H \boldsymbol{y}_m \right) \tag{1}$$

where $\boldsymbol{y}_m$ is the received pilot signal

$$\boldsymbol{y}_m = \sqrt{\rho_p} \sum_{k=1}^{K} h_{mk} \boldsymbol{\varphi} + \boldsymbol{n}_m \tag{2}$$

and ρ_p is the pilot signal-to-noise ratio (SNR) while $\boldsymbol{n}_m \in \mathbb{C}^{\tau_p}$ is the (normalized) noise vector with independent elements distributed as $\mathcal{CN}(0, 1)$. Also, we call

$$\psi_m = \beta_{mk} + \underbrace{\sum_{i=1, i \neq k}^{K} \beta_{mi}}_{\text{Interference from pilot-sharing UEs}} + \frac{1}{\rho_p}. \tag{3}$$

The estimates and estimation errors $\widetilde{h}_{mk} = h_{mk} - \widehat{h}_{mk}$ are independent random variables distributed as $\widehat{h}_{mk} \sim \mathcal{CN}(0, \gamma_{mk})$ with $\gamma_{mk} = \beta_{mk}^2/\psi_m$ and $\widetilde{h}_{mk} \sim \mathcal{CN}(0, \varepsilon_{mk})$ with $\varepsilon_{mk} = \beta_{mk} - \gamma_{mk}$. The interference generated by the pilot-sharing UEs in (3) is known as pilot contamination, e.g., [2]. It reduces the estimation quality, and makes the estimates correlated with $\widehat{h}_{mi} = \frac{\beta_{mi}}{\beta_{mk}} \widehat{h}_{mk}$. Both effects deteriorate performance but only the second one is responsible of the coherent interference (e.g., [9]). To perform coherent processing at multiple APs, knowledge of $\boldsymbol{h}_k$ is necessary. This is obtained as $\widehat{\boldsymbol{h}}_k = [\widehat{h}_{1k}, \ldots, \widehat{h}_{Mk}]^T \in \mathbb{C}^M$ and is distributed as $\widehat{\boldsymbol{h}}_k \sim \mathcal{CN}(\boldsymbol{0}_M, \boldsymbol{\Gamma}_k)$ with $\boldsymbol{\Gamma}_k = \boldsymbol{R}_k \boldsymbol{\Psi}^{-1} \boldsymbol{R}_k$ and $\boldsymbol{\Psi} = \text{diag}(\psi_1, \ldots, \psi_M)$ leading to $\boldsymbol{\Gamma}_k = \text{diag}(\gamma_{1k}, \ldots, \gamma_{Mk})$.

III. Uplink Signal Combining

We address now the UL combining for both centralized and distributed processing. The optimal MMSE scheme is set as the upper benchmark that our proposals will be compared to. In the centralized scenario, each AP m acts only as a relay that forwards its received signals to the CPU, which performs both channel estimation and data detection. Assuming error-free fronthaul links, the joint UL data signal $\boldsymbol{y}^{\text{ul}}$ available at the CPU is mathematically equivalent to the signal model of a single-cell system where the correlated fading $\boldsymbol{h}_k$ has a diagonal spatial covariance matrix. At the CPU, the UE k data signal estimate is obtained as $\widehat{s}_k^{\text{ul}} = \boldsymbol{v}_k^H \boldsymbol{y}^{\text{ul}}$ where $\boldsymbol{v}_k \in \mathbb{C}^M$ is the centralized combiner. With MMSE channel estimation, an achievable spectral efficiency (SE) of UE k is $\text{SE}_k^{\text{ul}} = \frac{\tau_u}{\tau_c} \mathbb{E}\left\{\log_2\left(1 + \text{SINR}_k^{\text{ul}}\right)\right\}$ where the expectation is with respect to channel realizations and the effective signal-to-interference-plus-noise ratio (SINR) is [7, Sec. 5.1.1]

$$\text{SINR}_k^{\text{ul}} = \frac{|\boldsymbol{v}_k^H \widehat{\boldsymbol{h}}_k|^2}{\boldsymbol{v}_k^H \left(\sum\limits_{i=1, i \neq k}^{K} \widehat{\boldsymbol{h}}_i \widehat{\boldsymbol{h}}_i^H + \boldsymbol{Z} \right) \boldsymbol{v}_k} \tag{4}$$

with $\boldsymbol{Z} = \text{diag}(z_1, \ldots, z_M)$ being diagonal with elements

$$z_m = \sum_{i=1}^{K} \varepsilon_{mi} + \frac{1}{\rho_u}. \tag{5}$$

This SE is valid for any combiner $\boldsymbol{v}_k$ but requires the channel estimates $\widehat{\boldsymbol{h}}_k$ and estimation errors $\widetilde{\boldsymbol{h}}_k = [\widetilde{h}_{1k}, \ldots, \widetilde{h}_{Mk}]^T$ to be independent. This condition is satisfied with the MMSE channel estimator. An alternative bound that can be applied along with any channel estimator is the so-called use-and-then-forget (UatF).[1] This yields $\text{SE}_k^{\text{ul,UatF}} = \frac{\tau_u}{\tau_c} \log_2(1+\text{SINR}_k^{\text{ul,UatF}})$, where $\text{SINR}_k^{\text{ul,UatF}}$ is given by [7, Sec. 5.1.2]

$$\frac{\left|\mathbb{E}\left\{|\boldsymbol{v}_k^H \boldsymbol{h}_k|\right\}\right|^2}{\sum\limits_{i=1}^{K} \mathbb{E}\left\{\left|\boldsymbol{v}_k^H \boldsymbol{h}_i\right|^2\right\} - \left|\mathbb{E}\left\{|\boldsymbol{v}_k^H \boldsymbol{h}_k|\right\}\right|^2 + \frac{1}{\rho_u} \mathbb{E}\left\{\|\boldsymbol{v}_k\|^2\right\}} \tag{6}$$

and expectations are with respect to the channel realizations. Intuitively, $\text{SE}_k^{\text{ul,UatF}}$ is smaller than SE_k^{ul} since it relies on a simplified implementation in which the channel estimates are not used at the CPU for signal detection.

Regarding the distributed network case, the MMSE channel estimates are computed locally at the APs and are used to obtain local estimates of UE data. This approach is more suitable for those use cases where low-latency is crucial. Let $v_{mk} \in \mathbb{C}$ denote the local coefficient that AP m uses for UE k. Then, the UE k local data estimate is $\widehat{s}_{mk}^{\text{ul}} = v_{mk}^* y_m^{\text{ul}}$, with y_m^{ul} the UL signal at AP m. Any coefficient v_{mk} can be adopted. Unlike a fully centralized network, however, AP m can only use its own local channel estimates for the design of v_{mk}. The local estimates $\{\widehat{s}_{mk}^{\text{ul}} : m = 1, \ldots, M\}$ are then sent to the CPU for final decoding. We assume that the final estimate $\widehat{s}_k^{\text{ul}}$ is computed by averaging the local estimates, i.e., $\widehat{s}_k^{\text{ul}} = \sum_{m=1}^{M} \widehat{s}_{mk}^{\text{ul}}$. In a distributed network, the CPU does not have knowledge of channel estimates and thus only the statistics can be utilized to estimate $\widehat{s}_k^{\text{ul}}$. Since $\text{SE}_k^{\text{ul,UatF}}$ does not rely on channel estimates, it can be used to compute an achievable SE in a distributed network [7, Sec. 5.2.1].

A. MMSE Combining

As in cellular mMIMO [5], [9], the SINR in (4) is a generalized Rayleigh quotient with respect to $\boldsymbol{v}_k$ and thus is maximized by the MMSE combining vector, i.e.

$$\boldsymbol{v}_k = \check{\boldsymbol{v}}_k = \left(\sum_{i=1}^{K} \widehat{\boldsymbol{h}}_i \widehat{\boldsymbol{h}}_i^H + \boldsymbol{Z} \right)^{-1} \widehat{\boldsymbol{h}}_k. \tag{7}$$

[1]The name comes from the fact that the channel estimates are used for computing receive combining vectors and then effectively forgotten before signal detection takes place.

We will now analyze the asymptotic behavior of $\mathrm{SE}_k^{\mathrm{ul}}$ with the MMSE as $M \to \infty$. We assume that the *global* covariance matrices $\{\boldsymbol{R}_k : k = 1, \ldots, K\}$ are asymptotically linearly independent, as analytically defined in [2]. This assumption is fairly realistic since each entry of $\boldsymbol{R}_k$ is subject to a different channel gain. This is a key difference between the CF architecture and traditional single-cell mMIMO systems that allows CF systems to achieve unbounded capacity even with single-antenna APs. Under this condition, and building on [2, Th. 1], the next proposition follows.

Proposition 1. *If MMSE combining is used and $\{\boldsymbol{R}_k : k = 1, \ldots, K\}$ are asymptotically linearly independent, then SE_k^{ul} increases unboundedly as $M \to \infty$.*

Proof: It follows easily from [2, App. B] since the centralized case is basically a single-cell mMIMO network with diagonal spatial correlation matrices. ■

B. Reduced-Complexity MMSE Combining

Although the MMSE combining vector $\check{\boldsymbol{v}}_k$ is optimal and achieves unbounded capacity (as in cellular mMIMO), it requires to compute the $M \times M$ matrix inverse in (7) in every coherence block, which may be too computationally demanding when the network size is large. Note this translates into a computational complexity that grows with M^3. To overcome this issue, we follow the same approach in [6], and consider the asymptotic regime. However, in contrast to a distributed cellular mMIMO network, we consider the asymptotic regime in CF mMIMO systems to be achieved by letting $M \to \infty$. This consideration invalidates the straightforward use of those derivations available in [6] in a CF scenario. Nevertheless, as shown in the Appendix, we can propose an alternative combiner $\boldsymbol{v}_k = \bar{\boldsymbol{v}}_k = [\bar{v}_{1k}, \ldots, \bar{v}_{Mk}]^T$ with $\bar{v}_{mk}$ given by

$$\bar{v}_{mk} = \frac{1}{\sum\limits_{i=1}^{K} \varepsilon_{mi} + \frac{1}{\rho_u}} \sum_{i=1}^{K} b_{ki} \widehat{h}_{mi}, \tag{8}$$

where b_{ki} is the (k,i)th element of the matrix $\left(\boldsymbol{B} + \frac{1}{M}\boldsymbol{I}_K\right)^{-1}$. The entries of $\boldsymbol{B}$ are given by

$$[\boldsymbol{B}]_{ki} = \frac{1}{M} \sum_{m=1}^{M} \frac{\beta_{mk}\beta_{mi}}{\left(\sum\limits_{i=1}^{K} \beta_{mi} + \frac{1}{\rho_p}\right)\left(\sum\limits_{i=1}^{K} \varepsilon_{mi} + \frac{1}{\rho_u}\right)}. \tag{9}$$

Since $\check{\boldsymbol{v}}_k$ reduces to $\bar{\boldsymbol{v}}_k$ in the limiting regime $M \to \infty$, it follows that also $\bar{\boldsymbol{v}}_k$ achieves unbounded capacity in accordance to Proposition 1.[2] Compared to $\check{\boldsymbol{v}}_k$, however, the computational complexity required by $\bar{\boldsymbol{v}}_k$ is much lower. To see this, we use $\widehat{h}_{mk}$ to obtain

$$\bar{v}_{mk} = \bar{\varsigma}_{mk} \left(\frac{1}{\sqrt{\rho_p}} \boldsymbol{\varphi}^H \boldsymbol{y}_m\right) \tag{10}$$

where

$$\bar{\varsigma}_{mk} = \frac{1}{\left(\sum\limits_{i=1}^{K} \beta_{mi} + \frac{1}{\rho_p}\right)\left(\sum\limits_{i=1}^{K} \varepsilon_{mi} + \frac{1}{\rho_u}\right)} \sum_{i=1}^{K} b_{ki}\beta_{mi}. \tag{11}$$

[2]In practice, M will not be infinite. Numerical results will show that (8) works well for practical numbers of APs.

Under the assumption that coefficients $\{\bar{\varsigma}_{mk} : m = 1, \ldots, M\}$ are available, from (8) the complexity required for the computation of $\bar{\boldsymbol{v}}_k = [\bar{v}_{1k}, \ldots, \bar{v}_{Mk}]^T$ scales linearly with M, rather than as M^3 as with MMSE combining. Therefore, we refer to $\bar{\boldsymbol{v}}_k$ as reduced-complexity MMSE (RC-MMSE). If one considers a centralized CF mMIMO network with limited computational resources, the RC-MMSE combiner seems to be the most suitable approach available in the literature[3]. Likewise, notice that, if the large-scale coefficients are locally available at AP m (which is a reasonable assumption since they change every several coherence intervals and can be broadcasted by the CPU), $\bar{v}_{mk}$ can be locally computed. This implies that, unlike the MMSE combiner $\check{\boldsymbol{v}}_k$, it can also be adopted in a distributed network. In this case, however, the UatF bound must be used to compute an achievable SE since signal detection takes place without channel knowledge in a distributed network.

C. Generalized Maximum Ratio Combining

The MMSE combiner is obtained as a matrix transformation of $\widehat{\boldsymbol{h}}_k$. Inspired by this, we now assume that $\boldsymbol{v}_k = \boldsymbol{W}_k \widehat{\boldsymbol{h}}_k$ where $\boldsymbol{W}_k$ is an arbitrary matrix to be optimized. Unlike (7), we assume that this optimization can only be done on the basis of channel statistics (rather than of channel estimates). The question is how to optimally design $\boldsymbol{W}_k$ in order to not incur a significant SE loss. Following [6], [10], and considering again the asymptotic regime in CF mMIMO, we can use the SINR in (6) from the UatF bound to obtain:

$$\frac{\left|\mathrm{tr}(\boldsymbol{W}_k^H \boldsymbol{R}_k)\right|^2}{\sum\limits_{i=1}^{K} \left|\mathrm{tr}\left(\boldsymbol{\Psi}^{-1}\boldsymbol{R}_k \boldsymbol{W}_k^H \boldsymbol{R}_i\right)\right|^2 + \mathrm{tr}\left(\boldsymbol{W}_k \boldsymbol{\Gamma}_k \boldsymbol{W}_k^H \boldsymbol{U}_k\right)} \tag{12}$$

with $\boldsymbol{U}_k = \sum_{i=1}^{K} \boldsymbol{R}_i + \frac{1}{\rho_u}\boldsymbol{I}_M$ being diagonal. The matrix $\boldsymbol{W}_k$ that maximizes (12) is given by (e.g., [9, Sec. VII])

$$\boldsymbol{W}_k = \boldsymbol{U}_k^{-1} \left(\sum_{i=1}^{K} a_{ki} \boldsymbol{R}_i\right) \boldsymbol{R}_k^{-1} \tag{13}$$

where a_{ki} is the (k,i)th element of the matrix $\left(\boldsymbol{A} + \frac{1}{M}\boldsymbol{I}_K\right)^{-1}$. The entries of $\boldsymbol{A}$ are given by

$$[\boldsymbol{A}]_{ki} = \frac{1}{M} \sum_{m=1}^{M} \frac{\beta_{mk}\beta_{mi}}{\left(\sum\limits_{i=1}^{K} \beta_{mi} + \frac{1}{\rho_p}\right)\left(\sum\limits_{i=1}^{K} \beta_{mi} + \frac{1}{\rho_u}\right)}. \tag{14}$$

Plugging (13) into $\boldsymbol{v}_k = \boldsymbol{W}_k \widehat{\boldsymbol{h}}_k$ yields $\boldsymbol{v}_k = \tilde{\boldsymbol{v}}_k = [\tilde{v}_{1k}, \ldots, \tilde{v}_{Mk}]^T$ with (thanks to the diagonal structure of matrices)

$$\tilde{v}_{mk} = \frac{1}{\sum\limits_{i=1}^{K} \beta_{mi} + \frac{1}{\rho_u}} \sum_{i=1}^{K} a_{ki} \widehat{h}_{mi}. \tag{15}$$

We refer to (15) as generalized maximum ratio (GMR) combining. As for MMSE, we can analyze its asymptotic behavior as $M \to \infty$. With [6], [10], the next proposition follows.

[3]Though not shown here due to the lack of space, the RC-MMSE combiner outperforms the proposed local MMSE (L-MMSE) schemes in [5] under full pilot reuse.

TABLE I: Simulation parameters

Parameter	Value
Carrier frequency	2 GHz
Path loss exponent (α)	3.76
Shadowing standard deviation (σ_χ)	10
Noise figure at receivers	7 dB
DL transmission power	200 mW
UL transmission power	100 mW
Channel coherence interval length (τ_c)	200 samples
Training phase interval length (τ_p)	1 sample
Payload phase interval length ($\tau_d = \tau_u$)	$(\tau_c \quad \tau_p)/2$ samples

Proposition 2. *If GMR combining is used and* $\{\boldsymbol{R}_k : k = 1, \ldots, K\}$ *are asymptotically linearly independent, then* $SE_k^{ul,UatF}$ *increases unboundedly as* $M \to \infty$.

This proposition shows that GMR has the same scaling behavior of MMSE and RC-MMSE as $M \to \infty$. As with RC-MMSE, this is achieved by using only the channel statistics. Moreover, GMR has the same complexity of RC-MMSE as it follows by plugging $\widehat{h}_{mk}$ into (15) to obtain

$$\tilde{v}_{mk} = \tilde{\varsigma}_{mk} \left(\frac{1}{\sqrt{\rho_p}} \boldsymbol{\varphi}^H \boldsymbol{y}_m \right) \tag{16}$$

where

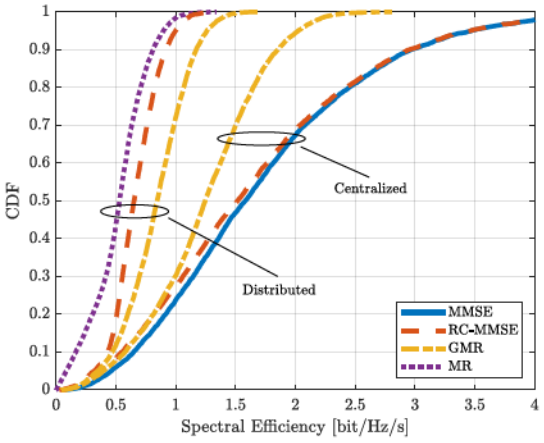
$$\tilde{\varsigma}_{mk} = \frac{1}{\left(\sum_{i=1}^{K} \beta_{mi} + \frac{1}{\rho_p} \right) \left(\sum_{i=1}^{K} \beta_{mi} + \frac{1}{\rho_u} \right)} \sum_{i=1}^{K} a_{ki} \beta_{mi}. \tag{17}$$

The key difference with respect to RC-MMSE in (8) is the scaling factor $\tilde{\varsigma}_{mk}$, which does not depend on the variances $\{\varepsilon_{mi} : i = 1, \ldots, K\}$ of the channel estimation errors but on the large-scale fading coefficients $\{\beta_{mi} : i = 1, \ldots, K\}$. This is because (15) maximizes $\mathrm{SE}_k^{\mathrm{ul,UatF}}$, which does not take into account the imperfect CSI in the decoding process. As RC-MMSE, GMR can also be used in a decentralized network. Whatsmore, anytime a distributed scenario with global statistics knowledge is considered, GMR should be used since it is the optimal scheme for MR-based processing.

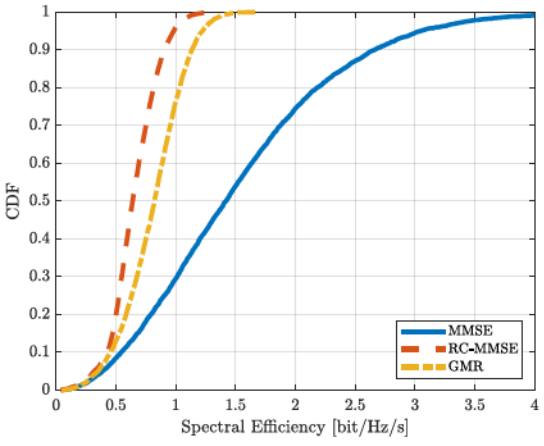
IV. Downlink Signal Precoding

We now consider the DL and design the precoding schemes for centralized and decentralized networks. As in the UL, the most advanced DL implementation is a fully centralized operation, where the AP only act as relays that transmit signals generated by the CPU. A distributed operation is also possible in the DL where the CPU encodes the DL data signals $\{s_k^{\mathrm{dl}} : k = 1, \ldots, K\}$ and sends them to the AP, which selects the precoding coefficients on the basis of local estimates. This is the key difference between the two operation modes in the DL. The information available at the UE for signal detection is the same with both implementations. Since there are no DL pilots, in both cases UEs have no knowledge of channel estimates and must rely on statistics. An achievable SE can be computed for the two operation modes by using the channel hardening bound. This yields $\mathrm{SE}_k^{\mathrm{dl}} = \frac{\tau_d}{\tau_c} \log_2(1 + \mathrm{SINR}_k^{\mathrm{dl}})$ with [7, Sec. 6.1.1]

$$\mathrm{SINR}_k^{\mathrm{dl}} = \frac{\left| \mathbb{E}\left\{ |\boldsymbol{h}_k^H \boldsymbol{w}_k| \right\} \right|^2}{\sum_{i=1}^{K} \mathbb{E}\left\{ |\boldsymbol{h}_k^H \boldsymbol{w}_i|^2 \right\} - \left| \mathbb{E}\left\{ |\boldsymbol{h}_k^H \boldsymbol{w}_k| \right\} \right|^2 + \frac{1}{\rho_d}}. \tag{18}$$



(a) UL.



(b) DL.

Fig. 1: CDF of average SE per UE for $K = 20$.

As seen, $\mathrm{SE}_k^{\mathrm{dl}}$ depends on the precoding vectors $\{\boldsymbol{w}_i : i = 1, \ldots, K\}$ of all UEs. This stands in contrast to the UL SE and makes optimal precoding design hard. A common heuristic approach relies on the UL-DL duality [7, Sec. 6.1.2], which holds between the UatF bound and hardening bound. Motivated by this duality, we select the DL precoding vectors as $\boldsymbol{w}_k = \frac{\boldsymbol{v}_k}{\sqrt{\mathbb{E}\{|\boldsymbol{v}_k^H \boldsymbol{v}_k|\}}}$. Precoding vectors can be chosen on the basis of RC-MMSE or GMR for either centralized or distributed networks. On the other hand, the MMSE precoder can only be used with a centralized implementation.

V. Performance Analysis

Up next the system performance is shown for the proposed schemes. The simulation parameters are reported in Table I. APs will be randomly deployed within a square coverage area of side D. The large-scale fading coefficients are computed as

$$\beta_{mk}[\mathrm{dB}] = -35.3 + 10\alpha \log_{10}(d_{mk}) + \chi_{mk} - \sigma_n^2 \tag{19}$$

where α is the path loss exponent, d_{mk} is the distance between AP m and UE k, $\chi_{mk} \sim \mathcal{N}(0, \sigma_\chi^2)$ is the shadow fading component and σ_n^2 the noise variance.

The cumulative distribution function (CDF) for the UL average SE per UE is shown in Fig. 1a. The number of

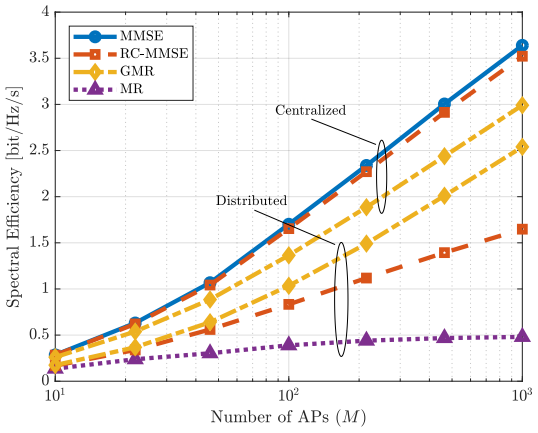


Fig. 2: Average UL SE per UE for $K = 5$.

APs is set to $M = 100$ and $K = 20$ randomly distributed UEs are dropped with $D = 250$ m. Both centralized and distributed versions of the RC-MMSE and GMR combiners are considered. As upper and lower benchmarks, the centralized MMSE and distributed MR combiners are also presented. This latter scheme is computed by letting the entries of $\boldsymbol{v}_k$ be $v_{mk} = \widehat{h}_{mk}$. Regarding centralized schemes, the proposed RC-MMSE scheme nearly matches the optimal combiner. This proves this new scheme is a very suitable alternative for lowering the required computational resources with little performance loss under severe pilot contamination. On the other side, the GMR shows a considerable gap. We can notably point out that for low SNR both RC-MMSE and GMR match each other. This is due to the fact that the RC-MMSE combiner relies on having good channel estimates, while GMR relies on channel coefficients instead. Focusing on the distributed approaches, the GMR combiner outperforms the RC-MMSE. This is reasonable since GMR is designed for maximization of the UatF bound.

The DL segment is shown in Fig. 1b for the same setup. Power allocation coefficients with MMSE and RC-MMSE are selected following a network-wide approach by letting $\eta_{mk} = \eta_k = 1/K$, while these are selected as $\eta_{mk} = \sqrt{\beta_{mk}} / \sum_{i=1}^{K} \sqrt{\beta_{mi}}$ for the distributed GMR scheme. Thanks to the UL-DL duality, the channel hardening bound gives nearly the same DL performance as the UatF bound in UL. This proves that RC-MMSE is unsuitable for a scenario in which only the statistics of channels are available.

Fig. 2 shows the asymptotic behavior of the average UL SE per UE for $K = 4$ with increasing M. The UEs are located in the center of the area being 5 m apart from each other with $D = 1000$ m. The optimal MMSE shows the best performance and confirms that CF mMIMO networks can indeed achieve unbounded capacity in the asymptotic regime (as stated in Proposition 1). Likewise, both RC-MMSE and GMR centralized combiners SE grows unboundedly with M. In agreement to Proposition 2, the distributed approach of GMR shows the same asymptotic behavior, though the UatF bound considerably underestimates its performance. Lastly, the poor SE showed by the RC-MMSE proves that it is not an adequate scheme under statistical CSI.

VI. Conclusions

This letter focused on pilot decontamination in CF mMIMO. Building on linear independence of correlation matrices, the RC-MMSE combining was proposed as an alternative to centralized MMSE that almost matches the optimal performance and is asymptotically optimal. This scheme turns out to be a more than fitting approach for centralized computationally-limited networks. The GMR alternative scheme was also derived showing the same asymptotic behavior but maximizing the UatF. This approach is the upper-bound of MR-based CF networks. Both schemes are valid for a finite number of APs and have linear increasing complexity (in contrast to the optimal MMSE). By the UL-DL duality, the same performance can be achieved in DL where the GMR alternative serves as an optimal approach for MR-based precoders.

Appendix

We define $\widehat{\mathbf{H}} = [\widehat{\mathbf{h}}_1, \ldots, \widehat{\mathbf{h}}_K] \in \mathbb{C}^{M \times K}$ and use the matrix inversion lemma to get (after multiplying and dividing by M)

$$\check{\mathbf{v}}_k \triangleq \frac{1}{M} \mathbf{Z}^{-1} \widehat{\mathbf{H}} \left(\frac{1}{M} \widehat{\mathbf{H}}^H \mathbf{Z}^{-1} \widehat{\mathbf{H}} + \frac{1}{M} \mathbf{I}_K \right)^{-1} \mathbf{e}_k \tag{20}$$

from which $\check{v}_{mk}$ is obtained as (by neglecting the factor $1/M$, since it does not affect the SINR in (4))

$$\check{v}_{mk} \triangleq \frac{1}{z_m} \widehat{\mathbf{h}}_m^T \left(\frac{1}{M} \widehat{\mathbf{H}}^H \mathbf{Z}^{-1} \widehat{\mathbf{H}} + \frac{1}{M} \mathbf{I}_K \right)^{-1} \mathbf{e}_k \tag{21}$$

with $\widehat{\mathbf{h}}_m = [\widehat{h}_{m1}, \ldots, \widehat{h}_{mK}]^T \in \mathbb{C}^K$. By exploiting the fact that, as $M \to \infty$,

$$\frac{1}{M} \left[\widehat{\mathbf{H}}^H \mathbf{Z}^{-1} \widehat{\mathbf{H}} \right]_{\ell j} = \frac{1}{M} \widehat{\mathbf{h}}_\ell^H \mathbf{Z}^{-1} \widehat{\mathbf{h}}_j \asymp \frac{1}{M} \mathrm{tr} \left(\mathbf{R}_j \boldsymbol{\Psi}^{-1} \mathbf{R}_\ell \mathbf{Z}^{-1} \right)$$

the combiner $\bar{v}_{mk}$ defined in (8) follows.

References

[1] T. L. Marzetta, "Noncooperative Cellular Wireless with Unlimited Numbers of Base Station Antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, November 2010.

[2] E. Björnson, J. Hoydis, and L. Sanguinetti, "Massive MIMO Has Unlimited Capacity," *IEEE Trans. Wireless Commun.*, vol. 17, no. 1, pp. 574–590, 2018.

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

[4] E. Nayebi, A. Ashikhmin, T. L. Marzetta, H. Yang, and B. D. Rao, "Precoding and Power Optimization in Cell-Free Massive MIMO Systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, 2017.

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

[6] L. Sanguinetti, E. Björnson, and J. Hoydis, "Fundamental Asymptotic Behavior of (Two-User) Distributed Massive MIMO," in *Proc. 2018 IEEE Global Commun. Conf. (GLOBECOM)*, 2018, pp. 1–6.

[7] Ö. T. Demir, E. Björnson, and L. Sanguinetti, "Foundations of User-Centric Cell-Free Massive MIMO," *Foundations and Trends in Signal Processing*, vol. 14, no. 3-4, pp. 162–472, 2021. [Online]. Available: http://dx.doi.org/10.1561/2000000109

[8] S. Buzzi and C. D'Andrea, "Cell-Free Massive MIMO: User-Centric Approach," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, 2017.

[9] L. Sanguinetti, E. Björnson, and J. Hoydis, "Toward Massive MIMO 2.0: Understanding Spatial Correlation, Interference Suppression, and Pilot Contamination," *IEEE Trans. Commun.*, vol. 68, no. 1, 2020.

[10] D. Neumann, T. Wiese, M. Joham, and W. Utschick, "A Bilinear Equalizer for Massive MIMO Systems," *IEEE Trans. Signal Process.*, vol. 66, no. 14, pp. 3740–3751, 2018.