Scalable Cell-Free Massive MIMO Systems: Impact of Hardware Impairments

Anastasios Papazafeiropoulos, Emil Björnson, Pandelis Kourtessis, Symeon Chatzinotas, John M. Senior

Abstract

Scalable cell-free CF (SCF) massive multiple-input-multiple-output (mMIMO) systems is a promising technology to cover the demands for higher data rates and increasing number of users in fifth generation (5G) networks and beyond. According to this concept, a large number of distributed access points (APs) communicates with the users in the network by means of joint coherent transmission while facing the main challenges against standard CF mMIMO systems being their high fronthaul load and computational complexity. Given that the cost-efficient deployment of such large networks requires low-cost transceivers being prone to unavoidable hardware imperfections, in this work, we focus on their impact on the advantageous SCF mMIMO systems by means of a general model accounting for both additive and multiplicative hardware impairments (HWIs). Notably, the scalability, depending on the time-variant characteristics of the network, is clearly affected by means of HWIs being time-varying. There is no other work in the literature studying the phase noise (PN) in CF mMIMO systems or in general any HWIs in SCF mMIMO systems. Hence, we derive upper and lower bounds on the uplink capacity accounting for HWIs. Especially, the lower bound is derived in closed-form by means of the theory of deterministic equivalents (DEs) and after obtaining the optimal hardware-aware (HA) partial minimum mean-squared error (PMMSE) combiner. Among the interesting findings, we observe that separate local oscillators (SLOs) outperform a common LO (CLO) architecture, and the additive transmit distortion degrades more the performance than the receive distortion.

Index Terms

Cell-free massive MIMO systems, user-centric 5G networks, transceiver hardware impairments, MMSE processing, capacity bounds.

A. Papazafeiropoulos is with the Communications and Intelligent Systems Research Group,University of Hertfordshire, Hatfield AL10 9AB, U. K. and with SnT at the University of Luxembourg, Luxembourg. E. Björnson is is with the Department of Electrical Engineering (ISY), Linköping University, 58183 Linköping, Sweden. P. Kourtessis and John M.Senior are with the Communications and Intelligent Systems Research Group, University of Hertfordshire, Hatfield AL10 9AB, U. K. S. Chatzinotas is with the SnT at the University of Luxembourg, Luxembourg. E-mails: tapapazaf@gmail.com, emil.bjornson@liu.se, {p.kourtessis,j.m.senior}@herts.ac.uk, symeon.chatzinotas@uni.lu.

I. INTRODUCTION

The radical new concept of "massive multiple-input-multiple-output (mMIMO)", initially proposed by Marzetta [1], has become a mainstream and mature technology [2], [3] coping with the increasing demand for high data rates in the fifth generation (5G) wireless networks. In particular, mMIMO systems improve by at least $10\times$ the spectral efficiency (SE) of previous generation networks with simple signal processing by just equipping the existing cell-sites with more antennas [1]. Notably, the mMIMO paradigm appears under two design approaches with regards to its basic characteristic, being the antennas deployment: a co-located compact large antenna array or a geographically distributed array with a large number of access points (APs) exploiting coherent processing to serve a smaller number of users.

Although the latter design, combining the benefits of network MIMO/coordinated multipoint (CoMP) (see [4], [5]) and mMIMO systems, presents a significant theoretical potential, it also appears practical limitations due to huge computational complexity and fronthaul signaling. Recently, the cell-free (CF) mMIMO infrastructure has been proposed in [6] as a promising solution by disappearing the cell boundary concept from the network MIMO architecture while enriching the operating regime with many more single or multiple antenna access points (APs) than user equipments (UEs). Specifically, a large number of APs, connected to a central processing unit (CPU), serves jointly the UEs while reaping the benefits of network MIMO and mMIMO systems. Hence, CF mMIMO systems outperform small-cell systems and co-located cellular architectures by providing uniformly good quality of service and better coverage with higher SE due to lower path-losses, increased macrodiversity, channel hardening, favorable propagation, and suppressed interference. Recently, [7] suggested a taxonomy with four different implementations of CF mMIMO systems and that a centralized implementation with minimum mean-squared error (MMSE) combining outperforms maximum-ratio combining (MRC) with reduced fronthaul signaling.

Unfortunately, the majority of papers, concerning CF mMIMO systems, assume that all APs are connected to one CPU, being responsible to undertake the coordination and processing of the signals of all UEs [6], [7]. Consequently, the computational burden and fronthaul capacity grow further with the number of UEs and indicate that the primitive model of CF mMIMO systems is not scalable. Interestingly, [8], [9] suggested a user-centric approach, where each UE is not served by all APs but a subset providing the best channel conditions. Thus, every UE connects to different subsets, or equivalently, each AP cooperates with different APs to serve different

UEs. However, in practice, the system characteristics such as the channels and node locations vary with time. Taking this into account, in [10], dynamic cooperation clustering (DCC) [11] has been applied to achieve scalable CF (SCF) mMIMO systems.

The prior literature of mMIMO systems assumed perfect hardware, however, practical applications necessitate studying the impact of the inevitable hardware impairments (HWIs). In particular, a cost-efficient implementation of a large number of antenna elements in co-located or distributed layouts, which would make the mMIMO technology practically attractive, should take into consideration of inexpensive hardware, being more energy-consuming and prone to HWIs such as In-phase/Quadrature-phase (I/Q)-imbalance [12], oscillator phase noise (PN) [13]–[15], and high power amplifier non-linearities [16]. Even if calibration schemes and compensation algorithms are normally utilized at the transmitter and receiver, respectively, a certain amount of distortions, known as residual HWIs, remains and can be categorized into additive and multiplicative distortions. The model of additive HWIs includes additive Gaussian noises at both the transmitter and receiver, expressing the aggregate effect of many impairments and has been grounded based on its analytical tractability and experimental validation [17], [18]. The second category includes the PN, which is expressed in terms of the phase drifts emerging from the local oscillators (LOs) and appearing as multiplicative factors to the channels [14]. Notably, the PN accumulates within the channel coherence time. Despite the importance of HWIs only a few works have addressed their impact in the case of CF mMIMO systems [19]–[22] while none work has accounted for the effect of PN. Notably, [23], concerning massive MIMO with distributed arrays, which is essentially the same as CF massive MIMO systems, has studied additive HWIs and PN but perfect hardware was assumed at the user devices. In fact, there is no work, investigating any impact of HWIs at SCF mMIMO systems.

A. Motivation

This paper is incited by the following observations: 1) HWIs are unavoidable in realistic wireless communications and their negligence results in inconsistency between theory and practice, 2) mMIMO is a more attractive technology if their implementation is based on cheap hardware being more susceptible on HWIs, 3) CF mMIMO systems is a promising architecture but its feasibility depends on its scalability in large networks with increasing number of UEs, 4) recently, it was shown that CF mMIMO systems outperform small cells and cellular systems if a centralized version of MMSE decoding is applied which also reduces the fronthaul signaling compared to

distributed systems, 5) previous works in CF mMIMO systems, assuming MMSE decoding, did not obtain any closed-form expressions but relied only on Monte-Carlo simulations, 6) channel state information (CSI) at the transmitter, being of paramount importance in CF mMIMO systems, is drastically affected by the presence of HWIs, and 7) there is no prior work investigating the impact of HWIs in emerging SCF mMIMO systems. Actually, the PN has not been studied at all even in the case of standard CF mMIMO systems, where a few existing works focused only on a single type of HWIs [19]–[21]. Notably, since HWIs are time-varying, they affect further SCF systems by means of the DCC being time-dependent.

These observations suggest that the investigation of HWIs in SCF mMIMO systems with MMSE combining in terms of closed-form expressions by accounting for a general model involving both additive and multiplicative HWIs is of great practical interest.

B. Contributions

The main contributions are summarized as follows.

- Contrary to existing works [19]–[22] which have studied only certain kinds of HWIs in CF mMIMO systems, we introduce a general model for HWIs including both additive and multiplicative (PN) distortions as well as we consider the more novel architecture of SCF mMIMO systems, being scalable with the number of UEs as realistic scenarios demand. Note that there is no work addressing PN at CF mMIMO systems or in general any HWIs at SCF mMIMO systems.
- Previous works in CF mMIMO systems that applied MMSE combining presented only simulations while, in this work, we employ MMSE combining with imperfect CSI and obtain closed-form expressions by means of the deterministic equivalent (DE) analysis. Notably, there are no other works providing DE expressions for SCF mMIMO systems.
- We derive an upper bound on the uplink capacity and the corresponding optimal MMSE combiner. Next, we obtain the hardware- aware (HA) MMSE combiner by minimizing the sum mean square error (MSE), and derive the corresponding DE lower bound on the uplink capacity. For the sake of comparison, we also consider the hardware-unaware (HU) MMSE decoders, the MMSE decoders in the absence of HWIs, and MRC decoding.
- We shed light on the impact of HWIs on the uplink achievable SE of SCF mMIMO systems. Hence, we show how the PN degrades the system performance. Specifically, it is shown that separate local oscillators (SLOs) and common LO (CLO) designs behave differently,

i.e., the achievable SE with SLOs appears superior against the CLO setting. Moreover, we quantify the degradation of the system due to additive HWIs, where the transmit distortion has more severe impact than the receive one. In addition, we show the increase in the gap between SLOs and CLO deployments as the number of APs increases.

Paper outline: Section II presents the system model of a SCF mMIMO system in terms of fundamental design parameters. In addition, a description of the HWIs is provided. In Section III, the scalable channel estimation with HWIs is described. In Section IV, we elaborate on the uplink data transmission and obtain an upper and a lower bound on the uplink capacity as well as the HA-PMMSE decoder. Section V presents the DE of the lower bound on the uplink capacity with HWIs. The numerical results are placed in Section VI, while Section VII summarises the paper.

Notation: Vectors and matrices are denoted by boldface lower and upper case symbols. The symbols $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{\mathsf{H}}$ express the transpose and Hermitian transpose operators, while the expectation operator is denoted by $\mathbb{E}[\cdot]$. The notations $\mathbb{C}^{M\times 1}$ and $\mathbb{C}^{M\times N}$ refer to complex M-dimensional vectors and $M \times N$ matrices, respectively. The notation $\xrightarrow[M\to\infty]{\text{a.s.}}$ denotes almost sure convergence as $M \to \infty$, and the notation $a_n \asymp b_n$ expresses the equivalence relation $a_n - b_n \xrightarrow[M\to\infty]{\text{a.s.}} 0$ with a_n and b_n being two infinite sequences. Finally, $\mathbf{b} \sim \mathcal{CN}(\mathbf{0}, \mathbf{\Sigma})$ represents a circularly symmetric complex Gaussian variable with covariance matrix $\mathbf{\Sigma}$

II. System Model

We consider a CF network architecture including M APs and K UEs randomly distributed over a geographic area with M and K being large. The APs, each equipped with N antennas, are connected to central processing units by means of backhaul links in an arbitrary way [24], [25]. The UEs are equipped with a single-antenna. All communications take place in the same time-frequency resources. In particular, coherent joint transmission and reception are enabled among the APs.

A. Channel Model

We consider a time-varying narrowband channel that is divided into coherence blocks and we employ the standard TDD protocol where the length of each coherence interval/block is τ_{c} channel uses being equal to the product of the coherence bandwidth B_{c} in Hz and the coherence time T_{c} in s [3]. Specifically, this block accounts for τ channel uses for an uplink training phase, involving with the channel estimation, as well as $\tau_{\mathrm{u}} = \zeta(\tau_{\mathrm{c}} - \tau)$ and τ_{d} channel uses for the

uplink and downlink data transmission phases with $\zeta \leq 1$ accounting for the UL payload fraction transmission.

During the transmission in each coherence block, the channel vector between the mth AP and the kth UE, denoted by $\mathbf{h}_{mk} \in \mathbb{C}^N$, is fixed and exhibits flat-fading. Especially, the random channel vector is expressed by means of an independent correlated Rayleigh fading distribution as

$$\mathbf{h}_{mk} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_{mk}\right), \qquad m = 1, \ldots, M \text{ and } k = 1, \ldots, K \tag{1}$$

where the complex Gaussian distribution models the small-scale fading and $\mathbf{R}_{mk} \in \mathbb{C}^{N \times N}$ is a deterministic Hermitian-symmetric positive semi-definite correlation matrix being able to describe certain realistic effects such as shadowing and geometric pathloss with large-scale fading coefficient $\beta_{mk} = \mathrm{tr}\left(\mathbf{R}_{mk}\right)/N$. Notably, we assume that the set $\{\mathbf{R}_{mk}\}$ is assumed to be known by the network elements whenever needed by means of practical methods (see e.g., [26], [27]).

B. Basic Scalability Guidelines

The network-wide implementation of the original CF mMIMO systems is impractical in terms of cost and complexity in the case of large deployments with large K. To make this technology feasible and attractive, it has to be scalable, in particular, regarding the number of UEs [10], Hence, therein, the authors examined the scalability issue and provided a corresponding definition by letting $K \to \infty$. Speficically, they defined a CF mMIMO network as scalable when: *i)* the signal processing for channel estimation, *ii)* the signal processing for data reception and transmission, *iii)* the fronthaul signaling for data and CSI sharing, and *iv)* the power control optimization have finite complexity and resource requirements per AP when $K \to \infty$.

In particular, each AP serves a set of UEs with constant cardinality as the total number of UEs $K \to \infty$, in order to satisfy the scalability conditions above. Especially, the set of UEs served by AP m is defined as [10]

$$\mathcal{D}_m = \{i : \ \mathrm{tr}\left(\mathbf{D}_{mi}\right) \geq 1, \ i \in \{1, \ldots, K\}\}, \tag{2}$$

where $\mathbf{D}_{mi} \in \mathbb{C}^{N \times N}$ denotes a diagonal matrix defining which APs and UEs communicate with each other according to the DCC framework [11]. These matrices enable a unified analysis with original CF mMIMO system being one of the many architectures that could be described by this framework.

C. Hardware Impairments

Although the majority of papers in the CF mMIMO literature assumes ideal transceiver hardware, physical transceiver implementations such as the converters and the oscillators undergo to HWIs and inevitably distort the signal. Despite the application of compensation and mitigation algorithms, HWIs remain; for example, due to the limited modeling accuracy of the various system parameters. The choice of specific hardware quality is crucial for the overall cost and power consumption. Especially, attractive solutions for large deployments such as the CF mMIMO technology under study should take into account for cheap hardware that results in low cost and power consumption. Otherwise, the deployment of this technology would be prohibitive in terms of cost and energy efficiency. Hence, it is of paramount importance to incorporate in the analysis of CF mMIMO systems the impact of the hardware constraints.

The residual hardware impairments can be classified into three separate categories being $1)$ a multiplicative distortion due to phase shift of the signal, $2)$ additive distortions at both the transmitter and the receiver with powers proportional to the transmit and the received signal, respectively, and $3)$ a channel-independent amplified thermal noise.

1) Multiplicative distortion: The multiplicative distortion of the channel expresses time-dependendent random phase-drifts known as PN, and it is induced during the up-conversion of the baseband signal to passband and vice-versa by multiplying the signal with the LO's output. Based on standard references concerning the PN [14], [28], this distortion at the nth channel use can be expressed by a discrete-time independent Wiener process. Especially, the PNs at the LOs of the jth antenna of the mth AP and kth UE are modeled as

$$\phi_{m,n}^{j} = \phi_{m,n-1}^{j} + \delta_{n}^{\phi} \tag{3}$$

$$\varphi_{k,n} = \varphi_{k,n-1} + \delta_{n}^{\varphi}, \tag{4}$$

where $\delta_{n}^{\phi} \sim \mathcal{N}(0, \sigma_{\phi}^{2})$ and $\delta_{n}^{\varphi} \sim \mathcal{N}(0, \sigma_{\varphi}^{2})$ [14], [28]–[30]. The increment variance of the PN process can be expressed by [29]

$$\sigma_{i}^{2} = 4\pi^{2} f_{\mathrm{c}}^{2} c_{i} T_{\mathrm{s}}, \qquad i = \phi, \varphi \tag{5}$$

where T_{s}, c_i, and f_{c} are the symbol interval, a constant dependent on the oscillator, and the carrier frequency, respectively. Note that the total PN process from the kth user and the jth antenna of the mth AP is $\theta_{mk,n}^{(j)} = \phi_{m,n}^{j} + \varphi_{k,n}$ while $\mathbf{\Theta}_{mk,n} \triangleq \mathrm{diag}\left\{ e^{i\theta_{mk,n}^{(1)}}, \ldots, e^{i\theta_{mk,n}^{(N)}} \right\}$ denotes the corresponding total PN matrix at the nth channel use. In fact, this generic matrix accounts for

the general scenario (non-synchronous operation), where we have SLOs at each antenna justifying the independence among the PN processes [14], [30]. In the special case of the synchronous operation, all PN processes $\theta_{mk,n}^{(j)}$ are identical for all $j = 1, \ldots, N$. The matrix degenerates to $\Theta_{mk,n} \triangleq e^{i\theta_{mk,n}^{(1)}} \mathbf{I}_N$ when there is just one CLO connected to all antennas of an AP. In our analysis, we focus on both SLOs and CLO scenarios.

Henceforth, we will use the following lemma providing the expectation of the complex exponential of the phase drift $\Delta\phi = \phi_{n_2} - \phi_{n_1}$ of a general PN process ϕ_n with increment variance σ_ϕ^2 during the time interval $\Delta n = n_2 - n_1$ where n_1, n_2 are two different time slots. Obviously, the phase drift is a zero-mean Gaussian variable with a variance proportional to Δn distributed as $\Delta\phi \sim \mathcal{N}\left(0, \sigma_\phi^2 \Delta n\right)$, i.e., the variance increases with time.

Lemma 1: The expectation of the complex exponential of the phase drift $\Delta\phi_n$ occured during Δn is

$$\mathbb{E}\left\{e^{-j\Delta\phi}\right\} = e^{-\frac{\sigma_\phi^2}{2}\Delta n}. \tag{6}$$

Proof: The proof is straightforward since this expectation is basically the characteristic function of the Gaussian variable $\Delta\phi$ having zero mean and variance σ_ϕ^2. ■

D. Additive distortion

In real systems, unavoidable HWIs emerge due to the imperfect compensation of the quantization noise in the analog-to-digital converters (ADCs) at the receiver, the I/Q imbalance, etc. [17], [18]. As a result, both the transmit and receive signals are distorted during the transmission and reception processing, respectively. For example, at the transmitter side, a mismatch appears between the generated signal and the signal intended to be transmitted.

Measurement results have shown that the conditional additive distortion noise for the ith link, given the channel realizations, is modeled as Gaussian distributed with average power proportional to the average signal power [18]. The Gaussianity is based on the consideration that this distortion results by means of the aggregate contribution of many impairments. Hence, the impairments at different antenna branches are modeled as mutually uncorrelated Gaussian random variables. Moreover, this distortion depends on time since it takes a new realization for each data signal being time-dependent itself. Especially, in the uplink case, the additive transceiver distortions are

expressed in terms of conditional distributions as

$$\delta_{\mathrm{t},n}^{k} \sim \mathcal{CN}\left(0, \Lambda_{n}^{k}\right) \tag{7}$$

$$\boldsymbol{\delta}_{\mathrm{r},n}^{m} \sim \mathcal{CN}\left(\mathbf{0}, \boldsymbol{\Upsilon}_{n}^{m}\right), \tag{8}$$

where $\Lambda_{n}^{k} = \kappa_{\mathrm{t}}^{2} p_k$ and $\boldsymbol{\Upsilon}_{n}^{m} = \kappa_{\mathrm{r}_m}^{2} \sum_{i=1}^{K} p_i \mathrm{diag}\left(|h_{mi}^{(1)}|^2, \ldots, |h_{mi}^{(N)}|^2\right)$ with p_k being the transmit power from UE k while the proportionality parameters κ_{t}^{2} and $\kappa_{\mathrm{r}_m}^{2}$ describe the severity of the residual impairments at the transmitter and the receiver side. In applications, these parameters are met as the error vector magnitude (EVM) [31].

E. Amplified thermal noise

Certain components such as the low noise amplifiers and mixers induce an amplification of the thermal noise denoted as $\boldsymbol{\xi}_{m,n}$ that manifests as an increase of the variance. In particular, this amplified receiver noise is modeled as Gaussian distributed with zero-mean and variance $\xi_{m,n}\mathbf{I}_N$ with $\xi_{m,n} \geq \sigma^2$ where σ^2 is the corresponding parameter of the actual thermal noise [23]. In addition, it is worthwhile to mention that it is independent of the UE channels. Also, note that the amplified thermal noise is time-dependent taking different random realizations over time because it consists of samples obtained from a white noise process that has passed by some amplified "filter" [32].

III. Scalable Channel Estimation with Hardware Impairments

In practice, the APs do not have perfect CSI but estimate the channel. Especially, in the area of mMIMO, and subsequently, CF mMIMO, the APs estimate their channels by employing in their design TDD operation which includes an uplink training phase with pilot symbols [6].

We assume that the training phase takes place at time 0. Moreover, we denote $\mathcal{S} \subset \{1, \ldots, K\}$ a subset of the UEs that transmit one of the τ-length mutually orthogonal training sequences, i.e., $\boldsymbol{\omega}_k \triangleq [\omega_{k,1}, \ldots, \omega_{k,\tau}]^{\mathsf{T}} \in \mathbb{C}^{\tau \times 1}$ with $\rho_{\mathrm{p}} = \mathbb{E}\left[|\omega_{k,n}|^2\right], \forall k, n$, with p_{p} being the common average transmit power per UE while the sequences among different UEs are mutually orthogonal, in order that each UE has a unique pilot sequence not shared with any other UE. Moreover, according to [10, Assumption 1], each AP servers at most one UE per per pilot sequence which implies that $|\mathcal{D}_m| \leq \tau$ and

$$\mathbf{D}_{mk} = \begin{cases} \mathbf{I}_N & \text{if } k \in \mathcal{D}_m \\ \mathbf{0} & \text{if } k \notin \mathcal{D}_m. \end{cases} \tag{9}$$

Notably, the independence of τ from K agrees with the requirement for scalability.

Thus, during the training phase, the $\mathbb{C}^{N\times 1}$ received signal by the mth AP at time n, coming from the subset $\mathcal{S}$ of UEs and including the hardware impairments, is given by

$$\mathbf{y}^{\mathrm{p}}_{m,n} = \sum_{i\in\mathcal{S}} \mathbf{\Theta}_{mi,n}\mathbf{h}_{mi}\left(\omega_{i,n} + \delta^{i}_{\mathrm{t},n}\right) + \boldsymbol{\delta}^{m}_{\mathrm{r},n} + \boldsymbol{\xi}_{m,n}, \tag{10}$$

where we have incorporated the HWIs, i.e., $\mathbf{\Theta}_{mi,n} = \mathrm{diag}\left\{e^{j\theta^{(1)}_{i,n}}, \ldots, e^{j\theta^{(N)}_{i,n}}\right\}$ is the phase noise because of the LOs of the mth AP and UE k at time n, $\delta^{i}_{\mathrm{t},n} \sim \mathcal{CN}\left(0, \kappa^{2}_{\mathrm{t}}\rho_{\mathrm{p}}\right)$ is the additive transmit distortion, $\boldsymbol{\delta}^{m}_{\mathrm{r},n} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{\Upsilon}^{m}_{n}\right)$ is the additive receive distortion with $\mathbf{\Upsilon}^{m}_{n} = \kappa^{2}_{\mathrm{r}_m}\sum_{i=1}^{\mathcal{S}}\rho_i \mathrm{diag}\left(|h^{(1)}_{mi}|^2, \ldots, |h^{(N)}_{mi}|^2\right)$, and $\boldsymbol{\xi}_{m,n}$ is the spatially amplified Gaussian thermal noise matrix at the mth AP.

Remark 1: The received signal depends on both multiplicative and additive HWIs[1]. In particular, only the distortions of the UEs in $\mathcal{S}$ degrade the reception during the training phase and will affect the estimated channel obtained below.

Proposition 1: The linear minimum mean-square error (LMMSE) estimate of the channel of UE k during the training phase is given by

$$\begin{aligned}\hat{\mathbf{h}}_{mk,n} &= \mathbb{E}\left[\mathbf{h}_{mk,n}\boldsymbol{\psi}^{\mathsf{H}}\right]\left(\mathbb{E}\left[\boldsymbol{\psi}\boldsymbol{\psi}^{\mathsf{H}}\right]\right)^{-1}\boldsymbol{\psi} \\ &= \left(\boldsymbol{\omega}^{\mathsf{H}}_{k}\mathbf{\Delta}^{\mathrm{tr}}_{k,n} \otimes \mathbf{R}_k\right)\mathbf{\Sigma}^{-1}\boldsymbol{\psi},\end{aligned} \tag{12}$$

where

$$\mathbf{\Delta}^{\mathrm{tr}}_{k,n} \triangleq \mathrm{diag}\left\{e^{-\frac{\sigma^2_{\varphi}+\sigma^2_{\phi}}{2}n}, \ldots, e^{-\frac{\sigma^2_{\varphi}+\sigma^2_{\phi}}{2}|n-\tau|}\right\} \tag{13}$$

$$\mathbf{\Sigma} \triangleq \sum_{j\in\mathcal{S}} \mathbf{X}_j \otimes \mathbf{R}_j + \xi_{m,n}\mathbf{I}_{\tau N}, \tag{14}$$

$$\mathbf{X}_j \triangleq \tilde{\mathbf{X}}_j + \kappa^{2}_{\mathrm{r}_m}\mathbf{D}_{|\boldsymbol{\omega}_j|^2}, \tag{15}$$

$$\mathbf{D}_{|\boldsymbol{\omega}_j|^2} \triangleq \mathrm{diag}\left(|\omega_{j,1}|^2, \ldots, |\omega_{j,\tau}|^2\right), \tag{16}$$

$$\left[\tilde{\mathbf{X}}_j\right]_{u,v} \triangleq \left(\omega_{j,u}\omega^{*}_{j,v} + \kappa^{2}_{\mathrm{t_{UE}}}\right)p e^{-\frac{\sigma^2_{\varphi}+\sigma^2_{\phi}}{2}|u-v|}. \tag{17}$$

[1]For the sake of a better presentation of the transceiver HWIs, herein, we provide the ideal uplink transmission of a SCF mMIMO system without any imperfections during training. Thus, in such case, the received signal at the mth UE at a given channel use n is expressed as

$$\mathbf{y}^{\mathrm{p}}_{m,n} = \sum_{i\in\mathcal{S}} \mathbf{h}_{mi}\omega_{i,n} + \mathbf{w}_{m,n}, \tag{11}$$

where $\mathbf{w}_{m,n} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2\mathbf{I}_N\right)$.

Proof: The proof follows the same steps with Theorem 1 in [23] but it is more general since the proposed model includes the transmit HWIs $\delta_{\mathrm{t},n}^{k}$ and requires some extra algebraic manipulations. Also, the summation in (14) is over the sum of participating UEs and not all UEs. The proof is omitted due to limited space. ■

Based on the property of orthogonality of LMMSE estimation, the current channel at the end of the training phase is given by

$$\tilde{\mathbf{h}}_{mk,\tau} = \hat{\mathbf{h}}_{mk,\tau} + \tilde{\mathbf{e}}_{mk,\tau}, \tag{18}$$

where $\hat{\mathbf{h}}_{mk,\tau}$ and $\tilde{\mathbf{e}}_{mk,\tau}$ have zero mean and variances $\boldsymbol{\Phi}_{mk} = \left(\boldsymbol{\omega}_{k}^{\mathsf{H}} \boldsymbol{\Delta}_{k,n}^{\mathrm{tr}} \otimes \mathbf{R}_{k}\right) \boldsymbol{\Sigma}^{-1}$ and $\tilde{\mathbf{R}}_{mk} = \mathbf{R}_{mk} - \boldsymbol{\Phi}_{mk}$.

Remark 2: It is worthwhile to mention that $\hat{\mathbf{h}}_{mk,\tau}$ and $\mathbf{e}_{mk,\tau}$ are neither independent nor jointly complex Gaussian vectors contrary to conventional estimation theory concerning independent Gaussian noise but they are uncorrelated and each of them has zero mean. The reasoning relies on the fact that the effective distortion noises, e.g., $\boldsymbol{\Theta}_{mi,n}\mathbf{h}_{mi}\delta_{\mathrm{t},n}^{i}$ are not Gaussian since they appear as products between two Gaussian variables. In other words, we have not derived the optimal MMSE estimator but the suboptimal LMMSE which, in general, is accompanied by a little performance loss [3].

Remark 3: In the special case of ideal hardware, the LMMSE estimator becomes the optimal MMSE estimator.

Clearly, the channel estimate in (12) depends on time n, which means that any decoders $\mathbf{v}_{k,n}$ should be computed in every symbol interval of the data transmission phase but this is computationally prohibitive. For this reason, we obtain the decoders for one symbol interval $n = \tau + 1$ by means of the channel estimate at time τ, and then, we apply them during the whole data transmission phase. Specifically, we denote $\mathbf{v}_{k,n} = f\left(k, \tau\right)$ for $k = \{1, \ldots, K\}$, where $f\left(\cdot\right)$ is a function of the channel estimates, dependent on τ, since they are obtained during the uplink transmission phase.

IV. Uplink Data Transmission with HWIs

The focal point of this section is the derivation of the achievable SE of a practical SCF mMIMO setup with HWIs. Specifically, we consider a fully centralized processing architecture, where the APs, acting basically as relays, forward all their signals (received pilot and data signals) to the CPU. In particular, the CPU handles both the channel estimation and data detection. In the

uplink detection phase, let the mth AP receive its signal but delegate its detection to the CPU. For this reason, the M APs have to send their received signals $\{\mathbf{y}_{m,n} : m = 1, \ldots, M\}$ to the CPU. In such case and by setting $W = MN^2$, the received signal by the CPU at time n can be written in a compact form as

$$\mathbf{y}_n = \sum_{i=1}^{K} \mathbf{\Theta}_{i,n}\mathbf{h}_i \left(s_{i,n} + \delta_{\mathrm{t},n}^i\right) + \boldsymbol{\delta}_{\mathrm{r},n} + \boldsymbol{\xi}_n, \tag{19}$$

where $s_{i,n} \in \mathbb{C}$ is the transmit signal from UE i with power ρ_i, $\mathbf{y}_n = \left[\mathbf{y}_{1,n}^{\mathsf{T}} \cdots \mathbf{y}_{M,n}^{\mathsf{T}}\right]^{\mathsf{T}} \in \mathbb{C}^W$ and $\boldsymbol{\xi}_n = \left[\boldsymbol{\xi}_{1,n}^{\mathsf{T}} \cdots \boldsymbol{\xi}_{M,n}^{\mathsf{T}}\right]^{\mathsf{T}} \in \mathbb{C}^W$ are block vectors while $\mathbf{h}_{i,n} = \left[\mathbf{h}_{i1,n}^{\mathsf{T}} \cdots \mathbf{h}_{iM,n}^{\mathsf{T}}\right]^{\mathsf{T}} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_i\right)$ is the concatenated channel vector from all APs with $\mathbf{R}_i = \mathrm{diag}\left(\mathbf{R}_{i1}, \ldots, \mathbf{R}_{iM}\right) \in \mathbb{C}^{W \times W}$ being the block diagonal spatial correlation matrix by assuming that the channel vectors of different APs are independently distributed. In a similar way, we have the estimated channel and estimated error covariance matrices given by $\mathbf{\Phi}_i = \mathrm{diag}\left(\mathbf{\Phi}_{i1}, \ldots, \mathbf{\Phi}_{iM}\right) \in \mathbb{C}^{W \times W}$ and $\mathbf{\Delta}_i = \mathbf{R}_i - \mathbf{\Phi}_i \in \mathbb{C}^{W \times W}$, respectively. Moreover, regarding the HWIs, we have the PN block diagonal matrix $\mathbf{\Theta}_i = \mathrm{diag}\left(\mathbf{\Theta}_{i1}, \ldots, \mathbf{\Theta}_{iM}\right) \in \mathbb{C}^{W \times W}$, the transmit additive distortion $\delta_{\mathrm{t},n}^i$ from the ith UE, and the receive additive distortion block vector $\boldsymbol{\delta}_{\mathrm{r},n} = \left[\boldsymbol{\delta}_{\mathrm{r},n}^{1\,\mathsf{T}} \cdots \boldsymbol{\delta}_{\mathrm{r},n}^{M\,\mathsf{T}}\right]^{\mathsf{T}} \in \mathbb{C}^W$. Note that $\boldsymbol{\xi}_n \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{F}_{\boldsymbol{\xi}_n}\right)$ with $\mathbf{F}_{\boldsymbol{\xi}_n} = \mathrm{diag}\left(\xi_{1,n}\mathbf{I}_N, \ldots, \xi_{M,n}\mathbf{I}_N\right)$[3].

According to the DCC framework [11], although all APs receive the signals from all UEs, only a subset of the APs contributes to signal detection. Thus, the CPU estimates $s_{k,n}$ by means of (19) as

$$\begin{aligned}\hat{s}_{k,n} &= \sum_{m=1}^{M} \mathbf{v}_{mk,n}^{\mathsf{H}} \mathbf{D}_{mk} \mathbf{y}_m \\ &= \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{h}_{k,n} s_k + \sum_{i=1, i\neq k}^{K} \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{h}_{i,n} s_i + \sum_{i=1}^{K} \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{h}_{i,n} \delta_{\mathrm{t},n}^i + \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k (\boldsymbol{\delta}_{\mathrm{r},n} + \boldsymbol{\xi}_n)\end{aligned} \tag{20}$$

with $\mathbf{D}_k = \mathrm{diag}\left(\mathbf{D}_{1k}, \ldots, \mathbf{D}_{Mk}\right) \in \mathbb{C}^{W \times W}$ being a block-diagonal matrix.

[2]Given that a CF massive MIMO architecture assumes that the number of APs is very large, we assume that W, defined as the product of the number of APs with the number of antennas per AP and expressing the total number of antennas at the system, is very large too. Also, the ratio of the number of UEs to the total number of antennas W, denoted by $\beta = K/W$, is assumed constant. Hence, the proposed design is eligible for application of the DE tools used below.

[3]Henceforth, we assume that all APs are constructed with the same hardware, and thus, are degraded from the same impairments, e.g., $\xi_{i,n}^i = \xi_n \forall i$, in order to facilitate the analysis. However, the results can be easily extended to the general case where different APs have different hardware parameters. A similar assumption holds for the hardware parameters across the UEs.

Aiming at the study of performance of the realistic aspect of the SCF mMIMO to reveal the fundamental impact of non-ideal hardware, we focus on upper and lower bounds since the ergodic capacity is unknown for this system model. Given that the channels change with time n, the general procedure for both bounds consists of obtaining one SE for each channel use n in the transmission phase, and then, taking the average of the SE over time as in [14].

A. Upper bound

Following a procedure similar to [33, Lem. 1], we assume that the uplink pilots provide each AP with perfect CSI. Then, we infer that single-stream Gaussian signaling is optimal since all the additive noise terms (the transmit and receive distortion as well as the amplified thermal noise) are circularly symmetric complex Gaussian distributed and independent of the useful signals.

Proposition 2: An upper bound on the uplink capacity in a practical CF mMIMO system is given by

$$\mathrm{SE}_k^{\mathrm{up}} = \frac{1}{\tau_{\mathrm{c}}} \sum_{n=1}^{\tau_{\mathrm{c}}-\tau_{\mathrm{p}}} \mathbb{E}\left\{\log_2\left(1+\gamma_{k,n}^{\mathrm{up}}\left(\mathbf{v}_{k,n}\right)\right)\right\}, \qquad [\mathrm{bit/s/Hz}] \tag{21}$$

where the instantaneous SINR $\gamma_{k,n}^{\mathrm{up}}\left(\mathbf{v}_{k,n}\right)$ is given by

$$\gamma_{k,n}^{\mathrm{up}}(\mathbf{v}_{k,n}) = \rho_k \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \left(\sum_{i=1}^{K} \rho_i \left(\kappa_{\mathrm{t}}^2 \mathbf{h}_{i,n} \mathbf{h}_{i,n}^{\mathsf{H}} + \boldsymbol{\kappa}_{\mathrm{r}}^2 \mathbf{F}_{|\mathbf{h}_i|^2} \right) + \xi_n \mathbf{I}_W \right)^{-1} \mathbf{D}_k \mathbf{v}_{k,n} \tag{22}$$

with the receive combining vector being

$$\mathbf{v}_{k,n}^{\mathrm{MMSE,up}} = \left(\sum_{i=1}^{K} \kappa_{\mathrm{t}}^2 \rho_i \mathbf{h}_{i,n} \mathbf{h}_{i,n}^{\mathsf{H}} + \boldsymbol{\kappa}_{\mathrm{r}}^2 \sum_{i=1}^{K} \rho_i \mathbf{F}_{|\mathbf{h}_i|^2} + \xi_n \mathbf{I}_W \right)^{-1} \mathbf{D}_k \mathbf{h}_{k,n}, \tag{23}$$

where $\boldsymbol{\kappa}_{\mathrm{r}}^2 = \mathrm{diag}\left(\kappa_{\mathrm{r}_1}^2, \ldots, \kappa_{\mathrm{r}_m}^2\right)$ while $\mathbf{F}_{|\mathbf{h}_i|^2} = \mathrm{diag}\left(\mathbf{F}_{|\mathbf{h}_{1i}|^2}, \ldots, \mathbf{F}_{|\mathbf{h}_{Mi}|^2}\right)$ where $\mathbf{F}_{|\mathbf{h}_{mi}|^2} = \mathrm{diag}\left(|h_{mi}^{(1)}|^2, \ldots, |h_{mi}^{(N)}|^2\right)$ with $\mathbf{h}_{mi} = \left[h_{mi}^{(1)}, \ldots, h_{mi}^{(N)}\right]^{\mathsf{T}}$.

Proof: The proof is provided in Appendix A. ■

Note that the pre-log sum in (21) defines the fraction of samples per coherence block used for uplink data transmission. Clearly, the SINR-maximizing combiner in (23) is not a scalable solution since its complexity increases with K.

Following the approach in [10], we assume that (23) should include only the UEs which are served by the same APs as UE k because the interference afflicting UE k is basically a result of a small subset of the other UEs. The set of these users is defined as [10]

$$\mathcal{P}_k = \left\{ i : \mathbf{D}_k \mathbf{D}_i \neq \mathbf{0} \right\}. \tag{24}$$

Hence, the scalable partial MMSE (PMMSE) version of (23) can be written as

$$\mathbf{v}_{k,n}^{\mathrm{PMMSE,up}} = \rho_k \left(\sum_{i \in \mathcal{P}_k} \kappa_{\mathrm{t}}^2 \rho_i \mathbf{h}_{i,n} \mathbf{h}_{i,n}^{\mathsf{H}} + \boldsymbol{\kappa}_{\mathrm{r}}^2 \sum_{i \in \mathcal{P}_k} \rho_i \mathbf{F}_{|\mathbf{h}_i|^2} + \xi_n \mathbf{I}_W \right)^{-1} \mathbf{D}_k \mathbf{h}_{k,n}, \tag{25}$$

which does not scale with K. It is worthwhile to mention that in the special case where all APs, serving user k, transmit to the same set of UEs, we obtain $|\mathcal{P}_k| = \tau_{\mathrm{p}}$ [10].

This decoding vector, based on the uplink SINR with perfect CSI, is not indicated because it will not result in representative extraction of conclusions since it is obtained by an upper bound. The proposition in the following subsection presents the uplink MSE which will allow the derivation of the practical optimal decoder by taking into account the inevitable imperfect CSI.

B. Optimal scalable MMSE combiner

Herein, we derive the HA-PMMSE combiner by means of the minimization of the MSE.

Lemma 2: The uplink MSE in the case of SCF mMIMO systems with HWI for UE k, conditioned on the imperfect channel estimates $\hat{\mathbf{H}}_n = \left[\hat{\mathbf{h}}_{1,n} \dots \hat{\mathbf{h}}_{K,n} \right]$, is given by

$$\mathrm{MSE}_{k,n} = \mathrm{tr}\left(\mathbf{C}_{k,n} \right), \tag{26}$$

where $\mathbf{C}_{k,n}$ is the error covariance matrix given by

$$\begin{aligned} \mathbf{C}_{k,n} = \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \left(\sum_{i=1}^{K} \rho_i \left(1 + \kappa_{\mathrm{t}}^2 \right) \left(\hat{\mathbf{h}}_{i,\tau} \hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}} + \boldsymbol{\Delta}_i \right) + \boldsymbol{\kappa}_{\mathrm{r}}^2 \sum_{i=1}^{K} \rho_i \left(\mathbf{F}_{|\hat{\mathbf{h}}_i|^2} + \mathbf{F}_{|\boldsymbol{\Delta}_i|^2} \right) + \xi_n \mathbf{I}_W \right) \mathbf{D}_k \mathbf{v}_{k,n} \\ - \rho_k \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \hat{\mathbf{h}}_{k,\tau} - \rho_k \hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}} \mathbf{D}_k \mathbf{v}_{k,n} + \rho_k. \end{aligned} \tag{27}$$

Proof: The proof is provided in Appendix B. ■

Taking advantage of Lemma 2, we formulate the optimization problem for minimization of the sum MSE as

$$\begin{aligned} \min_{p_i, \mathbf{v}_{i,n} \forall i} \quad & \sum_{i=1}^{K} \mathrm{MSE}_{k,n} \\ & p_i \geq 0, \ \forall i, \end{aligned} \tag{28}$$

where MSE_k is given by (26). The following proposition provides the optimal scalable combiner by minimizing $\mathrm{MSE}_{k,n}$[4].

[4]It is worthwhile to mention that the seminal works in CF mMIMO systems developed power control algorithms with increasing complexity with the number of UEs, being unfeasible in practice.

Proposition 3: The HA-PMMSE receive combining vector at time n, taking into account account for the transceiver HWIs in SCF massive MIMO systems, is written as

$$\mathbf{v}_{k,n}^{\mathrm{HA-PMMSE}} = \rho_k \mathbf{D}_k^{-1} \left(\sum_{i \in \mathcal{P}_k} \rho_i \left(1 + \kappa_{\mathrm{t}}^2\right) \hat{\mathbf{h}}_{i,0} \hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}} + \boldsymbol{\kappa}_{\mathrm{r}}^2 \sum_{i \in \mathcal{P}_k} \rho_i \mathbf{F}_{|\hat{\mathbf{h}}_i|^2} + \tilde{\boldsymbol{\Delta}} + \xi_n \mathbf{I}_W \right)^{-1} \hat{\mathbf{h}}_{k,\tau}, \quad (29)$$

where $\tilde{\boldsymbol{\Delta}} = \sum_{i \in \mathcal{P}_k} \tilde{\boldsymbol{\Delta}}_i$ with $\tilde{\boldsymbol{\Delta}}_i$ being the covariance matrices of the HWIs and the estimation error, which is given by

$$\tilde{\boldsymbol{\Delta}}_i = \rho_i \left(1 + \kappa_{\mathrm{t}}^2\right) \boldsymbol{\Delta}_i + \boldsymbol{\kappa}_{\mathrm{r}}^2 \rho_i \mathbf{F}_{|\boldsymbol{\Delta}_i|^2}. \quad (30)$$

Proof: The derivation of the HA-PMMSE relies on the minimization of the $\mathrm{MSE}_{k,n}$. Taking the derivative of the expression in (49) with respect to $\mathbf{v}_{k,n}$ and setting it to zero, we result in the optimal decoder given by (29) after taking into consideration the scalability as described in (25) . ■

Substitution of (29) into (28) results in the optimization problem depending only on the transmit powers p_i. After finding the transmit powers, the optimal decoder is given by means of (29). However, this is a quite hard problem. Since the focal point of this work is the scalable implementation of CF mMIMO systems, we have to apply a scalable power control algorithm. Heuristic scalable algorithms for CF mMIMO systems are found in [8], [34], [35] but their comparative evaluation is out of the scope of this work. Herein, for the sake of exposition, we resort to the basic equal power allocation which is a scalable solution, i.e., $p_i = p \ \forall i$ [10]. In such a case, we obtain

$$\mathbf{v}_{k,n}^{\mathrm{HA-PMMSE}} = \mathbf{D}_k^{-1} \left(\sum_{i \in \mathcal{P}_k} \left(1 + \kappa_{\mathrm{t}}^2\right) \hat{\mathbf{h}}_{i,\tau} \hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}} + \boldsymbol{\kappa}_{\mathrm{r}}^2 \sum_{i \in \mathcal{P}_k} \mathbf{F}_{|\hat{\mathbf{h}}_i|^2} + \tilde{\boldsymbol{\Delta}} + \frac{\xi_n}{\rho} \mathbf{I}_W \right)^{-1} \hat{\mathbf{h}}_{k,\tau}. \quad (31)$$

For the sake of comparison, we consider the conventional MMSE decoder $\mathbf{v}_{k,n}^{\mathrm{HU-MMSE}}$ and scalable MMSE decoder $\mathbf{v}_{k,n}^{\mathrm{HU-PMMSE}}$ in CF mMIMO systems studied in [7] and [10], respectively. Hence, the benchmark decoders are

$$\mathbf{v}_{k,n}^{\mathrm{HU-MMSE}} = \left(\sum_{i=1}^{K} \hat{\mathbf{h}}_{i,\tau} \hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}} + \frac{1}{\rho} \mathbf{I}_W \right)^{-1} \hat{\mathbf{h}}_{k,\tau}, \quad (32)$$

$$\mathbf{v}_{k,n}^{\mathrm{HU-PMMSE}} = \left(\sum_{i \in \mathcal{P}_k} \mathbf{D}_k \hat{\mathbf{h}}_{i,\tau} \hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}} \mathbf{D}_k + \frac{1}{\rho} \mathbf{I}_W \right)^{-1} \mathbf{D}_k \hat{\mathbf{h}}_{k,\tau}. \quad (33)$$

Remark 4: Clearly, the decoders, described by (32) and (33), do not include the hardware impairments. Specifically, comparing (33) with (31), the latter expression contains the additional term $\sum_{i\in\mathcal{P}_k}\left(1+\kappa_{\mathrm{t}}^2\right)\hat{\mathbf{h}}_{i,\tau}\hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}}+\boldsymbol{\kappa}_{\mathrm{r}}^2\sum_{i\in\mathcal{P}_k}\mathbf{F}_{|\hat{\mathbf{h}}_i|^2}+\tilde{\boldsymbol{\Delta}}$ which corresponds to the HWIs and the channel estimation error due to pilot contamination by means of κ_{t}, $\boldsymbol{\kappa}_{\mathrm{r}}$ and $\boldsymbol{\Delta}_i$, respectively. Also, (31) includes the amplified thermal noise. Note that (32) refers to the most basic expression that do not account for neither HWI nor scalability design.

C. Lower bound

Since the derivation of lower bounds is in general of higher importance than upper bounds such as the bound described by Proposition 2, herein, we focus on the derivation of an accurate lower bound by following the common analysis in mMIMO theory. In particular, according to [36], we rewrite the received signal in terms of the average effective channel $\mathbb{E}\left\{\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Theta}_{k,n}\mathbf{h}_k s_k\right\}$. Hence, we have

$$\begin{aligned}\hat{s}_{k,n}=&\mathbb{E}\left\{\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Theta}_{k,n}\mathbf{h}_k s_k\right]\}+\left(\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Theta}_{k,n}\mathbf{h}_k s_k-\mathbb{E}\left\{\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Theta}_{k,n}\mathbf{h}_k s_k\right\}\right)\\&+\sum_{i\neq k}^{K}\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}s_i+\sum_{i=1}^{K}\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}\delta_{\mathrm{t},n}^i+\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k(\boldsymbol{\delta}_{\mathrm{r},n}+\boldsymbol{\xi}_n)\,.\end{aligned}\tag{34}$$

Proposition 4: The uplink average SE for user k is lower bounded by

$$\mathrm{SE}_k^{\mathrm{lo}}=\frac{1}{\tau_{\mathrm{c}}}\sum_{n=1}^{\tau_{\mathrm{c}}-\tau_{\mathrm{p}}}\log_2\left(1+\tilde{\gamma}_{k,n}^{\mathrm{lo}}\right),\tag{35}$$

where $\tilde{\gamma}_{k,n}^{\mathrm{lo}}=\frac{S_{k,n}}{I_{k,n}}$ with $S_{k,n}$ and $I_{k,n}$ being the desired signal power and the interference plus noise power, respectively, which are defined as

$$S_{k,n}=\rho_k|\mathbb{E}\left\{\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{k,n}\right\}|^2\tag{36}$$

$$\begin{aligned}I_{k,n}=&\,\mathrm{Var}\left\{\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{k,n}\right\}+\sum_{i\neq k}^{K}\rho_i\mathbb{E}\left\{|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}|^2\right\}+\sigma_{\mathrm{t}}^2\\&+\sigma_{\mathrm{r}}^2+\xi_n\mathbb{E}\left\{\|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\|^2\right\},\end{aligned}\tag{37}$$

where $\sigma_{\mathrm{t}}^2=\sum_{i=1}^{K}\kappa_{\mathrm{t}}^2\rho_i\mathbb{E}\left\{|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}|^2\right\}$ and $\sigma_{\mathrm{r}}^2=\boldsymbol{\kappa}_{\mathrm{r}}^2\mathbb{E}\left\{\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\left(\sum_{i=1}^{K}\rho_i\mathbf{F}_{|\mathbf{h}_i|^2}\right)\mathbf{D}_k\mathbf{v}_{k,n}\right\}$ are the variances of the components corresponding to the additive hardware impairments at the transmit (UE k) and receive side (output of the MMSE decoder), respectively.

Proof: The proof is provided in Appendix C. ■

V. Deterministic Equivalent Analysis

The theory of DEs concerns derivations in the asymptotic limit $K, W \to \infty$ while their ratio $K/W = \beta$ is fixed and the number of antennas per AP N is finite, i.e., $M \to \infty$. Also, we assume similar assumptions to [37, Assump. A1-A3] concerning the covariance matrices under study. Their main advantage is that the corresponding results are tight approximations even for moderate system dimensions, e.g., 8×8 [38]. Also, the DE expressions make lengthy Monte-Carlo simulations unnecessary. Hence, given that CF massive MIMO systems consider a large number of APs makes W quite large, and validates us to employ the DE analysis, in order to obtain the DE uplink achievable rate and extract useful conclusions of significant importance since they can describe CF systems of a practical size.

The DE of the SINR $\tilde{\gamma}_{k,n}^{\mathrm{lo}}$, provided by Proposition 4, obeys to $\tilde{\gamma}_{k,n}^{\mathrm{lo}} - \bar{\gamma}_{k,n}^{\mathrm{lo}} \xrightarrow[M\to\infty]{\text{a.s.}} 0$, while the DE of the rate of user k, relied on the dominated convergence [39] and the continuous mapping theorem [40], is given by

$$\mathrm{SE}_k^{\mathrm{lo}} - \overline{\mathrm{SE}}_k^{\mathrm{lo}} \xrightarrow[N\to\infty]{\text{a.s.}} 0, \tag{38}$$

where $\overline{\mathrm{SE}}_k^{\mathrm{lo}} = \frac{1}{T_{\mathrm{c}}} \sum_{n=1}^{T_{\mathrm{c}}-\tau} \log_2(1 + \bar{\gamma}_{k,n}^{\mathrm{lo}})$. Taking into account for Proposition 3, we have $\mathbf{v}_{k,n} = \mathbf{v}_{k,n}^{\mathrm{HA-PMMSE}}$ while we rewrite the HA-PMMSE decoder at time n as

$$\mathbf{v}_{k,n}^{\mathrm{HA-PMMSE}} = \mathbf{D}_k^{-1} \boldsymbol{\Sigma} \hat{\mathbf{h}}_{k,\tau}, \tag{39}$$

where $\boldsymbol{\Sigma}^{-1} = \sum_{i \in \mathcal{P}_k} \rho_i \left(1 + \kappa_{\mathrm{t}}^2\right) \hat{\mathbf{h}}_{i,\tau} \hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}} + \boldsymbol{\kappa}_{\mathrm{r}}^2 \sum_{i \in \mathcal{P}_k} \rho_i \mathbf{F}_{|\hat{\mathbf{h}}_i|^2} + \tilde{\boldsymbol{\Delta}} + \alpha W \xi_n \mathbf{I}_W$ with a being a regularization scaled by W, in order to make expressions converge to a constant, as W, $K \to \infty$. A similar regularization takes place during the simulations for the benchmark decoders given by (32) and (33), respectively.

Theorem 1: The uplink DE of the SINR of user k at time n with HA-PMMSE decoding in SCF massive MIMO systems, accounting for imperfect CSI and transceiver HWIs, is given by

$$\bar{\gamma}_{k,n}^{\mathrm{lo}} = \frac{|\bar{\sigma}_{\mathrm{PN}}|^2 \tilde{\delta}_k^2}{\boldsymbol{\kappa}_{\mathrm{t}}^2 \tilde{\sigma}_{\mathrm{PN}} \tilde{\delta}_k^2 + \tilde{\sigma}_{\mathrm{PN}} \tilde{\zeta}_{ki} + \left(1 + \boldsymbol{\kappa}_{\mathrm{t}}^2\right) \sum_{i \neq k}^{K} \frac{\rho_i}{\rho_k} \tilde{\mu}_{ki} + \boldsymbol{\kappa}_{\mathrm{r}}^2 \eta_k^{'} \sum_{i=1}^{K} \frac{p_i}{p_k} \mathbf{e}_i + \frac{\tilde{\mathbf{e}}_k^{'}}{p_k}} \tag{40}$$

with $\sigma_{\mathrm{PN}} = \frac{1}{W} \operatorname{tr}\left(\widetilde{\boldsymbol{\Theta}}_{k,n}\right)$, $\bar{\sigma}_{\mathrm{PN}} = \mathbb{E}_{\boldsymbol{\Theta}}\{\sigma_{\mathrm{PN}}\}$, $\tilde{\sigma}_{\mathrm{PN}} = \mathbb{E}_{\boldsymbol{\Theta}}\{|\sigma_{\mathrm{PN}}|^2\}$, $\delta_k = \frac{\left(1+\kappa_{\mathrm{t_{UE}}}^2\right)}{W} \operatorname{tr} \boldsymbol{\Phi}_k \mathbf{T}$, $\tilde{\delta}_k = \frac{1}{W} \operatorname{tr} \boldsymbol{\Phi}_k \mathbf{T}$, $\tilde{\zeta}_{ki} = \frac{1}{W^2} \operatorname{tr}\left(\left(\mathbf{R}_{mk} - \boldsymbol{\Phi}_{mk}\right) \mathbf{T}_k^{'}\right)$, $\mathbf{e}_i = \frac{1}{W} \operatorname{tr}\left(\mathbf{R}_i\right)$, $\zeta_{ki} = \frac{1}{W^2} \operatorname{tr}\left(\mathbf{R}_i \mathbf{T}_k^{'}\right)$, $\nu_{ki} = \frac{1}{W} \operatorname{tr}\left(\boldsymbol{\Phi}_i \mathbf{T}\right)$, $\mu_{ki} = \frac{1}{W^2} \operatorname{tr}\left(\boldsymbol{\Phi}_i \mathbf{T}_k^{'}\right)$, $\tilde{\mu}_{ki} = \zeta_{ki} + \frac{|\nu_{ki}|^2 \mu_{ki}}{(1+\delta_i)^2} - 2\tilde{\sigma}_{\mathrm{PN}} \mathrm{Re}\left\{\frac{\nu_{ki}^* \mu_{ki}}{(1+\delta_i)}\right\}$, $\eta_k^{'} = \frac{1}{W} \operatorname{tr} \mathbf{T}_k^{'}$, and $\tilde{\mathbf{e}}_k^{'} = \frac{1}{W^2} \operatorname{tr}\left(\mathbf{T}_k^{''}\right)$, where

$*$ $\mathbf{T} = \left(\sum_{i=1}^{\mathcal{S}} p_i \left(\frac{1+\kappa_{\mathrm{t}_{\mathrm{UE}}}^2}{W(1+\delta_i)} \mathbf{\Phi}_i + \frac{\kappa_{\mathrm{t}}^2}{W} \mathbf{I}_W \circ \mathbf{\Phi}_i \right) + \frac{1}{W} \tilde{\mathbf{\Delta}} + \alpha \xi_n \mathbf{I}_W \right)^{-1}$,

$*$ $\mathbf{T}_k^{'} = \mathbf{T}\mathbf{\Phi}_k\mathbf{T} + \sum_{i=1}^{K} \frac{\delta_i^{'} \mathbf{T}\mathbf{R}_i\mathbf{T}}{W(1+\delta_i)^2}$,

$*$ $\mathbf{T}_k^{''} = \mathbf{T}^2 + \sum_{i=1}^{K} \frac{\delta_i \mathbf{T}\mathbf{R}_i\mathbf{T}}{W(1+\delta_i)^2}$,

$*$ $\boldsymbol{\delta}^{'} = (\mathbf{I}_K - \mathbf{F})^{-1}$ $[\mathbf{F}]_{k,i} = \frac{1}{W^2(1+\delta_i)} \operatorname{tr}\left(\mathbf{R}_k \mathbf{T}\mathbf{R}_i\mathbf{T}\right)$ $[\mathbf{f}]_k = \frac{1}{W} \operatorname{tr}\left(\mathbf{R}_k\mathbf{T}\mathbf{\Phi}_k\mathbf{T}\right)$ $q \in \{1, \ldots, K\}$,

while all the covariance matrices are assumed with uniformly bounded spectral norms with respect to W.

Proof: The proof of Theorem 1 is given in Appendix D. ■

Remark 5: The term σ_{PN}, capturing the effect of PN variation between the transmission and training phases, is given by [30]

$$\sigma_{\mathrm{PN}} = \frac{1}{W} \sum_{i=1}^{W} e^{i\left(\theta_{mk,n}^{(i)} - \theta_{mk,\tau}^{(i)}\right)} \tag{41}$$

$$\xrightarrow[W\to\infty]{} \begin{cases} e^{-\frac{\sigma_\theta^2}{2}\Delta_n} & \text{SLO setup} \\ e^{-j\left(\theta_{mk,n} - \theta_{mk,\tau}\right)} & \text{CLO setup.} \end{cases} \tag{42}$$

In the last step, the law of large numbers has been applied. The expression of $\bar{\sigma}_{\mathrm{PN}}$ is identical for both scenarios and equal to $e^{-\sigma_\theta^2\Delta_n}$. However, in the case of $\tilde{\sigma}_{\mathrm{PN}}$, we obtain

$$\tilde{\sigma}_{\mathrm{PN}} = \begin{cases} e^{-\sigma_\theta^2\Delta_n} & \text{SLO setup} \\ 1 & \text{CLO setup,} \end{cases} \tag{43}$$

where it is shown that the variances decreases from 1 to $e^{-\sigma_\theta^2\Delta_n}$ when we have CLO and SLO, respectively. In particular, in the case of SLO, $\tilde{\sigma}_{\mathrm{PN}}$ hardens to a value depending on σ_θ^2 and Δ_n while there is no averaging in the case of CLO. Hence, the deterministic SINR $\bar{\gamma}_{k,n}^{\mathrm{lo}}$ is lower in the case of CLO and higher when we employ SLOs.

VI. Numerical Results

This section aims at demonstrating representative numerical examples quantifying the performance of SCF mMIMO networks with inevitable HWIs. As a metric for study and comparison, we consider the achievable SE per UE, given by Theorem 1 and Proposition 2. In parallel, Monte-Carlo simulations, conducted for 10^3 independent channel realizations, verify the proposed analytical results and show their tightness. Specifically, although the analytical results rely on the assumption that $W, K \to \infty$, they coincide with the simulations even for finite values of W and K being of practical interest[5].

[5]This observation is already known in the literature concerning the DEs and supports their usefulness [37], [38], [41], [42].

A. Simulation Setup

We consider $M = 200$ APs with $N = 3$ antennas and $K = 40$ UEs in an 2×2 km^2 area and the 3GPP Urban Microcell model in [43, Table B.1.2.1-1] as a proper benchmark for CF massive MIMO systems[6]. In particular, according to this model assuming a 2 GHz carrier frequency, the large-scale fading coefficient is given by

$$\beta_{mk}[dB] = -30.5 - 36.7 \log_{10}\left(\frac{d_{mk}}{1 \text{ m}}\right) + F_{mk}, \tag{44}$$

where d_{mk} is the distance between AP m and UE k and $F_{mk} \sim \mathcal{CN}(0, 4^2)$ is the shadow fading. Note that shadowing terms between different UEs are correlated as $\mathbb{E}\{F_{mk}F_{ij}\} = 4^2 2^{-\delta_{kj}/9}$ when $m = i$, while if $m \neq i$, they are uncorrelated. The parameter δ_{kj} denotes the distance between UEs k and i. In addition, we assume that the coherence time and bandwidth are $T_{\mathrm{c}} = 2$ ms and $B_{\mathrm{c}} = 100$ kHz, respectively, i.e., the coherence block consists of 200 channel uses with $\tau_{\mathrm{p}} = 20$ and $\tau_{\mathrm{u}} = 180$ while $\zeta = 1$. Moreover, all UEs transmit with the same power in both uplink training and transmission phases given by $\rho = \rho_{\mathrm{p}} = 100$ mW while the thermal noise variance is $\sigma^2 = -174\mathrm{dBm/Hz}$.

Regarding the HWIs, we assume similar to [23], [44], [45] that the variance of PN is $\sigma_i^2 = 1.58 \cdot 10^{-4}$ by setting $f_{\mathrm{c}} = 2\mathrm{GHz}$, $T_{\mathrm{s}} = 10^{-7}\mathrm{s}$, and $c_i = 10^{-17}$ for $i = \phi, \varphi$ in (5). Also, we consider an analog-to-digital converter (ADC) quantizing the received signal to a b bit resolution. In such case, we have $\kappa_{\mathrm{t}} = \kappa_{\mathrm{r}} = 2^{-b}/\sqrt{1 - 2^{-2b}}$. Note that the trend in 5G networks is to employ low-precision ADCs [22]. For example, if $b = 2, 3, 4$, then $\kappa_{\mathrm{t}} = 0.258,\ 0.126, 0.062$. Moreover, we assume that $\xi_n = 1.6\sigma^2$ by considering a low noise amplifier with $\mathcal{F}$ being the noise amplification factor. If $\mathcal{F} = 2$ dB and $b = 3$ bits, we result in $\xi_n = \frac{F\sigma^2}{1-2^{-2b}} = 1.6\sigma^2$. The following figures focus on the impact of the PN and additive HWIs in CF mMIMO systems while the impact concerning the amplified thermal noise is not shown because it is is negligible and due to limited space.

The rate comparison between the MMSE and PMMSE receivers for different values of ρ, considering the impact of HWIs, is shown in Fig. 1. In particular, the receivers, describing

[6]In [7, Remark 4], it is explained that this propagation model corresponds better to the architecture design of CF massive MIMO systems than the established model for CF systems suggested initially in [6]. The main reason is that [6] uses the COST-Hata model which is suitable for macro-cells with APs being at least 1 km far from the UEs and at least 30 m above the ground. Obviously, these characteristics do not match the CF setting where the APs are very close to the UEs, and possibly, at a lower hight. Another important reason is that the model in [6] does not account for shadowing when the UE is closer than 50 m from an AP, while CF massive MIMO systems are more likely suggested for shorter distances.

the HWIs, are denoted as HA-MMSE and HA-PMMSE, respectively. Note that we include both scalable and conventional versions to show the loss in each case. Also, we provide the HU receivers. Clearly, the proposed HA receivers perform better than the standard MMSE and PMMSE receivers, which are hardware unaware. Hence, the increase of the achievable SE of HA-MMSE against HU-MMSE is 17%. In particular, at low SNR due to low quality of CSI, all decoders perform similarly while as the transmit SNR increases the gap between HA and HU decoders increases because the quality of CSI is improved and reduces the inter-user interference. As expected, all the rates saturate at high SNR due to pilot contamination and the dependence from the power of the additive HWIs as can be seen by (7), (8). It is shown that the performance loss between scalable and full MMSE is not significant and the PMMSE version should be preferable given its advantage. Especially, at 20 dB, the loss for the HA-MMSE receiver is 8%. In addition, we depict the upper bound on the capacity for both scalable and unscalable MMSE decoders. In the same figure and for the sake of reference, we have provided the cases of MRC with/without HWIs as well as ideal MMSE and PMMSE with no HWIs. The latter present the best performance but do not take into account the inevitable HWIs.

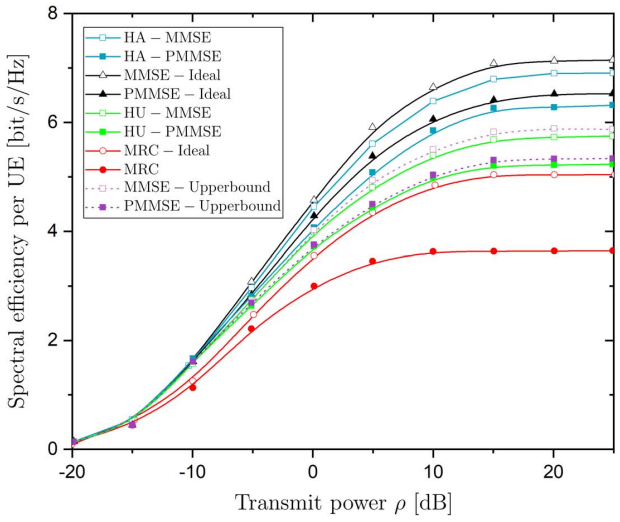

Fig. 1. Uplink achievable SE per UE versus the power ρ for SCF mMIMO systems with HA and HU MMSE and PMMSE decoding ($\sigma_i^2 = 1.58 \cdot 10^{-4}$ for $i = \phi, \varphi$, $\kappa_{\mathrm{t}} = \kappa_{\mathrm{r}} = 0.126$, and $\xi_n = 1.6\sigma^2$).

Keeping the additive HWIs equal to zero, Fig. 2 illustrates the variation of the achievable SE per user versus the time instance of the data transmission phase. In particular, in Fig. 4, the

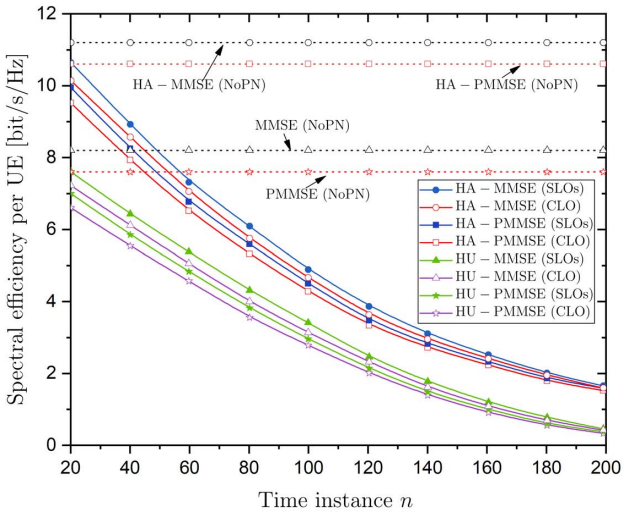

Fig. 2. Uplink achievable SE per UE versus the time instance n of the data transmission phase for SCF mMIMO systems with HA and HU MMSE and PMMSE decoding ($\sigma_i^2 = 1.58 \cdot 10^{-4}$ for $i = \phi, \varphi$, $\kappa_{\mathrm{t}} = \kappa_{\mathrm{r}} = 0$, and $\xi_n = 1.6\sigma^2$).

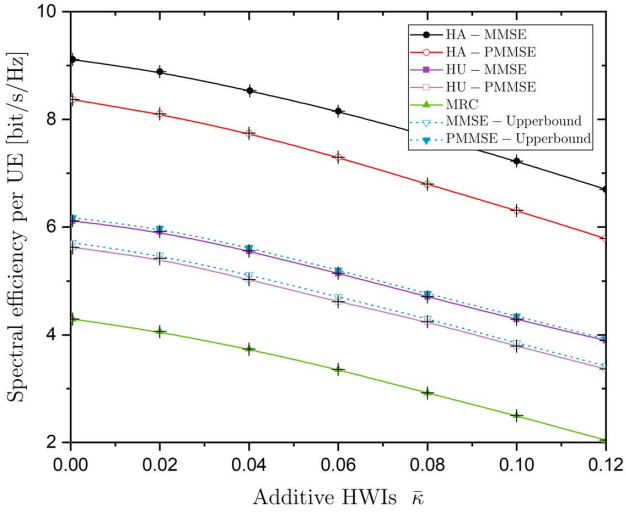

Fig. 3. Uplink achievable SE per UE versus $\bar{\kappa}$ for SCF mMIMO systems with HA and HU MMSE and PMMSE decoding ($\sigma_i^2 = 0$ for $i = \phi, \varphi$, $\kappa_{\mathrm{t}} = \kappa_{\mathrm{r}} = 0.126$, and $\xi_n = 1.6\sigma^2$).

achievable SE decreases as the time increases since the aggregate detrimental contribution from PN becomes higher. In addition, we depict the SEs of MMSE and PMMSE receivers and their HA versions with no PN. These are constant with respect to n since they do not depend on the PN being the only source of channel aging in this work. Next, in the case of both HA-MMSE and

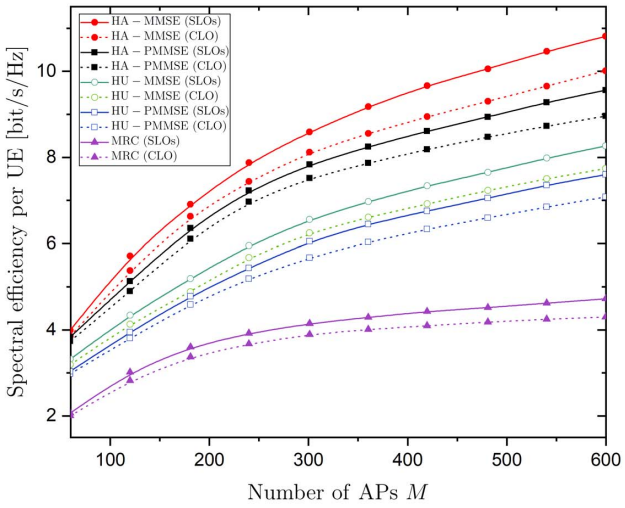

Fig. 4. Uplink achievable SE per UE versus the number of APs M for SCF mMIMO systems with HA and HU MMSE and PMMSE decoding ($\sigma_i^2 = 1.58 \cdot 10^{-4}$ for $i = \phi, \varphi$, $\kappa_{\mathrm{t}} = \kappa_{\mathrm{r}} = 0$, and $\xi_n = 1.6\sigma^2$).

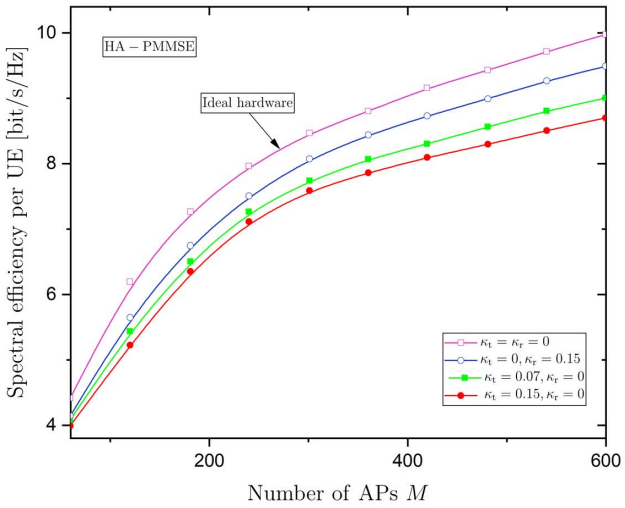

Fig. 5. Uplink achievable SE per UE versus the power ρ for SCF mMIMO systems with HA and HU MMSE and PMMSE decoding ($\sigma_i^2 = 1.58 \cdot 10^{-4}$ for $i = \phi, \varphi$, $\kappa_{\mathrm{t}} = \kappa_{\mathrm{r}} = 0.126$,and $\xi_n = 1.6\sigma^2$).

HA-PMMSE receivers, the SLOs configuration outperforms the CLO design since the distortions from different LOs average out in the large system limit described by (43) [15]. Moreover, the SEs for both HA-MMSE and HA-PMMSE receivers diminish with increasing n since the phase drift becomes higher and the corresponding impact of PN becomes quite destructive. At the end

of the data transmission, where $n = 200$, the degradation is so heavy that the performance with scalable and conventional CF mMIMO systems becomes almost identical for both cases of HA and HU decoders.

Fig. 3 provides the performance of the achievable SE versus $\bar{\kappa}$, where $\kappa_{\mathrm{t}} = \bar{\kappa} + 0.1$ and $\kappa_{\mathrm{r}} = \bar{\kappa} + 0.3$ while the impact of PN has not been considered. The figure starts from the case of no additive HWIs when $\bar{\kappa} = 0$ and ends to severe additive HWIs. Obviously, the higher the additive HWIs, the higher the degradation of the system performance becomes. Furthermore, it is evident that the practical SCF mMIMO systems perform very well since the proposed HA-PMMSE achieves just 9% less SE with HA-MMSE receivers when $\bar{\kappa} = 0.06$. When it comes to the HU decoders, the rates worsen with increasing RAHWIs as well. Regarding the difference between the HA-PMMSE and PMMSE decoders is significant since it is approximately 2.6 bit/s/Hz for all values of $\bar{\kappa}$. Also, we illustrate the superior performance of MMSE-style decoders against the MRC decoder as anticipated. Moreover, we provide the MMSE decoder given by Proposition 2. The "cross" symbol represents the simulation results and it is shown the provided tightness of the DE analytical expressions.

Fig. 4 provides the comparison of the SEs versus the number of APs M in the HA and HU cases for both scalable and conventional CF mMIMO systems. We consider the impact of PN while the additive HWIs are not taken into account. In all cases, $\overline{\mathrm{SE}}_k^{\mathrm{lo}}$ increases with M. Moreover, we observe that SLOs achieve a better SE than a CLO design. Actually, the performance gap between the SLOs and CLO design increases with M because the phase drifts are independent and in the large system limit they are averaged, which is a benefit paid as a trade of higher deployment cost [23], [32].

Fig. 5 depicts the impact of the additive HWIs on CF mMIMO systems with focus on the proposed HA-PMMSE combiner by showing the average achievable SE per UE versus the number of APs while the PN is assumed negligible. Obviously, the achievable SE increases M but the additive HWIs at the AP side and UE sides behave differently. Of course, the presence of HWIs degrades the performance. Specifically, the additive HWIs at the transmit side, i.e., at the UE, has a greater impact on the SE. For example, at $M = 400$ APs, the loss due to the receive distortion is only 4% while, in the case of the transmit distortion, the degradation is 11%. Hence, the quality of hardware at the transmitter side should be considered more during the design.

VII. Conclusion

In this paper, we investigated the impact of HWIs in SCF mMIMO systems. Specifically, we introduced a general molel with HWIs during the channel estimation and we showed how HWIs alter the CSI. Moreover, we derived the HA-PMMSE combiner accounting for HWIs and scalability. Next, we obtained upper and lower bounds on the SE. In particular, we derived the DE uplink SE, carrying the impact of HWIs, and we demonstrated the realistic performance of SCF mMIMO systems by varying the fundamental system parameters. Insightful conclusions as guidelides for practical implementation of SCF mMIMO systems.

Appendix A
Proof of Proposition 2

An upper bound on the uplink capacity is

$$R_k = \frac{1}{\tau_{\mathrm{p}}} \sum_{n=1}^{\tau_{\mathrm{c}}-\tau_{\mathrm{p}}} \mathbb{E}\left\{ \max_{\mathbf{v}_{k,n}:\|\mathbf{v}_{k,n}\|=1} \log_2\left(1+\gamma_{k,n}\left(\mathbf{v}_{k,n}\right)\right) \right\} \qquad [\mathrm{bit/s/Hz}]\,. \tag{45}$$

Relied on the assumptions of perfect CSI and that the multi-user interference is somehow canceled, we take advantage of the optimality of Gaussian signaling as in [46], [47] since all the additive noise terms including the transmit and receive distortion are circularly symmetric complex Gaussian distributed and independent of the useful signals. Thus, we obtain the uplink SINR by means of (20) as

$$\gamma_{k,n}(\mathbf{v}_{k,n}) = \frac{\rho_k \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{h}_{k,n} \mathbf{h}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{v}_{k,n}}{\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \left(\sum_{i=1}^{K} \kappa_{\mathrm{t}}^2 \rho_i \mathbf{h}_{i,n} \mathbf{h}_{i,n}^{\mathsf{H}} + \boldsymbol{\kappa}_{\mathrm{r}}^2 \sum_{i=1}^{K} \rho_i \mathbf{F}_{|\mathbf{h}_i|^2} + \xi_n \mathbf{I}_W \right) \mathbf{D}_k \mathbf{v}_{k,n}}, \tag{46}$$

where $\boldsymbol{\kappa}_{\mathrm{r}}^2$ and $\mathbf{F}_{|\mathbf{h}_i|^2}$ are defined in Proposition 2.

The maximization of the achievable rate is obtained by noticing that the logarithm is a monotonically increasing function. Actually, the optimization concerns a generalized Rayleigh quotient problem which is solved by (23) and achieves the upper bound given by 21 after plugging the maximizing combiner into (46).

APPENDIX B
PROOF OF LEMMA 2

The MSE definition between the estimated received signal, given by (20), and the transmit signal during the training phase, i.e., $n \in \{0, \tau\}$ gives

$$\mathrm{MSE}_{k,n} = \mathbb{E}\left\{ \|\hat{s}_{k,n} - s_{k,n}\|_2^2 \big| \hat{\mathbf{H}}_n \right\} \tag{47}$$

$$= \mathrm{tr}\left(\mathbb{E}\left\{ \left(\hat{s}_{k,n} - s_{k,n}\right) \left(\hat{s}_{k,n} - s_{k,n}\right)^{\mathsf{H}} \big| \hat{\mathbf{H}}_n \right\} \right) \tag{48}$$

$$= \mathrm{tr}\left(\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \left(\sum_{i=1}^{K} \rho_i \left(1 + \kappa_{\mathrm{t}}^2\right) \left(\hat{\mathbf{h}}_{i,n} \hat{\mathbf{h}}_{i,n}^{\mathsf{H}} + \boldsymbol{\Delta}_i \right) + \boldsymbol{\kappa}_{\mathrm{r}}^2 \sum_{i=1}^{K} \rho_i \left(\mathbf{F}_{|\hat{\mathbf{h}}_i|^2} + \mathbf{F}_{|\boldsymbol{\Delta}_i|^2} \right) + \xi_n \mathbf{I}_W \right) \mathbf{D}_k \mathbf{v}_{k,n} \right.$$

$$\left. - \rho_k \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \hat{\mathbf{h}}_{k,n} - p_k \hat{\mathbf{h}}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{v}_{k,n} + \rho_k \right), \tag{49}$$

where we have considered that $\mathrm{tr}\left(\mathbb{E}\left\{ \mathbf{h}_{i,n} \mathbf{h}_{i,n}^{\mathsf{H}} \big| \hat{\mathbf{h}}_{i,n} \right\} \right) = \mathrm{tr}\left(\hat{\mathbf{h}}_{i,n} \hat{\mathbf{h}}_{i,n}^{\mathsf{H}} + \boldsymbol{\Delta}_i \right)$ with $\boldsymbol{\Delta}_i = \mathbb{E}\left\{ \tilde{\mathbf{h}}_{i,n} \tilde{\mathbf{h}}_{i,n}^{\mathsf{H}} | \hat{\mathbf{h}}_{i,n} \right\}$.

APPENDIX C
PROOF OF PROPOSITION 4

First, we follow the approach in [14], in order to compute the average achievable SE for each n in the transmission phase, and then, we obtain the average over these SEs. The achievable SE per UE in (35) is obtained by taking into account for the Gaussianity of the input symbols and by assuming a worst-case assumption regarding the computation of the mutual information [48, Theorem 1], where the inter-user interference and the distortion noises are treated as independent Gaussian noise. Moreover, relied on [36], we leverage the knowledge of $\mathbb{E}\left\{ \mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \boldsymbol{\Theta}_{k,n} \mathbf{h}_k s_k \right]\}$ for the detection while the deviation from the average effective channel gain is treated as as worst-case Gaussian noise. Thus, we obtain $\tilde{\gamma}^{\mathrm{lo}}$, where the expectation operator in the various terms is taken with respect to the channel vectors as well as the noise processes, and the proof is concluded.

APPENDIX D
PROOF OF THEOREM 1

First, we obtain the DE of the desired signal power given by (36). Specifically, we have

$$\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{h}_{k,n} = \hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}} \boldsymbol{\Sigma} \boldsymbol{\Theta}_{k,n} \mathbf{h}_k \tag{50}$$

$$= \frac{\frac{1}{W} \hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}} \boldsymbol{\Sigma}_k \widetilde{\boldsymbol{\Theta}}_{k,n} \left(\hat{\mathbf{h}}_{k,\tau} + \tilde{\mathbf{e}}_{k,\tau} \right)}{1 + \frac{q}{W} \hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}} \boldsymbol{\Sigma}_k \hat{\mathbf{h}}_{k,\tau}} \tag{51}$$

$$\asymp \frac{\frac{1}{W}\hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}}\boldsymbol{\Sigma}_k\widetilde{\boldsymbol{\Theta}}_{k,n}\hat{\mathbf{h}}_{k,\tau}}{1+\frac{q}{W}\hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}}\boldsymbol{\Sigma}_k\hat{\mathbf{h}}_{k,\tau}}, \tag{52}$$

where in (50), we have replaced the expression of the MMSE decoder given by (39). In the next equation, we have applied the matrix inversion lemma and used that $\mathbf{h}_k = \boldsymbol{\Theta}_{k,\tau}^{*}\mathbf{h}_{k,\tau}$ together with the orthogonality between the channel and its estimated version. Note that $q = \left(1+\kappa_{\mathrm{t_{UE}}}^2\right)\rho_k$, $\widetilde{\boldsymbol{\Theta}}_{k,n} = \boldsymbol{\Theta}_{k,n}\boldsymbol{\Theta}_{k,\tau}^{*}$ while $\boldsymbol{\Sigma}_k^{-1}$ is defined as

$$\begin{aligned}\boldsymbol{\Sigma}_k^{-1} &= \boldsymbol{\Sigma}^{-1} - \frac{q}{W}\hat{\mathbf{h}}_{k,\tau}\hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}} \qquad (53)\\ &= \sum_{\substack{i\in\mathcal{P}_k\\ i\neq k}}\frac{q}{W}\hat{\mathbf{h}}_{i,\tau}\hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}} + \frac{\boldsymbol{\kappa}_{\mathrm{r}}^2}{W}\sum_{i\in\mathcal{P}_k}\rho_i\mathbf{F}_{|\hat{\mathbf{h}}_i|^2} + \tilde{\boldsymbol{\Delta}} + \alpha\xi_n\mathbf{I}_W.\end{aligned}$$

In the numerator of (52), we have applied [49, Lem. B.26]. The derivation continues with the use of [49, Lem. 14.3] known as rank-1 perturbation lemma, [49, Lem. B.26], [50, Theorem 1] and [30, Lem. 9] as[7]

$$\begin{aligned}\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{k,n} &\asymp \frac{\frac{1}{W}\operatorname{tr}\left(\widetilde{\boldsymbol{\Theta}}_{k,n}\boldsymbol{\Phi}_k\mathbf{T}\right)}{1+\frac{q}{W}\operatorname{tr}\left(\boldsymbol{\Phi}_k\mathbf{T}\right)} \qquad (54)\\ &= \frac{\sigma_{\mathrm{PN}}\tilde{\delta}_k}{1+\delta_k}, \qquad (55)\end{aligned}$$

where $\sigma_{\mathrm{PN}} = \frac{1}{W}\operatorname{tr}\left(\widetilde{\boldsymbol{\Theta}}_{k,n}\right)$, $\tilde{\delta}_k = \frac{1}{W}\operatorname{tr}\boldsymbol{\Phi}_k\mathbf{T}$, and $\delta_k = \frac{q}{W}\operatorname{tr}\boldsymbol{\Phi}_k\mathbf{T}$. Note that in (55), we have made use of the commuting property among diagonal matrices inside a trace and [51, p. 207] The term, concerning the deviation from the average effective channel gain, becomes

$$\begin{aligned}\operatorname{Var}\left\{\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{k,n}\right\} &\asymp \frac{\frac{1}{W}\mathbb{E}\left\{|\hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}}\boldsymbol{\Sigma}_k\mathbf{e}_{k,n}|^2\right\}}{1+\delta_k} \qquad (56)\\ &= \frac{\frac{1}{W}\mathbb{E}\left\{|\hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}}\boldsymbol{\Sigma}_k\widetilde{\boldsymbol{\Theta}}_{k,n}\mathbf{e}_{k,\tau}|^2\right\}}{1+\delta_k} \qquad (57)\\ &\asymp \frac{\mathbb{E}\left\{|\sigma_{\mathrm{PN}}|^2\right\}\mathbb{E}\left\{\frac{1}{W^2}\operatorname{tr}\left(\boldsymbol{\Phi}_k\boldsymbol{\Sigma}_k\left(\mathbf{R}_{mk}-\boldsymbol{\Phi}_{mk}\right)\boldsymbol{\Sigma}_k\right)\right\}}{1+\delta_k} \qquad (58)\\ &\asymp \frac{\mathbb{E}\left\{|\sigma_{\mathrm{PN}}|^2\right\}\mathbb{E}\left\{\frac{1}{W^2}\operatorname{tr}\left(\left(\mathbf{R}_{mk}-\boldsymbol{\Phi}_{mk}\right)\mathbf{T}_k^{'}\right)\right\}}{1+\delta_k} \qquad (59)\end{aligned}$$

[7]It is worthwhile to mention that the diagonal matrix $\mathbf{F}_{|\hat{\mathbf{h}}_i|^2}$ inside the MMSE decoder is considered a deterministic matrix with entries in the diagonal elements the limits of the individual diagonal elements [32]. In particular, exploiting the uniform convergence $\limsup_W \max_{1\leq i\leq W}\left|\left[\hat{\mathbf{h}}_i\hat{\mathbf{h}}_i^{\mathsf{H}}\right]_{ww} - \left[\hat{\boldsymbol{\Phi}}_i\right]_{ww}\right| = 0$, we have $\left\|\frac{1}{W}\mathrm{diag}(\hat{\mathbf{h}}_i\hat{\mathbf{h}}_i^{H}) - \frac{1}{W}\operatorname{tr}\left(\mathrm{diag}\left(\hat{\boldsymbol{\Phi}}_i\right)\right)\right\| \xrightarrow[W\to\infty]{\text{a.s.}} 0$.

where in (56), we have used the matrix inversion lemma, [49, Lem. B.26], and [50, Theorem 1]. In (57), we have used that $\mathbf{e}_k = \mathbf{\Theta}^*_{k,\tau}\mathbf{e}_{k,\tau}$, while in (58), we have applied the rank-1 perturbation lemma, [49, Lem. B.26], and [30, Lem. 10]. and [49, Lem. B.26] again. The last step includes application of [37, Theorem 2].

Based on (37), the interference power of the kth UE is written as

$$|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}|^2 \asymp \left|\frac{\frac{1}{W}\hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}}\mathbf{\Sigma}_k\widetilde{\mathbf{\Theta}}_{i,n}\mathbf{h}_{i,\tau}}{1+\delta_k}\right|^2 \tag{60}$$

$$\asymp \frac{\hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}}\mathbf{\Sigma}_k\widetilde{\mathbf{\Theta}}_{i,n}\mathbf{h}_{i,\tau}\mathbf{h}_{i,\tau}^{\mathsf{H}}\widetilde{\mathbf{\Theta}}_{i,n}^{\mathsf{H}}\mathbf{\Sigma}_k\hat{\mathbf{h}}_{k,\tau}}{W^2\left(1+\delta_k\right)^2} \tag{61}$$

$$\asymp \frac{\mathbf{h}_{i,\tau}^{\mathsf{H}}\widetilde{\mathbf{\Theta}}_{i,n}^{\mathsf{H}}\mathbf{\Sigma}_k\mathbf{\Phi}_k\mathbf{\Sigma}_k\widetilde{\mathbf{\Theta}}_{i,n}\mathbf{h}_{i,\tau}}{W^2\left(1+\delta_k\right)^2} \tag{62}$$

$$\begin{aligned}\asymp \frac{1}{\left(1+\delta_k\right)^2}\Bigg(&\frac{1}{W^2}\mathbf{h}_{i,\tau}^{\mathsf{H}}\widetilde{\mathbf{\Theta}}_{i,n}^{\mathsf{H}}\mathbf{\Sigma}_{ki}\mathbf{\Phi}_k\mathbf{\Sigma}_{ki}\widetilde{\mathbf{\Theta}}_{i,n}\mathbf{h}_{i,\tau}\\ &+\frac{|\mathbf{h}_{i,\tau}^{\mathsf{H}}\mathbf{\Sigma}_{ki}\hat{\mathbf{h}}_{i,\tau}|^2\hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}}\widetilde{\mathbf{\Theta}}_{i,n}^{\mathsf{H}}\mathbf{\Sigma}_{ki}\mathbf{\Phi}_k\mathbf{\Sigma}_{ki}\widetilde{\mathbf{\Theta}}_{i,n}\hat{\mathbf{h}}_{i,\tau}}{W^4\left(1+\delta_i\right)^2}\\ &-2\mathrm{Re}\left\{\frac{\left(\hat{\mathbf{h}}_{i,\tau}^{\mathsf{H}}\mathbf{\Sigma}_{ki}\mathbf{h}_{i,\tau}\right)\left(\mathbf{h}_{i,\tau}^{\mathsf{H}}\mathbf{\Sigma}_{ki}\mathbf{\Phi}_k\mathbf{\Sigma}_{ki}\widetilde{\mathbf{\Theta}}_{i,n}\hat{\mathbf{h}}_{i,\tau}\right)}{W^3\left(1+\delta_i\right)^2}\right\}\Bigg)\end{aligned} \tag{63}$$

$$\asymp \frac{1}{\left(1+\delta_k\right)^2}\left(\zeta_{ki}+\frac{|\nu_{ki}|^2\mu_{ki}}{\left(1+\delta_i\right)^2}-2\mathrm{Re}\left\{\frac{\sigma_{\mathrm{PN}}\nu_{ki}^*\mu_{ki}}{\left(1+\delta_i\right)}\right\}\right), \tag{64}$$

where we have applied the matrix inversion lemma in (60), while in (61) and (62), we have used [49, Lem. B.26]. In (63), we have applied again the matrix inversion lemma, and in the last step, we have used the rank-1 perturbation lemma, [49, Lem. B.26], [50, Theorem 1], and [37, Theorem 2]. The last step includes application of [30, Lem. 10]. The definitions of the various parameters are given in the presentation of the theorem.

Below, the terms corresponding to the transmit and receive distortions are derived. Specifically, we have

$$\begin{aligned}\sigma_{\mathrm{t}}^2 &= \sum_{i=1}^{K}\kappa_{\mathrm{t}}^2\rho_i|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}|^2\\ &= \kappa_{\mathrm{t}}^2\left(\rho_k\mathbb{E}\left\{|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{k,n}|^2\right\}+\sum_{i\neq k}^{K}\rho_i\mathbb{E}\left\{|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}|^2\right\}\right)\\ &\asymp \frac{\kappa_{\mathrm{t}}^2}{\left(1+\delta_k\right)^2}\left(\rho_k|\sigma_{\mathrm{PN}}|^2\tilde{\delta}_k^2+\sum_{i\neq k}^{K}\rho_i\left(\zeta_{ki}+\frac{|\nu_{ki}|^2\mu_{ki}}{\left(1+\delta_i\right)^2}-2\mathrm{Re}\left\{\frac{\sigma_{\mathrm{PN}}\nu_{ki}^*\mu_{ki}}{\left(1+\delta_i\right)}\right\}\right)\right),\end{aligned} \tag{65}$$

where, in (65), we have substituted (55) and 64. Also, the deterministic σ_{r}^2 as $W \to \infty$ becomes

$$\sigma_{\mathrm{r}}^2 = \boldsymbol{\kappa}_{\mathrm{r}}^2 \mathbb{E}\left\{\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k (\mathbf{I}_W \circ \mathbf{H}_n \mathbf{P} \mathbf{H}_n^{\mathsf{H}})\, \mathbf{D}_k \mathbf{v}_{k,n}\right\} \tag{66}$$

$$\asymp \frac{\boldsymbol{\kappa}_{\mathrm{r}}^2}{W\left(1+\delta_k\right)^2} \mathbb{E}\left\{\mathrm{tr}\left((\mathbf{I}_W \circ \mathbf{H}_n \mathbf{P} \mathbf{H}_n^{\mathsf{H}})\, \boldsymbol{\Sigma}_k \boldsymbol{\Phi}_k \boldsymbol{\Sigma}_k\right)\right\} \tag{67}$$

$$\asymp \frac{\boldsymbol{\kappa}_{\mathrm{r}}^2 \eta_k^{'}}{W\left(1+\delta_k\right)^2} \sum_{i=1}^{K} \rho_i \,\mathrm{tr}\left(\mathbf{R}_i\right), \tag{68}$$

where in (66) we have written the diagonal matrix by means of a hadamard product. Next, we have exploited the freeness between $\mathbf{v}_{k,n}\mathbf{v}_{k,n}^{\mathsf{H}}$ and the diagonal matrix $\mathbf{I}_W \circ \mathbf{H}\mathbf{H}^{\mathsf{H}}$. Also, we have applied [49, Lem. B.26] [50, Theorem 1], [37, Theorem 2], and [51, p. 207] while we have set $\eta_k^{'} = \frac{1}{W}\,\mathrm{tr}\,\mathbf{T}_k^{'}$.

The DE of the last term of (37), corresponding to the amplified thermal noise contribution, becomes

$$\|\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k\|^2 \asymp \frac{\frac{1}{W^2} \hat{\mathbf{h}}_{k,\tau}^{\mathsf{H}} \boldsymbol{\Sigma}_k^2 \hat{\mathbf{h}}_{k,\tau}}{\left(1+\delta_k\right)^2} \tag{69}$$

$$\asymp \frac{\frac{1}{W^2}\,\mathrm{tr}\left(\boldsymbol{\Sigma}_k^2 \boldsymbol{\Phi}_k\right)}{\left(1+\delta_k\right)^2} \tag{70}$$

$$= \frac{\frac{1}{W^2}\,\mathrm{tr}\left(\mathbf{T}_k^{''}\right)}{\left(1+\delta_k\right)^2}, \tag{71}$$

where in (69) and (70), we have applied the matrix inversion lemma and [49, Lem. B.26], respectively. In the last step, we have used the rank-1 perturbation lemma as well as [37, Theorem 2] and [50, Theorem 1].

By substituting (55), (59), (64), (65), (68), and (71) into (36) and (37), the DE SINR is derived and the proof is concluded.

REFERENCES

[1] T. 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] T. L. Marzetta *et al.*, *Fundamentals of Massive MIMO*. Cambridge University Press, 2016.

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

[4] S. Venkatesan, A. Lozano, and R. Valenzuela, "Network MIMO: Overcoming intercell interference in indoor wireless systems." Citeseer, 2007, pp. 83–87.

[5] R. Irmer *et al.*, "Coordinated multipoint: Concepts, performance, and field trial results," *IEEE Commun. Mag.*, vol. 49, no. 2, pp. 102–111, 2011.

[6] H. Q. Ngo *et al.*, "Cell-free massive MIMO versus small cells," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1834–1850, 2017.

[7] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *accepted in IEEE Trans. Wireless Commun.*, 2020.

[8] S. Buzzi and C. D'Andrea, "Cell-free massive MIMO: User-centric approach," *IEEE Wireless Commun. Lett.*, vol. 6, no. 6, pp. 706–709, 2017.

[9] S. Buzzi *et al.*, "User-centric 5G cellular networks: Resource allocation and comparison with the cell-free massive MIMO approach," *IEEE Trans. Wireless Commun.*, 2019.

[10] E. Björnson and L. Sanguinetti, "Scalable cell-free massive MIMO systems," *arXiv preprint arXiv:1908.03119*, 2019.

[11] E. Björnson *et al.*, "Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission," *IEEE Trans. Signal Process.*, vol. 59, no. 12, pp. 6086–6101, 2011.

[12] J. Qi and S. Aïssa, "Analysis and compensation of I/Q imbalance in MIMO transmit-receive diversity systems," *IEEE Trans. Commun.*, vol. 58, no. 5, pp. 1546–1556, 2010.

[13] H. Mehrpouyan *et al.*, "Joint estimation of channel and oscillator phase noise in MIMO systems," *IEEE Trans. Signal Process.*, vol. 60, no. 9, pp. 4790–4807, Sept 2012.

[14] A. Pitarokoilis, S. Mohammed, and E. Larsson, "Uplink performance of time-reversal MRC in massive MIMO systems subject to phase noise," *IEEE Trans. Wireless Commun.*, vol. 14, no. 2, pp. 711–723, Feb 2015.

[15] A. K. Papazafeiropoulos, "Impact of general channel aging conditions on the downlink performance of massive MIMO," *IEEE Trans. Veh. Tech.*, vol. 66, no. 2, pp. 1428–1442, Feb 2017.

[16] J. Qi and S. Aïssa, "On the power amplifier nonlinearity in MIMO transmit beamforming systems," *IEEE Trans. Commun.*, vol. 60, no. 3, pp. 876–887, 2012.

[17] T. Schenk, *RF imperfections in high-rate wireless systems: impact and digital compensation*. Springer Science & Business Media, 2008.

[18] C. Studer, M. Wenk, and A. Burg, "MIMO transmission with residual transmit-RF impairments," in *ITG/IEEE Work. Smart Ant. (WSA)*. IEEE, 2010, pp. 189–196.

[19] J. Zhang *et al.*, "Performance analysis and power control of cell-free massive MIMO systems with hardware impairments," *IEEE Access*, vol. 6, pp. 55 302–55 314, 2018.

[20] X. Zhang *et al.*, "Secure communications over cell-free massive MIMO networks with hardware impairments," *IEEE Syst. J.*, pp. 1–12, 2019.

[21] H. Masoumi and M. J. Emadi, "Performance analysis of cell-free massive MIMO system with limited fronthaul capacity and hardware impairments," *IEEE Trans. Wireless Commun.*, pp. 1–1, 2019.

[22] X. Hu *et al.*, "Cell-free massive MIMO systems with low resolution ADCs," *IEEE Trans. Commun.*, vol. 67, no. 10, pp. 6844–6857, Oct 2019.

[23] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with non-ideal arbitrary arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, vol. 14, no. 8, pp. 4353–4368, Aug 2015.

[24] S. Perlman and A. Forenza, "An introduction to pcell," *Artemis Networks white paper*, 2015.

[25] G. Interdonato *et al.*, "Ubiquitous cell-free massive MIMO communications," *EURASIP J. Wireless Commun. Net.*, vol. 2019, no. 1, p. 197, 2019.

[26] D. Neumann, M. Joham, and W. Utschick, "Covariance matrix estimation in massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 6, pp. 863–867, 2018.

[27] K. Upadhya and S. A. Vorobyov, "Covariance matrix estimation for massive MIMO," *IEEE Signal Process. Lett.*, vol. 25, no. 4, pp. 546–550, 2018.

[28] A. Demir, A. Mehrotra, and J. Roychowdhury, "Phase noise in oscillators: A unifying theory and numerical methods for characterization," *IEEE Trans. Circuits Syst. I*, vol. 47, no. 5, pp. 655–674, May 2000.

[29] D. Petrovic, W. Rave, and G. Fettweis, "Effects of phase noise on OFDM systems with and without PLL: Characterization and compensation," *IEEE Trans. Commun.*, vol. 55, no. 8, pp. 1607–1616, Aug 2007.

[30] R. Krishnan *et al.*, "Linear massive MIMO precoders in the presence of phase noise-A large-scale analysis," *IEEE Trans. Veh. Tech.*, vol. 65, no. 5, pp. 3057–3071, 2016.

[31] H. Holma and A. Toskala, *LTE for UMTS: Evolution to LTE-Advanced*, Wiley, Ed., 2011.

[32] A. Papazafeiropoulos, B. Clerckx, and T. Ratnarajah, "Rate-splitting to mitigate residual transceiver hardware impairments in massive MIMO systems," *IEEE Trans. Veh. Tech.*, vol. 66, no. 9, pp. 8196–8211, 2017.

[33] E. Björnson *et al.*, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, 2014.

[34] E. Nayebi *et al.*, "Precoding and power optimization in cell-free massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4445–4459, 2017.

[35] G. Interdonato, P. Frenger, and E. G. Larsson, "Scalability aspects of cell-free massive MIMO," in *IEEE International Conference on Communications (ICC)*, May 2019, pp. 1–6.

[36] M. Medard, "The effect upon channel capacity in wireless communications of perfect and imperfect knowledge of the channel," *IEEE Trans. Inf. Theory*, vol. 46, no. 3, pp. 933–946, May 2000.

[37] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Select. Areas Commun.*, vol. 31, no. 2, pp. 160–171, February 2013.

[38] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge University Press, 2011.

[39] P. Billingsley, *Probability and measure*, 3rd ed. John Wiley & Sons, Inc., 2008.

[40] A. W. van der Vaart, *Asymptotic statistics (Cambridge series in statistical and probabilistic mathematics)*. New York: Cambridge University Press, 2000.

[41] K. Truong and R. Heath, "Effects of channel aging in massive MIMO systems," *IEEE/KICS J. Commun. Net., Special Issue on massive MIMO*, vol. 15, no. 4, pp. 338–351, Aug 2013.

[42] A. K. Papazafeiropoulos and T. Ratnarajah, "Deterministic equivalent performance analysis of time-varying massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 10, pp. 5795–5809, 2015.

[43] 3GPP, "Further advancements for E-UTRA physical layer aspects (Release 9)," 3GPP TS 36.814, Tech. Rep., 2017.

[44] A. Papazafeiropoulos and T. Ratnarajah, "Towards a realistic assessment of multiple antenna HCNs: Residual additive transceiver hardware impairments and channel aging," *IEEE Trans. Veh. Tech.*, vol. 66, no. 10, pp. 9061–9073, Oct 2017.

[45] A. Papazafeiropoulos *et al.*, "Nuts and bolts of a realistic stochastic geometric analysis of mmWave HetNets: Hardware impairments and channel aging," *IEEE Trans. Veh. Tech.*, vol. 68, no. 6, pp. 5657–5671, June 2019.

[46] E. Telatar, "Capacity of multi-antenna gaussian channels," *Europ. Trans. on Telecom.*, vol. 10, no. 6, pp. 585–595, 1999.

[47] E. Björnson *et al.*, "Capacity limits and multiplexing gains of MIMO channels with transceiver impairments," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 91–94, 2013.

[48] B. Hassibi and B. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inform. Theory*, vol. 49, no. 4, pp. 951–963, April 2003.

[49] Z. Bai and J. W. Silverstein, *Spectral analysis of large dimensional random matrices*. Springer, 2010, vol. 20.

[50] S. Wagner *et al.*, "Large system analysis of linear precoding in correlated MISO broadcast channels under limited feedback," *IEEE Trans. Inform. Theory*, vol. 58, no. 7, pp. 4509–4537, July 2012.

[51] T. Tao, *Topics in random matrix theory*. American Mathematical Soc., 2012, vol. 132.