Scalable Cell-Free Massive MIMO Systems With Hardware Impairments

A. K. Papazafeiropoulos*,†, Emil Björnson♮, P. Kourtessis*, S. Chatzinotas†, John M. Senior*
*Communications and Intelligent Systems Research Group, University of Hertfordshire, Hatfield, U. K.
♮Department of Electrical Engineering (ISY), Linköping University, 58183 Linköping, Sweden
†SnT - securityandtrust.lu, University of Luxembourg, Luxembourg
Email: tapapazaf@gmail.com, emil.bjornson@liu.se, p.kourtessis@herts.ac.uk, symeon.chatzinotas@uni.lu

***Abstract*—Despite the deleterious effect of hardware impairments (HWIs) on wireless systems, most prior works in cell-free (CF) massive multiple-input-multiple-output (mMIMO) systems have not accounted for their impact. In particular, the effect of phase noise (PN) has not been investigated at all in CF systems. Moreover, there is no work investigating HWIs in scalable CF (SCF) mMIMO systems, encountering the prohibitively demanding fronthaul requirements of large networks with many users. Hence, we derive the uplink spectral efficiency (SE) under HWIs with minimum mean-squared error (MMSE) combining in closed-form by means of the deterministic equivalent (DE) analysis. Notably, previous works, accounted for MMSE decoding, studied the corresponding SE only by means of simulations. Numerical results illustrate the performance loss due to HWIs and result in insightful conclusions.**

***Index Terms*—Cell-free massive MIMO systems, user-centric 5G networks, transceiver hardware impairments, MMSE processing, achievable spectral efficiency.**

I. Introduction

Recently, cell-free (CF) massive multiple-input-multiple-output (mMIMO) systems have been proposed as a promising technology for fifth-generation (5G) and beyond networks [1]. This novel architecture reaps the benefits of network MIMO [2] and mMIMO systems [3] by assuming a large number of distributed access points (APs) that are connected to a central processing unit (CPU) and serve jointly all user equipments (UEs) [1]. Thus, CF mMIMO systems enjoy lower path-losses, increased macro-diversity, and under certain conditions [4], they take advantage of favorable propagation and channel hardening.

However, the high computational burden and fronthaul capacity, undertaken by the CPU which coordinates and processes all the signals in canonical CF mMIMO systems, increases with the number of UEs [1], [5]. The result is an unfeasible technology in the large system size limit [6]. Contrary to canonical CF mMIMO systems, the authors in [7] proposed a user-centric design, where each UE is served by a subset of APs (and not all APs) providing the best channel conditions, in order to decrease complexity. Moreover, by taking into account for Poisson point process distributed APs, in [8], the coverage probablilty for CF mMIMO systems was derived, while in [9], the optimal energy efficiency was obtained in closed form. Furthermore, although previous works mentioned the outperformance of CF mMIMO systems against small cells (SCs), the authors in [5] showed that this behaviour holds only in the case of a centralized version of minimum mean-squared error (MMSE) combining, which also decreases the fronthaul signaling.

In parallel, the majority of works in CF mMIMO systems have assumed ideal hardware while, in practice, hardware impairments (HWIs) exist and are inevitable [10], [11]. Moreover, given that an attractive implementation of CF mMIMO in terms of cost is a prerequisite, the large number of components involved, requires the use of cheap hardware to keep in affordable limits the total expenditure. Unfortunately, cheaper hardware is equivalent to more severe HWIs that should be taken into account during the analysis. Disregarding the importance of HWIs, the corresponding research for CF mMIMO systems is limited [12]–[15]. Especially, no work has investigated the effect of phase noise (PN).

This work covers the arising need for the assessment of the impact of HWIs on scalable CF (SCF) mMIMO systems when MMSE combining is applied. In fact, to the best of our knowledge, this is the first paper studying the effects of HWIs in SCF systems. Especially, this is the only work investigating the impact of PN even in canonical CF mMIMO systems. Specifically, we present a new insightful expression for the average achievable spectral efficiency (SE) with MMSE combining under the presence of HWIs. Compared to the existing literature, we pursue a deterministic equivalent (DE) analysis [16], which provides a closed-form analytical expression of the SE. Note that previous works with MMSE combining in CF mMIMO systems have presented only simulation results.

The remainder of this paper is organized as follows: Section II presents the system model of SCF mMIMO systems with a description of the HWIs. In Section III, the scalable estimation with HWIs is provided. Section IV describes the uplink data transmission accounting for the scalability of CF mMIMO systems while Section V presents the DE of the average SE with HWIs. In Section VI, the numerical results are provided, and Section VII concludes the paper.

II. System Model

We consider a CF network architecture including M APs which are randomly distributed over a geographic area and serve jointly K UEs in the same time-frequency resources with

M and K being large. Each AP is equipped with N antennas while the UEs have a single-antenna. The APs are connected to central processing units (CPUs) by means of backhaul links in an arbitrary way [5], [17].

We assume a time-varying narrowband channel, where the $N \times 1$ flat-fading channel vector between the mth AP and the kth UE is expressed by means of an independent correlated Rayleigh fading distribution as

$$\mathbf{h}_{mk} \sim \mathcal{CN}\left(\mathbf{0}, \mathbf{R}_{mk}\right), \quad (1)$$

where $\mathbf{R}_{mk} \in \mathbb{C}^{N \times N}$ is a deterministic Hermitian-symmetric positive semi-definite correlation matrix describing shadowing and geometric pathloss with large-scale fading coefficient $\beta_{mk} = \mathrm{tr}\left(\mathbf{R}_{mk}\right)/N$. Moreover, we employ the standard TDD protocol with $\tau_{\mathrm{c}} = B_{\mathrm{c}} T_{\mathrm{c}}$ being the length of each coherence block, where B_{c} in Hz and T_{c} in s are the coherence bandwidth the coherence time, respectively. Also, this block includes τ channel uses for an uplink training phase while $\tau_{\mathrm{u}} = \zeta\left(\tau_{\mathrm{c}} - \tau\right)$ channel uses and τ_{d} channel uses are considered for the uplink and downlink data transmission phases with $\zeta \leq 1$ describing the payload UL fraction transmission. It is assumed that $\mathbf{h}_{mk}$ is assumed fixed during the coherence time.

A. Main Scalability Guidelines

A large deployment of the original CF mMIMO systems with large K is unfeasible due to the high cost and complexity of the fronthaul. For this reason, [6] suggested that this architecture should be scalable with increasing K which can be achieved when the following conditions are met. Specifically, *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 should have finite complexity and resource requirements per AP when $K \to \infty$.

According to the framework in [6], even if the total number of UEs K increases indefinitely, each AP should serve a set of UEs with constant cardinality, defined by as

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

where $\mathbf{D}_{mi} \in \mathbb{C}^{N \times N}$ denote the diagonal matrices determining which APs and UEs communicate with each other based on the dynamic cooperation clustering (DCC) framework [18]. These matrices can establish a unified framework describing many architectures with original CF mMIMO systems being a special case.

B. Hardware Impairments

We consider a general model with the following HWIs.

1) Multiplicative distortion: It expresses time-dependent random phase-drifts known as PN, induced during the up-conversion of the baseband signal to passband and vice-versa by multiplying the signal with the LO's output. The total PN process from the kth user to the jth antenna of the mth AP at the nth channel use is $\theta_{mk,n}^{(j)} = \phi_{m,n}^{j} + \varphi_{k,n}$, where $\phi_{m,n}^{j}$ and $\varphi_{k,n}$ are discrete-time independent Wiener processes $\phi_{m,n}^{j} = \mathcal{CN}\left(\phi_{m,n-1}^{j}, \sigma_n^{\phi}\right)$ and $\varphi_{m,n}^{j} = \mathcal{CN}\left(\varphi_{m,n-1}^{j}, \sigma_n^{\varphi}\right)$ with $\sigma_i^2 = 4\pi^2 f_{\mathrm{c}}^2 c_i T_{\mathrm{s}}$ for $i = \phi, \varphi$ [19]. Note that T_{s}, c_i, and f_{c} are the symbol interval, a constant dependent on the oscillator, and the carrier frequency, respectively. Hence, we denote the total PN at the nth channel use $\mathbf{\Theta}_{mk,n} \triangleq \mathrm{diag}\Big\{e^{i\theta_{mk,n}^{(1)}}, \ldots, e^{i\theta_{mk,n}^{(N)}}\Big\}$, where we have separate LOs (SLOs) at each antenna justifying the independence among the PN processes [20]. In the special case of one common local oscillator (CLO) connected to all antennas of an AP, the PN matrix degenerates to $\Theta_{mk,n} \triangleq e^{i\theta_{mk,n}}\mathbf{I}_N$.

C. Additive distortion

Measurement results have shown that the transmit and receive signals are distorted during the transmission and reception processing, respectively [10], [11]. The additive distortion noises are Gaussian distributed as a result of the aggregate contribution of many impairments and modeled as [21]

$$\delta_{\mathrm{t},n}^{k} \sim \mathcal{CN}\left(0, \kappa_{\mathrm{t}}^2 p_k\right), \quad (3)$$

$$\boldsymbol{\delta}_{\mathrm{r},n}^{m} \sim \mathcal{CN}\Big(\mathbf{0}, \kappa_{\mathrm{r}_m}^2 \sum_{i=1}^{K} p_i \mathrm{diag}\left(|h_{mi}^{(1)}|^2, \ldots, |h_{mi}^{(N)}|^2\right)\Big), \quad (4)$$

with the proportionality parameters κ_{t}^2 and $\kappa_{\mathrm{r}_m}^2$ describing the severity of the residual impairments at the transmitter and the receiver side.

D. Amplified thermal noise

Certain components such as the low noise amplifiers result in an amplification of the thermal noise 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 [21][1].

III. Scalable Channel Estimation with Hardware Impairments

In CF mMIMO systems, TDD operation takes place, i.e., the APs estimate their channels during an uplink training phase by means of pilot symbols [1].

According to [6, Assumption 1], each AP servers at most one UE per pilot sequence which implies that $|\mathcal{D}_m| \leq \tau$ and

$$\mathbf{D}_{mk} = \begin{cases} \mathbf{I}_N & \text{if } k \in \mathcal{D}_m \\ \mathbf{0} & \text{if } k \notin \mathcal{D}_m \end{cases}. \quad (5)$$

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

Moreover, let $\mathcal{S} \subset \{1, \ldots, K\}$ be a subset of the UEs that transmit one of the τ-length mutually orthogonal training sequences, i.e., $\boldsymbol{\omega}_k \triangleq [\omega_{k,1}, \ldots, \omega_{k,\tau}] \in \mathbb{C}^{\tau \times 1}$ with $p_{\mathrm{p}} = |\omega_{k,n}|^2, \forall k, n$, while the sequences among different UEs are mutually orthogonal.

[1]For the sake of exposition, we have assumed that all HWIs are the same across APs and users, respectively.

Hence, during this phase, the $\mathbb{C}^{N\times 1}$ received signal by the mth AP at time n from a UE in $\mathcal{S}$, is given by

$$\mathbf{y}^{\mathrm{p}}_{m,n} = \sum_{i\in\mathcal{S}} \mathbf{\Theta}_{mi,n}\mathbf{h}_{mi}\left(\omega_{i,n}+\delta^{i}_{\mathrm{t},n}\right)+\boldsymbol{\delta}^{m}_{\mathrm{r},n}+\boldsymbol{\xi}_{m,n}, \quad (6)$$

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, $\delta^{i}_{\mathrm{t},n} \sim \mathcal{CN}\left(0,\kappa^2_{\mathrm{t}}p_{\mathrm{p}}\right)$ 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 $\mathbf{\Upsilon}^{m}_{n} = \kappa^2_{\mathrm{r}_m}\sum_{i=1}^{\mathcal{S}} p_i \mathrm{diag}\left(|h^{(1)}_{mi}|^2,\ldots,|h^{(N)}_{mi}|^2\right)$, and $\boldsymbol{\xi}_{m,n}$ is the amplified Gaussian thermal noise matrix at the mth AP [2]. Clearly, only the distortions of the UEs in $\mathcal{S}$ degrade the reception.

Proposition 1: The linear minimum mean-square error (LMMSE) estimate of the effective channel of UE k $\mathbf{h}_{mk,n} = \mathbf{\Theta}_{mk,n}\mathbf{h}_{mk}$ at any channel use $n\in\{1,\ldots,\tau_{\mathrm{c}}\}$ is given by

$$\begin{aligned}\hat{\mathbf{h}}_{mk,n} &= \mathbb{E}\left[\mathbf{h}_{mk,n}\boldsymbol{\psi}^{\mathsf{H}}_{m}\right]\left(\mathbb{E}\left[\boldsymbol{\psi}_m\boldsymbol{\psi}^{\mathsf{H}}_{m}\right]\right)^{-1}\boldsymbol{\psi}_m\\ &= \left(\boldsymbol{\omega}^{\mathsf{H}}_{k}\mathbf{\Lambda}_{k,n}\otimes\mathbf{R}_{mk}\right)\mathbf{Q}^{-1}_{m}\boldsymbol{\psi}_m, \quad (8)\end{aligned}$$

where

$$\boldsymbol{\psi}_m = \left[\mathbf{y}^{\mathsf{T}}_{m,1},\ldots,\mathbf{y}^{\mathsf{T}}_{m,\tau}\right]^{\mathsf{T}}, \quad (9)$$

$$\mathbf{\Lambda}_{k,n} \triangleq \mathrm{diag}\left\{e^{-\frac{\sigma^2_{\varphi}+\sigma^2_{\phi}}{2}|n-1|},\ldots,e^{-\frac{\sigma^2_{\varphi}+\sigma^2_{\phi}}{2}|n-\tau|}\right\}, \quad (10)$$

$$\mathbf{Q}_m \triangleq \sum_{j\in\mathcal{S}}\mathbf{X}_j\otimes\mathbf{R}_{mj}+\xi_m\mathbf{I}_{\tau N}, \quad (11)$$

$$\mathbf{X}_j \triangleq \tilde{\mathbf{X}}_j+\kappa^2_{\mathrm{r}_m}\mathbf{L}_{|\boldsymbol{\omega}_j|^2}, \quad (12)$$

$$\mathbf{L}_{|\boldsymbol{\omega}_j|^2} \triangleq \mathrm{diag}\left(|\omega_{j,1}|^2,\ldots,|\omega_{j,\tau}|^2\right), \quad (13)$$

$$\left[\tilde{\mathbf{X}}_j\right]_{u,v} \triangleq \left(1+\kappa^2_{\mathrm{t_{UE}}}\right)\omega_{j,u}\omega^{*}_{j,v}e^{-\frac{\sigma^2_{\varphi}+\sigma^2_{\phi}}{2}|u-v|}. \quad (14)$$

Proof: The proof follows the same steps with Theorem 1 in [21] but it is more general since the proposed model includes the transmit HWIs $\delta^{k}_{\mathrm{t},n}$ and requires some extra algebraic manipulations. Also, this estimated channel accounts for the scalability design since the summation in (11) is over the sum of participating UEs and not all UEs. The proof is omitted due to limited space. ■

Thus, the current channel at the nth channel use $(n\in\{1,\ldots,\tau_{\mathrm{c}}\})$ is given by

$$\mathbf{h}_{mk,n} = \hat{\mathbf{h}}_{mk,n}+\tilde{\mathbf{h}}_{mk,n}, \quad (15)$$

where $\hat{\mathbf{h}}_{mk,n}$ and $\tilde{\mathbf{h}}_{mk,n}$ have zero mean and variances $\mathbf{\Phi}_{mk} = \left(\boldsymbol{\omega}^{\mathsf{H}}_{k}\mathbf{\Lambda}_{k,n}\otimes\mathbf{R}_{mk}\right)\mathbf{Q}^{-1}_{m}\left(\mathbf{\Lambda}^{\mathsf{H}}_{k,n}\boldsymbol{\omega}_k\otimes\mathbf{R}_{mk}\right)$ and $\tilde{\mathbf{R}}_{mk} = \mathbf{R}_{mk}-\mathbf{\Phi}_{mk}$.

As a remark, note that $\hat{\mathbf{h}}_{mk,n}$ and $\mathbf{e}_{mk,n}$ are neither independent nor jointly complex Gaussian vectors but they are uncorrelated and each of them has zero mean because the effective distortion noises, e.g., $\mathbf{\Theta}_{mi,n}\mathbf{h}_{mi}\delta^{i}_{\mathrm{t},n}$ are not Gaussians since they appear as products between two Gaussian variables. The LMMSE estimator results in the optimal MMSE estimator when HWIs are neglected while the performance loss between the LMMSE and MMSE estimators is in general little [22].

[2]For the sake of a better presentation of the HWIs, the received signal at the mth AP at a given channel use n with no HWIs is expressed as

$$\mathbf{y}^{\mathrm{p}}_{m,n} = \sum_{i\in\mathcal{S}}\mathbf{h}_{mi}\omega_{i,n}+\mathbf{w}_{m,n}, \quad (7)$$

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

IV. Uplink Data Transmission

We consider a fully centralized processing architecture, where the CPU handles both the channel estimation and data detection. Hence, the M APs delegate their received signals $\{\mathbf{y}_{m,n}: m=1,\ldots,M\}$ to the CPU for detection. The received signal by the CPU at time n can be written in a compact form as

$$\mathbf{y}_n = \sum_{i=1}^{K}\mathbf{\Theta}_{i,n}\mathbf{h}_i\left(s_{i,n}+\delta^{i}_{\mathrm{t},n}\right)+\boldsymbol{\delta}_{\mathrm{r},n}+\boldsymbol{\xi}_n, \quad (16)$$

where $W=MN$, $s_{i,n}\in\mathbb{C}$ is the transmit signal from UE i with power p_i, $\mathbf{y}_n = \left[\mathbf{y}^{\mathsf{T}}_{1,n}\cdots\mathbf{y}^{\mathsf{T}}_{M,n}\right]^{\mathsf{T}}\in\mathbb{C}^{W}$ and $\boldsymbol{\xi}_n = \left[\boldsymbol{\xi}^{\mathsf{T}}_{1,n}\cdots\boldsymbol{\xi}^{\mathsf{T}}_{M,n}\right]^{\mathsf{T}}\in\mathbb{C}^{W}$ are block vectors while $\mathbf{h}_{i,n} = \left[\mathbf{h}^{\mathsf{T}}_{i1,n}\cdots\mathbf{h}^{\mathsf{T}}_{iM,n}\right]^{\mathsf{T}}\sim\mathcal{CN}\left(\mathbf{0},\mathbf{R}_i\right)$ is the concatenated channel vector from all APs with $\mathbf{R}_i = \mathrm{diag}\left(\mathbf{R}_{i1},\ldots,\mathbf{R}_{iM}\right)\in\mathbb{CN}^{W\times W}$ 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 $\mathbf{\Phi}_i = \mathrm{diag}\left(\mathbf{\Phi}_{i1},\ldots,\mathbf{\Phi}_{iM}\right)\in\mathbb{CN}^{W\times W}$ and $\mathbf{\Delta}_i = \mathbf{R}_i-\mathbf{\Phi}_i\in\mathbb{CN}^{W\times W}$, respectively. Moreover, regarding the HWIs, we have the phase noise block diagonal matrix $\mathbf{\Theta}_i = \mathrm{diag}\left(\mathbf{\Theta}_{i1},\ldots,\mathbf{\Theta}_{iM}\right)\in\mathbb{CN}^{W\times W}$, the transmit additive distortion $\delta^{i}_{\mathrm{t},n}$ from the ith UE, the receive additive distortion block vector $\boldsymbol{\delta}_{\mathrm{r},n} = \left[\boldsymbol{\delta}^{1\,\mathsf{T}}_{\mathrm{r},n}\cdots\boldsymbol{\delta}^{M\,\mathsf{T}}_{\mathrm{r},n}\right]^{\mathsf{T}}\in\mathbb{C}^{W}$. Note that $\boldsymbol{\xi}_n\sim\mathcal{CN}\left(\mathbf{0},\mathbf{F}_{\boldsymbol{\xi}_n}\right)$ with $\mathbf{F}_{\boldsymbol{\xi}_n} = \mathrm{diag}\left(\xi_{1,n}\mathbf{I}_N,\ldots,\xi_{M,n}\mathbf{I}_N\right)$.

According to the DCC framework [18], the CPU estimates $s_{k,n}$ by means of (16) as

$$\hat{s}_{k,n} = \sum_{m=1}^{M}\mathbf{v}^{\mathsf{H}}_{mk,n}\mathbf{D}_{mk}\mathbf{y}_m \quad (17)$$

with $\mathbf{D}_k = \mathrm{diag}\left(\mathbf{D}_{k1},\ldots,\mathbf{D}_{kM}\right)\in\mathbb{C}^{W\times W}$ being a block-diagonal matrix denoting a subset of the APs that contribute to signal detection.

A. Achievable SE

According to [23], we express the received signal in terms of the average effective channel $\mathbb{E}\left\{\mathbf{v}^{\mathsf{H}}_{k,n}\mathbf{D}_k\mathbf{\Theta}_{k,n}\mathbf{h}_k s_k\right\}$ as

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

Proposition 2: The achievable uplink SE for user k is given by

$$\mathrm{SE}_k = \frac{1}{\tau_\mathrm{c}} \sum_{n=1}^{\tau_\mathrm{c}-\tau_\mathrm{p}} \log_2\left(1+\gamma_{k,n}\right), \tag{18}$$

where $\gamma_{k,n} = \frac{S_{k,n}}{I_{k,n}}$ with $S_{k,n}$ and $I_{k,n}$ being the desired signal power and the interference plus noise power, respectively, which are defined as

$$S_{k,n} = p_k |\mathbb{E}\left\{\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{h}_{k,n}\right\}|^2 \tag{19}$$

$$I_{k,n} = \mathrm{Var}\left\{\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{h}_{k,n}\right\} + \sum_{i\neq k}^{K} p_i \mathbb{E}\left\{|\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{h}_{i,n}|^2\right\} + \sigma_\mathrm{t}^2 + \sigma_\mathrm{r}^2 + \xi_n \mathbb{E}\left\{\|\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k\|^2\right\}, \tag{20}$$

where $\sigma_\mathrm{t}^2 = \sum_{i=1}^{K} \kappa_\mathrm{t}^2 p_i \mathbb{E}\left\{|\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \mathbf{h}_{i,n}|^2\right\}$ and $\sigma_\mathrm{r}^2 = \mathbb{E}\left\{\mathbf{v}_{k,n}^{\mathsf{H}} \mathbf{D}_k \boldsymbol{\kappa}_\mathrm{r}^2 \left(\sum_{i=1}^{K} p_i \mathbf{F}_{|\mathbf{h}_i|^2}\right) \mathbf{D}_k \mathbf{v}_{k,n}\right\}$ with $\boldsymbol{\kappa}_\mathrm{r}^2 = \mathbf{I}_N \otimes \mathrm{diag}\left(\kappa_{\mathrm{r}_1}^2, \ldots, \kappa_{\mathrm{r}_M}^2\right)$ 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 A. ■

V. Deterministic Equivalent Analysis

The DE of the SINR $\tilde{\gamma}_{k,n}$, provided by Proposition 2, obeys to $\gamma_{k,n} - \bar{\gamma}_{k,n} \xrightarrow[M\to\infty]{\text{a.s.}} 0$, and the DE of the rate of user k, relied on the dominated convergence and the continuous mapping theorem [16], is given by

$$\mathrm{SE}_k - \overline{\mathrm{SE}}_k \xrightarrow[N\to\infty]{\text{a.s.}} 0 \tag{21}$$

where $\overline{\mathrm{SE}}_k = \frac{1}{T_\mathrm{c}} \sum_{n=1}^{T_\mathrm{c}-\tau} \log_2(1+\bar{\gamma}_{k,n})$. We write the scalable MMSE (SMMSE) decoder at time n as [6]

$$\mathbf{v}_{k,n}^{\mathrm{SMMSE}} = \boldsymbol{\Sigma} \mathbf{D}_k \hat{\mathbf{h}}_{k,n}, \tag{22}$$

where $\boldsymbol{\Sigma}^\dagger = \sum_{i\in\mathcal{P}_k} \mathbf{D}_k \hat{\mathbf{h}}_{i,n} \hat{\mathbf{h}}_{i,n}^{\mathsf{H}} \mathbf{D}_k + \bar{\mathbf{Z}}_k$ with $\bar{\mathbf{Z}}_k = \mathbf{D}_k \left(\sum_{i\in\mathcal{P}_k} \tilde{\mathbf{R}}_i + \frac{1}{\rho}\mathbf{I}_W\right) \mathbf{D}_k$, and a being a regularization scaled by W, in order to make expressions converge to a constant, as W, $K \to \infty$. The notation $(\cdot)^\dagger$ expresses the pseudo-inverse operator, and it is used instead of the inverse because the terms inside the parentheses may not be strictly positive definite [6]. Note that (22) includes 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 [6]

$$\mathcal{P}_k = \left\{i : \mathbf{D}_k \mathbf{D}_i \neq \mathbf{0}\right\}. \tag{23}$$

Theorem 1: The uplink DE of the SINR of user k at time n with MMSE decoding in SCF massive MIMO systems, accounting for imperfect CSI and HWIs, is given by (24) at the bottom of next page with $\delta_k = \frac{1}{W} \mathrm{tr}\, \mathbf{D}_k \boldsymbol{\Phi}_k \mathbf{D}_k \mathbf{T}$, $\tilde{\zeta}_{ki} = \frac{1}{W^2} \mathrm{tr}\left(\mathbf{D}_k\left(\mathbf{R}_k - \boldsymbol{\Phi}_k\right)\mathbf{D}_k \mathbf{T}_k^{'}\right)$, $e_i = \frac{1}{W} \mathrm{tr}\left(\mathbf{D}_k^2 \boldsymbol{\kappa}_\mathrm{r}^2 \mathbf{R}_i\right)$, $\zeta_{ki} = \frac{1}{W^2} \mathrm{tr}\left(\mathbf{D}_k \mathbf{R}_i \mathbf{D}_k \mathbf{T}_k^{'}\right)$, $\nu_{ki} = \frac{1}{W} \mathrm{tr}\left(\mathbf{D}_k \boldsymbol{\Phi}_i \mathbf{D}_k \mathbf{T}\right)$, $\mu_{ki} = \frac{1}{W^2} \mathrm{tr}\left(\mathbf{D}_k \boldsymbol{\Phi}_i \mathbf{D}_k \mathbf{T}_k^{'}\right)$, $\tilde{\mu}_{ki} = \zeta_{ki} + \frac{|\nu_{ki}|^2 \mu_{ki}}{(1+\delta_i)^2} - 2\mathrm{Re}\left\{\frac{\nu_{ki}^* \mu_{ki}}{(1+\delta_i)}\right\}$, $\eta_k^{'} = \frac{1}{W} \mathrm{tr}\, \mathbf{T}_k^{'}$, and $\tilde{e}_k^{'} = \frac{1}{W^2} \mathrm{tr}\left(\mathbf{T}_k^{''}\right)$, where

- $$\mathbf{T} = \left(\sum_{i\in\mathcal{P}_k} \mathbf{D}_k \left(\frac{1}{W(1+\delta_i)} \boldsymbol{\Phi}_i + \tilde{\mathbf{R}}_i + \alpha\xi \mathbf{I}_W\right) \mathbf{D}_k + \frac{1}{W} \tilde{\boldsymbol{\Delta}}\right)^\dagger,$$
- $\mathbf{T}_k^{'} = \mathbf{T}\mathbf{D}_k \boldsymbol{\Phi}_k \mathbf{D}_k \mathbf{T} + \sum_{i=1}^{K} \frac{\delta_i^{'} \mathbf{T}\mathbf{D}_k \boldsymbol{\Phi}_i \mathbf{D}_k \mathbf{T}}{W(1+\delta_i)^2}$,
- $\mathbf{T}_k^{''} = \mathbf{T}\mathbf{D}_k \mathbf{T} + \sum_{i=1}^{K} \frac{\delta_i \mathbf{T}\mathbf{D}_k \boldsymbol{\Phi}_i \mathbf{D}_k \mathbf{T}}{W(1+\delta_i)^2}$,
- $\boldsymbol{\delta}^{'} = \left(\mathbf{I}_K - \mathbf{F}\right)^{-1}$ with

$$[\mathbf{F}]_{k,i} = \frac{1}{W^2(1+\delta_i)} \mathrm{tr}\left(\mathbf{D}_k \boldsymbol{\Phi}_k \mathbf{D}_k \mathbf{T} \mathbf{D}_k \boldsymbol{\Phi}_i \mathbf{D}_k \mathbf{T}\right),$$
$$[\mathbf{f}]_k = \frac{1}{W} \mathrm{tr}\left(\mathbf{D}_k \boldsymbol{\Phi}_k \mathbf{D}_k \mathbf{T} \mathbf{D}_k \boldsymbol{\Phi}_k \mathbf{D}_k \mathbf{T}\right).$$

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

VI. Numerical Results

We consider the SE per UE, given by Theorem 1. In addition, we perform Monte-Carlo simulations by means of 10^4 independent channel realizations to show the tightness of the proposed analytical result[3].

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 [25, Table B.1.2.1-1] as a proper benchmark for CF massive MIMO systems[4]. In particular, according to this model assuming a 2 GHz carrier frequency, the large-scale fading coefficient is given by

$$\beta_{mk}[dB] = -30.5 - 36.7 \log_{10}\left(\frac{d_{mk}}{1\ \mathrm{m}}\right) + F_{mk}, \tag{25}$$

where d_{mk} is the distance between AP m and UE k and $F_{mk} \sim \mathcal{CN}\left(0, 4^2\right)$ 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

[3]We observe the coincidence between the analytical and simulation results even practical values of K and W, which agrees with literature and shows the usefulness of DEs [16], [24].

[4]In [5, 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 [1]. The main reason is that [1] 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 [1] 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.

$$\bar{\gamma}_{k,n}^{\mathrm{lo}} = \frac{\delta_k^2}{\kappa_{\mathrm{t}}^2 \delta_k^2 + \tilde{\zeta}_{ki} + \left(1+\kappa_{\mathrm{t}}^2\right)\sum_{i \neq k}^{K} \frac{\rho_i}{\rho_k}\tilde{\mu}_{ki} + \eta_k^{'} \sum_{i=1}^{K} \frac{p_i}{p_k} e_i + \frac{\tilde{e}_k^{'}}{p_k}} \tag{24}$$

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 $p_k = p_{\mathrm{p}} = 100$ mW while the thermal noise variance is $\sigma^2 = -174\mathrm{dBm/Hz}$.

Regarding the HWIs, we assume similar to [21], [26] 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 $i = \phi, \varphi$. Also, we have 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 [15]. For example, if $b = 2, 3, 4$, then $\kappa_{\mathrm{t}} = 0.258,\ 0.126, 0.062$. Moreover, we assume that $\xi_{m,n} = 1.6\sigma^2$ 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 $\xi_{m,n} = \frac{F\sigma^2}{1-2^{-2b}} = 1.6\sigma^2$.

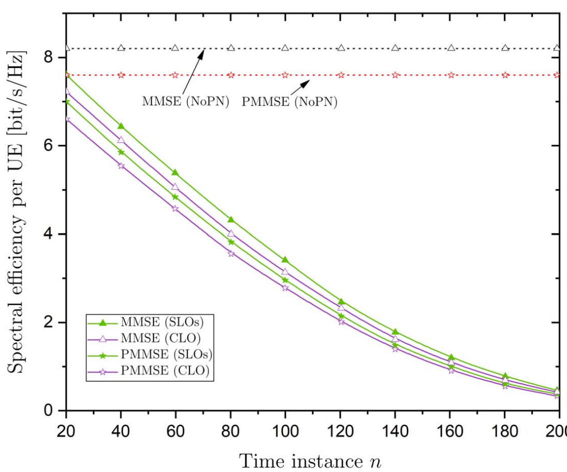

Fig. 1. Uplink SE per UE versus the time instance n of the data transmission phase for SCF mMIMO systems. Both scalable and conventional MMSE combiners are depicted.

Keeping the additive HWIs equal to zero, Fig. 1 illustrates the variation of the achievable SE per user versus the time instance of the data transmission phase. In particular, in Fig. 1, the SE decreases as the time increases since the aggregate detrimental contribution from PN becomes higher. Note that we depict the SEs of both MMSE and PMMSE receivers with and without PN. The latter are constant with respect to n since they do not depend on the PN being the only source of channel aging in this work. Moreover, the SLOs configuration outperforms the CLO design since the distortions from different LOs average out in the large system limit described by [21], [27]. The SEs for both MMSE and PMMSE receivers diminish with increasing n since

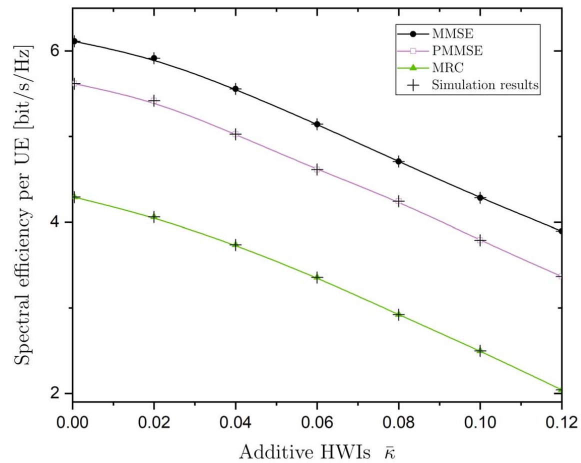

Fig. 2. Uplink SE per UE versus the additive HWIs $\bar{\kappa}$ for SCF mMIMO systems. Both scalable and conventional MMSE combiners are depicted together with simulation results.

the phase drift becomes higher and the corresponding impact of PN becomes quite destructive. Note that (18) is obtained by averaging over all time instances.

Fig. 2 provides the performance of the 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 PMMSE achieves just 8% less of the SE with MMSE receivers when $\bar{\kappa} = 0.06$. Also, we illustrate the outperformance of MMSE decoders against the maximum ratio combining (MRC) decoder as anticipated. The "cross" symbol represents the simulation results and it is shown the provided tightness of the DE analytical results.

VII. Conclusion

This paper presented a thorough investigation of the impact of HWIs on a new framework of CF mMIMO systems being scalable in large networks with many users. Specifically, we derived the uplink SE with HWIs and MMSE decoding in closed form by employing the theory of DEs. We depicted the quantification of the degradation due to HWIs on the SE by varying the quality of HWIs. The validation and tightness of the analytical expression were shown by means of simulations for practical system dimensions.

Appendix A
Proof of Proposition 2

First, we follow the approach in [20], in order to compute the SE for each n in the transmission phase, and then, we obtain the average over these SEs. The SE per UE in (18) 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 [28, Theorem 1], where the inter-user interference and the distortion noises are treated as independent Gaussian noises. Moreover, relied on [23], 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 worst-case Gaussian noise. Thus, we obtain the $\gamma_{k,n}$, 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 B
Proof of Theorem 1

First, we obtain the DE of the desired signal power given by (19). Specifically, we have

$$\begin{aligned}\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{k,n} &= \hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}\mathbf{D}_k\mathbf{h}_{k,n} &(26)\\ &= \frac{\frac{1}{W}\hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\left(\hat{\mathbf{h}}_{k,n}+\tilde{\mathbf{h}}_{k,n}\right)}{1+\frac{1}{W}\hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\hat{\mathbf{h}}_{k,n}} &(27)\\ &\asymp \frac{\frac{1}{W}\hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\hat{\mathbf{h}}_{k,n}}{1+\frac{1}{W}\hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\hat{\mathbf{h}}_{k,n}}, &(28)\end{aligned}$$

where in (26), we have replaced the expression of the MMSE decoder given by (22). In the next equation, we have applied the matrix inversion lemma. Note that $\boldsymbol{\Sigma}_k^{\dagger}$ is defined as

$$\begin{aligned}\boldsymbol{\Sigma}_k^{\dagger} &= \boldsymbol{\Sigma}^{\dagger}-\frac{p_k}{W}\mathbf{D}_k\hat{\mathbf{h}}_{k,n}\hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k &(29)\\ &= \mathbf{D}_k\Big(\sum_{\substack{i\in\mathcal{P}_k\\ i\neq k}}\frac{p_k}{W}\hat{\mathbf{h}}_{i,n}\hat{\mathbf{h}}_{i,n}^{\mathsf{H}}+\alpha\xi\mathbf{I}_W\Big)\mathbf{D}_k+\bar{\mathbf{Z}}_k.\end{aligned}$$

In the numerator of (28), we have applied [29, Lem. B.26]. The derivation continues with the use of [29, Lem. 14.3], [29, Lem. B.26], and [30, Theorem 1] as

$$\begin{aligned}\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{k,n} &\asymp \frac{\frac{1}{W}\operatorname{tr}\left(\mathbf{D}_k\boldsymbol{\Phi}_k\mathbf{D}_k\mathbf{T}\right)}{1+\frac{1}{W}\operatorname{tr}\left(\mathbf{D}_k\boldsymbol{\Phi}_k\mathbf{D}_k\mathbf{T}\right)} &(30)\\ &= \frac{\delta_k}{1+\delta_k}, &(31)\end{aligned}$$

where $\delta_k = \frac{1}{W}\operatorname{tr}\boldsymbol{\Phi}_k\mathbf{T}$. 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 [31]. 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 $\left\|\frac{1}{W}\operatorname{diag}(\hat{\mathbf{h}}_i\hat{\mathbf{h}}_i^H)-\frac{1}{W}\operatorname{tr}\left(\operatorname{diag}\left(\hat{\boldsymbol{\Phi}}_i\right)\right)\right\| \xrightarrow[W\to\infty]{\text{a.s.}} 0$. Note that in (31), we have made use of the commuting property among diagonal matrices inside a trace and [32, p. 207] The term, concerning the deviation from the average effective channel gain, becomes

$$\begin{aligned}\operatorname{Var}\left\{\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{k,n}\right\} &\asymp \frac{\frac{1}{W}\mathbb{E}\left\{|\hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\mathbf{h}_{k,n}|^2\right\}}{1+\delta_k} &(32)\\ &\asymp \frac{\mathbb{E}\left\{\frac{1}{W^2}\operatorname{tr}\left(\mathbf{D}_k\boldsymbol{\Phi}_k\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\left(\mathbf{R}_k-\boldsymbol{\Phi}_k\right)\mathbf{D}_k\boldsymbol{\Sigma}_k\right)\right\}}{1+\delta_k} &(33)\\ &\asymp \frac{\mathbb{E}\left\{\frac{1}{W^2}\operatorname{tr}\left(\mathbf{D}_k\left(\mathbf{R}_k-\boldsymbol{\Phi}_k\right)\mathbf{D}_k\mathbf{T}_k^{'}\right)\right\}}{1+\delta_k} &(34)\end{aligned}$$

where in (32), we have used the matrix inversion lemma, [29, Lem. B.26], and [30, Theorem 1]. In (33), we have applied the rank-1 perturbation lemma, [29, Lem. B.26], and [29, Lem. B.26] again. The last step includes application of [33, Theorem 2]. Based on (20), the interference power of the kth UE is

$$\begin{aligned}|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}|^2 &\asymp \left|\frac{\frac{1}{W}\hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\mathbf{h}_{i,n}}{1+\delta_k}\right|^2 &(35)\\ &\asymp \frac{\hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\mathbf{h}_{i,n}\mathbf{h}_{i,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\hat{\mathbf{h}}_{k,n}}{W^2\left(1+\delta_k\right)^2} &(36)\\ &\asymp \frac{\mathbf{h}_{i,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\boldsymbol{\Phi}_k\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k\mathbf{h}_{i,n}}{W^2\left(1+\delta_k\right)^2} &(37)\\ &\asymp \frac{1}{\left(1+\delta_k\right)^2}\Bigg(\frac{1}{W^2}\mathbf{h}_{i,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_{ki}\mathbf{D}_k\boldsymbol{\Phi}_k\mathbf{D}_k\boldsymbol{\Sigma}_{ki}\mathbf{D}_k\mathbf{h}_{i,n}\\ &\quad+\frac{|\mathbf{h}_{i,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_{ki}\mathbf{D}_k\hat{\mathbf{h}}_{i,n}|^2\hat{\mathbf{h}}_{i,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_{ki}\mathbf{D}_k\boldsymbol{\Phi}_k\mathbf{D}_k\boldsymbol{\Sigma}_{ki}\mathbf{D}_k\hat{\mathbf{h}}_{i,n}}{W^4\left(1+\delta_i\right)^2}\\ &\quad-2\mathrm{Re}\Bigg\{\frac{\left(\hat{\mathbf{h}}_{i,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_{ki}\mathbf{D}_k\mathbf{h}_{i,n}\right)}{W^3\left(1+\delta_i\right)^2}\\ &\quad\times\left(\mathbf{h}_{i,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_{ki}\mathbf{D}_k\boldsymbol{\Phi}_k\mathbf{D}_k\boldsymbol{\Sigma}_{ki}\mathbf{D}_k\hat{\mathbf{h}}_{i,n}\right)\Bigg\}\Bigg) &(38)\\ &\asymp \frac{1}{\left(1+\delta_k\right)^2}\left(\zeta_{ki}+\frac{|\nu_{ki}|^2\mu_{ki}}{\left(1+\delta_i\right)^2}-2\mathrm{Re}\left\{\frac{\nu_{ki}^*\mu_{ki}}{\left(1+\delta_i\right)}\right\}\right), &(39)\end{aligned}$$

where we have applied the matrix inversion lemma in (35), while in (36) and (37), we have used [29, Lem. B.26]. In (38), we have applied again the matrix inversion lemma, and in the last step, we have used the rank-1 perturbation lemma, [29, Lem. B.26], [30, Theorem 1], and [33, Theorem 2]. The definitions of the various parameters are given in the presentation of the theorem. Also, we have

$$\begin{aligned}\sigma_{\mathrm{t}}^2 &= \sum_{i=1}^{K}\kappa_{\mathrm{t}}^2\rho_i|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}|^2\\ &= \kappa_{\mathrm{t}}^2\left(\rho_k\mathbb{E}\left\{|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{k,n}|^2\right\}+\sum_{i\neq k}^{K}\rho_i\mathbb{E}\left\{|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\mathbf{h}_{i,n}|^2\right\}\right)\\ &\asymp \frac{\kappa_{\mathrm{t}}^2}{(1+\delta_k)^2}\left(\rho_k\delta_k^2+\sum_{i\neq k}^{K}\rho_i\left(\zeta_{ki}+\frac{|\nu_{ki}|^2\mu_{ki}}{(1+\delta_i)^2}-2\mathrm{Re}\left\{\frac{\nu_{ki}^*\mu_{ki}}{(1+\delta_i)}\right\}\right)\right), &(40)\end{aligned}$$

which is obtained by substituting (31) and 39. Also, the deterministic σ_{r}^2 as $W \to \infty$ becomes

$$\sigma_{\mathrm{r}}^2 = \mathbb{E}\left\{\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\kappa}_{\mathrm{r}}^2(\mathbf{I}_W \circ \mathbf{H}_n\mathbf{P}\mathbf{H}_n^{\mathsf{H}})\,\mathbf{D}_k\mathbf{v}_{k,n}\right\} \tag{41}$$

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

$$\asymp \frac{\eta_k^{'}}{W\left(1+\delta_k\right)^2}\sum_{i=1}^{K}\rho_i\,\mathrm{tr}\left(\mathbf{D}_k^2\boldsymbol{\kappa}_{\mathrm{r}}^2\mathbf{R}_i\right), \tag{42}$$

where in (41) 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 [29, Lem. B.26] [30, Theorem 1], [33, Theorem 2], and [32, p. 207] while we have set $\eta_k^{'} = \frac{1}{W}\,\mathrm{tr}\,\mathbf{T}_k^{'}$.
The DE of the last term of (20), corresponding to the amplified thermal noise contribution, becomes

$$\|\mathbf{v}_{k,n}^{\mathsf{H}}\mathbf{D}_k\|^2 \asymp \frac{\frac{1}{W^2}\hat{\mathbf{h}}_{k,n}^{\mathsf{H}}\mathbf{D}_k\boldsymbol{\Sigma}_k\mathbf{D}_k^2\boldsymbol{\Sigma}_k\mathbf{D}_k\hat{\mathbf{h}}_{k,n}}{\left(1+\delta_k\right)^2} \tag{43}$$

$$\asymp \frac{\frac{1}{W^2}\,\mathrm{tr}\left(\boldsymbol{\Sigma}_k\mathbf{D}_k^2\boldsymbol{\Sigma}_k\mathbf{D}_k\boldsymbol{\Phi}_k\mathbf{D}_k\right)}{\left(1+\delta_k\right)^2} \tag{44}$$

$$= \frac{\frac{1}{W^2}\,\mathrm{tr}\left(\mathbf{T}_k^{''}\right)}{\left(1+\delta_k\right)^2}, \tag{45}$$

where in (43) and (44), we have applied the matrix inversion lemma and [29, Lem. B.26], respectively. In the last step, we have used the rank-1 perturbation lemma as well as [33, Theorem 2] and [30, Theorem 1].
By substituting (31), (34), (39), (40), (42), and (45) into (19) and (20), the DE SINR is derived and the proof is concluded.

REFERENCES

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

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

[3] 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.

[4] Z. Chen and E. Björnson, "Channel hardening and favorable propagation in cell-free massive MIMO with stochastic geometry," *IEEE Trans. Commun.*, vol. 66, no. 11, pp. 5205–5219, 2018.

[5] E. Björnson and L. Sanguinetti, "Making cell-free massive MIMO competitive with MMSE processing and centralized implementation," *arXiv preprint arXiv:1903.10611*, 2019.

[6] ——, "Scalable cell-free massive MIMO systems," *arXiv preprint arXiv:1908.03119*, 2019.

[7] 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.

[8] A. Papazafeiropoulos *et al.*, "Performance analysis of cell-free massive MIMO systems: A stochastic geometry approach," *IEEE Trans. Veh. Tech.*, vol. 69, no. 4, pp. 3523–3537, 2020.

[9] ——, "Optimal energy efficiency in cell-free massive MIMO systems: A stochastic geometry approach," in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)*, 2020, pp. 1–7.

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

[11] 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.

[12] 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.

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

[14] 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.

[15] 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.

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

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

[18] E. Bjornson *et al.*, "Optimality properties, distributed strategies, and measurement-based evaluation of coordinated multicell OFDMA transmission," *IEEE Transactions on Signal Processing*, vol. 59, no. 12, pp. 6086–6101, 2011.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 3GPP, "Further advancements for E-UTRA physical layer aspects (Release 9)," 3GPP TS 36.814, Tech. Rep., 2017.

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

[27] 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.

[28] 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.

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

[30] 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.

[31] 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.

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

[33] 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.