Uplink/Downlink Duality in Multi-Cell MU-MIMO Systems with Hardware Impairments

Shahram Zarei, Wolfgang Gerstacker, and Robert Schober
Friedrich-Alexander Universität Erlangen-Nürnberg, Erlangen, Germany
Email: {shahram.zarei, wolfgang.gerstacker, robert.schober}@fau.de

***Abstract*—In this paper, we propose an uplink/downlink duality framework for multi-cell multi-user multiple-input multiple-output (MU-MIMO) systems with residual hardware impairments (HWIs) at the base station and the user terminals. By employing the proposed uplink/downlink duality framework, complex downlink optimization problems can be transformed to equivalent dual uplink problems which are easier to solve. Thereby, for the same total transmit power, the same per-user signal-to-interference-plus-noise ratios (SINRs) are achieved in downlink and uplink. We apply the proposed uplink/downlink duality, and derive a multi-cell HWI aware minimum mean square error (MCHWA-MMSE) precoder. Our simulation results show that the proposed MCHWA-MMSE precoder achieves a substantially higher sum rate than conventional MMSE and conjugate beamforming precoders.**

I. Introduction

With the increasing demand for higher data rates, multiple-input multiple-output (MIMO) systems have attracted much attention over the last decade. Today, MIMO technology is a key element of many modern wireless communication standards including Long Term Evolution (LTE) and worldwide interoperability for microwave access (WiMAX).

An emerging research field in wireless communications are so-called massive MIMO systems [1]. Massive MIMO systems employ a large number of antennas, e.g., one hundred or more at the base station (BS), and achieve very high spectral and energy efficiencies [2]. Moreover, in massive MIMO systems, the transmit power of the BS and the user terminals (UTs) can be decreased by increasing the number of antennas at the BS [2]. These and other desirable features render massive MIMO a promising technology for future wireless communication systems.

In this paper, we consider the downlink of a massive multi-cell multi-user MIMO (MU-MIMO) system with residual hardware impairments (HWIs) at the BSs and at the UTs. A simpler form of the considered system, i.e., the downlink single-cell MU-MIMO system with ideal hardware embodies a vector Gaussian broadcast channel (GBC) whose capacity region can be achieved by dirty paper coding (DPC). The capacity region of the vector GBC was derived by exploiting the concept of uplink/downlink (UL/DL) duality which was introduced in [3]. More generally, the UL/DL duality can be exploited to transform difficult downlink optimization problems into simpler dual uplink problems [4]. After solving the simpler uplink problem, the precoder and power allocation for the downlink can be calculated from the uplink detection matrices and power allocation such that the same per-user signal-to-interference-plus-noise ratios (SINRs) are achieved in uplink and downlink. The SINR UL/DL duality framework in [3], [4] has been extended to mean square error (MSE) UL/DL duality in [5], where the authors show that by employing the precoding and detection matrices appropriately and under the same total power constraint, the same MSE region in the downlink can be achieved as in the dual uplink. In [6], the authors show that in a multi-cell MU-MIMO system, the Lagrangian dual problem of the weighted transmit power minimization problem with SINR constraints can be rewritten as an equivalent dual uplink problem, which can be solved more easily than the original downlink problem.

The system model considered in [3]–[6] assumes ideal hardware (H/W) components and the only impairment is the additive white Gaussian noise (AWGN). However, in practice, the AWGN-based system model may be overly optimistic, since it does not take HWIs into account, which exist in all physical implementations. Recently, a significant amount of research has been dedicated to the study of the impact of residual HWIs (i.e., HWIs which remain after applying appropriate compensation measures) on the performance of MU-MIMO systems. In particular, it has been shown that residual HWIs can be modeled by an additive Gaussian impairment, whose variance depends on the useful signal power [7]–[10]. This model has also been experimentally validated, c.f. [8], [11]. One of the earliest works, which adopts this new system model to investigate the performance of massive MIMO systems with residual HWIs is [12]. Here, the authors provide capacity bounds for the downlink and uplink of massive MIMO systems with residual HWIs. Another related work is [13], where the authors derive analytical expressions for the asymptotic achievable sum rate of matched filter (MF) and conventional minimum mean square error (MMSE) detectors in uplink massive MIMO systems with residual HWIs, and also present an HWI aware MMSE detector. Moreover, in our recent work [14], we have proposed an UL/DL duality framework for single-cell MU-MIMO systems with residual HWIs, and derived a HWI aware precoder. In [14], using results from random matrix theory, we have also provided analytical expressions for the asymptotic downlink power allocation in the large system limit which only depend on the channel statistics.

In this paper, we derive an UL/DL duality framework for multi-cell MU-MIMO systems with residual HWIs at the BSs and the UTs. We extend our proposed SINR UL/DL duality to MSE UL/DL duality. Using the derived UL/DL duality, difficult downlink optimization problems can be transformed into simpler uplink problems. As an example, we use the proposed MSE UL/DL duality theorem to derive a multi-cell HWI aware MMSE (MCHA-MMSE) precoder and the corresponding power allocation in the downlink based on the respective uplink counterparts such that the total sum-MSE is minimized, while achieving the same per-user SINR and per-user MSE in downlink and in uplink for the same total transmit power. Our simulation results show that the proposed MCHA-

MMSE precoder achieves substantially higher sum rates than the conventional MMSE and conjugate beamforming (BF) precoders, which do not take multi-cell interference and HWIs into account.

The remainder of this paper is organized as follows. In Section II, the system model is presented and benchmark schemes are introduced. We develop the proposed UL/DL duality framework for multi-cell MU-MIMO systems with residual HWIs in Section III. In Section IV, the MCHA-MMSE precoder is derived, and in Section V, numerical results are provided. Finally, conclusions are drawn in Section VI.

Notation: Boldface lower and upper case letters represent column vectors and matrices, respectively. $\mathbf{I}_K$ denotes the $K \times K$ identity matrix and $[\mathbf{A}]_{k,:}$, $[\mathbf{A}]_{:,l}$, and $[\mathbf{A}]_{k,l}$ stand for the kth row, the lth column, and the element in the kth row and the lth column of matrix $\mathbf{A}$, respectively. $(\cdot)^*$ denotes the complex conjugate and $\mathrm{tr}(\cdot)$, $(\cdot)^{\mathsf{T}}$, and $(\cdot)^{\mathsf{H}}$ represent the trace, transpose, and Hermitian transpose of a matrix, respectively. $\mathbb{E}\{\cdot\}$ stands for the expectation operator and $\mathcal{CN}(\mathbf{u}, \mathbf{\Phi})$ denotes the circular symmetric complex Gaussian distribution with mean vector $\mathbf{u}$ and covariance matrix $\mathbf{\Phi}$. Moreover, $\mathbf{A} \circ \mathbf{B}$ represents the element-wise product of matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathrm{diag}(a_1, \ldots, a_K)$ denotes a diagonal matrix with $a_1, \ldots, a_K$ on its main diagonal, and "a.s." stands for "almost surely".

II. System Model and Benchmark Schemes

In this section, the considered multi-cell MU-MIMO system model with residual HWIs is introduced and two benchmark schemes are presented.

A. System Model

A downlink multi-cell MU-MIMO system with universal frequency reuse is considered. In our system model, there are L cells and in each cell, one BS with N antennas simultaneously serves K single-antenna UTs. Here, N is assumed to be very large and the ratio of the number of UTs to the number of BS antennas is denoted by $\beta = K/N$. The independent and identically distributed (i.i.d.) zero-mean complex Gaussian data symbols intended for the transmission to the UTs in the jth cell are stacked into vector $\mathbf{d}_j = [d_{j1}, \ldots, d_{jK}]^{\mathsf{T}}$, $\mathbb{E}\left\{\mathbf{d}_j \mathbf{d}_j^{\mathsf{H}}\right\} = \mathbf{I}_K$, where d_{jk} is the data symbol of the kth UT in the jth cell. The vector of the stacked received data symbols of the UTs in the jth cell is given by

$$\hat{\mathbf{d}}_j^{\mathrm{DL}} = \sum_{l=1}^{L} \mathbf{P}_j^{-1/2} \mathbf{\Xi}_j \mathbf{G}_{lj}^{\mathsf{H}} \left(\mathbf{V}_l \mathbf{P}_l^{1/2} \mathbf{d}_l + \boldsymbol{\epsilon}_l\right) + \mathbf{P}_j^{-1/2} \mathbf{\Xi}_j \left(\boldsymbol{\mu}_j + \mathbf{z}_j\right), \tag{1}$$

where $\mathbf{G}_{lj} = \left[\mathbf{g}_{lj1} \ldots \mathbf{g}_{ljK}\right] \in \mathbb{C}^{N \times K}$ and $\mathbf{V}_l = [\mathbf{v}_{l1} \ldots \mathbf{v}_{lK}] \in \mathbb{C}^{N \times K}$ with $\mathbf{g}_{ljk}$ and $\mathbf{v}_{lk}$ being the channel vector between the kth UT in the jth cell and the lth BS and the unit-norm precoding vector for the kth UT at the lth BS, respectively. In this work, we assume a block flat fading channel. We further assume a correlated channel model, i.e., $\mathbf{g}_{ljk} = \tilde{\mathbf{R}}_{ljk} \mathbf{h}_{ljk}$, where $\mathbf{h}_{ljk} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$, and $\mathbf{R}_{ljk} = \mathbb{E}\{\mathbf{g}_{ljk} \mathbf{g}_{ljk}^{\mathsf{H}}\} = \tilde{\mathbf{R}}_{ljk} \tilde{\mathbf{R}}_{ljk}^{\mathsf{H}}$ represents the channel covariance matrix of the kth UT in the jth cell. Furthermore, we adopt the channel correlation model used in [15], [16], i.e., $\tilde{\mathbf{R}}_{ljk} = \alpha_{ljk} \mathbf{A}$, where α_{ljk} models the path loss between the kth UT in the jth cell and the lth BS, and $\mathbf{A}$ models the BS antenna correlation. Here, α_{ljk} is assumed to be equal to one for $j = l$ (direct gain), and η for $j \neq l$ (cross gain). In this paper, we assume that the BS employs a uniform linear array (ULA), and adopt the ULA channel correlation model used in [15], [16]. Accordingly, we have $\mathbf{A} = [\mathbf{B} \; \mathbf{0}_{N \times N-M}]$ with $\mathbf{0}_{N \times (N-M)}$ and M being an $N \times (N-M)$ all-zero matrix and the number of dimensions of the antenna's physical model, respectively. Correspondingly, we adopt $\mathbf{B} = [\mathbf{b}(\phi_1), \ldots, \mathbf{b}(\phi_M)]$, where the steering vector $\mathbf{b}(\phi_m)$ is defined as [15], [16]

$$\mathbf{b}(\phi_m) = \frac{1}{\sqrt{M}} \left[1, \ldots, e^{-2\pi i \lambda (N-1) \sin(\phi_m)}\right]^{\mathsf{T}}, m \in \{1, \ldots, M\}, \tag{2}$$

where $\phi_m = -\pi/2 + (m-1)\pi/M$ is the mth angle of arrival (AoA), and λ is the antenna spacing in wavelength, respectively. Moreover, in (1), $\mathbf{P}_l = \mathrm{diag}(p_{l1}, \ldots, p_{lK})$ represents the downlink power allocation matrix, where p_{lk} is the power allocated to the kth UT in the lth cell. Here, we consider a sum transmit power constraint at each BS, i.e., $\mathrm{tr}(\mathbf{P}_l) \leq K\rho_{\mathrm{DL}}$, $\forall l \in \{1, \ldots, L\}$, where we define ρ_{DL} to be the SNR in the downlink. In addition, $\mathbf{\Xi}_j = \mathrm{diag}(\xi_{j1}, \ldots, \xi_{jK})$, where ξ_{jk} is used to optimize the MSE of the kth UT in the jth cell in the downlink, wich we denote by $\mathrm{MSE}_{jk}^{\mathrm{DL}}$. As mentioned in Section I, in this paper, we adopt the HWI model from [12], where the residual HWI at each antenna branch is modeled as a mutually uncorrelated Gaussian random variable, whose variance is proportional to the average signal power at that antenna. Thus, in (1), the stacked vector of HWIs in the transmit chain of the lth BS is modelled by $\boldsymbol{\epsilon}_l \sim \mathcal{CN}\left(\mathbf{0}, \kappa_{\mathrm{BT}}^2 \mathbf{I}_N \circ \left(\mathbf{V}_l \mathbf{P}_l \mathbf{V}_l^{\mathsf{H}}\right)\right)$, where κ_{BT} is a parameter, which reflects the amount of residual HWI at the transmitter chain of the BSs. Similarly, the residual HWI at each UT can be modeled by an independent Gaussian random variable, whose variance is proportional to the average received power [12]. Hence, in (1), the stacked vector of residual HWIs at all UTs in the jth cell is modeled by $\boldsymbol{\mu}_j = [\mu_{j1}, \ldots, \mu_{jK}]^{\mathsf{T}}$, where $\mu_{jk} \sim \mathcal{CN}\left(0, \sum_{l=1}^{L} \sum_{q=1}^{K} \kappa_{\mathrm{UR}}^2 p_{lq} \left|\mathbf{g}_{ljk}^{\mathsf{H}} \mathbf{v}_{lq}\right|^2\right)$ is the residual HWI in the receiver chain of the kth UT in the jth cell, where κ_{UR}^2 represents the amount of residual HWI at the receiver chain of the UTs. Furthermore, $\mathbf{z}_j \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ represents the stacked vector of the additive white Gaussian noise (AWGN) at the UTs in the jth cell. Here, we assume without loss of generality that the residual HWI parameter at the receiver and transmitter chains of all UTs and BSs are identical, respectively. According to (1), the received symbol at the kth UT in the jth cell is given by

$$\hat{d}_{jk}^{\mathrm{DL}} = \xi_{jk} \mathbf{g}_{jjk}^{\mathsf{H}} \mathbf{v}_{jk} d_{jk} + \sum_{\substack{l=1 \\ (l,q) \neq (j,k)}}^{L} \sum_{q=1}^{K} \xi_{jk} \sqrt{\frac{p_{lq}}{p_{jk}}} \mathbf{g}_{ljk}^{\mathsf{H}} \mathbf{v}_{lq} d_{lq} + \frac{\xi_{jk}}{\sqrt{p_{jk}}} \left(\sum_{l=1}^{L} \mathbf{g}_{ljk}^{\mathsf{H}} \boldsymbol{\epsilon}_l + \mu_{jk} + z_{jk}\right). \tag{3}$$

Hence, the signal-to-interference-plus-noise ratio (SINR) at the kth UT in the jth cell in the downlink is defined as

$$\mathrm{SINR}_{jk}^{\mathrm{DL}} \triangleq \frac{p_{jk}\left|\mathbf{g}_{jjk}^{\mathsf{H}}\mathbf{v}_{jk}\right|^2}{\sum\limits_{\substack{l=1\\(l,q)\neq(j,k)}}^{L}\sum\limits_{q=1}^{K} p_{lq}\left|\mathbf{g}_{ljk}^{\mathsf{H}}\mathbf{v}_{lq}\right|^2+\sigma_{\epsilon_{jk}}^2+\sigma_{\mu_{jk}}^2+1}, \quad (4)$$

where $\sigma_{\epsilon_{jk}}^2$ and $\sigma_{\mu_{jk}}^2$ are given by

$$\sigma_{\epsilon_{jk}}^2 = \sum_{l=1}^{L}\kappa_{\mathrm{BT}}^2\mathbf{g}_{ljk}^{\mathsf{H}}\left(\mathbf{I}_N\circ\left(\mathbf{V}_l\mathbf{P}_l\mathbf{V}_l^{\mathsf{H}}\right)\right)\mathbf{g}_{ljk}, \quad (5)$$

$$\sigma_{\mu_{jk}}^2 = \sum_{l=1}^{L}\sum_{q=1}^{K}\kappa_{\mathrm{UR}}^2 p_{lq}\left|\mathbf{g}_{ljk}^{\mathsf{H}}\mathbf{v}_{lq}\right|^2. \quad (6)$$

B. Benchmark Schemes

BF and conventional MMSE precoders are the most commonly used linear precoders for downlink massive MIMO systems [1], [16], [17]. Thus, in this paper, we consider BF and MMSE precoders as benchmark schemes, and compare their performance with that of the proposed MCHA-MMSE precoder. The BF and MMSE precoders at the jth BS are given by, respectively,

$$\mathbf{V}_j^{\mathrm{BF}} = \zeta_j^{\mathrm{BF}}\mathbf{G}_{jj}, \quad (7)$$

$$\mathbf{V}_j^{\mathrm{MMSE}} = \zeta_j^{\mathrm{MMSE}}\left(\mathbf{G}_{jj}\mathbf{G}_{jj}^{\mathsf{H}}+\frac{1}{\rho_{\mathrm{DL}}}\mathbf{I}_N\right)^{-1}\mathbf{G}_{jj}, \quad (8)$$

where ζ_j^{BF} and ζ_j^{MMSE} are normalization factors, which ensure that the total transmit power constraints $\mathrm{tr}\left(\mathbf{V}_j^{\mathrm{BF}}\left(\mathbf{V}_j^{\mathrm{BF}}\right)^{\mathsf{H}}\right)=K$ and $\mathrm{tr}\left(\mathbf{V}_j^{\mathrm{MMSE}}\left(\mathbf{V}_j^{\mathrm{MMSE}}\right)^{\mathsf{H}}\right)=K$ are met. As can be observed from (7) and (8), the BF and MMSE precoders take neither the multi-cell interference nor the residual HWIs into account which leads to performance degradation.

III. UL/DL Duality in Multi-Cell MU-MIMO Systems with Residual HWIs

In this section, we propose an UL/DL duality framework for multi-cell MU-MIMO systems with residual HWIs. First, in Theorem 1, an SINR UL/DL duality framework is presented. Then, in Corollary 2, we extend the SINR duality presented in Theorem 1 to MSE duality.

Theorem 1: For the downlink multi-cell MU-MIMO system defined in Section II, an equivalent dual uplink system exists, whose channel and detection matrices are given by $\mathbf{G}_{jl}, \forall j, l\in\{1,\ldots,L\}$ and $\tilde{\mathbf{U}}_j=\mathbf{\Xi}_j\mathbf{V}_j^{\mathsf{H}}$, respectively, where matrices $\mathbf{V}_j,\forall j$, have unit norm columns and diagonal matrix $\mathbf{\Xi}_j$ contains the norms of the rows of matrix $\tilde{\mathbf{U}}_j$. In particular, the stacked vector of the detected symbols in the jth cell in the dual uplink system is given by

$$\begin{aligned}\hat{\mathbf{d}}_j^{\mathrm{UL}} =&\check{\mathbf{P}}_j^{-1/2}\mathbf{\Xi}_j\mathbf{V}_j^{\mathsf{H}}\sum_{l=1}^{L}\left(\mathbf{G}_{jl}\left(\check{\mathbf{P}}_l^{1/2}\mathbf{d}_l+\check{\boldsymbol{\mu}}_l\right)\right)\\ &+\check{\mathbf{P}}_j^{-1/2}\mathbf{\Xi}_j\mathbf{V}_j^{\mathsf{H}}\left(\check{\boldsymbol{\epsilon}}_j+\mathbf{z}_j\right),\end{aligned} \quad (9)$$

where $\check{\mathbf{P}}_l=\mathrm{diag}\left(\check{p}_{l1},\ldots,\check{p}_{lK}\right)$ represents the UTs' transmit powers in the lth cell in the uplink with $\check{p}_{lk}$ being the transmit power of the kth UT in the lth cell. Moreover, $\check{\boldsymbol{\epsilon}}_j\sim\mathcal{CN}\left(\mathbf{0},\sum_{l=1}^{L}\kappa_{\mathrm{BT}}^2\mathbf{I}_N\circ\left(\mathbf{G}_{jl}\check{\mathbf{P}}_l\ \mathbf{G}_{jl}^{\mathsf{H}}\right)\right)$ and $\check{\boldsymbol{\mu}}_l=\left[\check{\mu}_{l1},\ldots,\check{\mu}_{lK}\right]^{\mathsf{T}}$, where $\check{\mu}_{lk}\sim\mathcal{CN}\left(0,\kappa_{\mathrm{UR}}^2\check{p}_{lk}\right)$, represent the dual UL equivalents of residual HWIs at the BSs and UTs, respectively. If the stacked power allocation vector of all UTs in all cells in the downlink, i.e., $\mathbf{p}=\left[\mathbf{p}_1^{\mathsf{T}}\ldots\mathbf{p}_L^{\mathsf{T}}\right]^{\mathsf{T}}$, where $\mathbf{p}_j=\left[p_{j1},\ldots,p_{jK}\right]^{\mathsf{T}}, j\in\{1,\ldots,L\}$, is chosen as

$$\mathbf{p}=\left(\mathbf{I}_{KL}-\mathrm{diag}\left(\mathbf{a}\right)\cdot\mathbf{A}^{\mathsf{T}}\right)^{-1}\mathbf{a}, \quad (10)$$

identical SINRs in downlink and in uplink can be achieved, i.e., $\mathrm{SINR}_{jk}^{\mathrm{DL}}=\mathrm{SINR}_{jk}^{\mathrm{UL}},\forall j,k$, for the same sum powers, i.e., $\sum_{j=1}^{L}\sum_{k=1}^{K}p_{jk}=\sum_{j=1}^{L}\sum_{k=1}^{K}\check{p}_{jk}$. Here, the elements of vector $\mathbf{a}=\left[a_1,\ldots,a_{KL}\right]^{\mathsf{T}}$ in (10) are defined as

$$[\mathbf{a}]_{(j-1)K+k}\triangleq\frac{\mathrm{SINR}_{jk}^{\mathrm{UL}}}{\left(1+\mathrm{SINR}_{jk}^{\mathrm{UL}}\right)\left|\mathbf{v}_{jk}^{\mathsf{H}}\mathbf{g}_{jjk}\right|^2},\forall j,\forall k, \quad (11)$$

where $\mathrm{SINR}_{jk}^{\mathrm{UL}}$ is the SINR of the kth UT in the jth cell in the uplink, and is given by

$$\mathrm{SINR}_{jk}^{\mathrm{UL}}\triangleq\frac{\check{p}_{jk}\left|\mathbf{v}_{jk}^{\mathsf{H}}\mathbf{g}_{jjk}\right|^2}{\sum\limits_{\substack{l=1\\(l,q)\neq(j,k)}}^{L}\sum\limits_{q=1}^{K}\check{p}_{lq}\left|\mathbf{v}_{jk}^{\mathsf{H}}\mathbf{g}_{jlq}\right|^2+\check{\sigma}_{\epsilon_{jk}}^2+\check{\sigma}_{\mu_{jk}}^2+1}, \quad (12)$$

where $\check{\sigma}_{\epsilon_{jk}}^2$ and $\check{\sigma}_{\mu_{jk}}^2$ are defined as

$$\check{\sigma}_{\epsilon_{jk}}^2=\mathbf{v}_{jk}^{\mathsf{H}}\left(\sum_{l=1}^{L}\kappa_{\mathrm{BT}}^2\left(\mathbf{I}_N\circ\left(\mathbf{G}_{jl}\check{\mathbf{P}}_l\mathbf{G}_{jl}^{\mathsf{H}}\right)\right)\right)\mathbf{v}_{jk}, \quad (13)$$

$$\check{\sigma}_{\mu_{jk}}^2=\sum_{l=1}^{L}\sum_{q=1}^{K}\kappa_{\mathrm{UR}}^2\check{p}_{lq}\left|\mathbf{v}_{jk}^{\mathsf{H}}\mathbf{g}_{jlq}\right|^2. \quad (14)$$

Moreover, the elements of matrix $\mathbf{A}\in\mathbb{R}^{KL\times KL}$ are given by

$$\begin{aligned}[\mathbf{A}]_{(j-1)K+k,(l-1)K+q}\triangleq\mathbf{v}_{jk}^{\mathsf{H}}\Big(&\left(1+\kappa_{\mathrm{UR}}^2\right)\mathbf{g}_{jlq}\mathbf{g}_{jlq}^{\mathsf{H}}\\ &+\kappa_{\mathrm{BT}}^2\mathbf{I}_N\circ\left(\mathbf{g}_{jlq}\mathbf{g}_{jlq}^{\mathsf{H}}\right)\Big)\mathbf{v}_{jk},\forall j,l,\forall k,q.\end{aligned} \quad (15)$$

Proof: Please refer to Appendix A. ■

Corollary 1: For a single-cell MU-MIMO system with ideal H/W, i.e., with $L=1$ and $\kappa_{\mathrm{BT}}^2=\kappa_{\mathrm{UR}}^2=0$, Theorem 1 reduces to a special case, where $\mathrm{SINR}_k^{\mathrm{UL}}$ and matrix $\mathbf{A}$ are given by

$$\mathrm{SINR}_k^{\mathrm{UL}}\triangleq\frac{\check{p}_k\left|\mathbf{v}_k^{\mathsf{H}}\mathbf{g}_k\right|^2}{\sum\limits_{q=1,q\neq k}^{K}\check{p}_q\left|\mathbf{v}_k^{\mathsf{H}}\mathbf{g}_q\right|^2+1}, \quad (16)$$

$$[\mathbf{A}]_{k,q}\triangleq\left|\mathbf{v}_k^{\mathsf{H}}\mathbf{g}_q\right|^2, \quad (17)$$

where we have omitted the cell indices for the sake of notation simplicity. Note that, as expected, the UL/DL duality theorem for single-cell MU-MIMO systems with ideal H/W presented in Corollary 1 is identical to the UL/DL theorem in [3].

As can be observed from (12), the SINR expression of the kth UT in the jth cell in the dual uplink system model depends only on the detection vector of the kth UT in the jth cell. This makes the design of the detection vectors in the uplink much simpler than the design of the precoding vectors in the original downlink problem, where the received signals at the UTs are coupled with respect to the precoding vectors. Now, we extend the results in Theorem 1 to MSE UL/DL duality, and provide the result in the following corollary.

Corollary 2: The DL and the dual UL multi-cell MU-MIMO systems with residual HWIs as defined in Theorem 1 have identical per-user MSEs, i.e., $\mathrm{MSE}_{jk}^{\mathrm{DL}} = \mathrm{MSE}_{jk}^{\mathrm{UL}}, \forall j \in \{1, \ldots, L\}, k \in \{1, \ldots, K\}$, where $\mathrm{MSE}_{jk}^{\mathrm{DL}}$ and $\mathrm{MSE}_{jk}^{\mathrm{UL}}$ are given by

$$\mathrm{MSE}_{jk}^{\mathrm{DL}} = \mathbb{E}\left\{\|\hat{d}_{jk}^{\mathrm{DL}} - d_{jk}\|^2\right\} = \xi_{jk}^2 \left|\mathbf{g}_{jjk}^{\mathsf{H}}\mathbf{v}_{jk}\right|^2 - 2\xi_{jk}\Re\left\{\mathbf{g}_{jjk}^{\mathsf{H}}\mathbf{v}_{jk}\right\} + \frac{\xi_{jk}^2}{p_{jk}}\left(1 + \sigma_{\epsilon_{jk}}^2 + \sigma_{\mu_{jk}}^2 + \sum_{\substack{l=1\\(l,q)\neq(j,k)}}^{L}\sum_{q=1}^{K} p_{lq}\left|\mathbf{g}_{ljk}^{\mathsf{H}}\mathbf{v}_{lq}\right|^2\right) + 1, \tag{18}$$

$$\mathrm{MSE}_{jk}^{\mathrm{UL}} = \mathbb{E}\left\{\|\hat{d}_{jk}^{\mathrm{UL}} - d_{jk}\|^2\right\} = \xi_{jk}^2 \left|\mathbf{v}_{jk}^{\mathsf{H}}\mathbf{g}_{jjk}\right|^2 - 2\xi_{jk}\Re\left\{\mathbf{v}_{jk}^{\mathsf{H}}\mathbf{g}_{jjk}\right\} + \frac{\xi_{jk}^2}{\check{p}_{jk}}\left(1 + \check{\sigma}_{\epsilon_{jk}}^2 + \check{\sigma}_{\mu_{jk}}^2 + \sum_{\substack{l=1\\(l,q)\neq(j,k)}}^{L}\sum_{q=1}^{K} \check{p}_{lq}\left|\mathbf{v}_{jk}^{\mathsf{H}}\mathbf{g}_{jlq}\right|^2\right) + 1. \tag{19}$$

Proof: Please refer to Appendix B. ■

IV. Multi-Cell HWI Aware Precoding

In this section, a MCHA-MMSE precoder is derived by exploiting the proposed MSE UL/DL duality framework for multi-cell MU-MIMO systems with residual HWIs presented in Corollary 2. The optimization problem for minimization of the sum-MSE in the downlink is formulated as follows:

$$\begin{aligned} \min_{\mathbf{P}_j, \mathbf{V}_j, \, \boldsymbol{\Xi}_j} \quad & \sum_{j=1}^{L}\sum_{k=1}^{K} \mathrm{MSE}_{jk}^{\mathrm{DL}}, \\ \text{subject to:} \quad & \sum_{j=1}^{L}\sum_{k=1}^{K} p_{jk} \leq KL\rho_{\mathrm{DL}}, \\ & p_{jk} \geq 0, \forall j, k, \\ & \mathbf{v}_{jk}^{\mathsf{H}}\mathbf{v}_{jk} = 1, \forall j, k. \end{aligned} \tag{20}$$

Next, in Section IV-A, we apply the UL/DL duality framework from Corollary 2 to the downlink optimization problem in (20) to obtain the dual uplink optimization problem, which is easier to solve.

A. Multi-Cell HWI Aware Detection in the Dual UL

Now, we apply Corollary 2 and transform the downlink optimization problem in (20) into its dual uplink equivalent. The equivalent uplink optimization problem is given by

$$\begin{aligned} \min_{\check{\mathbf{P}}_j, \tilde{\mathbf{U}}_j, \forall j} \quad & \sum_{j=1}^{L}\sum_{k=1}^{K} \mathrm{MSE}_{jk}^{\mathrm{UL}} \\ \text{subject to:} \quad & \sum_{j=1}^{L}\sum_{k=1}^{K} \check{p}_{jk} \leq KL\rho_{\mathrm{DL}}, \\ & \check{p}_{jk} \geq 0, \forall j, k, \end{aligned} \tag{21}$$

where $\tilde{\mathbf{U}}_j = \boldsymbol{\Xi}_j\mathbf{V}_j^{\mathsf{H}}$, and $\mathrm{MSE}_{jk}^{\mathrm{UL}}$ is given by (19). As can be seen from (21), the dual uplink optimization problem has only a total transmit power constraint, and is therefore easier to solve compared to the original downlink problem in (20). Note that in (21), the normalization matrix $\boldsymbol{\Xi}_j$ is absorbed into matrix $\tilde{\mathbf{U}}_j$. Moreover, the constraint $\mathbf{v}_{jk}^{\mathsf{H}}\mathbf{v}_{jk} = 1$ is implicitly met in (21), since $\mathbf{V}_j = \tilde{\mathbf{U}}_j^{\mathsf{H}}\boldsymbol{\Xi}_j^{-1}$, where $\boldsymbol{\Xi}_j$ by definition contains the norms of the rows of $\tilde{\mathbf{U}}_j$ on its main diagonal. Here, in order to focus on the precoder design, we assume a uniform power allocation in the uplink, i.e., $\check{p}_{jk} = \rho_{\mathrm{DL}}, \forall j, k$. Moreover, since the detector matrix at the jth BS has only impact on the MSEs of the UTs in the jth cell, the sum-MSEs in different cells can be minimized individually. This leads to the following unconstrained optimization problem

$$\min_{\tilde{\mathbf{U}}_j} \quad \sum_{k=1}^{K} \mathrm{MSE}_{jk}^{\mathrm{UL}}, \forall j. \tag{22}$$

In the following theorem, we present the optimal detection vectors for uniform power allocation in the dual uplink system.

Theorem 2: The solution to the optimization problem in (21) for a fixed power allocation is given by

$$\tilde{\mathbf{U}}_j^{\mathrm{MCHA}} = \frac{1}{N}\mathbf{G}_{jj}^{\mathsf{H}}\left(\frac{1+\kappa_{\mathrm{UR}}^2}{N}\sum_{m=1}^{L}\mathbf{G}_{jm}\mathbf{G}_{jm}^{\mathsf{H}} + \frac{\kappa_{\mathrm{BT}}^2}{N}\mathbf{I}_N \circ \left(\sum_{m=1}^{L}\mathbf{G}_{jm}\mathbf{G}_{jm}^{\mathsf{H}}\right) + \frac{1}{N\rho_{\mathrm{DL}}}\mathbf{I}_N\right)^{-1}. \tag{23}$$

Proof: Please refer to Appendix C. ■

Applying the UL/DL duality in Theorem 1, the MCHA-MMSE precoding vector of the kth UT at the jth BS is given by $\mathbf{v}_{jk}^{\mathrm{MCHA}} = \left(\tilde{\mathbf{u}}_{jk}^{\mathrm{MCHA}}\right)^{\mathsf{H}} / \|\tilde{\mathbf{u}}_{jk}^{\mathrm{MCHA}}\|, \forall j, k$, where $\tilde{\mathbf{u}}_{jk}^{\mathrm{MCHA}}$ is the kth row of matrix $\tilde{\mathbf{U}}_j^{\mathrm{MCHA}}$ which is given in (23). Next, the uplink SINRs can be calculated by (12) after substituting $\mathbf{v}_{jk} = \mathbf{v}_{jk}^{\mathrm{MCHA}}$, and subsequently the downlink power allocation is determined by (10).

Remark 1: Comparing (23) and (8), and considering $\mathbf{V}_j^{\mathrm{MCHA}} = \left(\tilde{\mathbf{U}}_j^{\mathrm{MCHA}}\right)^{\mathsf{H}}\left(\boldsymbol{\Xi}_j^{\mathrm{MCHA}}\right)^{-1}$, where $\boldsymbol{\Xi}_j^{\mathrm{MCHA}}$ is a diagonal matrix, which contains the norms of the rows of $\tilde{\mathbf{U}}_j^{\mathrm{MCHA}}$ on its main diagonal, it can be seen that the expression for the MCHA-MMSE precoder contains two additional terms compared to the conventional MMSE precoder. The first term, $\sum_{m=1, m\neq j}^{L}\mathbf{G}_{jm}\mathbf{G}_{jm}^{\mathsf{H}}/N$, contains information regarding the channels between the UTs in other cells and the considered BS, and is exploited by the MCHA-MMSE precoder to suppress multi-cell interference. The second term, $\kappa_{\mathrm{UR}}^2\sum_{m=1}^{L}\mathbf{G}_{jm}\mathbf{G}_{jm}^{\mathsf{H}}/N + \kappa_{\mathrm{BT}}^2\mathbf{I}_N \circ \left(\sum_{m=1}^{L}\mathbf{G}_{jm}\mathbf{G}_{jm}^{\mathsf{H}}/N\right)$, accounts for the HWI.

In order to further evaluate the performance of the proposed MCHA-MMSE precoder, we also consider a multi-cell aware but hardware unaware MMSE (MCAHU-MMSE) precoder which is given by

$$\mathbf{V}_j^{\mathrm{MCAHU}} = \left(\tilde{\mathbf{U}}_j^{\mathrm{MCAHU}}\right)^{\mathsf{H}}\left(\boldsymbol{\Xi}_j^{\mathrm{MCAHU}}\right)^{-1}, \tag{24}$$

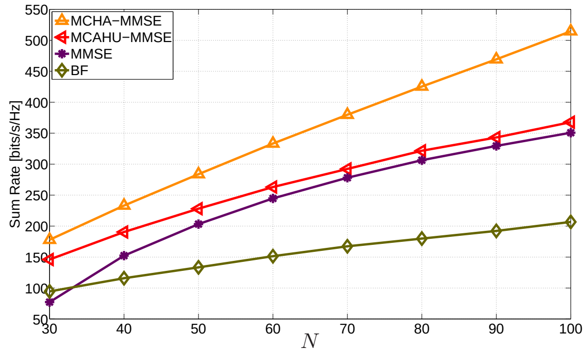

Fig. 1. Sum rate vs. number of BS antennas N for $K = 20$ and $\kappa_{\text{BT}} = \kappa_{\text{UR}} = 0.03$.

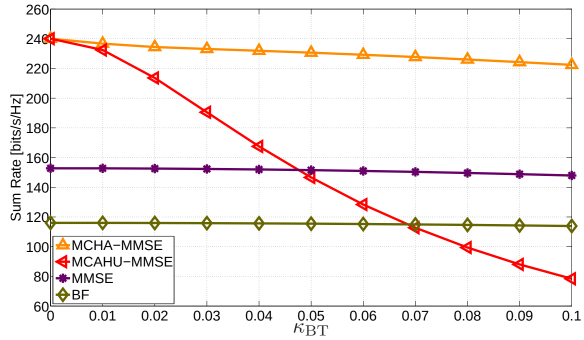

Fig. 2. Sum rate vs. κ_{BT} for $N = 40$, $K = 20$, and $\kappa_{\text{UR}} = \kappa_{\text{BT}}$.

where $\boldsymbol{\Xi}_j^{\text{MCAHU}}$ is a diagonal matrix, which contains the norms of the rows of $\tilde{\mathbf{U}}_j^{\text{MCAHU}}$ on its main diagonal, and $\tilde{\mathbf{U}}_j^{\text{MCAHU}}$ is obtained by setting $\kappa_{\text{UR}}^2 = \kappa_{\text{BT}}^2 = 0$ in (23), and is given by

$$\tilde{\mathbf{U}}_j^{\text{MCAHU}} = \frac{1}{N}\mathbf{G}_{jj}^{\text{H}}\left(\frac{1}{N}\sum_{m=1}^{L}\mathbf{G}_{jm}\mathbf{G}_{jm}^{\text{H}} + \frac{1}{N\rho_{\text{DL}}}\mathbf{I}_N\right)^{-1}. \tag{25}$$

The performance metric used in this paper is the network-wide ergodic achievable sum rate, which is given by

$$R = \sum_{j=1}^{L}\sum_{k=1}^{K}\mathbb{E}\left\{\log_2\left(1 + \text{SINR}_{jk}^{\text{DL}}\right)\right\}, \tag{26}$$

where the expectation is approximated by averaging over a sufficient number of channel realizations, and $\text{SINR}_{jk}^{\text{DL}}$ is given by (4) after replacing $\mathbf{V}_j$ by $\mathbf{V}_j^{\text{BF}}$, $\mathbf{V}_j^{\text{MMSE}}$, $\mathbf{V}_j^{\text{MCAHU}}$ and $\mathbf{V}_j^{\text{MCHA}}$ for the BF, MMSE, MCAHU-MMSE, and MCHA-MMSE precoders, respectively.

V. Numerical Results

In order to evaluate the performance of the proposed MCHA-MMSE precoder, Monte-Carlo simulations have been performed. Here, we assume a system consisting of $L = 7$ cells, where in each cell, one BS serves $K = 20$ UTs. The cross gain is assumed to be $\eta = 0.3$, and the transmit SNR is set to $\rho_{\text{DL}} = 20$ dB. Moreover, we adopt similar antenna correlation parameters as in [16]. In particular, we assume that the number of physical paths is equal to $M = N$, and the normalized antenna spacing is $\lambda = 0.5$.

In Fig. 1, the ergodic achievable sum rate of the proposed MCHA-MMSE precoder as a function of N is compared to that of the MCAHU-MMSE, conventional MMSE, and the BF precoders. In this simulation, we adopt similar residual HWI parameters at the BS and the UTs as in [12]. Accordingly, we have $\kappa_{\text{BT}} = \kappa_{\text{UR}} = 0.03$. As can be seen, the MCHA-MMSE precoder achieves substantially higher sum rates than all other investigated precoders. For example, for $N = 100$, the MCHA-MMSE precoder achieves an almost 40% higher sum rate than the MCAHU-MMSE and the conventional MMSE precoders. In particular, for increasing number of BS antennas N, the gap between the MCHA-MMSE precoder and all other considered precoders increases, too. From Fig. 1, it can also be observed that the performance gap between the MCAHU-MMSE precoder and the conventional MMSE precoder decreases with increasing number of BS antennas N. This is due to the fact that for increasing N, the orthogonality of the channel vectors between the UTs in the neighboring cells and the BS under consideration increases, which leads to less multi-cell interference.

In Fig. 2, the sum rate performance of the investigated precoders for $N = 40$ as a function of HWI parameter κ_{BT} is depicted. It can be seen that with increasing κ_{BT}, the sum rate of the proposed MCHA-MMSE precoder decreases only slightly, whereas the performance of the MCAHU-MMSE precoder decreases rapidly. Consequently, the performance gap between the MCHA-MMSE precoder and the MCAHU-MMSE precoder increases with increasing κ_{BT}, too. From Fig. 2, it can also be seen that the MCHA-MMSE precoder achieves considerably higher sum rates than the conventional MMSE and BF precoders for the entire range of κ_{BT}. For example, for $\kappa_{\text{BT}} = 0.03$, the MCHA-MMSE precoder achieves 53% higher sum rate than the conventional MMSE precoder. Surprisingly, for large values of κ_{BT}, the MCAHU-MMSE precoder performs even worse than the conventional MMSE precoder. This is due to the fact that, in contrast to the conventional MMSE precoder, the MCAHU-MMSE precoder uses all available degrees of freedom in an effort to suppress the multi-cell interference, which makes it more sensitive to mismatches.

VI. Conclusion

We presented an uplink/downlink duality framework for multi-cell MU-MIMO systems with residual HWIs. We showed that if the power allocation, precoding, and detection matrices are chosen properly, under the same total transmit power constraint, the same per-user SINR and per-user MSE as in the downlink can be achieved in the dual uplink. We used the proposed uplink/downlink duality framework to transform the network-wide sum-MSE minimization problem in the downlink to its uplink equivalent, and presented a MCHA-MMSE precoder, which takes the multi-cell interference and HWI into account. Our simulation results showed that the proposed MCHA-MMSE precoder achieves considerably higher sum rates than the multi-cell aware HWI unaware MMSE, the conventional MMSE, and the BF precoders.

Appendix A - Proof of Theorem 1

The downlink BS HWI component in (5) can be rewritten as

$$
\begin{aligned}
\sigma_{\epsilon_{jk}}^2 &= \sum_{l=1}^{L} \kappa_{\mathrm{BT}}^2 \mathbf{g}_{ljk}^{\mathsf{H}} \left(\mathbf{I}_N \circ \left(\mathbf{V}_l \mathbf{P}_l \mathbf{V}_l^{\mathsf{H}} \right) \right) \mathbf{g}_{ljk} \\
&= \sum_{l=1}^{L} \kappa_{\mathrm{BT}}^2 \mathbf{g}_{ljk}^{\mathsf{H}} \left(\mathbf{I}_N \circ \left(\sum_{q=1}^{K} p_{lq} \mathbf{v}_{lq} \mathbf{v}_{lq}^{\mathsf{H}} \right) \right) \mathbf{g}_{ljk} \\
&= \sum_{l=1}^{L} \sum_{n=1}^{N} \sum_{q=1}^{K} \kappa_{\mathrm{BT}}^2 p_{lq} \left| [\mathbf{G}_{lj}]_{n,k} \right|^2 \left| [\mathbf{V}_l]_{n,q} \right|^2 \\
&= \sum_{l=1}^{L} \sum_{q=1}^{K} \kappa_{\mathrm{BT}}^2 p_{lq} \mathbf{v}_{lq}^{\mathsf{H}} \left(\mathbf{I}_N \circ \left(\mathbf{g}_{ljk} \mathbf{g}_{ljk}^{\mathsf{H}} \right) \right) \mathbf{v}_{lq}. \qquad (27)
\end{aligned}
$$

Next, we rewrite the downlink UT HWI component in (6) as

$$
\sigma_{\mu_{jk}}^2 = \sum_{l=1}^{L} \sum_{q=1}^{K} \kappa_{\mathrm{UR}}^2 p_{lq} \mathbf{v}_{lq}^{\mathsf{H}} \mathbf{g}_{ljk} \mathbf{g}_{ljk}^{\mathsf{H}} \mathbf{v}_{lq}. \qquad (28)
$$

Taking into account (15), (27), and (28), and performing straightforward algebraic operations, the expression in (4) can be reformulated in the following compact form [18]

$$
\left(\mathbf{I}_{KL} - \mathrm{diag}\left(\mathbf{b} \right) \cdot \mathbf{A}^{\mathsf{T}} \right) \mathbf{p} = \mathbf{b}, \qquad (29)
$$

where the elements of $\mathbf{b} = [b_1, \ldots, b_{KL}]^{\mathsf{T}}$ are given by

$$
[\mathbf{b}]_{(j-1)K+k} \triangleq \frac{\mathrm{SINR}_{jk}^{\mathrm{DL}}}{\left(1 + \mathrm{SINR}_{jk}^{\mathrm{DL}}\right) \left| \mathbf{g}_{jjk}^{\mathsf{H}} \mathbf{v}_{jk} \right|^2}, \forall j, \forall k. \qquad (30)
$$

Now, considering (11) and (15), and performing a similar procedure as for (29), the expression in (12) can be rewritten as the following matrix-vector form

$$
\left(\mathbf{I}_{KL} - \mathrm{diag}\left(\mathbf{a} \right) \cdot \mathbf{A} \right) \check{\mathbf{p}} = \mathbf{a}, \qquad (31)
$$

where $\check{\mathbf{p}} = \left[\check{\mathbf{p}}_1^{\mathsf{T}} \ldots \check{\mathbf{p}}_L^{\mathsf{T}} \right]^{\mathsf{T}}$ is the stacked vector of power allocation of all UTs in all cells in the uplink with $\check{\mathbf{p}}_j = \left[\check{p}_{j1}, \ldots, \check{p}_{jK} \right]^{\mathsf{T}}, \forall j \in \{1, \ldots, L\}$ being the power allocation of the UTs in the jth cell. Comparing (11) and (30), it can be concluded that identical individual SINRs in the downlink and in the uplink can be achieved if and only if $\mathbf{a} = \mathbf{b}$. In the following, we show that the condition $\mathbf{a} = \mathbf{b}$ also leads to identical sum powers in the downlink and the uplink. According to (10), the sum transmit power of all UTs in all cells in the DL is given by [18]

$$
\begin{aligned}
\sum_{j=1}^{L} \sum_{k=1}^{K} p_{jk} &= \mathbf{1}^{\mathsf{T}} \left(\mathbf{I}_{KL} - \mathrm{diag}(\mathbf{a}) \cdot \mathbf{A}^{\mathsf{T}} \right)^{-1} \mathbf{a} = \mathbf{1}^{\mathsf{T}} \left(\mathbf{\Lambda}_a - \mathbf{A}^{\mathsf{T}} \right)^{-1} \mathbf{1} \\
&= \mathbf{1}^{\mathsf{T}} \left(\mathbf{\Lambda}_b - \mathbf{A}^{\mathsf{T}} \right)^{-1} \mathbf{1} = \mathbf{1}^{\mathsf{T}} \left(\left(\mathbf{\Lambda}_b - \mathbf{A}^{\mathsf{T}} \right)^{-1} \right)^{\mathsf{T}} \mathbf{1} \\
&= \mathbf{1}^{\mathsf{T}} \left(\left(\mathbf{I}_{KL} - \mathrm{diag}\left(\mathbf{b} \right) \cdot \mathbf{A} \right)^{\mathsf{T}} \right)^{-1} \mathbf{b} = \sum_{j=1}^{L} \sum_{k=1}^{K} \check{p}_{jk},
\end{aligned}
$$

where $\mathbf{1} \in \mathbb{R}^{KL}$ is an all-one column vector and $\mathbf{\Lambda}_a = \mathrm{diag}\left(1/a_1, \ldots, 1/a_{KL} \right)$, $\mathbf{\Lambda}_b = \mathrm{diag}\left(1/b_1, \ldots, 1/b_{KL} \right)$, and we applied $\mathbf{\Lambda}_a = \mathbf{\Lambda}_b$. This completes the proof.

Appendix B - Proof of Corollary 2

The second term on the right hand side of (18) can be rewritten as

$$
\begin{aligned}
&\frac{\xi_{jk}^2}{p_{jk}} \left(1 + \sigma_{\epsilon_{jk}}^2 + \sigma_{\mu_{jk}}^2 + \sum_{\substack{l=1 \\ (l,q) \neq (j,k)}}^{L} \sum_{q=1}^{K} p_{lq} \left| \mathbf{g}_{ljk}^{\mathsf{H}} \mathbf{v}_{lq} \right|^2 \right) = \frac{\xi_{jk}^2 \left| \mathbf{g}_{jjk}^{\mathsf{H}} \mathbf{v}_{jk} \right|^2}{\mathrm{SINR}_{jk}^{\mathrm{DL}}} \\
&= \frac{\xi_{jk}^2 \left| \mathbf{v}_{jk}^{\mathsf{H}} \mathbf{g}_{jjk} \right|^2}{\mathrm{SINR}_{jk}^{\mathrm{UL}}} = \frac{\xi_{jk}^2}{\check{p}_{jk}} \left(1 + \check{\sigma}_{\epsilon_{jk}}^2 + \check{\sigma}_{\mu_{jk}}^2 + \sum_{\substack{l=1 \\ (l,q) \neq (j,k)}}^{L} \sum_{q=1}^{K} \check{p}_{lq} \left| \mathbf{v}_{jk}^{\mathsf{H}} \mathbf{g}_{jlq} \right|^2 \right), \qquad (32)
\end{aligned}
$$

where we exploited (4), (12), and the equality of the per-user SINRs in the DL and the dual UL from Theorem 1. Comparing (32) with (18) and (19) completes the proof.

Appendix C - Proof of Theorem 2

After defining the error vector $\mathbf{e}_j = \hat{\mathbf{d}}_j^{\mathrm{UL}} - \mathbf{d}_j$, substituting $\check{\mathbf{P}}_j = \rho_{\mathrm{DL}} \mathbf{I}_K$ and $\tilde{\mathbf{U}}_j = \mathbf{\Xi}_j \mathbf{V}_j^{\mathsf{H}}$ into (9), and using the properties $\mathbb{E}\left\{ \mathbf{d}_j^{\mathsf{H}} \mathbf{d}_l \right\} = 0, \forall j \neq l$, $\mathbb{E}\left\{ \mathbf{d}_j^{\mathsf{H}} \mathbf{z}_l \right\} = 0, \forall j, l$, the sum-MSE in the jth cell in the uplink can be formulated as

$$
\begin{aligned}
\mathbb{E}\left\{ \mathrm{tr}\left(\mathbf{e}_j \mathbf{e}_j^{\mathsf{H}} \right) \right\} = \mathrm{tr} \Bigg(\tilde{\mathbf{U}}_j \Bigg(&\left(1 + \kappa_{\mathrm{UR}}^2 \right) \sum_{l=1}^{L} \mathbf{G}_{jl} \mathbf{G}_{jl}^{\mathsf{H}} \\
&+ \kappa_{\mathrm{BT}}^2 \sum_{l=1}^{L} \mathbf{I}_N \circ \left(\mathbf{G}_{jl} \mathbf{G}_{jl}^{\mathsf{H}} \right) + \frac{1}{\rho_{\mathrm{DL}}} \mathbf{I}_N \Bigg) \tilde{\mathbf{U}}_j^{\mathsf{H}} - 2 \tilde{\mathbf{U}}_j \mathbf{G}_{jj} + \mathbf{I}_K \Bigg). \qquad (33)
\end{aligned}
$$

Next, we take the derivative of the expression in (33) with respect to $\tilde{\mathbf{U}}_j^*$ and set it to zero to obtain the optimal detection matrix:

$$
\begin{aligned}
\frac{\partial}{\partial \tilde{\mathbf{U}}_j^*} \mathbb{E}\left\{ \mathrm{tr}\left(\mathbf{e}_j \mathbf{e}_j^{\mathsf{H}} \right) \right\} = \tilde{\mathbf{U}}_j \Bigg(&\left(1 + \kappa_{\mathrm{UR}}^2 \right) \sum_{l=1}^{L} \mathbf{G}_{jl} \mathbf{G}_{jl}^{\mathsf{H}} \\
&+ \kappa_{\mathrm{BT}}^2 \sum_{l=1}^{L} \mathbf{I}_N \circ \left(\mathbf{G}_{jl} \mathbf{G}_{jl}^{\mathsf{H}} \right) + \frac{1}{\rho_{\mathrm{DL}}} \mathbf{I}_N \Bigg) - \mathbf{G}_{jj}^{\mathsf{H}} \overset{!}{=} \mathbf{0}. \qquad (34)
\end{aligned}
$$

Performing straightforward algebraic operations, the optimal detection matrix in (23) is obtained. This completes the proof.

References

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

[2] H. Ngo, E. Larsson, and T. Marzetta, "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Feb. 2013.

[3] P. Viswanath and D. N. C. Tse, "Sum Capacity of the Vector Gaussian Broadcast Channel and Uplink-Downlink Duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.

[4] M. Schubert and H. Boche, "A Unifying Theory for Uplink and Downlink Multiuser Beamforming," in *Int. Zurich Seminar Broadband Commun., Access, Transmission, Networking*, 2002.

[5] S. Shi, M. Schubert, and H. Boche, "Downlink MMSE Transceiver Optimization for Multiuser MIMO Systems: Duality and Sum-MSE Minimization," *IEEE Trans. Signal Process.*, vol. 55, no. 11, pp. 5436–5446, Nov. 2007.

[6] H. Dahrouj and W. Yu, "Coordinated Beamforming for the Multicell Multi-Antenna Wireless System," *IEEE Trans. Wireless Commun.*, vol. 9, no. 5, pp. 1748–1759, May 2010.

[7] T. Schenk, *RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation*. Springer-Verlag, 2008.

[8] C. Studer, M. Wenk, and A. Burg, "MIMO Transmission with Residual Transmit-RF Impairments," in *Int. ITG Workshop on Smart Antennas (WSA)*, Feb. 2010.

[9] E. Björnson, M. Matthaiou, and M. Debbah, "A New Look at Dual-Hop Relaying: Performance Limits with Hardware Impairments," *IEEE Trans. Commun.*, vol. 61, no. 11, pp. 4512–4525, Nov. 2013.

[10] X. Zhang, M. Matthaiou, M. Coldrey, and E. Björnson, "Impact of Residual Transmit RF Impairments on Training-Based MIMO Systems," *IEEE Trans. Commun.*, vol. 63, no. 8, pp. 2899–2911, Aug. 2015.

[11] P. Zetterberg, "Experimental Investigation of TDD Reciprocity-Based Zero-Forcing Transmit Precoding," *EURASIP J. Adv. Signal Process.*, vol. 2011, Article ID 137541.

[12] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO Systems with Non-Ideal Hardware: Energy Efficiency, Estimation, and Capacity Limits," *IEEE Trans. Inf. Theory*, vol. 60, no. 11, pp. 7112–7139, Nov. 2014.

[13] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with Non-Ideal Arbitrary Arrays: Hardware Scaling Laws and Circuit-Aware Design." [Online]. Available: http://arxiv.org/abs/1409.0875.

[14] S. Zarei, W. Gerstacker, and R. Schober, "Uplink/Downlink Duality in Massive MIMO Systems with Hardware Impairments," *Submitted to IEEE Int. Conf. Commun. (ICC)*, 2016.

[15] H. Ngo, E. Larsson, and T. Marzetta, "The Multicell Multiuser MIMO Uplink with Very Large Antenna Arrays and a Finite-Dimensional Channel," *IEEE Trans. Commun.*, vol. 61, no. 6, pp. 2350–2361, Jun. 2013.

[16] 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. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.

[17] S. Wagner, R. Couillet, M. Debbah, and D. T. M. Slock, "Large System Analysis of Linear Precoding in Correlated MISO Broadcast Channels Under Limited Feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 7, pp. 4509–4537, Jul. 2012.

[18] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*. Cambridge University Press, 2011.