



**QUEEN'S
UNIVERSITY
BELFAST**

IQ Imbalance in Multiuser Systems: Channel Estimation and Compensation

Kolomvakis, N., Matthaiou, M., & Coldrey, M. (2016). IQ Imbalance in Multiuser Systems: Channel Estimation and Compensation. *IEEE Transactions on Communications*, 64(7).
<https://doi.org/10.1109/TCOMM.2016.2558186>

Published in:
IEEE Transactions on Communications

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
[Link to publication record in Queen's University Belfast Research Portal](#)

Publisher rights

Copyright 2016 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works.

General rights

Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy

The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

IQ Imbalance in Multiuser Systems: Channel Estimation and Compensation

Nikolaos Kolomvakis, *Student Member, IEEE*, Michail Matthaiou, *Senior Member, IEEE*,
and Mikael Coldrey, *Member, IEEE*

Abstract—In this paper, we consider the uplink of a single-cell multi-user single-input multiple-output (MU-SIMO) system with in-phase and quadrature-phase imbalance (IQI). Particularly, we investigate the effect of receive (RX) IQI on the performance of MU-SIMO systems with large antenna arrays employing maximum-ratio combining (MRC) receivers. In order to study how IQI affects channel estimation, we derive a new channel estimator for the IQI-impaired model and show that the higher the value of signal-to-noise ratio (SNR) the higher the impact of IQI on the spectral efficiency (SE). Moreover, a novel pilot-based joint estimator of the augmented MIMO channel matrix and IQI coefficients is described and then, a low-complexity IQI compensation scheme is proposed which is based on the IQI coefficients' estimation and it is independent of the channel gain. The performance of the proposed compensation scheme is analytically evaluated by deriving a tractable approximation of the ergodic SE assuming transmission over Rayleigh fading channels with large-scale fading. Furthermore, we investigate how many MSs should be scheduled in massive multiple-input multiple-output (MIMO) systems with IQI and show that the highest SE loss occurs at the optimal operating point. Finally, by deriving asymptotic power scaling laws, and proving that the SE loss due to IQI is asymptotically independent of the number of BS antennas, we show that massive MIMO is resilient to the effect of RX IQI.

I. INTRODUCTION

Massive multiple-input multiple-output (MIMO) systems are built from an excessive number of antenna elements and show great promise for mobile wireless technologies [1], [2]. Extra antennas help to focus energy into small regions of space and bring huge improvements in throughput when compared to conventional MIMO systems with small number of antennas [3]. Other benefits of massive MIMO include: extensive use of inexpensive low-power components, simplest linear receivers e.g. maximum-ratio combining (MRC), become nearly optimal [4], [5]. The focusing of energy also reduces the effects of inter-cell interference, while the effect of channel estimation

Manuscript received 27 November 2015; revised 15 March 2016; accepted 13 April 2016. This work was presented in part at the IEEE International Conference on Communications (ICC), London, U.K., June 2015. This work was supported by the Swedish Governmental Agency for Innovation Systems (VINNOVA) within the VINN Excellence Center Chase. The associate editor coordinating the review of this paper and approving it for publication was X. Gao.

N. Kolomvakis is with the Department of Signals and Systems, Chalmers University of Technology, 412 96, Gothenburg, Sweden (email: nikolaos.kolomvakis@chalmers.se).

M. Matthaiou is with the School of Electronics, Electrical Engineering and Computer Science, Queen's University Belfast, Belfast, U.K. (email: m.matthaiou@qub.ac.uk).

M. Coldrey is with Ericsson Research, Gothenburg, Sweden (email: mikael.coldrey@ericsson.com).

errors in individual receiver chains vanishes on the average [6].

The excessive degrees-of-freedom offered by massive MIMO may be used to limit performance degradation in the event of failure of individual antenna units. This characteristic of massive MIMO creates new opportunities in terms of the accuracy of radio-frequency (RF) front-ends. Hardware accuracy constraints can be relaxed, thus allowing the deployment of lower-quality (inexpensive) components on future massive base stations (BSs), compared to today's examples [3]. The use of low-quality hardware is desirable in order to make massive MIMO an economically sustainable technological shift, or its total deployment cost will scale with the number of RF front-ends and components. Unfortunately, these low-quality RF components are more prone to hardware imperfections, such as phase noise [7] and in-phase and quadrature-phase imbalance (IQI), which refers to the mismatch between the I and Q branches, i.e., the mismatch between the real and imaginary parts of the complex signal. Several stages in the receiver structure can contribute to the IQI, e.g., errors in the nominal 90° phase shift between the local oscillator signals used for down-conversion of the I and Q signals, as well as, the difference in amplitude transfer between the total I and Q branches.

Another source of IQI is the limited accuracy of analog hardware, such as finite tolerance of capacitors and transistors [8]. This leads to a degradation in the overall performance and, therefore, to a deteriorated user experience. The performance of single-cell uplink massive multi-user single-input multiple-output (MU-SIMO) systems has been well investigated in the literature [4], [9], [10]; however, the impact of IQ imperfections have been scarcely studied so far, especially in the massive MIMO context. Although several IQI compensation algorithms have been proposed, especially for the case of single-input single-output (SISO) systems [11]–[13], little is still known for the case of massive MIMO systems. To the best of our knowledge the only relevant works are [8], [14]–[18]. Specifically, [8] studied the sensitivity of massive antenna arrays to RF IQI by elaborating on the design of single-user beamforming schemes; moreover, [14] proposed augmented spatial post-processing linear minimum-mean-square-error (LMMSE) filters for mitigating the effect of IQI in the uplink receiver. Finally, [16]–[18] proposed widely linear precoding algorithms for massive MIMO systems with IQI. However, none of these works proposed practical channel estimation and low complexity IQI compensation schemes, while a detailed performance analysis is also missing from

the literature.

Motivated by the above discussion, we study the performance of a single-cell uplink massive MU-SIMO system under the presence of IQI at the BS side, while we assume that the transmit IQ branches at the MSs are perfect. Although in real implementations hardware impairments are not negligible, there is a long body of literature on transmit IQI compensation and, even elimination [12], [19]. Therefore, it is reasonable to assume that IQI affects only the low-cost branches of the massive BS.¹

The paper makes the following contributions: we first assume that the BS does not have knowledge of the IQI and, thus, we evaluate analytically the impact of IQI on channel estimation by deriving a LMMSE channel estimator for the IQI-impaired model. Next, we propose a novel low-complexity pilot-based joint estimator of the propagation MIMO channel and the IQI coefficients at the receiver. After that, and in order to mitigate the effect of IQI, we propose a low-complexity IQI compensation scheme; this scheme can be implemented individually on each antenna branch and it is independent of the channel matrix (depends only on the estimation of the IQI coefficients). Based on our analytical results, it is shown that the proposed IQI compensation scheme is able to eliminate IQI when the knowledge of the IQI parameters is perfect. Furthermore, we study the performance of the proposed compensation scheme by deriving an approximation of the ergodic SE of MRC receivers. Moreover, it is demonstrated that when the number of BS antennas, N , grows without bound, we can reduce the transmit power of each user proportionally to $1/N$ if the BS has perfect channel state information (CSI), and proportionally to $1/\sqrt{N}$ if the BS has imperfect CSI; interestingly, in both cases the extra interference introduced by the IQI is significantly suppressed in the large-antenna limit. Finally, we show that by increasing the number of BS antennas, the impact of IQI is reduced with a logarithmic scaling.

This paper is organized as follows. In Section II, the system model is presented. In section III, the LMMSE channel estimator for the IQI-impaired model is derived. In section IV, we introduce and analyze the proposed joint estimator of the channel and the IQI coefficients. In Section V, we present the proposed IQI compensation scheme, while in Section VI, we derive analytical expressions for the achievable rates. Numerical results are provided in Section VII, and conclusions are drawn in Section VIII.

Notation: The superscripts $(\cdot)^H$, $(\cdot)^T$, $(\cdot)^{-1}$, $(\cdot)^*$ and $(\cdot)^{-*}$ stand for the conjugate-transpose, transpose, inverse, conjugate and conjugate-inverse respectively. The operators $\text{tr}\{\cdot\}$, $\text{vec}\{\cdot\}$, $\mathbb{E}\{\cdot\}$ and \otimes denote the trace of a matrix, the vectorization of a matrix, which converts the matrix into a column vector, the expectation and the Kronecker product correspondingly. For a matrix \mathbf{A} , $[\mathbf{A}]_{ij}$ denotes its entry in the i th row and j th column of \mathbf{A} . Moreover, we define the

¹Direct-conversion radio (DCR) architectures are widely used in wireless communication systems and it is a good candidate for massive MIMO [8]. One of the DCR RF imperfections is IQI due to analogue circuit sensitivity to component variations, resulting in unavoidable errors in the I and Q branches due to temperature variations and process mismatches.

operator $|\mathbf{A}|^2 \triangleq \mathbf{A}\mathbf{A}^H$. Finally, $\|\cdot\|_F$ refers to the Frobenius norm, \mathbf{I}_n denotes the $n \times n$ identity matrix, and $\mathbf{0}_n$ represents the $n \times n$ zero matrix.

II. SYSTEM AND IQ IMBALANCE MODELS

We consider the uplink of a single-cell MU-SIMO system, which includes a BS equipped with N antennas communicating with K single-antenna mobile stations (MSs). The $N \times 1$ received vector at the BS is

$$\mathbf{r} = \sqrt{\rho_u} \mathbf{Y} \mathbf{x} + \mathbf{w} \quad (1)$$

where \mathbf{Y} is the $N \times K$ channel matrix that characterizes the propagation environment, \mathbf{x} is a zero-mean circularly symmetric complex Gaussian $K \times 1$ vector (i.e. $\mathbb{E}\{\mathbf{x}\mathbf{x}^T\} = \mathbf{0}$) of independent, unit-power symbols transmitted simultaneously by the K MSs, with the average transmit power of each MS being ρ_u . Finally, $\mathbf{w} \sim \mathcal{CN}(0, \sigma_w^2 \mathbf{I})$ is the additive white Gaussian noise (AWGN).

More specifically, \mathbf{Y} models the composite propagation channel affected by small-scale fading, geometric attenuation and log-normal shadow fading. Its elements $[\mathbf{Y}]_{nk}$ are given by

$$[\mathbf{Y}]_{nk} = [\mathbf{H}]_{nk} \sqrt{\beta_k} \quad (2)$$

where $[\mathbf{H}]_{nk} \sim \mathcal{CN}(0, 1)$, is the small-scale channel coefficient from the k th user to the n th antenna element. The term β_k models geometric attenuation and shadow fading between the k -th MS and BS. The large-scale fading is modeled via $\beta_k = \zeta_k / d_k^\alpha$, where ζ_k is the lognormal shadowing with variance σ^2 . Finally, the term d_k is the reference distance between the BS and the k -th MS, and α is the path loss exponent. We can alternatively express \mathbf{Y} as follows

$$\mathbf{Y} = \mathbf{H} \mathbf{D}^{1/2} \quad (3)$$

where \mathbf{D} is a $K \times K$ diagonal matrix, whose diagonal elements are given by $[\mathbf{D}]_{kk} = \beta_k$.

A. RF IQ Imbalance

The RF front-ends of the MSs are assumed to be perfect and we focus on the impact of the IQI at the BS. The IQI can be modeled in either a symmetrical or asymmetrical fashion. In the symmetrical method, each branch (I and Q) experiences half of the phase and amplitude errors, see e.g. [16], [19]. In the asymmetrical method, the I branch is modeled to be ideal and the errors are modeled in the Q branch. Nevertheless, it has been easily verified that the symmetrical model can be obtained from the asymmetrical one by a linear transformation with a rotation matrix and a scaling factor.

In this paper, we consider the asymmetrical IQI model for our analysis. Note that this model has been extensively used in the majority of relevant works (see e.g. [8], [14], [20], [21]). The RX IQI can be modeled for the n -th ($n = 1, \dots, N$) individual antenna element on the baseband equivalent level as [20]

$$r_{\text{imb},n} = K_{1,n} r_n + K_{2,n} r_n^* \quad (4)$$

where r_n is the baseband equivalent signal under ideal IQ matching. The IQI coefficients $K_{1,n}$ and $K_{2,n}$ are of the

form $K_{1,n} \triangleq (1 + g_n e^{-j\phi_n})/2$ and $K_{2,n} \triangleq (1 - g_n e^{j\phi_n})/2$, where g_n and ϕ_n denote the RX amplitude and phase mismatch, respectively. For perfect IQ matching, the imbalance parameters are $g_n = 1$ and $\phi_n = 0$; thus, in this case, we have $K_{1,n} = 1$ and $K_{2,n} = 0$. The antenna array (4) can be rewritten as

$$\mathbf{r}_{\text{imb}} = \mathbf{K}_1 \mathbf{r} + \mathbf{K}_2 \mathbf{r}^* \quad (5)$$

where \mathbf{K}_1 and \mathbf{K}_2 are $N \times N$ diagonal matrices whose the n th diagonal entries are $K_{1,n}$ and $K_{2,n}$ respectively. Thus, the matrices \mathbf{K}_1 and \mathbf{K}_2 are referring to the IQI coefficients of each parallel receiver branch.

The IQI-impaired version of (1) is given by

$$\mathbf{r}_{\text{imb}} = \sqrt{\rho_u} \mathbf{K}_1 \mathbf{Y} \mathbf{x} + \sqrt{\rho_u} \mathbf{K}_2 \mathbf{Y}^* \mathbf{x}^* + \tilde{\mathbf{w}} \quad (6)$$

where $\tilde{\mathbf{w}} \triangleq \mathbf{K}_1 \mathbf{w} + \mathbf{K}_2 \mathbf{w}^*$.

We assume that both the IQI coefficients \mathbf{K}_1 and \mathbf{K}_2 , and the small-scale channel gain \mathbf{H} are unknown at the receiver. Next, we seek to develop a channel estimation scheme for the IQI-impaired model in (6).

III. EFFECTIVE CHANNEL ESTIMATION WITHOUT IQI COMPENSATION

In general, full knowledge of the channel gain of each transmit-receive antenna link is required in MIMO systems for implementing MRC [22]. In this section, we assume that the BS does not have knowledge of the IQI and therefore we consider that MRC receivers detect the transmitted signal \mathbf{x} , while \mathbf{x}^* is treated as interference. On this basis, we derive a LMMSE estimator for the effective channel: $\mathbf{Y}_1 \triangleq \mathbf{K}_1 \mathbf{Y}$.

It is assumed that the channels are quasi-static, which implies that the channel gain matrix remains invariant in each frame but may vary from frame to frame. Under this assumption, pilot symbols are inserted at the beginning of each transmit frame in order to perform channel estimation.

Let \mathbf{S}_p denote the $K \times \tau$ transmitted pilot symbol matrix, where τ is the length of channel training.² Then, the equivalent MIMO signal model for pilot symbol transmission can be expressed as

$$\begin{aligned} \mathbf{R}_{\text{imb},p} &= \mathbf{K}_1 (\sqrt{\rho_p} \mathbf{Y} \mathbf{S}_p + \mathbf{W}_p) + \mathbf{K}_2 (\sqrt{\rho_p} \mathbf{Y} \mathbf{S}_p + \mathbf{W}_p)^* \\ &= \sqrt{\rho_p} \mathbf{\Omega}_{\text{imb}} \mathbf{Z}_p + \tilde{\mathbf{W}}_p \end{aligned} \quad (7)$$

where $\mathbf{R}_{\text{imb},p}$ represents the $N \times \tau$ received signal matrix during pilot transmission, \mathbf{W}_p refers to the $N \times \tau$ noise matrix and we define $\mathbf{Y}_2 \triangleq \mathbf{K}_2 \mathbf{Y}^*$ and the power of each pilot symbol $\rho_p \triangleq \tau \rho_u$. Furthermore, we define $\mathbf{\Omega}_{\text{imb}} \triangleq [\mathbf{Y}_1 \ \mathbf{Y}_2] \in \mathbb{C}^{N \times 2K}$, $\mathbf{Z}_p \triangleq [\mathbf{S}_p^T \ \mathbf{S}_p^H]^T \in \mathbb{C}^{2K \times \tau}$, and $\tilde{\mathbf{W}}_p \triangleq \mathbf{K}_1 \mathbf{W}_p + \mathbf{K}_2 \mathbf{W}_p^*$.

By vectorizing the received signal in (7) and applying the property [23]: $\text{vec}(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A}) \text{vec}(\mathbf{B})$, the received training signal of our system can be expressed as

$$\text{vec}(\mathbf{R}_{\text{imb},p}) = \sqrt{\tau \rho_u} (\mathbf{Z}_p^T \otimes \mathbf{I}_N) \text{vec}(\mathbf{\Omega}_{\text{imb}}) + \text{vec}(\tilde{\mathbf{W}}_p)$$

where, it can be verified that,

$$\text{vec}(\tilde{\mathbf{W}}_p) = (\mathbf{I}_\tau \otimes \mathbf{K}_1) \text{vec}(\mathbf{W}_p) + (\mathbf{I}_\tau \otimes \mathbf{K}_2) \text{vec}(\mathbf{W}_p^*)$$

²Henceforth, the subscript p denotes transmission during the training phase.

and

$$\text{vec}(\mathbf{\Omega}_{\text{imb}}) = \begin{bmatrix} (\mathbf{D}^{1/2} \otimes \mathbf{K}_1) \text{vec}(\mathbf{H}) \\ (\mathbf{D}^{1/2} \otimes \mathbf{K}_2) \text{vec}(\mathbf{H}^*) \end{bmatrix}.$$

Then, we apply the results of [24, Chapter 15.8] to derive the estimation error $\mathcal{E} = \text{vec}(\mathbf{\Omega}_{\text{imb}}) - \text{vec}(\hat{\mathbf{\Omega}}_{\text{imb}})$ whose mean is zero and its covariance matrix \mathbf{C}_{MMSE} is given by:

$$\begin{aligned} \mathbf{C}_{\text{MMSE}} &= \left(\tau \rho_u (\mathbf{Z}_p^* \otimes \mathbf{I}_N) \mathbf{C}_{w_p}^{-1} (\mathbf{Z}_p^T \otimes \mathbf{I}_N) + \mathbf{C}_{\mathbf{\Omega}_{\text{imb}}}^{-1} \right)^{-1} \\ &= \left(\frac{\tau \rho_u}{\sigma_w^2} (\mathbf{Z}_p^* \mathbf{Z}_p^T) \otimes (|\mathbf{K}_1|^2 + |\mathbf{K}_2|^2) + \mathbf{C}_{\mathbf{\Omega}_{\text{imb}}}^{-1} \right)^{-1} \end{aligned} \quad (8)$$

where $\mathbf{C}_{w_p} \triangleq \mathbb{E} \{ \text{vec}(\tilde{\mathbf{W}}_p) \text{vec}(\tilde{\mathbf{W}}_p)^H \}$ and $\mathbf{C}_{\mathbf{\Omega}_{\text{imb}}} \triangleq \mathbb{E} \{ \text{vec}(\mathbf{\Omega}_{\text{imb}}) \text{vec}(\mathbf{\Omega}_{\text{imb}})^H \}$ denote the covariance matrices of $\text{vec}(\tilde{\mathbf{W}}_p)$ and $\text{vec}(\mathbf{\Omega}_{\text{imb}})$ respectively. It can be shown that

$$\mathbf{C}_{\mathbf{\Omega}_{\text{imb}}} = \begin{bmatrix} \mathbf{D} \otimes |\mathbf{K}_1|^2 & \mathbf{0}_{NK} \\ \mathbf{0}_{NK} & \mathbf{D} \otimes |\mathbf{K}_2|^2 \end{bmatrix}$$

and $\mathbf{C}_{w_p} = \sigma_w^2 \mathbf{I}_\tau \otimes (|\mathbf{K}_1|^2 + |\mathbf{K}_2|^2)$.

Remark 1: We assume that the statistics, i.e., covariance matrices, of both $\mathbf{\Omega}_{\text{imb}}$ and $\tilde{\mathbf{W}}_p$ are perfectly known at the receiver in order to perform the LMMSE estimator. This is a reasonable assumption since the channel statistics change over a slower time scale.³

Therefore, the total MSE becomes $\sigma_{\mathcal{E}}^2 \triangleq \mathbb{E} \{ \|\mathcal{E}\|_F^2 \} = \text{tr} \{ \mathbf{C}_{\text{MMSE}} \}$. In [25, Lemma 1] it is shown that for the positive-definite matrix $\mathbf{C}_{\text{MMSE}}^{-1}$ with (m, n) th entry $[\mathbf{C}_{\text{MMSE}}^{-1}]_{mn}$, it holds that

$$\begin{aligned} \sigma_{\mathbf{\Omega}_{\text{imb}}}^2 &= \text{tr} \{ \mathbf{C}_{\text{MMSE}} \} \\ &\geq \sum_{n=1}^{2NK} \frac{1}{[\mathbf{C}_{\text{MMSE}}^{-1}]_{nn}} \end{aligned}$$

where the equality is obtained if and only if $\mathbf{C}_{\text{MMSE}}^{-1}$ is diagonal. Given the fact that the matrices $\mathbf{C}_{\mathbf{\Omega}_{\text{imb}}}$ and \mathbf{C}_{w_p} are diagonal, it is easy to verify that \mathbf{C}_{MMSE} is a diagonal matrix, if and only if, the training sequences are orthogonal between the i th and j th MS. Mathematically speaking, the following condition needs to be satisfied⁴

$$\mathbf{Z}_p \mathbf{Z}_p^H = \mathbf{I}_{2K}. \quad (9)$$

that is $\mathbf{S}_p \mathbf{S}_p^H = \mathbf{I}_K$ and $\mathbf{S}_p \mathbf{S}_p^T = \mathbf{0}_K$. For instance, this can be achieved by choosing the rows of DFT matrices. Therefore, the MSE $\sigma_{\mathbf{\Omega}_{\text{imb}}}^2$ is minimized if and only if the training sequences are orthogonal between different MSs. It can be derived from (8) that the minimum MSE of the (n, k) th ($1 \leq n \leq N$, $1 \leq k \leq K$) entry of the effective channel \mathbf{Y}_1 is given by

$$\tilde{\sigma}_{nk}^2 = \left(\frac{1}{\beta_k |\mathbf{K}_1|_n|^2} + \frac{2\tau \rho_u}{(1 + g_n^2) \sigma_w^2} \right)^{-1}, \quad \tau \geq 2K. \quad (10)$$

³Despite the fact that the covariance matrix $\mathbf{C}_{\mathbf{\Omega}_{\text{imb}}}$ is known, i.e. we know $\mathbf{D} \otimes |\mathbf{K}_1|^2$, the large-scale fading coefficient, β_k , as well as the phase of \mathbf{K}_1 are unknown at the receiver. Thus, the IQI coefficients cannot be estimated directly from $\mathbf{C}_{\mathbf{\Omega}_{\text{imb}}}$ such that an IQI estimation algorithm is proposed in Section IV-B.

⁴This result is consistent with the classical results on MIMO channel estimation in spatially white noise channels [26].

Notice that the effective channel estimation is deteriorated only if the amplitude mismatch is $g_n > 1$. Interestingly, we can observe that when the amplitude mismatch is $g_n < 1$ or there is phase mismatch (i.e. $\phi_n \neq 0$), then the MSE in (10) is lower than the one with perfect IQ matching (i.e. $g_n = 1$ and $\phi_n = 0$). In this case, intuitively, the variance of the entries of the effective channel becomes smaller than the ones of the case without IQI, and thus, the LMMSE performs better. However, it is worth mentioning that in order to achieve the MSE in (10), *the training sequence length needs to be twice of that in the case with ideal IQ branches*. This can be a crucial drawback especially when the number of MSs becomes very high, thereby reducing significantly the duration of data transmission within a frame interval.

By the orthogonality principle of LMMSE estimators [24], each element of $\hat{\mathbf{Y}}_1$ has variance, $\hat{\beta}_{1nk}$ ($1 \leq n \leq N, 1 \leq k \leq K$), given by

$$\begin{aligned} \hat{\beta}_{1nk} &= \beta_{1nk} - \tilde{\sigma}_{nk}^2 \\ &= \frac{2\tau\rho_u\beta_k^2|\mathbf{K}_1|_n^4}{2\tau\rho_u\beta_k|\mathbf{K}_1|_n^2 + (g_n^2 + 1)\sigma_w^2} \end{aligned} \quad (11)$$

where $\beta_{1nk} = |\mathbf{K}_1|_n^2\beta_k$ represents the variance of the (n, k) th element of \mathbf{Y}_1 .

IV. JOINT ESTIMATION OF CHANNEL AND IQI COEFFICIENTS

In contrast to Section III, where the channel estimation was blind to the IQI coefficients, we herein introduce a novel pilot-based IQI estimation scheme. By estimating the IQI coefficients \mathbf{K}_1 and \mathbf{K}_2 , we are able to develop an IQI compensation scheme which is proposed in the next section. First, the propagation channel \mathbf{Y} is estimated by decoupling it from the IQI coefficients and, then, based on the estimated channel, an estimator for the IQI coefficients is developed.

A. Channel Estimation

The estimation of the channel \mathbf{Y} is of particular importance for two reasons: i) it is an intermediate step in order to estimate the IQI coefficients and ii) it is required for implementing MRC reception after applying the IQI compensation scheme. Following the technique in [27], the channel estimation can account for IQI coefficients. From (7) and making use of the property $\mathbf{K}_1 + \mathbf{K}_2^* = \mathbf{I}_N$, it is easy to verify that

$$\mathbf{R}_{\text{imb},p} + \mathbf{R}_{\text{imb},p}^* = \sqrt{\tau\rho_u}\Omega\mathbf{Z}_p + \mathbf{W}'_p \quad (12)$$

where $\Omega \triangleq [\mathbf{Y} \ \mathbf{Y}^*]$ and $\mathbf{W}'_p \triangleq \mathbf{W}_p + \mathbf{W}_p^*$. The entries of \mathbf{W}'_p are i.i.d. real Gaussian random variables, each with a $\mathcal{N}(0, 2\sigma_w^2)$ distribution. We can now see that the received signal in (12) is independent of the IQI coefficients but the additive Gaussian noise variance is twice of that in the case with ideal IQ branches. Following the channel estimation technique in Section III, the LMMSE estimation error for the k th MS is given by:

$$\begin{aligned} \sigma_k^2 &= \left(\frac{1}{\beta_k} + \frac{\tau\rho_u}{2\sigma_w^2} \right)^{-1} \\ &= \frac{2\beta_k\sigma_w^2}{2\sigma_w^2 + \tau\rho_u\beta_k}, \quad 1 \leq k \leq K. \end{aligned} \quad (13)$$

By the orthogonality principle of LMMSE estimators, each element of the k th column of $\hat{\mathbf{Y}}$ has variance, $\hat{\beta}_k$, and it is given by

$$\begin{aligned} \hat{\beta}_k &= \beta_k - \sigma_k^2 \\ &= \frac{\tau\rho_u\beta_k^2}{2\sigma_w^2 + \tau\rho_u\beta_k}, \quad 1 \leq k \leq K. \end{aligned} \quad (14)$$

Then, denoting \mathbf{E}_Y as the channel estimation error, the relationship between \mathbf{Y} and $\hat{\mathbf{Y}}$ can be expressed as

$$\mathbf{Y} = \hat{\mathbf{Y}} + \mathbf{E}_Y \quad (15)$$

where each element of \mathbf{E}_Y is zero-mean complex Gaussian distributed random variable, and all entries in its k th column have variance σ_k^2 .

B. Estimation of IQI coefficients

In this section, based on the estimated channel $\hat{\mathbf{Y}}$, we derive a novel least-square (LS) estimator of \mathbf{K}_1 and \mathbf{K}_2 . A precise estimator of the IQI coefficients is necessary in order for the proposed IQI compensation scheme to be efficient. The IQI coefficients are estimated individually on each antenna element based on the received pilot symbols. Decoupling the estimation problem on each antenna element, renders its implementation simple, especially, when the number of antennas becomes large. More specifically, the received signal for the n th ($n = 1, \dots, N$) individual antenna element during the pilot transmission can be expressed as

$$\begin{aligned} r_{n,p} &= \sqrt{\tau\rho_u}K_{1,n}[\mathbf{A}]_{np} \\ &\quad + \sqrt{\tau\rho_u}K_{2,n}[\mathbf{A}^*]_{np} + \tilde{w}_n, \quad 1 \leq p \leq \tau \end{aligned}$$

where

$$\begin{aligned} [\mathbf{A}]_{np} &\triangleq \sum_{k=1}^K [\mathbf{Y}]_{nk}[\mathbf{S}_p]_{kp} \\ &\stackrel{(15)}{=} \sum_{k=1}^K ([\hat{\mathbf{Y}}]_{nk} + [\mathbf{E}_Y]_{nk})[\mathbf{S}_p]_{kp} \\ &= \underbrace{\sum_{k=1}^K [\hat{\mathbf{Y}}]_{nk}[\mathbf{S}_p]_{kp}}_{\triangleq [\hat{\mathbf{A}}]_{np}} + \underbrace{\sum_{k=1}^K [\mathbf{E}_Y]_{nk}[\mathbf{S}_p]_{kp}}_{\triangleq [\mathbf{E}]_{np}} \end{aligned} \quad (16)$$

and $\tilde{w}_n \triangleq K_{1,n}w_n + K_{2,n}w_n^*$.

After the n th antenna element receives all transmitted pilot symbols, $\{r_{n,p}\}_{p=1}^\tau$, the pilot-based received signal vector $\mathbf{r}_n = [r_{n,1}, \dots, r_{n,\tau}]^T$ can be written as

$$\mathbf{r}_n = \sqrt{\tau\rho_u}\hat{\mathbf{A}}_n\boldsymbol{\kappa}_n + \sqrt{\tau\rho_u}\mathbf{E}_n\boldsymbol{\kappa}_n + \tilde{\mathbf{w}}_n \quad (17)$$

where $\boldsymbol{\kappa}_n \triangleq [K_{1,n} \ K_{2,n}]^T$ and

$$\mathcal{X}_n \triangleq \begin{bmatrix} [\mathcal{X}]_{n1} & [\mathcal{X}^*]_{n1} \\ \vdots & \vdots \\ [\mathcal{X}]_{n\tau} & [\mathcal{X}^*]_{n\tau} \end{bmatrix} \in \mathbb{C}^{\tau \times 2} \quad (18)$$

where $\mathcal{X} \triangleq \{\hat{\mathbf{A}}, \mathbf{E}\}$. The LS estimate $\hat{\boldsymbol{\kappa}}_n$ is given by

$$\hat{\boldsymbol{\kappa}}_n = \frac{1}{\sqrt{\tau\rho_u}}(\hat{\mathbf{A}}_n^H \hat{\mathbf{A}}_n)^{-1} \hat{\mathbf{A}}_n^H \mathbf{r}_n. \quad (19)$$

Next, we simplify the expression of $\hat{\kappa}_n$ by providing the following proposition.

Proposition 1: Let $\mathbf{W}_n \triangleq \hat{\mathbf{A}}_n^H \hat{\mathbf{A}}_n \in \mathbb{C}^{2 \times 2}$. Then, assuming that the uplink pilots sequence \mathbf{Z}_p satisfies the condition (9) we have that

$$[\mathbf{W}_n]_{11} = [\mathbf{W}_n]_{22} = \sum_{k=1}^K \left| [\hat{\mathbf{Y}}]_{nk} \right|^2 \quad (20)$$

and

$$[\mathbf{W}_n]_{12} = [\mathbf{W}_n]_{21} = 0. \quad (21)$$

Proof: See Appendix I. ■

Proposition 1 shows that \mathbf{W}_n is a scaled identity matrix which makes its inverse simple to compute. Therefore, assuming that (9) is satisfied, the LS estimate $\hat{\kappa}_n$ can be rewritten as

$$\hat{\kappa}_n = \frac{\hat{\mathbf{A}}_n^H \mathbf{r}_n}{\sqrt{\tau \rho_u} \sum_{k=1}^K \left| [\hat{\mathbf{Y}}]_{nk} \right|^2}, \quad \tau \geq 2K, \quad (22)$$

and the resulting MSE is provided by the following proposition:

Proposition 2: Assuming that the IQI coefficients' vector κ_n is estimated by the LS estimator in (22), then the MSE of each coefficient is given by

$$\sigma_{\kappa_{n,i}}^2 \triangleq \mathbb{E}\{ \|K_{i,n} - \hat{K}_{i,n}\|_2^2 \} \quad (23)$$

$$= \mathbb{E} \left\{ \left| K_{i,n} \right|^2 \frac{\sum_{k=1}^K \sigma_k^2 \left| [\hat{\mathbf{Y}}]_{nk} \right|^2}{\left(\sum_{k=1}^K \left| [\hat{\mathbf{Y}}]_{nk} \right|^2 \right)^2} + \frac{\| \kappa_n \|^2}{\tau \rho_u \sum_{k=1}^K \left| [\hat{\mathbf{Y}}]_{nk} \right|^2} \right\}, \quad i = 1, 2, \quad (24)$$

and it can be upper bounded by

$$\sigma_{\kappa_{n,i}}^2 \leq \alpha_i \mathbb{E} \left\{ \frac{1}{\sum_{k=1}^K \left| [\hat{\mathbf{Y}}]_{nk} \right|^2} \right\} \quad (25)$$

where σ_k^2 expresses the channel estimation error in (13) and $\alpha_i \triangleq |K_{i,n}|^2 \sigma_{\max}^2 + \| \kappa_n \|^2 / \tau \rho_u$, where $\sigma_{\max}^2 \triangleq \max_{1 \leq k \leq K} \{ \sigma_k^2 \}$.

Proof: See Appendix II. ■

According to the upper bound (25), the number of MSs, is a crucial parameter in the performance of the estimator. Notice that when $N, K \rightarrow \infty$, while $N/K = c > 1$, then $\sigma_{\kappa_{n,i}}^2 \rightarrow 0$. This implies that, when the BS serves a large number of MSs (i.e. $K > 10$), then the estimation of \mathbf{K}_1 and \mathbf{K}_2 can be nearly perfect. Therefore, the IQI estimator can be very efficient in a MU-SIMO regime where tens of users are served.

Unfortunately, the exact analytic derivation of the expression (24) imposes several mathematical challenges, since the first negative moment of (24) does not exist [28]. Therefore, we now provide a tractable tight approximation via the following proposition:

Proposition 3: Assuming that $[\hat{\mathbf{Y}}]_{nk}$ ($1 \leq k \leq K$) are K independent, zero-mean circular symmetric complex Gaussian

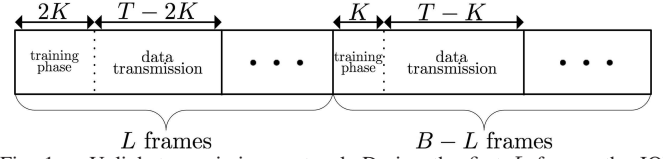


Fig. 1. Uplink transmission protocol: During the first L frames the IQI coefficients: \mathbf{K}_1 , \mathbf{K}_2 and the channel \mathbf{Y} are jointly estimated with pilot length $\tau = 2K$. After estimating the IQI coefficients, and employing the IQI compensation scheme the channel can be estimated directly from (37) during the next $B - L$ frames with pilot length $\tau = K$.

random variables with variance $\hat{\beta}_k$, then the MSE of the LS estimate $\hat{\kappa}_n$ in (22) can be approximated as

$$\sigma_{\kappa_{n,i}}^2 \approx \alpha_i \sum_{l=1}^K \frac{\Gamma(0, \frac{\alpha}{\hat{\beta}_l})}{\hat{\beta}_l \prod_{j=1, j \neq l}^K (1 - \frac{\hat{\beta}_j}{\hat{\beta}_l})}, \quad i = 1, 2 \quad (26)$$

where $\alpha > 0$ is a constant that can be arbitrary small, and $\Gamma(\cdot, \cdot)$ is the upper incomplete gamma function defined as

$$\Gamma(s, x) \triangleq \int_x^\infty t^{s-1} e^{-t} dt.$$

Proof: See Appendix III. ■

Given the fact that IQI is a slowly varying process, we assume that \mathbf{K}_1 and \mathbf{K}_2 remain constant over a sequence of B frames. Therefore, it is sufficient to utilize the IQI estimator in (22) only during the first L ($L \ll B$) frames as shown in Fig. 1. The IQI coefficient \mathbf{K}_1 (or \mathbf{K}_2) is estimated and then, \mathbf{K}_2 (or \mathbf{K}_1) can be estimated according to the property $\hat{\mathbf{K}}_2 = \mathbf{I}_N - \hat{\mathbf{K}}_1^*$. Therefore, by denoting as $\Psi^{(l)}$ the estimation error of \mathbf{K}_i ($i = 1, 2$) during the l th ($1 \leq l \leq L$) frame and given the fact that $\mathbf{K}_1 + \mathbf{K}_2^* = \mathbf{I}_N$, the estimated parameters $\hat{\mathbf{K}}_1^{(l)}$, $\hat{\mathbf{K}}_2^{(l)}$ can be written as

$$\hat{\mathbf{K}}_1^{(l)} = \mathbf{K}_1^{(l)} + \Psi^{(l)} \quad (27)$$

$$\hat{\mathbf{K}}_2^{(l)} = \mathbf{K}_2^{(l)} - \Psi^{*(l)} \quad (28)$$

where $\Psi^{(l)}$ is a diagonal matrix which entries are modeled as complex Gaussian distributed random variables, with zero mean and the n th diagonal element has variance $\sigma_{\kappa_{n,i}}^2$.

We can estimate \mathbf{K}_1 and \mathbf{K}_2 with better precision by averaging out the estimation error as follows:

$$\hat{\mathbf{K}}_1 \triangleq \frac{1}{L} \sum_{l=1}^L \hat{\mathbf{K}}_1^{(l)} = \mathbf{K}_1 + \Psi \quad (29)$$

$$\hat{\mathbf{K}}_2 \triangleq \frac{1}{L} \sum_{l=1}^L \hat{\mathbf{K}}_2^{(l)} = \mathbf{K}_2 - \Psi^* \quad (30)$$

where Ψ is a diagonal matrix with zero mean entries and the n th diagonal element has variance $\sigma_{\kappa_{n,i}}^2 / L$.

V. IQ IMBALANCE COMPENSATION

The interference caused by IQI motivates the need for compensation schemes. In this section, we propose an IQI compensation scheme which is only based on the estimated $\hat{\mathbf{K}}_1$ and $\hat{\mathbf{K}}_2$. Its advantage is that it can be implemented individually on each antenna element and thus it has very low complexity implementation at the receiver.

Combining the received signal in (5) with the corresponding expression for $\mathbf{r}_{\text{imb}}^*$, the augmented received signal is given by

$$\mathbf{R}_{\text{imb}} = \Phi \mathbf{R} \quad (31)$$

where $\mathbf{R}_{\text{imb}} \triangleq [\mathbf{r}_{\text{imb}}^T \mathbf{r}_{\text{imb}}^H]^T$, $\Phi \triangleq \begin{bmatrix} \mathbf{K}_1 & \mathbf{K}_2 \\ \mathbf{K}_2^* & \mathbf{K}_1^* \end{bmatrix} \in \mathbb{C}^{2N \times 2N}$ and $\mathbf{R} \triangleq [\mathbf{r}^T \mathbf{r}^H]^T$ is the augmented ideal received signal (i.e. without IQI).

Given the estimated IQI coefficients, we now define the compensation matrix Φ_{comp} as follows⁵

$$\Phi_{\text{comp}} \triangleq \begin{bmatrix} \widehat{\mathbf{K}}_1^* \widehat{\mathbf{K}}_2^{-1} & -\mathbf{I}_N \\ -\mathbf{I}_N & \widehat{\mathbf{K}}_1 \widehat{\mathbf{K}}_2^{-*} \end{bmatrix}. \quad (32)$$

Note that the proposed IQI compensation scheme is independent of the channel matrix gain, which makes its implementation very feasible. Moreover, since \mathbf{K}_2 is diagonal, its inverse requires simply to compute the inverse of N scalar parameters (with complexity $\mathcal{O}(N)$). Thus, Φ_{comp} can be computed offline with very low implementation complexity since it contains only $4N$ non-zero entries. Substituting (29) and (30) into the matrix Φ , it can be decomposed as

$$\Phi = \underbrace{\begin{bmatrix} \widehat{\mathbf{K}}_1 & \widehat{\mathbf{K}}_2 \\ \widehat{\mathbf{K}}_2^* & \widehat{\mathbf{K}}_1^* \end{bmatrix}}_{\triangleq \widehat{\Phi}} + \underbrace{\begin{bmatrix} -\Psi & \Psi^* \\ \Psi & -\Psi^* \end{bmatrix}}_{\triangleq \Phi_\psi}. \quad (33)$$

Multiplying the compensation matrix in (32) with the received signal model in (31), using the decomposition in (33), we have the following compensated received signal

$$\mathbf{R}_{\text{comp}} = \Phi_{\text{comp}} \mathbf{R}_{\text{imb}} \quad (34)$$

$$\begin{aligned} &= \Phi_{\text{comp}} \widehat{\Phi} \mathbf{R} + \underbrace{\Phi_{\text{comp}} \Phi_\psi \mathbf{R}}_{\triangleq \mathbf{W}_\psi} \\ &= \Phi_c \mathbf{r}_n + \mathbf{W}_\psi \end{aligned} \quad (35)$$

where $\Phi_c \triangleq \Phi_{\text{comp}} \widehat{\Phi} = \begin{bmatrix} \mathbf{K}_c & \mathbf{0}_{N \times N} \\ \mathbf{0}_{N \times N} & \mathbf{K}_c^* \end{bmatrix} \in \mathbb{C}^{2N \times 2N}$, and $\mathbf{K}_c \triangleq \widehat{\mathbf{K}}_1^* \widehat{\mathbf{K}}_2^{-1} \widehat{\mathbf{K}}_1 - \widehat{\mathbf{K}}_2^*$. Finally, $\mathbf{W}_\psi \triangleq [\mathbf{w}_\psi^T \mathbf{w}_\psi^H]^T$, where $\mathbf{w}_\psi = \widehat{\mathbf{K}}_2^{-1} (\Psi^* \mathbf{r}^* - \Psi \mathbf{r})$. It is worth mentioning that the matrix Φ_c , and \mathbf{W}_ψ in (35) represent the effective channel gain and the interference due to the imperfect estimation of \mathbf{K}_1 and \mathbf{K}_2 , respectively after IQI compensation.

The received signal vector in (35) can be written as $\mathbf{R}_{\text{comp}} = [\mathbf{r}_{\text{comp}}^T \mathbf{r}_{\text{comp}}^H]^T$, where \mathbf{r}_{comp} is the $N \times 1$ desired received vector after IQI compensation, which can be expanded as

$$\mathbf{r}_{\text{comp}} = \mathbf{K}_c \mathbf{r} + \mathbf{w}_\psi. \quad (36)$$

The final stage of the compensation is to recover the ideal received signal from (36). Therefore, we conclude that

$$\hat{\mathbf{r}}_{\text{comp}} = \mathbf{K}_c^{-1} \mathbf{r}_{\text{comp}} = \mathbf{r} + \tilde{\mathbf{w}}_\psi \quad (37)$$

where $\tilde{\mathbf{w}}_\psi \triangleq \mathbf{K}_c^{-1} \mathbf{w}_\psi = \mathbf{G} \left(\Psi^* \mathbf{r}^* - \Psi \mathbf{r} \right)$, represents the interference due to the imperfect knowledge of \mathbf{K}_1 and

⁵Note that, in practice, both \mathbf{K}_1 and \mathbf{K}_2 are always full-rank, since all antenna elements are IQI-impaired.

\mathbf{K}_2 . Notice that each element of the diagonal matrix $\mathbf{G} \triangleq \left(\widehat{\mathbf{K}}_2 \mathbf{K}_c \right)^{-1}$ is given by $[\mathbf{G}]_{nn} = \frac{1}{g_n \cos(\phi_n)}$ ($n = 1, \dots, N$). If the error of the IQI parameters' estimation is zero, then $\tilde{\mathbf{w}}_\psi = 0$ and the ideal received signal \mathbf{r} is obtained.

Remark 2: Notice that after estimating \mathbf{K}_1 and \mathbf{K}_2 , the propagation MIMO channel \mathbf{Y} can be estimated directly from (37) during the last $B - L$ frames as shown in Fig. 1. In this case, the training sequence length becomes at least $\tau \geq K$.

Note that the compensated signal \mathbf{r}_{comp} in (34) can be expressed as

$$\mathbf{r}_{\text{comp}} = \widehat{\mathbf{K}}_1^* \widehat{\mathbf{K}}_2^{-1} \mathbf{r}_{\text{imb}} - \mathbf{r}_{\text{imb}}^*.$$

Given the fact that the matrices $\widehat{\mathbf{K}}_1^* \widehat{\mathbf{K}}_2^{-1}$ and \mathbf{K}_c are diagonal, the compensated signal $\hat{r}_{\text{comp},n}$ for the n th individual antenna element is given by

$$\hat{r}_{\text{comp},n} = [\mathbf{K}_c^{-1} \widehat{\mathbf{K}}_1^* \widehat{\mathbf{K}}_2^{-1}]_{nn} r_{\text{imb},n} - [\mathbf{K}_c^{-1}]_{nn} r_{\text{imb},n}^*$$

which means that the proposed compensation scheme can be implemented parallel on each antenna element and, thus, it has very low complexity implementation.

VI. PERFORMANCE ANALYSIS

In this section, we derive a tractable analytical approximation of the ergodic SE and provide the power scaling laws for MRC receivers for the case with IQI but without compensation.⁶ Finally, we show that the SE loss due to IQI reaches a saturation point, as the number of BS antennas grows without bound.

Proposition 4: The exact SE (bits/s/Hz) of the k th MS for MRC receivers, R_k^{mrc} , after employing the proposed compensation scheme, is approximated by

$$R_k^{\text{wic}} \approx \log_2 \left(1 + \frac{\rho_u N \beta_k}{\rho_u \sum_{i=1, i \neq k}^K \beta_i + \rho_u \varepsilon_y + \varepsilon_\psi + 1} \right) \quad (38)$$

where $\varepsilon_y \triangleq \sum_{i=1}^K \sigma_i^2$ and

$$\varepsilon_\psi \triangleq 2(2\rho_u \beta_k + \rho_u \sum_{i \neq k} \beta_i + \frac{\sigma_k^2}{2} + 1) \bar{G},$$

where $\bar{G} \triangleq \frac{1}{N} \sum_{n=1}^N \left(\frac{\sigma_{\kappa_n, i}}{g_n \cos(\phi_n)} \right)^2$.

Proof: See Appendix IV. \blacksquare

Note that the expression in (38) is an approximation, which depends on the inter-user interference, the channel estimation error ε_y and the interference ε_ψ , which stems from the IQI coefficients' estimation error. We now turn our attention to the power scaling laws and provide the following simplified expressions for the case with IQI but without compensation.

Proposition 5: Assume that the BS has perfect CSI and that the transmit power of each MS is scaled with N according to $\rho_u = \frac{E_u}{N}$, where E_u is fixed. Then,

$$R_{k,P}^{\text{iqi}} \rightarrow \log_2 (1 + \beta_k E_u \lambda_P), \quad N \rightarrow \infty \quad (39)$$

⁶Henceforth, the superscript *ideal* stands for the ideal case without IQI, while the superscripts *wic* and *iqi* stand for the case with IQI employing the proposed compensation scheme and the one with IQI but without compensation respectively.

where $\lambda_P \triangleq \frac{(\frac{1}{N} \sum_{i=1}^N |\mathbf{K}_1|_n|^2)^2}{\frac{1}{N} \sum_{i=1}^N (|\mathbf{K}_1|_n|^4 + |\mathbf{K}_2|_n|^2)}$ is a bounded constant, depending on the IQI parameters. Notice that, when the IQ branches are perfectly matched then $\lambda_P = 1$.

Proof: Assuming perfect CSI, the proof can be easily derived from Appendix V. ■

Proposition 5 shows that with perfect CSI and the presence of IQI at the BS and a large N , the performance of a MU-SIMO system with N antennas at the BS and a transmit power per MS of E_u/N is equal to the performance of a single-input single output (SISO) system with transmit power $E_u \beta_k \lambda_P$.

Proposition 6: Assume that the BS has imperfect CSI, obtained by LMMSE estimation from uplink pilots, and that the transmit power of each MS is $\rho_u = \frac{E_u}{\sqrt{N}}$, where E_u is fixed. Then,

$$R_{k,IP}^{iq1} \rightarrow \log_2(1 + \tau \beta_k^2 E_u^2 \lambda_{IP}), \quad N \rightarrow \infty \quad (40)$$

where $\lambda_{IP} \triangleq \frac{(\frac{1}{N} \sum_{i=1}^N \tilde{\lambda}_n)^2}{\frac{1}{N} \sum_{i=1}^N |\mathbf{K}_1|_n|^4}$ and $\tilde{\lambda}_n \triangleq \frac{|\mathbf{K}_1|_n|^4}{|\mathbf{K}_1|_n|^2 + |\mathbf{K}_2|_n|^2}$.

Proof: See Appendix V. ■

Proposition 6 implies that with imperfect CSI and a large N -antenna array at the BS, the performance of a MU-SIMO system with IQI at the BS and the transmit power per MS to be E_u/\sqrt{N} is equal to the performance of a SISO link with transmit power $\tau \beta_k^2 E_u^2 \lambda_{IP}$. Most importantly, the above propositions showcase that massive MIMO with appropriate power scaling is resilient to IQI.

Finally, it is worth mentioning, that [15] studied a generalized Gaussian-type error model for the impact of residual hardware impairments (e.g. phase noise, quantization and IQ imbalance) at the BS. It was shown that in the massive MIMO regime residual hardware impairments vanish asymptotically. Similarly, we show, using a very different line of reasoning, that the SE loss due to IQI reaches asymptotically a saturation point which implies that massive MIMO systems are resilient to IQI. We will also show in the next section that unless proper compensation schemes are applied, the performance loss can be substantial when the number of antennas is finite.

Proposition 7: Assume that the BS has imperfect CSI (obtained by LMMSE estimation from uplink pilots), serving a finite number of MSs and that the transmit power $\rho_u = E_u$ of each MS is fixed. Then

$$\begin{aligned} \Delta R_{k,IP} &\triangleq R_{k,IP}^{ideal} - R_{k,IP}^{iq1} \\ &\rightarrow o(1), \quad N \rightarrow \infty. \end{aligned} \quad (41)$$

Proof: Proposition 7 can be easily derived by using the expression (67), the fact that $\log \alpha - \log \beta = \log \frac{\alpha}{\beta}$ and the Lindeberg-Lévy central limit theorem [4, Eq. (5)]. ■

Proposition 7 highlights that the SE loss due to IQI is asymptotically independent of the number of BS antennas. Furthermore, notice that $R_{k,IP}^{ideal} \rightarrow \infty$, in a logarithmic fashion, as the number of BS antennas grow without bound. Therefore, Proposition 7 implies that the ratio $\Delta R_{k,IP}/R_{k,IP}^{ideal} \rightarrow 0$, while $N \rightarrow \infty$. This result is of particular importance, since it shows that the effect of IQI becomes negligible by increasing the number of antennas. However, as will be shown via simulation, this property will kick in for a very large number of antennas.

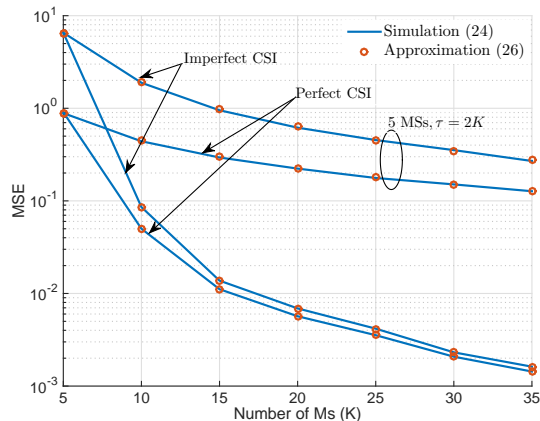


Fig. 2. MSE of \mathbf{K}_2 as a function of the number of MSs, K with $\tau = 2K$. In this example, the number of antennas is $N = 100$, the amplitude mismatch is $g = 1.2$, the transmit power per MS is $\rho_u = 10$ dB and the propagation channel parameters are $\sigma = 10$ dB and pathloss $\alpha = 3.8$.

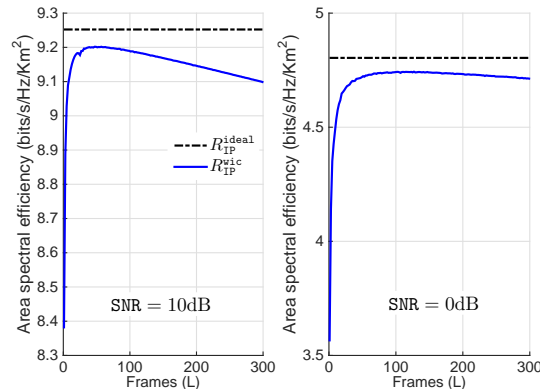


Fig. 3. Area spectral efficiency of MRC receivers as a function of frames L . In this example, $N = 20$, the number of frames that IQI remains invariant over is $B = 1000$, the propagation channel parameters are $\sigma = 10$ dB and pathloss $\alpha = 3.8$.

VII. NUMERICAL RESULTS

The analytical results are corroborated in this section by studying the uplink in a single-cell simulation scenario. We consider a circular cell with a radius of 1000m and assuming uniform MSs distribution, as well as, that no user is closer to the BS than 100m. The path loss exponent is $\alpha = 3.8$, the variance of the shadowing parameter is $\sigma^2 = 10$ dB and the phase mismatch is $\phi_n = 18^\circ$ ($\forall n = 1, \dots, N$) [20]. Hereafter, we assume identical amplitude mismatch across all antenna elements, i.e. $g_n = g \geq 1$, $\forall n = 1, \dots, N$. Furthermore, we assume that the transmitted data are modulated with OFDM. Here, we choose parameters that resemble those of the LTE standard, i.e., $T = 196$ is the number of OFDM symbols in a 1ms coherence interval [4]. Moreover, we assume that IQI coefficients remain invariant for a period of 1s, i.e., $B = 1000$ coherence intervals. Finally, we assume that the noise variance $\sigma_w^2 = 1$, such that $\rho_u \triangleq \text{SNR}$ can be defined as the normalized transmit SNR.

First, the IQI estimator in (22) is evaluated in Fig. 2. Note that the curves are based on the Monte-Carlo simulation of the expectation in (24), while the marker symbols correspond to

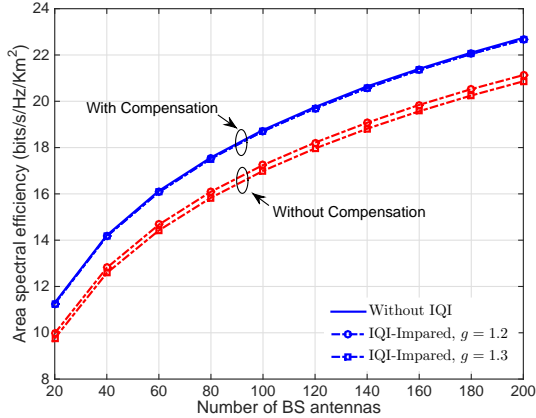


Fig. 4. Area spectral efficiency of MRC receivers with different estimation error of the IQI coefficients. In this example, $K = 10$, $\rho_u = 16$ dB and the propagation channel parameters are $\sigma = 10$ dB and pathloss $\alpha = 3.8$.

the analytical approximation in Proposition 3. It is observed that the approximation agrees with simulations in all cases.

We compare the MSE of a scheme with a fixed number of MSs, $K = 5$, and a scheme where the number of MSs is variable with $K \geq 5$, for perfect and imperfect CSI. Notice that the higher the number of MSs, the more training symbols need to be transmitted ($\tau \geq 2K$). Therefore, in order to compare these two schemes, we assume that both allocate the same number of pilots ($\tau = 2K$). In Fig. 2, we see that increasing the number of MSs, the MSE of $\hat{\mathbf{K}}_2$ is reduced faster than the scheme where the number of MSs is fixed $K = 5$. This behavior can be easily explained from the analytical expression in (26), by noticing that the summation decays exponentially with the number of MSs.

Moreover, we have seen in Section IV-B that the propagation channel estimation is of particular importance for the performance of the IQI estimator. Fig. 2 also depicts the MSE of $\hat{\mathbf{K}}_2$ considering both perfect and imperfect CSI of \mathbf{Y} . We see that for the scheme with higher number of MSs ($K \geq 10$), the MSE with imperfect CSI converges faster to that with perfect CSI. These results validate the importance of the number of MSs on the efficiency of the IQI estimator and show that the IQI estimator can be very efficient in a MU-SIMO topology where tens of users are served.

Next, we illustrate the proposed IQI compensation scheme utilizing the uplink transmission protocol in Fig. 1, and compare its performance with that of the case without IQI and the one with IQI but without compensation. On this basis, we define the total area SE (measurable in bits/s/Hz/km²) over B total frames, where IQI is invariant:

$$R_{\text{IP}}^{\text{wic}} = \left(\frac{T-K}{T} - \frac{LK}{BT} \right) \sum_{i=1}^K R_k^{\text{wic}}, \quad R_{\text{IP}}^A = \frac{T-K}{T} \sum_{i=1}^K R_{k,\text{IP}}^A$$

where $\mathcal{A} \triangleq \{\text{ideal}, \text{iqi}\}$ and T is one frame in symbols. Notice that due to the double length of pilots during the estimation of IQI coefficients, the SE for the case with IQI compensation is reduced by $\frac{LK}{BT}$ compared with the other ones.

Fig. 3, shows for different values of SNR the SE as a function of the frames L that the IQI coefficients' estimator

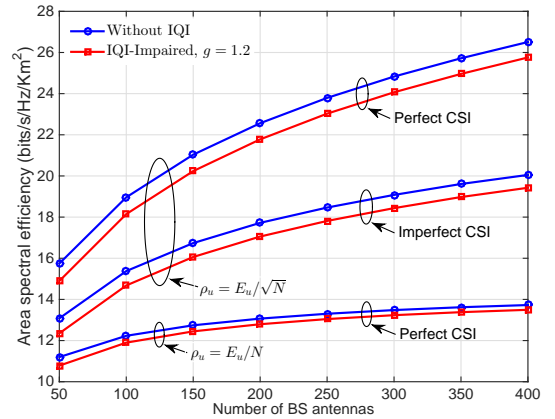


Fig. 5. Area spectral efficiency versus the number of BS antennas N for MRC processing at the receiver for the case without IQI and the one with IQI but without compensation. In this example, $K = 10$ MS are served simultaneously, the reference transmit power is $E_u = 20$ dB, and the propagation parameters is $\sigma = 10$ dB and $\alpha = 3.8$.

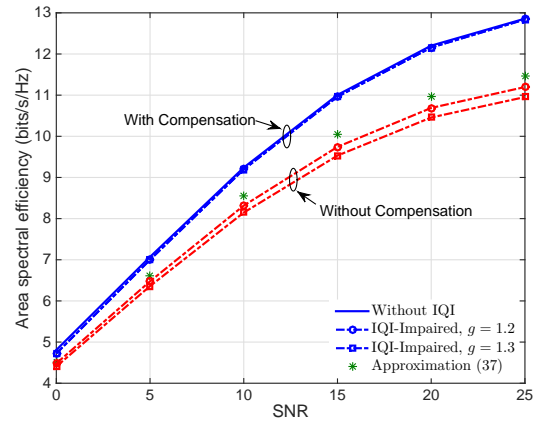


Fig. 6. Area spectral efficiency of MRC receivers as a function of SNR. In this example, $N = 20$, and the propagation channel parameters are $\sigma = 10$ dB and pathloss $\alpha = 3.8$.

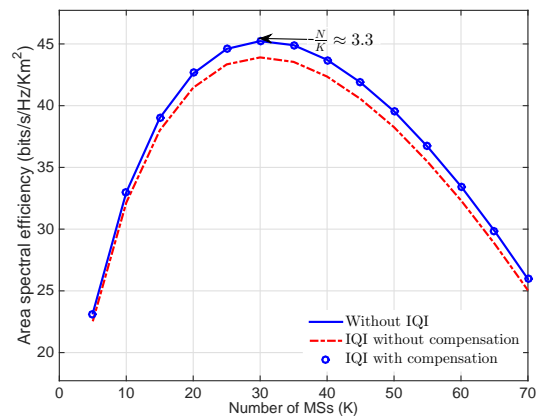


Fig. 7. Area spectral efficiency as a function of the number of MSs. In this example, the number of BS antennas is $N = 100$, the amplitude mismatch is $g = 1.4$, the transmit power per MS is $\rho_u = 16$ dB and the propagation channel parameters are $\sigma = 10$ dB and pathloss $\alpha = 3.8$.

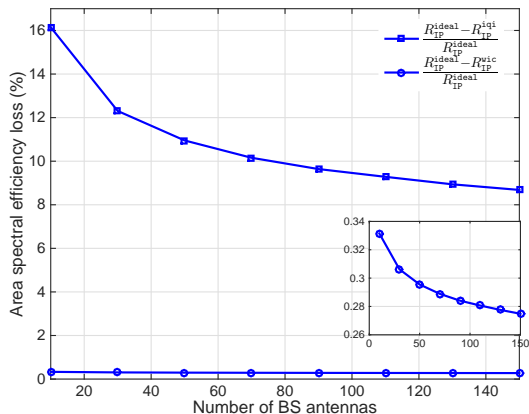


Fig. 8. Impact of IQI as a function of the number of BS antennas. In this example $K = 10$ MS are served simultaneously, the reference transmit power is $E_u = 16$ dB, the amplitude mismatch is $g = 1.3$ and the propagation parameters are $\sigma = 10$ dB and $\alpha = 3.8$.

utilizes. The IQI coefficients are estimated based on the expressions (29) and (30). As expected, there is a trade-off between the SE and the precision of the estimates \mathbf{K}_1 and \mathbf{K}_2 . Clearly, the lower SNR the more iterations are needed in order to achieve the highest SE. It is important to mention that for both cases of SNR the number of iterations needed to achieve the optimal SE is much lower than the total number of frames B where IQI remains invariant over. As a result, the SE loss of the compensation scheme due to IQI estimation is negligible, i.e. $\frac{LK}{BT} \approx 0$.

In Fig. 4, assuming imperfect CSI, we observe that, the performance degradation caused by the IQI highlights the need for compensation schemes. Notice that, by increasing the amplitude mismatch the proposed IQI compensation scheme yields almost the same SE. Intuitively, as we have shown in Proposition 4, after performing the compensation scheme, the achievable rate per MS is mainly degraded by the interference ε_ψ . Particularly, ε_ψ depends on the large-scale fading coefficients β_i ($i = 1, \dots, K$) and the bounded constant \bar{G} . By increasing the amplitude mismatch (and given that $-\pi/2 < \phi < \pi/2$), and assuming that the IQI coefficients' error is given by (26), \bar{G} will decrease converging asymptotically ($g \rightarrow \infty$) to $\frac{\gamma}{\cos^2(\phi)}$, where γ is a constant, which implies that the compensation scheme is intimately resilient to amplitude mismatches.

We next illustrate the power scaling laws. Fig. 5 compares the SE of the case without IQI and the one with IQI but without compensation versus the number of BS antennas. The power scaling of the MSs is $\rho_u = E_u/N$ and $\rho_u = E_u/\sqrt{N}$, for perfect and imperfect CSI at the receiver, respectively. As Proposition 5 implies, we see that the spectral efficiency of the case with IQI but without compensation converges to the asymptotic expression in (39). Interestingly, the effect of the IQI is significantly suppressed by increasing the number of BS antennas. Assuming perfect CSI with $\rho_u = E_u/\sqrt{N}$, the SE of both cases grows without bound when $N \rightarrow \infty$. On the other hand, assuming imperfect CSI, the SE of the case with IQI but without compensation converges slowly to the expression in (40). In this case, notice that the effect of IQI is

reduced slowly with the number of BS antennas (theoretically \sqrt{N} times slower than the convergence in (39)). These results confirm that when we can scale down the transmitted power of each user as $\rho_u = E_u/N$ for the perfect CSI case, and as $\rho_u = E_u/\sqrt{N}$ for the imperfect CSI case when N is large, then the interference due to the IQI is suppressed.

Fig. 6 demonstrates the SE as function of the SNR for different IQI coefficients in the cases with and without compensation. It can be observed that for low SNR, the compensation scheme cannot completely mitigate the residual IQI effects. Nevertheless, for low SNR the noise dominates the IQI and thus, the SE loss is negligible. However, for high SNR, the high SE loss (up to 13% when $g = 1.3$) highlights the need for IQI compensation. In this case, the proposed compensation scheme effectively suppress the residual IQI effects and it is easily observed that the SE after applying the compensation scheme almost cancels the effect of the IQI. Finally, we see that the approximation (38) is very tight with the exact SE for low and mid SNR.

In general, the number of BS antennas, N , is fixed in a deployment and not a variable, while the number of MSs, is the actual design parameter. A scheduling algorithm decides how many terminals are admitted in a certain coherence block, with the goal of maximizing some performance metric. We assume that the sum SE is the metric considered in the scheduler. Ignoring the large-scale fading, Fig. 7 shows this metric as a function of the number of scheduled MSs for a massive MIMO deployment with $N = 100$ BS antennas. The operating points that maximize the performance are marked and the corresponding values of the ratio N/K are indicated. We can easily observe that the optimized operating point is the same for both cases without IQI and the one with IQI but without compensation. We can see that the optimized operating point is the range $N/K < 10$ and thus, it is possible to let N and K be at the same order of magnitude for this particular scenario. Interestingly, the highest SE loss occurs at the optimal operating point. This can be explained by noticing from (66) that the higher the number of users the higher the interference due to IQI effect. However, by increasing the number of users beyond the optimal operating point, the overhead due to the length of pilots is dominant, and, thus, the effect of IQI becomes negligible.

Notice, that when the number of MSs is high, i.e. $K > 10$, the proposed compensation scheme curve coincides with the curve corresponding to no IQI; this implies that the IQI estimator of \mathbf{K}_1 and \mathbf{K}_2 become nearly perfect.

We finally demonstrate the importance of Proposition 7. Fig. 8 compares the SE loss of the case with IQI but without compensation and the one employing the proposed compensation scheme. In consistence with the observations in Fig. 4, we see that the proposed compensation scheme is not able to fully compensate for the total SE loss, however, its performance loss is negligible (less than 0.35%). Moreover, as Proposition 7 implies, notice that the IQI has diminishing impact by increasing the number of BS antennas.

To sum up, we can conclude that for a very large number of antennas the effect of IQI is reduced in a logarithmic fashion; yet, for practical number of antennas, the performance

degradation can be approximately as high as 16% (when the number of antennas is $N = 10$). Therefore, if we desire to achieve a SE loss less than 0.35%, then it is necessary to utilize the proposed compensation scheme.

VIII. CONCLUSIONS

In this paper, we analyzed the impact of IQI impairments at the BS by studying the uplink of a single-cell massive MIMO system. In particular, we derived a new LMMSE channel estimator for the IQI-impaired model and show that IQI can downgrade the performance of MU-SIMO systems. Next, we introduced and analyzed a novel pilot-based joint estimator of the propagation MIMO channel and IQI coefficients at the receiver. We showed that in a MU-SIMO regime, where tens of users are served simultaneously, the IQI estimator is very efficient. Moreover, we proposed a low-complexity IQI compensation scheme which is based only on the estimation of the IQI coefficients, and thus, its implementation can be done offline and individually on each antenna element. Based on this scheme, it was proved that IQI can be eliminated at the receiver as long as the IQI coefficients are perfectly known. Furthermore, we investigated how many MS should be scheduled in massive MIMO systems with IQI to maximize the SE for a fixed N . Interestingly, we found that the highest SE loss occurs at the optimal operating point. Finally, we showed, that by increasing the number of BS antennas without bound the effect of IQI can be reduced in a logarithmic fashion. These important observations showcase the resilience of massive MIMO to IQI imperfections.

APPENDIX I

From the structure (18) of $\hat{\mathbf{A}}_n$ we have

$$[\mathbf{W}_n]_{11} = [\mathbf{W}_n]_{22} = \sum_{i=1}^{\tau} |[\hat{\mathbf{A}}]_{ni}|^2 \quad (42)$$

and

$$[\mathbf{W}_n^*]_{12} = [\mathbf{W}_n]_{21} = \sum_{i=1}^{\tau} [\hat{\mathbf{A}}]_{ni}^2. \quad (43)$$

We first prove the expression (20). Using the definition of $[\hat{\mathbf{A}}]_{ni}$ in (16), the expression in (42) can be expanded as

$$\begin{aligned} \sum_{i=1}^{\tau} |[\hat{\mathbf{A}}]_{ni}|^2 &= \sum_{i=1}^{\tau} \sum_{k=1}^K \sum_{j=1}^K [\hat{\mathbf{Y}}^*]_{nk} [\hat{\mathbf{Y}}]_{nj} [\mathbf{S}_p^*]_{ki} [\mathbf{S}_p]_{ji} \\ &= \sum_{k=1}^K \sum_{j=1}^K [\hat{\mathbf{Y}}^*]_{nk} [\hat{\mathbf{Y}}]_{nj} \sum_{i=1}^{\tau} [\mathbf{S}_p^*]_{ki} [\mathbf{S}_p]_{ji} \end{aligned} \quad (44)$$

Then, it is easy to verify from the condition (9) that

$$\sum_{i=1}^{\tau} [\mathbf{S}_p^*]_{ki} [\mathbf{S}_p]_{ji} = \begin{cases} 1, & k = j \\ 0, & k \neq j \end{cases} \quad (45)$$

and therefore

$$\sum_{i=1}^{\tau} |[\hat{\mathbf{A}}]_{ni}|^2 = \sum_{k=1}^K |[\hat{\mathbf{Y}}^*]_{nk}|^2.$$

Similarly, it can be proved that the expression (43) equals (21).

APPENDIX II

Substituting the received signal (17) into the LS estimation of $\boldsymbol{\kappa}_n$ in (19), it is easy to see that the estimation error $\mathbf{n}_k \triangleq \hat{\boldsymbol{\kappa}}_n - \boldsymbol{\kappa}_n$, is given by

$$\begin{aligned} \mathbf{n}_k &= (\hat{\mathbf{A}}_n^H \hat{\mathbf{A}}_n)^{-1} \hat{\mathbf{A}}_n^H \mathbf{E}_n \boldsymbol{\kappa}_n \\ &\quad + \frac{1}{\sqrt{\tau} \rho_u} (\hat{\mathbf{A}}_n^H \hat{\mathbf{A}}_n)^{-1} \hat{\mathbf{A}}_n^H \tilde{\mathbf{w}}_n. \end{aligned}$$

Next, the expression of the \mathbf{n}_k is simplified by using the proposition 3 and providing the following corollary.

Corollary 1: Let $\mathbf{C}_n \triangleq \hat{\mathbf{A}}_n^H \mathbf{E}_n \in \mathbb{C}^{2 \times 2}$. Assuming that the uplink pilots sequence \mathbf{S} satisfies the condition (9) then,

$$[\mathbf{C}_n]_{11} = [\mathbf{C}_n^*]_{22} = \sum_{k=1}^K [\hat{\mathbf{Y}}]_{nk} [\mathbf{E}_Y^*]_{nk} \quad (46)$$

and

$$[\mathbf{C}_n]_{12} = [\mathbf{C}_n]_{21} = 0 \quad (47)$$

Proof: The proof is similar with Proposition's 3. ■
Then,

$$\mathbf{n}_k = \frac{\mathbf{C}_n \boldsymbol{\kappa}_n}{[\mathbf{W}_n]_{11}} + \frac{\hat{\mathbf{A}}_n^H \tilde{\mathbf{w}}_n}{\sqrt{\tau} \rho_u [\mathbf{W}_n]_{11}}. \quad (48)$$

The IQI coefficients' estimation error can be written as $\mathbf{n}_k = [\mathbf{n}_{k,1} \ \mathbf{n}_{k,2}]^T$, where $\mathbf{n}_{k,1}$ and $\mathbf{n}_{k,2}$ denote the estimation error of the IQI coefficients $K_{1,n}$ and $K_{2,n}$ respectively.

Therefore, the estimation error can be expressed as

$$\mathbf{n}_{k,i} = \frac{1}{[\mathbf{W}_n]_{11}} \left([\mathbf{C}_n]_{ii} K_{i,n} + \frac{[\hat{\mathbf{A}}_n^H \tilde{\mathbf{w}}_n]_i}{\sqrt{\tau} \rho_u} \right), \quad i = 1, 2 \quad (49)$$

where $[\hat{\mathbf{A}}_n^H \tilde{\mathbf{w}}_n]_i$ denotes the i th element of the column vector $\hat{\mathbf{A}}_n^H \tilde{\mathbf{w}}_n \in \mathbb{C}^{2 \times 1}$.

Given that \mathbf{A}_n , \mathbf{E}_n and $\tilde{\mathbf{w}}_n$ are uncorrelated, the MSE of the estimated $K_{i,n}$ ($i = 1, 2$) is given by

$$\begin{aligned} \sigma_{\kappa_i}^2 &= \mathbb{E}\{ \|\boldsymbol{\kappa}_{n,i} - \hat{\boldsymbol{\kappa}}_{n,i}\|^2 \} \\ &= \mathbb{E}\{ \mathbf{n}_{k,i} \mathbf{n}_{k,i}^* \} \\ &= \mathbb{E} \left\{ |K_{i,n}|^2 \frac{|[\mathbf{C}_n]_{11}|^2}{|[\mathbf{W}_n]_{11}|^2} \right\} + \mathbb{E} \left\{ \frac{\|\boldsymbol{\kappa}_n\|^2}{\tau \rho_u |[\mathbf{W}_n]_{11}|} \right\} \end{aligned} \quad (50)$$

$$= \mathbb{E} \left\{ |K_{i,n}|^2 \frac{\sum_{k=1}^K \sigma_k^2 |[\mathbf{Y}]_{nk}|^2}{\left(\sum_{k=1}^K |[\mathbf{Y}]_{nk}|^2 \right)^2} \right\} \quad (51)$$

$$+ \mathbb{E} \left\{ \frac{\|\boldsymbol{\kappa}_n\|^2}{\tau \rho_u \sum_{k=1}^K |[\mathbf{Y}]_{nk}|^2} \right\} \quad (52)$$

where the channel estimation error \mathbf{E}_Y and the thermal noise $\tilde{\mathbf{w}}_n$ are averaged out correspondingly. Finally, the MSE in (23) can be upper bounded by

$$\sigma_{\kappa_i}^2 \leq \alpha_i \mathbb{E} \left\{ \frac{1}{\sum_{k=1}^K |[\hat{\mathbf{Y}}]_{nk}|^2} \right\}$$

APPENDIX III

For ease of exposition we define the random variable $Q \triangleq \sum_{k=1}^K |[\mathbf{Y}]_{nk}|^2$. Then, Q has a generalized chi-squared distribution given as [29]

$$f(x; K, \beta_1, \dots, \beta_K) = \sum_{i=1}^K \frac{e^{-\frac{x}{\beta_i}}}{\beta_i \prod_{j=1, j \neq i}^K (1 - \frac{\beta_j}{\beta_i})}, \quad x \geq 0.$$

Therefore, the expectation of the inverse Q is given by

$$\begin{aligned} \mathbb{E} \left\{ \frac{1}{Q} \right\} &= \int_0^\infty \frac{1}{Q} f(Q; K, \beta_1, \dots, \beta_K) dQ \\ &= \sum_{i=1}^K \frac{1}{\beta_i \prod_{j=1, j \neq i}^K (1 - \frac{\beta_j}{\beta_i})} \underbrace{\int_0^\infty \frac{e^{-\frac{Q}{\beta_i}}}{Q} dQ}_{\Gamma(0,0)}. \end{aligned} \quad (53)$$

Notice that the incomplete Gamma function is not defined when $Q = 0$, since $\Gamma(0,0) \rightarrow \infty$. However, we practically have that $Q > 0$ and therefore, a constant α will be defined next such that $Q \geq \alpha$ and the cumulative distribution function (CDF) of Q , $F_Q(\cdot)$, satisfies the condition: $F(Q \leq \alpha) \approx 0$. Particularly, we want to have that

$$\begin{aligned} F_Q(Q \leq \alpha) &= \int_0^\alpha f(Q; K, \beta_1, \dots, \beta_K) dQ \\ &= \sum_{i=1}^K \frac{1 - e^{-\frac{\alpha}{\beta_i}}}{\prod_{j=1, j \neq i}^K (1 - \frac{\beta_j}{\beta_i})} \\ &\approx 0. \end{aligned} \quad (54)$$

By choosing a constant $\alpha > 0$ such that $\frac{\alpha}{\beta_i} \ll 1$, then $e^{-\frac{\alpha}{\beta_i}} \approx 1 - \frac{\alpha}{\beta_i}$. Substituting it into (54) we get

$$F_Q(Q \leq \alpha) \approx \alpha \sum_{i=1}^K \frac{1}{\underbrace{\prod_{j=1, j \neq i}^K (1 - \frac{\beta_j}{\beta_i})}_{\stackrel{[29]}{\approx} 0}} = 0.$$

Therefore, to ensure that $\frac{\alpha}{\beta_i} \ll 1$, α must be given by

$$\alpha = c\beta_{\min}, \quad c \ll 1$$

where c is an arbitrary constant and $\beta_{\min} \triangleq \min_{1 \leq k \leq K} \{\beta_i\}$, and thus, the expression (53) can be approximated as

$$\begin{aligned} \mathbb{E} \left\{ \frac{1}{Q} \right\} &\approx \sum_{i=1}^K \frac{1}{\beta_i \prod_{j=1, j \neq i}^K (1 - \frac{\beta_j}{\beta_i})} \int_\alpha^\infty \frac{e^{-\frac{Q}{\beta_i}}}{Q} dQ \\ &= \sum_{i=1}^K \frac{\Gamma(0, \alpha/\beta_i)}{\beta_i \prod_{j=1, j \neq i}^K (1 - \frac{\beta_j}{\beta_i})}. \end{aligned}$$

APPENDIX IV

Recalling that $\mathbf{Y} = \hat{\mathbf{Y}} + \mathbf{E}_Y$, we can rewrite the received signal in (37) as

$$\hat{\mathbf{r}}_c = \sqrt{\rho_u} \sum_{i=1}^K (\hat{\mathbf{y}}_i + \mathbf{e}_i) x_i + \mathbf{w} + \tilde{\mathbf{n}}_\psi \quad (55)$$

where $\hat{\mathbf{y}}_i$ and \mathbf{e}_i denote the i th column of the matrices $\hat{\mathbf{Y}}$ and \mathbf{E}_Y , respectively.

We denote the effective post-processing noise for the k th MS as

$$\tilde{\mathbf{n}}_k \triangleq \hat{\mathbf{y}}_k^H \hat{\mathbf{r}}_c - \sqrt{\rho_u} \|\hat{\mathbf{y}}_k\|^2 x_k \quad (56)$$

$$\begin{aligned} &= \sqrt{\rho_u} \sum_{i=1, i \neq k}^K \hat{\mathbf{y}}_k^H \hat{\mathbf{y}}_i x_i + \sqrt{\rho_u} \sum_{i=1}^K \hat{\mathbf{y}}_k^H (\mathbf{e}_i - \mathbf{G}\Psi \mathbf{y}_i) x_i \\ &+ \sqrt{\rho_u} \sum_{i=1}^K \hat{\mathbf{y}}_k^H \mathbf{G}\Psi^* \mathbf{y}_i^* x_i^* + \hat{\mathbf{y}}_k^H \mathbf{G} (\Psi^* \mathbf{w}^* - \Psi \mathbf{w}) + \hat{\mathbf{y}}_k^H \mathbf{w}. \end{aligned} \quad (57)$$

Given the estimated channel $\hat{\mathbf{y}}_k$ and the IQI estimation error, the variance, $\tilde{\sigma}_k^2 \triangleq \mathbb{E}\{\tilde{\mathbf{n}}_k \tilde{\mathbf{n}}_k^H\}$, of the post-processing noise is obtained as

$$\begin{aligned} \tilde{\sigma}_k^2 &= \rho_u \sum_{i=1, i \neq k}^K \|\hat{\mathbf{y}}_k^H \hat{\mathbf{y}}_i\|^2 + 2\rho_u \sum_{i=1, i \neq k}^K \beta_i \|\hat{\mathbf{y}}_k^H \tilde{\mathbf{G}}\|^2 \\ &+ 2\rho_u \sum_{i=1}^N [\tilde{\mathbf{G}}^2]_{ii} \|\hat{\mathbf{y}}_{ki}\|^4 + 2\|\hat{\mathbf{y}}_k^H \tilde{\mathbf{G}}\|^2 \\ &+ \sigma_k^2 \|\hat{\mathbf{y}}_k^H \tilde{\mathbf{G}}\|^2 + \rho_u \sum_{i=1}^K \sigma_i^2 \|\hat{\mathbf{y}}_k\|^2 + \|\hat{\mathbf{y}}_k\|^2 \end{aligned}$$

where $\tilde{\mathbf{G}} \triangleq \mathbf{G} \cdot \text{diag}(\sigma_{\kappa_{1,i}} \dots \sigma_{\kappa_{N,i}})$ ($i = 1, 2$). Then, the uplink achievable rate for the k th user, employing MRC receivers, is given by

$$R_k^{\text{wic}} = \mathbb{E} \left\{ \log_2 \left(1 + \frac{\rho_u \|\hat{\mathbf{y}}_k\|^4}{\tilde{\sigma}_k^2} \right) \right\}. \quad (58)$$

For MRC receivers, the exact achievable uplink rate of the k th user, R_k^{wic} , can be lower bounded as [4, Eq. (15)]: $R_k^{\text{wic}} \geq \hat{R}_k^{\text{wic}}$ where $\hat{R}_k^{\text{wic}} \triangleq \log_2 \left(1 + (\mathbb{E}\{\lambda_k^{\text{wic}}\})^{-1} \right)$ and λ_k^{wic} is the inverse signal-to-noise plus interference ratio

$$\lambda_k^{\text{wic}} \triangleq \frac{\tilde{\sigma}_k^2}{\rho_u \|\hat{\mathbf{y}}_k\|^4}.$$

The expectation of λ_k^{wic} can be evaluated as

$$\bar{\lambda}_k^{\text{wic}} \triangleq \mathbb{E} \left\{ \lambda_k^{\text{wic}} \right\} \quad (59)$$

$$= \mathbb{E} \left\{ \frac{\sum_{i=1, i \neq k}^K \|\hat{\mathbf{y}}_k^H \hat{\mathbf{y}}_i\|^2}{\|\hat{\mathbf{y}}_k\|^4} \right\} + \mathbb{E} \left\{ \frac{2 \sum_{i=1}^N [\tilde{\mathbf{G}}^2]_{ii} \|\hat{\mathbf{y}}_{ki}\|^4}{\|\hat{\mathbf{y}}_k\|^4} \right\} \quad (60)$$

$$+ \mathbb{E} \left\{ \frac{\alpha_\psi \|\hat{\mathbf{y}}_k^H \tilde{\mathbf{G}}\|^2}{\rho_u \|\hat{\mathbf{y}}_k\|^4} \right\} + \mathbb{E} \left\{ \frac{\alpha_e}{\rho_u \|\hat{\mathbf{y}}_k\|^2} \right\} \quad (61)$$

where $\alpha_e \triangleq 1 + \rho_u \sum_{i=1}^K \sigma_i^2$ and $\alpha_\psi \triangleq 2 \left(\rho_u \sum_{i \neq k} \beta_i + 1 \right) + \sigma_k^2$. It is well known [4], [30] that the expectation of the first term in (60) equals

$$\frac{\sum_{i=1, i \neq k}^K \hat{\beta}_i}{(N-1)\hat{\beta}_k^2} \quad (62)$$

while

$$\mathbb{E} \left\{ \frac{\alpha_e}{\rho_u \|\hat{\mathbf{y}}_k\|^2} \right\} = \frac{\alpha_e}{\rho_u \hat{\beta}_k (N-1)}. \quad (63)$$

Now, let $t_k \triangleq \|\hat{\mathbf{y}}_k^H \mathbf{G}\|^2 / \|\hat{\mathbf{y}}_k\|^4$. Then, the first term in (61) can be written as

$$\mathbb{E} \left\{ \frac{\alpha_\psi}{\rho_u} t_k \right\} = \frac{\alpha_\psi}{\rho_u} \sum_{m=1}^N [|\tilde{\mathbf{G}}|^2]_{mm} \mathbb{E} \left\{ \frac{|\hat{\mathbf{y}}_{km}|^2}{\|\hat{\mathbf{y}}_k\|^4} \right\}$$

where $\hat{\mathbf{y}}_{km}$ is the m th element of the column vector $\hat{\mathbf{y}}_k$.

We define $\|\hat{\mathbf{y}}_k^{(m)}\|^2 \triangleq \sum_{i=1, i \neq m}^N |\hat{\mathbf{y}}_k|^2$. For a sufficiently large number of antennas $\|\hat{\mathbf{y}}_k^{(m)}\|^4 \approx \|\hat{\mathbf{y}}_k\|^4$ and then $|\hat{\mathbf{y}}_{km}|^2$ and $\|\hat{\mathbf{y}}_k^{(m)}\|^4$ can be considered independent. Thus, its expected value is approximated as

$$\mathbb{E} \left\{ \frac{\alpha_\psi}{\rho_u} t_k \right\} \approx \frac{\alpha_\psi}{\rho_u} \sum_{m=1}^N \hat{\beta}_k [|\tilde{\mathbf{G}}|^2]_{mm} \mathbb{E} \left\{ \frac{1}{\|\hat{\mathbf{y}}_k^{(m)}\|^4} \right\}. \quad (64)$$

Similarly, the expectation of the second term in (60) can be approximated as

$$2 \sum_{m=1}^N 2\hat{\beta}_k^2 [|\tilde{\mathbf{G}}|^2]_{mm} \mathbb{E} \left\{ \frac{1}{\|\hat{\mathbf{y}}_k^{(m)}\|^4} \right\}. \quad (65)$$

The expectation term in (64) and (65) can be written as a scalar central complex Wishart matrix with $N - 1$ degrees of freedom [30, Eq. (2.9)]

$$\mathbb{E} \left\{ \frac{1}{\|\hat{\mathbf{y}}_k^{(m)}\|^4} \right\} = \frac{1}{\hat{\beta}_k^2 (N - 2)(N - 3)}.$$

Substituting (62)-(65) into (59) and assuming $N \gg 3$, the simplified approximation in (38) is obtained.

APPENDIX V

We can rewrite the received signal in (6) as

$$\mathbf{r}_{\text{imb}} = \sqrt{\rho_u} \sum_{i=1}^K (\hat{\mathbf{y}}_{1i} + \boldsymbol{\varepsilon}_i) x_i + \sqrt{\rho_u} \sum_{i=1}^K \mathbf{y}_{2i} x_i^* + \tilde{\mathbf{w}} \quad (66)$$

where $\mathbf{y}_{ik} \triangleq \mathbf{K}_i \mathbf{y}_k$ and $\boldsymbol{\varepsilon}_i$ denotes the effective channel estimation error. Therefore, the effective post-processing noise for the k th MS is given by

$$\begin{aligned} \hat{\mathbf{n}}_k &\triangleq \hat{\mathbf{y}}_{1k}^H \mathbf{r}_{\text{imb}} - \sqrt{\rho_u} \|\hat{\mathbf{y}}_{1k}\|^2 x_k \\ &= \sqrt{\rho_u} \sum_{i=1, i \neq k}^K \hat{\mathbf{y}}_{1k}^H \hat{\mathbf{y}}_{1i} x_i + \sqrt{\rho_u} \sum_{i=1}^K \hat{\mathbf{y}}_{1k}^H \boldsymbol{\varepsilon}_i x_i \\ &\quad + \sqrt{\rho_u} \sum_{i=1}^K \hat{\mathbf{y}}_{1k}^H \mathbf{K}_2 \mathbf{y}_i^* x_i^* + \hat{\mathbf{y}}_{1k}^H \tilde{\mathbf{w}}. \end{aligned}$$

Given the fact that the estimated channel $\hat{\mathbf{y}}_{1k}$ and its estimation error $\boldsymbol{\varepsilon}_i$ are uncorrelated, it can be shown that they are also statistically independent. Therefore, the variance, $\hat{\sigma}_k^2 \triangleq \mathbb{E}\{\hat{\mathbf{n}}_k \hat{\mathbf{n}}_k^*\}$, of the post-processing noise is obtained as

$$\begin{aligned} \hat{\sigma}_k^2 &= \rho_u \sum_{i=2}^K \|\hat{\mathbf{y}}_{1k}^H \hat{\mathbf{y}}_{1i}\|^2 + \rho_u \hat{\mathbf{y}}_{1k}^H \left(\sum_{i=1}^K \boldsymbol{\varepsilon}_i \right) \hat{\mathbf{y}}_{1k} \\ &\quad + \rho_u \sum_{i=1}^K \|\hat{\mathbf{y}}_{1k}^H \mathbf{K}_2 \mathbf{y}_i^*\|^2 + \|\hat{\mathbf{y}}_{1k}^H \mathbf{K}_1\|^2 + \|\hat{\mathbf{y}}_{1k}^H \mathbf{K}_2\|^2 \end{aligned}$$

where $\boldsymbol{\varepsilon}_i \triangleq \text{diag}(\tilde{\sigma}_{1,i}^2, \dots, \tilde{\sigma}_{N,i}^2)$. Then, the uplink achievable rate for the k th user is given by

$$R_{k,\text{IP}}^{\text{iqi}} = \mathbb{E} \left\{ \log_2 \left(1 + \frac{\rho_u \|\hat{\mathbf{y}}_{1k}\|^4}{\hat{\sigma}_k^2} \right) \right\}. \quad (67)$$

Substituting $p_u = E_u / \sqrt{N}$ and using the fact that each element of $\hat{\mathbf{y}}_{1k}^H$ is a RV with zero mean and variance $\hat{\beta}_{1nk}$ given by (11) along with

$$\frac{1}{N} \hat{\mathbf{y}}_{1k}^H \hat{\mathbf{y}}_{1k} \rightarrow \frac{1}{N} \sum_{n=1}^N \hat{\beta}_{1nk}$$

and

$$\frac{1}{N} \hat{\mathbf{y}}_{1k}^H \hat{\mathbf{y}}_{1k}^* \rightarrow 0, \quad \frac{1}{N} \hat{\mathbf{y}}_{1k}^H \hat{\mathbf{y}}_{1i} \rightarrow 0$$

as $N \rightarrow \infty$, we obtain Proposition 6.

REFERENCES

- [1] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–46, Jan. 2013.
- [2] 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.
- [3] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [5] 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.
- [6] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 172–179, Feb. 2013.
- [7] A. Pitarokoilis, S. K. Mohammed, and E. G. 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.
- [8] A. Hakkarainen, J. Werner, K. Dandekar, and M. Valkama, "Widely-linear beamforming and RF impairment suppression in massive antenna arrays," *J. Commun. Netw.*, vol. 15, no. 4, pp. 383–397, Aug. 2013.
- [9] Q. Zhang, S. Jin, K. K. Wong, H. Zhu, and M. Matthaiou, "Power scaling of uplink massive MIMO systems with arbitrary-rank channel means," *IEEE J. Sel. Areas Commun.*, vol. 8, no. 5, pp. 966–981, Oct. 2014.
- [10] B. M. Hochwald, T. L. Marzetta, and V. Tarokh, "Multiple-antenna channel hardening and its implications for rate feedback and scheduling," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1893–1909, Sept. 2004.
- [11] A. Tarighat, R. Bagheri, and A. H. Sayed, "Compensation schemes and performance analysis of IQ imbalances in OFDM receivers," *IEEE Trans. Signal Process.*, vol. 53, no. 8, pp. 3257–3268, Aug. 2005.
- [12] L. Anttila, M. Valkama, and M. Renfors, "Frequency-selective IQ mismatch calibration of wideband direct-conversion transmitters," *IEEE Trans. Circuits Syst. II, Exp. Briefs.*, vol. 55, no. 4, pp. 359–363, Apr. 2008.
- [13] M. Valkama, M. Renfors, and V. Koivunen, "Advanced methods for IQ imbalance compensation in communication receivers," *IEEE Trans. Signal Process.*, vol. 49, no. 10, pp. 2335–2344, Oct. 2001.
- [14] A. Hakkarainen, J. Werner, K. Dandekar, and M. Valkama, "Precoded massive MU-MIMO uplink transmission under transceiver IQ imbalance," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 320–326.
- [15] 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.
- [16] W. Zhang, R. C. de Lamare, and M. Chen, "Reduced-rank widely linear precoding in massive MIMO systems with IQ imbalance," in *Proc. EUSIPCO*, Sep. 2014, pp. 331–335.

- [17] S. Zarei, W. Gerstacker, and R. Schober, "I/Q imbalance aware widely-linear precoding for downlink massive MIMO systems," in *Proc. IEEE GLOBECOM*, Dec. 2014, pp. 301–307.
- [18] S. Zarei, W. H. Gerstacker, J. Aulin, and R. Schober, "I/Q imbalance aware widely-linear receiver for uplink multi-cell massive MIMO systems," [Online]. Available: <http://arxiv.org/abs/1506.06822>
- [19] J. Tubbax, B. Come, L. Van der Perre, S. Donnay, M. Moonen, and H. De Man, "Compensation of transmitter IQ imbalance for OFDM systems," in *Proc. IEEE ICASSP*, vol. 2, May 2004, pp. II–325–328.
- [20] T. Schenk, *RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation*. Springer Netherlands, 2008.
- [21] J. Li, M. Matthaiou, and T. Svensson, "I/Q imbalance in two-way AF relaying," *IEEE Trans. Commun.*, vol. 62, no. 7, pp. 2271–2285, July 2014.
- [22] S. W. Kim and Z. Wang, "Maximum ratio diversity combining receiver using single radio frequency chain and single matched filter," in *Proc. IEEE GLOBECOM*, Nov. 2007, pp. 4081–4085.
- [23] R. A. Horn and C. R. Johnson, *Topics in Matrix Analysis*. Cambridge University Press, 1991.
- [24] S. M. Kay, *Fundamentals of Statistical Signal Processing, Volume 1: Estimation theory*. Prentice Hall PTR, 1993.
- [25] S. Ohno and G. B. Giannakis, "Capacity maximizing MMSE-optimal pilots for wireless OFDM over frequency-selective block Rayleigh-fading channels," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 2138–2145, Sept. 2004.
- [26] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [27] 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, May 2010.
- [28] M. Matthaiou, C. Zhong, and T. Ratnarajah, "Novel generic bounds on the sum rate of MIMO ZF receivers," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4341–4353, Sept. 2011.
- [29] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1188–1204, Mar. 2008.
- [30] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, pp. 1–182, 2004. [Online]. Available: <http://dx.doi.org/10.1561/0100000001>



Michail Matthaiou (S'05–M'08–SM'13) was born in Thessaloniki, Greece in 1981. He obtained the Diploma degree (5 years) in Electrical and Computer Engineering from the Aristotle University of Thessaloniki, Greece in 2004. He then received the M.Sc. (with distinction) in Communication Systems and Signal Processing from the University of Bristol, U.K. and Ph.D. degrees from the University of Edinburgh, U.K. in 2005 and 2008, respectively. From September 2008 through May 2010, he was with the Institute for Circuit Theory and Signal Processing, Munich University of Technology (TUM), Germany working as a Postdoctoral Research Associate. He is currently a Senior Lecturer at Queen's University Belfast, U.K. after holding an Assistant Professor position at Chalmers University of Technology, Sweden. His research interests span signal processing for wireless communications, massive MIMO, hardware-constrained communications, and performance analysis of fading channels.

Dr. Matthaiou was the recipient of the 2011 IEEE ComSoc Best Young Researcher Award for the Europe, Middle East and Africa Region and a co-recipient of the 2006 IEEE Communications Chapter Project Prize for the best M.Sc. dissertation in the area of communications. He was co-recipient of the Best Paper Award at the 2014 IEEE International Conference on Communications (ICC) and was an Exemplary Reviewer for IEEE COMMUNICATIONS LETTERS for 2010. In 2014, he received the Research Fund for International Young Scientists from the National Natural Science Foundation of China. He currently serves as Senior Editor for IEEE COMMUNICATIONS LETTERS, an Associate Editor for the IEEE TRANSACTIONS ON COMMUNICATIONS, and was the Lead Guest Editor of the special issue on "Large-scale multiple antenna wireless systems" of the IEEE JOURNAL ON SELECTED AREAS IN COMMUNICATIONS. He is the chair of the Wireless Communications Symposium (WCS) at IEEE GLOBECOM 2016. He is an associate member of the IEEE Signal Processing Society SPCOM and SAM technical committees.



Mikael Coldrey received the M.Sc. degree in applied physics and electrical engineering from Linköping University, Linköping, Sweden, in 2000 and the Ph.D. degree in electrical engineering from Chalmers University of Technology, Gothenburg, Sweden, in 2006. He joined Ericsson Research in 2006, where he is currently a Master Researcher. He has been working with 4G research and for several years with 5G research. His main research interests include advanced antenna systems, channels, models, algorithms, and millimeter-wave communi-

cations for both radio access and wireless backhaul systems. Since 2012, he has been an Adjunct Associate Professor with Chalmers University of Technology.

PLACE
PHOTO
HERE

Nikolaos Kolomvakis (S'14) received his Diploma degree (5 years) in Electronic and Computer Engineering from the Technical University of Crete, Chania, Greece in 2010 and the M.Sc. degree in Electrical Engineering and Information Technology from the Swiss Federal Institute of Technology (ETH), Zurich, Switzerland in 2012. He is currently working toward the Ph.D. degree at the Department of Signals and Systems, Chalmers University of Technology, Gothenburg, Sweden. Since 2013, he has also been working as a research intern at Ericsson Research,

Gothenburg, Sweden. His research interests are in the area of massive MIMO, with special emphasis on hardware-constrained communications.