REDUCED-RANK WIDELY LINEAR PRECODING IN MASSIVE MIMO SYSTEMS WITH I/Q IMBALANCE

Wence Zhang, Rodrigo C. de Lamare† and Ming Chen**

*National Mobile Communications Research Lab. (NCRL), Southeast University, China
†Dept. of Electronics, University of York, UK / CETUC, PUC-Rio, Brazil
Emails: wencezhang@seu.edu.cn, rcdl500@york.ac.uk, chenming@seu.edu.cn

ABSTRACT

We present reduced-rank widely linear precoding algorithms for Massive MIMO systems with I/Q imbalance (IQI). With a large number of transmit antennas, the imperfection I/Q branches at the transmitter has a significant impact on the downlink performance. We develop linear precoding techniques using an equivalent real-valued model to mitigate IQI and multiuser interference. In order to reduce the computational complexity required by the matrix inverse, a widely linear reduced-rank precoding strategy based on the Krylov subspace (KS) is devised. Simulation results show that the proposed methods work well under IQI, and the KS precoding algorithm performs almost as well as the full-rank precoder while requiring much lower complexity.

Index Terms— widely linear precoding, Krylov subspace, I/Q imbalance, Massive MIMO

1. INTRODUCTION

Multiple-Input-Multiple-Output (MIMO) systems using multiple transmit and receive antennas can dramatically increase the capacity and reliability of wireless communication systems. In this context, Massive MIMO or large scale antenna systems have drawn great research interest recently [1, 2]. By employing a large number of transmit/receive antennas at the base station (BS), Massive MIMO systems exploit the excess degrees of freedom, boosting the mobile communication systems to provide significantly improved user experience in terms of both data rates [1] and energy efficiency [3]. Massive MIMO has been considered to be an enabling technique in future 5th Generation (5G) mobile communication systems [4].

Despite the obvious advantages of Massive MIMO systems, one of the main drawbacks of using a great number of antennas concerns hardware imperfections in the radio frequency (RF) links [5]. One of the most popular MIMO transmitter structure uses the in-phase and quadrature (I/Q) modulation [6]. The in-phase (I branch) and the quadrature (Q branch) components go through two mixers with a phase difference of $90°$. The so called I/Q imbalance (IQI) stems from the gain difference, or when the phase difference is not exactly $90°$ for the two branches. These differences can be either measured at the user equipment (UE) using channel estimation techniques [7], or measured at the BS following similar ideas like channel calibration methods in [8] using a reference transmitter or receiver with perfect I/Q branches.

The IQI in RF links actually introduces a non-circular data structure which can be exploited by widely linear processing. It is well known that widely linear processing at the receiver can obtain extra gain over linear minimum mean squared error (MMSE) design for non-circular signals [9–12]. The work in [5] uses widely linear design to improve uplink performance in Massive MIMO systems with IQI, based on the minimum variance distortionless response (MVDR) criterion. In the downlink, [13] extends the work in [14] and develops widely linear MMSE transceivers. However the results in [13] are only for single user MIMO without IQI.

In contrast to prior work which focused on precoding and estimation/compensation of IQI separately, in this paper, we devise widely linear precoding methods to overcome the impact of IQI in multi-user Massive MIMO systems and devise a reduced-rank signal processing technique to reduce the computational complexity. The main contributions of this work are summarized as follows. The linear matched filter (M-F), zero-forcing (ZF), MMSE precoding schemes are studied in the presence of IQI using an equivalent real-valued signal model. Specifically, we derive the widely linear MMSE precoding scheme, the expression of which turns out to be a pre-compensation for IQI. In order to reduce the computational complexity of widely linear precoders, a reduced-rank design based on the Krylov subspace (KS) is proposed. With only a few bases in the approximation, the KS precoding has a comparable performance to the widely linear MMSE scheme under IQI.

The rest of this paper is organized as follows. The system

The work of W. Zhang and M. Chen was supported in part by China Scholarship Council, National Nature Science Foundation of China (Nos. 61372106 & 61223001 & 61172077), National 863 High Technology Development Project (No. 2013AA013601), Key Special Project of National Science and Technology (No. 2013ZX03003006), Research Fund of National Mobile Communications Research Laboratory, Southeast University of 2014. The work of R. C. de Lamare was supported in part by China Research Scholarship, CNPq and Faperj.

model is given in Section II. In Section III, the widely linear MMSE and reduced-rank precoders are proposed with consideration of IQI. The numerical results are given in Section IV and conclusions are drawn in Section V.

Notation: vectors and matrices are represented in bold lowercase and capital letters, respectively; $(\bullet)^*$ denotes the complex conjugate operation; $(\bullet)^{\mathrm{T}}$, $(\bullet)^{\mathrm{H}}$, $\mathrm{Tr}(\bullet)$ denote the transpose and Hermitian transpose, trace of a matrix, respectively; $\|\bullet\|_{\mathrm{F}}$ denotes Frobenius norm; $\mathcal{CN}(\theta, \Sigma)$ denotes the circular symmetric complex Gaussian distribution with mean θ and covariance Σ; $\mathbf{I}_K$ denotes the K by K identity matrix; $\mathrm{diag}\{a_1, \cdots, a_K\}$ denotes a K by K diagonal matrix with diagonal entries given by $a_1, \cdots, a_K$; $\mathbb{E}(\bullet)$ denotes the expectation operation.

2. SYSTEM MODEL

Consider the downlink of a single cell Massive MIMO system with K single antenna users and N transmit antennas at the base station (BS), where $N > K$ so that excess degrees of freedom can be utilized.

The downlink channel of user k is represented by $\mathbf{h}_k = [h_{k1}, \cdots, h_{kN}]$, where h_{kn} is the channel gain from the n-th antenna at BS to the k-th user. The combined channel matrix of all users is given by $\mathbf{H} = \left[\mathbf{h}_1^{\mathrm{T}}, \cdots, \mathbf{h}_K^{\mathrm{T}}\right]^{\mathrm{T}}$. Then the received signal vector $\mathbf{y} \in \mathbb{C}^{K \times 1}$ at K users is given by

$$\mathbf{y} = \mathbf{H}P(\mathbf{s}) + \mathbf{n}. \tag{1}$$

where the signal vector $\mathbf{s} = [s_1, \cdots, s_K]^{\mathrm{T}}$ and s_k is the transmitted signal for user k; $\mathbf{n} = [n_1, \cdots, n_K]^{\mathrm{T}}$ is the noise vector, $\mathbf{n} \sim \mathcal{CN}\left(\mathbf{0}, \sigma^2 \mathbf{I}_K\right)$; $P(\bullet)$ is a $\mathbb{C}^K \to \mathbb{C}^N$ mapping performed by the precoder, e.g., $P(\mathbf{s}) = \sqrt{\lambda}\mathbf{H}^{\mathrm{H}}\mathbf{s}$ for MF and λ is a normalizing factor to maintain the power constraint, and is given by

$$\lambda = \frac{P_{\mathrm{Tx}}}{\mathbb{E}\left\{\|P(\mathbf{s})\|^2\right\}}, \tag{2}$$

in which P_{Tx} is the total transmit power. The parameter λ is used throughout the paper. For different precoding schemes, it is always defined as in (2).

The signal transmitted on the n-th antenna with IQI can be modeled as $\tilde{x} = a_{n1}x + a_{n2}x^*$ [7], where x is the baseband equivalent signal under ideal I/Q matching, and $a_{n1} = \cos(\theta_n/2) + jg_n \sin(\theta_n/2)$, $a_{n2} = g_n \cos(\theta_n/2) - j\sin(\theta_n/2)$. g_n and θ_n represent the relative gain and phase mismatch for I/Q branches, respectively. The ideal scenario implies $a_{n1} = 1$ and $a_{n2} = 0$. Let $\mathbf{A}_1 = \mathrm{diag}\{a_{11}, \cdots, a_{N1}\}$, $\mathbf{A}_2 = \mathrm{diag}\{a_{12}, \cdots, a_{N2}\}$, and then with consideration of IQI, (1) changes to

$$\mathbf{y} = \mathbf{H}\mathbf{A}_1 P(\mathbf{s}) + \mathbf{H}\mathbf{A}_2[P(\mathbf{s})]^* + \mathbf{n}. \tag{3}$$

As can be seen in (3), the precoded signal is contaminated by its complex conjugate. In MIMO systems, this contamination is a phenomenon that degrades the performance in general. However, in Massive MIMO systems the large number of antenna elements and their possibly cheaper manufacturing further increases the performance degradation. Therefore it is important to design precoding methods that take IQI into account.

Remarks: 1) For scenarios where users are equipped with multiple antennas, each antenna can be treated as a virtual user; 2) Power allocation among different users can be done by appropriate selection of the covariance matrix of $\mathbf{s}$, i.e., $\mathbf{R}_{\mathbf{ss}} = \mathbb{E}\{\mathbf{s}\mathbf{s}^{\mathrm{H}}\}$;3) No specific distribution of $\mathbf{H}$ is assumed in this paper, so that the channel could have any distribution of interest.

3. WIDELY LINEAR PRECODING UNDER IQI

3.1. Linear Precoding under IQI

In this subsection, we devise linear precoding schemes under IQI using three of the most popular design criteria, i.e. M-F, ZF and MMSE. For simplicity of expression, a mapping function of $\mathbb{C}^n \to \mathbb{R}^{2n}$ and $\mathbb{C}^{m \times p} \to \mathbb{R}^{2m \times 2p}$ is defined as

$$T(\mathbf{x}) = \begin{bmatrix}\mathrm{Re}(\mathbf{x})\\ \mathrm{Im}(\mathbf{x})\end{bmatrix}, T(\mathbf{X}) = \begin{bmatrix}\mathrm{Re}(\mathbf{X}) & -\mathrm{Im}(\mathbf{X})\\ \mathrm{Im}(\mathbf{X}) & \mathrm{Re}(\mathbf{X})\end{bmatrix}, \tag{4}$$

where $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ represent the real and imaginary parts of a vector or matrix, respectively. Some properties of this transformation are summarized in Lemma 1.

Lemma 1. *The following equations hold if the corresponding matrix or vector operation is valid:*

$$\begin{aligned} T(\mathbf{AB}) &= T(\mathbf{A})T(\mathbf{B}), & T(\mathbf{A}^{-1}) &= [T(\mathbf{A})]^{-1},\\ T(\mathbf{A}+\mathbf{B}) &= T(\mathbf{A}) + T(\mathbf{B}), & T(\mathbf{A}^{\mathrm{H}}) &= [T(\mathbf{A})]^{\mathrm{H}},\\ T(\mathbf{x}+\mathbf{y}) &= T(\mathbf{x}) + T(\mathbf{y}), & T(\mathbf{Ax}) &= T(\mathbf{A})T(\mathbf{x}). \end{aligned} \tag{5}$$

Proof. See [15]. ∎

Denote the linear precoding matrix by $\mathbf{P}$. Then the received signal vector with IQI is given by

$$\mathbf{y} = \mathbf{H}\mathbf{A}_1\mathbf{P}\mathbf{s} + \mathbf{H}\mathbf{A}_2\mathbf{P}^*\mathbf{s}^* + \mathbf{n}. \tag{6}$$

Applying the transformation $T(\bullet)$ to (6) gives

$$T(\mathbf{y}) = [T(\mathbf{H}\mathbf{A}_1) + T(\mathbf{H}\mathbf{A}_2)\mathbf{E}_N]T(\mathbf{P})T(\mathbf{s}) + T(\mathbf{n}), \tag{7}$$

where

$$\mathbf{E}_N = \begin{bmatrix}\mathbf{I}_N & \mathbf{0}_N\\ \mathbf{0}_N & -\mathbf{I}_N\end{bmatrix}. \tag{8}$$

Denote $\tilde{\mathbf{y}} = T(\mathbf{y})$, $\tilde{\mathbf{H}} = T(\mathbf{H}\mathbf{A}_1) + T(\mathbf{H}\mathbf{A}_2)\mathbf{E}_N$, $\tilde{\mathbf{P}} = T(\mathbf{P})$, $\tilde{\mathbf{s}} = T(\mathbf{s})$ and $\tilde{\mathbf{n}} = T(\mathbf{n})$, and then we have $\tilde{\mathbf{y}} = \tilde{\mathbf{H}}\tilde{\mathbf{P}}\tilde{\mathbf{s}} + \tilde{\mathbf{n}}$. Therefore, by treating the real and imaginary components as independent virtual users, the precoding matrices for transmit MF, ZF and MMSE precoding are given, respectively, by [16]

$$\begin{aligned}\tilde{\mathbf{P}}_{\text{L-MF}} &= \sqrt{\lambda}\tilde{\mathbf{H}}^{\mathrm{T}},\\ \tilde{\mathbf{P}}_{\text{L-ZF}} &= \sqrt{\lambda}\tilde{\mathbf{H}}^{\mathrm{T}}(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\mathrm{T}})^{-1},\\ \tilde{\mathbf{P}}_{\text{L-MMSE}} &= \sqrt{\lambda}\tilde{\mathbf{H}}^{\mathrm{T}}(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^{\mathrm{T}} + \mu_{\mathrm{L}}\mathbf{I}_{2K})^{-1},\end{aligned} \tag{9}$$

where $\mu_{\mathrm{L}} = \frac{K\sigma^2}{P_{\mathrm{Tx}}}$.

Remarks: Note that $\mathbf{E}_N$ and $\tilde{\mathbf{H}}$ do not have a structure as in (4), so that they cannot be inversely transformed to the complex-valued representation and consequently the same holds for $\tilde{\mathbf{P}}$. This implies a direct linear design for MF, ZF and MMSE using the signal model in (6) is generally inapplicable under IQI. Therefore, for traditional linear MF, ZF and MMSE precoding, the downlink performance will be significantly affected by IQI. However, we can still use $\tilde{\mathbf{P}}$ to precode $\tilde{\mathbf{s}}$ as $\tilde{\mathbf{x}} = \tilde{\mathbf{P}}\tilde{\mathbf{s}}$, and $\tilde{\mathbf{x}}$ can be inversely transformed. Actually this takes into account the non-circularity of the data because the real and imaginary components are separately processed for $\mathbf{s}$, but we still call it 'linear' due to the signal model structure in (6).

3.2. Widely Linear Precoding under IQI

For widely linear design, we focus on precoders that have the following structure:

$$P(\mathbf{s}) = \mathbf{P}_1\mathbf{s} + \mathbf{P}_2\mathbf{s}^* = \mathbf{P}_{\mathrm{a}}\mathbf{s}_{\mathrm{a}}, \tag{10}$$

where $\mathbf{P}_{\mathrm{a}} = [\mathbf{P}_1, \mathbf{P}_2]$, $\mathbf{s}_{\mathrm{a}} = [\mathbf{s}^{\mathrm{T}}, \mathbf{s}^{\mathrm{H}}]^{\mathrm{T}}$. The widely linear precoder processes both the original signal vector and its complex conjugate. Due to the augmented structure, i.e., $\mathbf{P}_{\mathrm{a}}$ and $\mathbf{s}_{\mathrm{a}}$, this precoding method is called 'widely linear' and the corresponding MMSE and ZF design are called WL-MMSE and WL-ZF in this paper, respectively.

Proposition 1. *For WL-MMSE without IQI, the received signal vector can be represented as* $\mathbf{y} = \mathbf{H}\mathbf{P}_{\mathrm{a}}\mathbf{s}_{\mathrm{a}} + \mathbf{n}$ *and the optimal precoder is given by*

$$\begin{cases}\mathbf{P}_1 = \sqrt{\lambda}(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \mu_{\mathrm{L}}\mathbf{I}_{\mathbf{N}})^{-1}\mathbf{H}^{\mathrm{H}}\\ \mathbf{P}_2 = \mathbf{0}_{K\times K}\end{cases}. \tag{11}$$

where $\mu_{\mathrm{L}} = \frac{K\sigma^2}{P_{\mathrm{Tx}}}$. *This is exactly the linear MMSE precoder.*

Proof. The sum mean square error (MSE) $\varepsilon_{\mathrm{WL}}$ of the received signal vector is given by

$$\begin{aligned}\varepsilon_{\mathrm{WL}} &= \mathbb{E}\left\{\left\|\alpha^{-1}\mathbf{y} - \mathbf{s}\right\|^2\right\}\\ &= \mathrm{Tr}\{\mathbf{H}\tilde{\mathbf{P}}_{\mathrm{a}}\mathbf{R}_{\mathbf{s}_{\mathrm{a}}\mathbf{s}_{\mathrm{a}}}\tilde{\mathbf{P}}_{\mathrm{a}}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}} - \mathbf{H}\tilde{\mathbf{P}}_{\mathrm{a}}\mathbf{R}_{\mathbf{s}_{\mathrm{a}}\mathbf{s}} + \mathbf{R}_{\mathbf{ss}}\\ &\quad - \mathbf{R}_{\mathbf{ss}_{\mathrm{a}}}\tilde{\mathbf{P}}_{\mathrm{a}}^{\mathrm{H}}\mathbf{H}^{\mathrm{H}} + \mathrm{Tr}\{\tilde{\mathbf{P}}_{\mathrm{a}}\mathbf{R}_{\mathbf{s}_{\mathrm{a}}\mathbf{s}_{\mathrm{a}}}\tilde{\mathbf{P}}_{\mathrm{a}}^{\mathrm{H}}\}\frac{\sigma^2}{P_{\mathrm{Tx}}}\mathbf{I}_K\},\end{aligned} \tag{12}$$

where $\alpha(\alpha \neq 0)$ is a gain of the received signal over the original transmitted signal, $\tilde{\mathbf{P}}_{\mathrm{a}} = \alpha^{-1}\mathbf{P}_{\mathrm{a}}$ and $\mathbf{R}_{\mathbf{ss}} = \mathbb{E}\{\mathbf{s}\mathbf{s}^{\mathrm{H}}\}$, $\mathbf{C}_{\mathbf{ss}} = \mathbb{E}\{\mathbf{s}\mathbf{s}^{\mathrm{T}}\}$, $\mathbf{R}_{\mathbf{ss}_{\mathrm{a}}} = \mathbb{E}\{\mathbf{s}\mathbf{s}_{\mathrm{a}}^{\mathrm{H}}\}$, $\mathbf{R}_{\mathbf{s}_{\mathrm{a}}\mathbf{s}} = \mathbb{E}\{\mathbf{s}_{\mathrm{a}}\mathbf{s}^{\mathrm{H}}\}$, $\mathbf{R}_{\mathbf{s}_{\mathrm{a}}\mathbf{s}_{\mathrm{a}}} = \mathbb{E}\{\mathbf{s}_{\mathrm{a}}\mathbf{s}_{\mathrm{a}}^{\mathrm{H}}\}$.

To minimize the sum MSE under some transmit power constraint is equivalent to solving the following unconstrained optimization problem [17]

$$\tilde{\mathbf{P}}_{\mathrm{a}} = \arg\min_{\tilde{\mathbf{P}}_{\mathrm{a}}} \varepsilon_{\mathrm{WL}}. \tag{13}$$

Taking the derivative of $\varepsilon_{\mathrm{WL}}$ with respect to $\tilde{\mathbf{P}}_{\mathrm{a}}$ yields

$$\tilde{\mathbf{P}}_{\mathrm{a}} = (\mathbf{H}^{\mathrm{H}}\mathbf{H} + \mu_{\mathrm{L}}\mathbf{I}_{\mathbf{N}})^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{R}_{\mathbf{ss}_{\mathrm{a}}}\mathbf{R}_{\mathbf{s}_a\mathbf{s}_a}^{-1}, \tag{14}$$

where $\mu_{\mathrm{L}} = K\sigma^2/P_{\mathrm{Tx}}$. Because

$$\mathbf{R}_{\mathbf{s}_{\mathrm{a}}\mathbf{s}_{\mathrm{a}}} = \begin{bmatrix}\mathbf{R}_{\mathbf{ss}} & \mathbf{C}_{\mathbf{ss}}\\ \mathbf{C}_{\mathbf{ss}}^* & \mathbf{R}_{\mathbf{ss}}^*\end{bmatrix} = \begin{bmatrix}\mathbf{R}_{\mathbf{ss}_{\mathrm{a}}}\\ \mathbf{R}_{-}\end{bmatrix}, \tag{15}$$

we have

$$\begin{aligned}\tilde{\mathbf{P}}_{\mathrm{a}} &= (\mathbf{H}^{\mathrm{H}}\mathbf{H} + \mu_{\mathrm{L}}\mathbf{I}_{\mathbf{N}})^{-1}\mathbf{H}^{\mathrm{H}}\mathbf{R}_{\mathbf{ss}_{\mathrm{a}}}\mathbf{R}_{\mathbf{s}_{\mathrm{a}}\mathbf{s}_{\mathrm{a}}}^{-1}\\ &= (\mathbf{H}^{\mathrm{H}}\mathbf{H} + \mu_{\mathrm{L}}\mathbf{I}_{\mathbf{N}})^{-1}\mathbf{H}^{\mathrm{H}}[\mathbf{I}_K, \mathbf{0}_{K\times K}],\end{aligned} \tag{16}$$

which leads to (11) and λ is chosen to satisfy the power constraint in (2). ■

Remarks: Proposition 1 shows that when there is no IQI in the RF link, the optimal precoder is the linear MMSE precoder. When IQI exists, however, the optimal precoder has a widely linear structure, as shown in Proposition 2.

Proposition 2. *With IQI and assuming* $\mathbf{A}_1$ *and* $\mathbf{A}_2$ *are invertible, the optimal WL_MMSE precoder is given by*

$$\begin{cases}\mathbf{P}_1 = \sqrt{\lambda}[\mathbf{A}_1 - \mathbf{A}_2(\mathbf{A}_1^*)^{-1}\mathbf{A}_2{}^*]^{-1}\mathbf{G}\\ \mathbf{P}_2^* = \sqrt{\lambda}[\mathbf{A}_2 - \mathbf{A}_1(\mathbf{A}_2^*)^{-1}\mathbf{A}_1{}^*]^{-1}\mathbf{G}\end{cases}, \tag{17}$$

and $\mathbf{G} = (\mathbf{H}^{\mathrm{H}}\mathbf{H} + \mu_{\mathrm{L}}\mathbf{I}_N)^{-1}\mathbf{H}^{\mathrm{H}} = \mathbf{H}^{\mathrm{H}}(\mathbf{H}\mathbf{H}^{\mathrm{H}} + \mu_{\mathrm{L}}\mathbf{I}_K)^{-1}$.

Proof. Substituting (10) into (3) gives

$$\mathbf{y} = \mathbf{H}[\mathbf{A}_1\mathbf{P}_1 + \mathbf{A}_2\mathbf{P}_2^*, \mathbf{A}_1\mathbf{P}_2 + \mathbf{A}_2\mathbf{P}_1^*]\begin{bmatrix}\mathbf{s}\\ \mathbf{s}^*\end{bmatrix} + \mathbf{n}. \tag{18}$$

According to Proposition 1, the optimal $\mathbf{P}_1$ and $\mathbf{P}_2$ satisfy

$$\begin{cases}\mathbf{A}_1\mathbf{P}_1 + \mathbf{A}_2\mathbf{P}_2^* = \sqrt{\lambda}(\mathbf{H}^{\mathrm{H}}\mathbf{H} + \mu_{\mathrm{L}}\mathbf{I}_N)^{-1}\mathbf{H}^{\mathrm{H}}\\ \mathbf{A}_1\mathbf{P}_2 + \mathbf{A}_2\mathbf{P}_1^* = \mathbf{0}_{N\times K}\end{cases}. \tag{19}$$

Solving (19) gives (17). ■

Similar results can be extended to ZF precoding by setting the parameter μ_{L} to zero (high SNR regime). We simply give the results here without proof. The optimal WL-ZF precoder is given by

$$\begin{cases}\mathbf{P}_1 = \sqrt{\lambda}[\mathbf{A}_1 - \mathbf{A}_2(\mathbf{A}_1^*)^{-1}\mathbf{A}_2{}^*]^{-1}\mathbf{H}^{\mathrm{H}}(\mathbf{H}\mathbf{H}^{\mathrm{H}})^{-1}\\ \mathbf{P}_2^* = \sqrt{\lambda}[\mathbf{A}_2 - \mathbf{A}_1(\mathbf{A}_2^*)^{-1}\mathbf{A}_1{}^*]^{-1}\mathbf{H}^{\mathrm{H}}(\mathbf{H}\mathbf{H}^{\mathrm{H}})^{-1}\end{cases}, \tag{20}$$

Remarks: The widely linear precoder structure in (10) is able to cope with IQI, as shown in Proposition 2. This WL-MMSE is actually a pre-compensation for the hardware imperfection. Moreover, because $\mathbf{A}_1$ and $\mathbf{A}_2$ are diagonal matrix, the additional complexity of this widely linear precoder is very small.

Table 1. Conjugate Gradient Method of calculating $\mathbf{R}^{-1}\mathbf{s}$.

Inputs	$\mathbf{H}$, s, I, ξ, μ_L for MMSE precoder
Outputs	$\mathbf{x}$
1	Let $\mathbf{x}_0 = \mathbf{0}$, $\mathbf{r}_0 = \mathbf{s}$, $\mathbf{d}_0 = \mathbf{r}_0$;
2 For $i = 0, \cdots, I-1$, do:	
2-1	Compute $\mathbf{t} = \mathbf{H}(\mathbf{H}^\mathrm{H}\mathbf{d}_i)$ for ZF; or $\mathbf{t} = \mathbf{H}(\mathbf{H}^\mathrm{H}\mathbf{d}_i) + \mu_\mathrm{L}\mathbf{d}_i$ for MMSE;
2-2	Compute $\alpha_i = \frac{\mathbf{r}_i^\mathrm{H}\mathbf{r}_i}{\mathbf{d}_i^\mathrm{H}\mathbf{t}}$;
2-3	Update estimates $\mathbf{x}_{i+1} = \mathbf{x}_i + \alpha_i\mathbf{d}_i$;
2-4	Update residual $\mathbf{r}_{i+1} = \mathbf{r}_i - \alpha_i\mathbf{t}$;
2-5	If $\|\mathbf{r}_{i+1}\|_\mathrm{F}^2 \leqslant \xi$, break;
2-6	Compute $\beta_i = \frac{\mathbf{r}_{i+1}^\mathrm{H}\mathbf{r}_{i+1}}{\mathbf{r}_i^\mathrm{H}\mathbf{r}_i}$;
2-7	Update search direction $\mathbf{d}_{i+1} = \mathbf{r}_{i+1} + \beta_i\mathbf{d}_i$;

Table 2. Complexity of MMSE/ZF and KS precoding.

Operation	**MMSE/ZF**	**KS**
$\mathbf{HH}^\mathrm{H}$	$K^2(8N-2)$	–
$(\bullet)^{-1} \times \mathbf{s}$	$16K^3$	$I(16NK + 46K - 2N)$
$\mathbf{H}^\mathrm{H} \times (\bullet)$	$8NK - 2N$	$8NK - 2N$

3.3. Reduced rank precoding based on Krylov subspace

It is computationally expensive to calculate the inverse of a large dimensional matrix $\mathbf{R} = \mathbf{HH}^\mathrm{H}$ for ZF, or $\mathbf{R} = \mathbf{HH}^\mathrm{H} + \mu_\mathrm{L}\mathbf{I}_K$ for MMSE. In this work, we propose Krylov subspace (KS) precoding methods to reduce the computational complexity.

The basic idea is to search for $\mathbf{x}^* = \mathbf{R}^{-1}\mathbf{s}$ with minimum error in the I-dimensional Krylov subspace with respect to $\mathbf{R}$ and $\mathbf{s}$, $\mathcal{K}_I = \mathrm{span}\{\mathbf{s}, \mathbf{Rs}, \cdots, \mathbf{R}^{I-1}\mathbf{s}\}$, such that $\mathbf{x}_I = \sum_{i=0}^{I-1} c_i\mathbf{R}^{I-1}\mathbf{s}$, where c_i 's are scalar projection coefficients of $\mathbf{x}_I$ in $\mathcal{K}_I$. The approximation error $\mathbf{e}_I$ is given by

$$\mathbf{e}_I = \mathbf{x}^* - \mathbf{x}_I = \left[\mathbf{R}^{-1} - \sum\nolimits_{i=0}^{I-1} c_i\mathbf{R}^{I-1}\mathbf{s}\right]\mathbf{s}, \qquad (21)$$

and is minimized in the matrix-norm $\|\mathbf{e}_I\|_\mathbf{R}^2 = \mathbf{e}_I^\mathrm{H}\mathbf{R}\mathbf{e}_I$. The Conjugate Gradient method in [18] is used in this work and the algorithm is summarized in Table 1.

To analyze the complexity, the FLOPs required for different precoding techniques are calculated. Each complex product counts for 6 FLOPs and complex addition counts for 2 FLOPs. Therefore multiplication of $N \times K$ and $K \times M$ complex matrices needs $NM(8K-2)$ FLOPs. A matrix inverse requires $16K^3 - 8K^2 + 2K$ FLOPs. The complexity comparison of MMSE and KS precoding is summarized in Table 2. For simplicity, small terms ($\mathcal{O}(1)$ and addition of identity matrix) is omitted. It can be seen that when K is comparable to N, the computational complexity using KS can be reduced from $\mathcal{O}(K^3)$ to $\mathcal{O}(K^2)$.

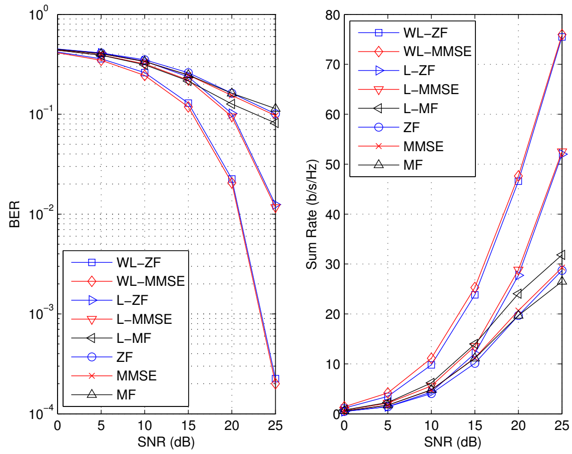

Fig. 1. Performance comparison of different precoding techniques

4. NUMERICAL RESULTS

In this section, the performance comparison of different precoding methods with IQI is carried out through Monte-Carlo simulation. The number of transmit antennas and users are $N = 100$ and $K = 20$, respectively. Without loss of generality the IQI parameters for all antennas are set to $g_\mathrm{r} = 1$ dB and $\theta_\mathrm{r} = 2°$. To evaluate the bit error rate (BER) performance, QPSK is considered as the modulation scheme. For the sake of fair comparison, we take IQI into account while calculating λ to satisfy the power constraint. A Rayleigh fading channel is used in the simulation and the channel gain is normalized, i.e., $\mathbb{E}\{\|\mathbf{H}_{ij}\|^2\} = \frac{1}{\sqrt{N}}$. SNR is defined as $\frac{P_\mathrm{Tx}}{\sigma^2}$. In all figures, MF, ZF, and MMSE stand for the traditional linear precoding schemes without consideration of IQI, and the real valued precoding methods with IQI are denoted by L-MF, L-ZF, and L-MMSE, respectively. The widely linear cases are denoted by WL-ZF and WL-MMSE.

Figure 1 shows the performance comparison in terms of BER and sum rate for three groups of precoding schemes under IQI. As can be seen in the figure, IQI has a remarkable impact on downlink performance in Massive MIMO systems. For real valued precoding methods which takes IQI into consideration, there is a significant gain in the high signal to noise ratio (SNR) region. Regarding widely linear design, WL-ZF is slightly worse than WL-MMSE, however, both methods dramatically outperform the other precoding schemes, with a gain of approximately 5-10 dB when BER is 10^{-1} and sum rate is 30 bit/s/Hz.

Figure 2 gives simulation results for WL-MMSE and KS precoding with $I = 1, 2, 3$, respectively. It can be seen that even for $I = 2$, the performance of KS is comparable to WL-MMSE and a gap of 2 dB is observed in the high SNR region. But KS has a much lower computational complexity (reduced by almost 75% in this example). When $I = 3$, the curves

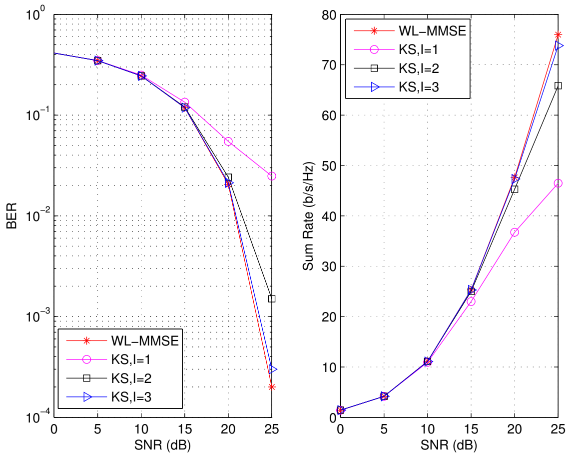

Fig. 2. Performance of KS precoding with $I = 1, 2, 3$

for KS almost coincide with WL-MMSE. The parameter I should be chosen to meet a satisfactory tradeoff between performance and complexity, and $I = 2, 3$ is reasonable in this example.

5. CONCLUSIONS

In this paper, widely linear precoding methods are devised for Massive MIMO systems to overcome the impact of I/Q imbalance. The proposed precoding algorithms correspond to a pre-compensation for I/Q imperfection. In order to reduce the complexity of calculating inverses of large dimensional matrices, we propose a Krylov subspace based method, which performs almost the same as MMSE/ZF precoders but with a much lower complexity. Simulation results show that the proposed methods work very well under IQI, outperforming existing designs.

REFERENCES

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

[2] R. C. de Lamare, "Massive MIMO Systems: Signal Processing Challenges and Future Trends," *URSI Radio Science Bulletin*, 2013.

[3] H. Q. Ngo, et al, "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[4] C.-X. Wang, et al, "Cellular architecture and key technologies for 5G wireless communication networks," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 122–130, Feb. 2014.

[5] A. Hakkarainen, et al, "Widely-Linear Beamforming and RF Impairment Suppression in Massive Antenna Arrays," *Journal of Communications and Networks*, vol. 15, no. 4, pp. 383–397, 2013.

[6] A.A. Abidi, "Direct-conversion radio transceivers for digital communications," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 12, pp. 1399–1410, 1995.

[7] A. Tarighat and A. Sayed, "Joint compensation of transmitter and receiver impairments in OFDM systems," *IEEE Transactions on Wireless Communications*, vol. 6, no. 1, pp. 240–247, Jan. 2007.

[8] R. Rogalin, et al, "Hardware-impairment compensation for enabling distributed large-scale MIMO," *2013 Information Theory and Applications Workshop (ITA)*, pp. 1–10, Feb. 2013.

[9] B. Picinbono and P. Chevalier, "Widely linear estimation with complex data," *IEEE Transactions on Signal Processing*, vol. 43, no. 8, pp. 2030–2033, 1995.

[10] P. Chevalier, et al, "Optimal widely linear MVDR beamforming for noncircular signals," in *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*. Apr. 2009, pp. 3573–3576, IEEE.

[11] N. Song, et al, "Adaptive Widely Linear Reduced-Rank Interference Suppression Based on the Multistage Wiener Filter," *IEEE Transactions on Signal Processing*, vol. 60, no. 8, pp. 4003–4016, Aug. 2012.

[12] Y. Zeng, et al, "Transmit Optimization With Improper Gaussian Signaling for Interference Channels," *IEEE Transactions on Signal Processing*, vol. 61, no. 11, pp. 2899–2913, June 2013.

[13] F. Sterle, "Widely Linear MMSE Transceivers for MIMO Channels," *IEEE Transactions on Signal Processing*, vol. 55, no. 8, pp. 4258–4270, Aug. 2007.

[14] H. Sampath, et al, "Generalized linear precoder and decoder design for MIMO channels using the weighted MMSE criterion," *IEEE Transactions on Communications*, vol. 49, no. 12, pp. 2198–2206, 2001.

[15] E. Telatar, "Capacity of Multi-antenna Gaussian Channels," *European transactions on telecommunications*, , no. 2, pp. 1–28, 1999.

[16] M. Joham, et al, "Linear transmit processing in MIMO communications systems," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2700–2712, Aug. 2005.

[17] A. Dabbagh and D. Love, "Multiple antenna MMSE based downlink precoding with quantized feedback or channel mismatch," *IEEE Transactions on Communications*, vol. 56, no. 11, pp. 1859–1868, Nov. 2008.

[18] L. N Trefethen and D. Bau III, *Numerical linear algebra*, vol. 50, Siam, 1997.