

Abstract—This work considers the precoding problem in massive multiuser multiple-input-multiple-output (MU-MIMO) systems equipped with low-resolution digital-to-analog converters (DACs). In previous literature on this topic, it is commonly assumed that the channel state information (CSI) is perfectly known. However, in practical applications the CSI is inevitably contaminated by noise. In this paper, we propose, for the first time, an eigen-inference (EI) precoding scheme to mitigate the deterioration caused by imperfect CSI for coarsely quantized massive MU-MIMO systems. First, we provide some theoretical analysis based on the block random matrix theory, which is generally applicable to any linear precoding scheme. To facilitate the analysis, we use Girkos Hermitization trick to consider an augment block symmetric channel matrix (BSCA). We derive the empirical distribution of the eigenvalues of the BSCA and establish the limiting spectra distribution connection between the true BSCA and its noisy observation. Then, based on these results, we propose an EI-based moments matching method for CSI-related noise level estimation and a rotation invariant estimation method for CSI reconstruction. Based on the cleaned (reconstructed) CSI, the quantized precoding problem is tackled via the Bussgang theorem and the Lagrangian multiplier method. Numerical simulations have been provided to demonstrate the effectiveness of the proposed precoder.

Index Terms—Massive MU-MIMO, Low-Resolution DACs, Eigen-Inference Precoding, Imperfect CSI, Block Random Matrix Theory.

I. INTRODUCTION

RECENT years have witnessed an increasing interest in massive MU-MIMO systems, in which the BS can simultaneously serve a large number of users equipments (UEs) by utilizing hundreds (even thousands) of transmit antennas [1]–[3]. As the number of antennas goes to infinity, the performance of classical precoders, e.g., maximum ratio transmission (MRT), zero-forcing (ZF) and water-filling (WF) [4], approach to that of optimal nonlinear ones [5]. On the other hand, gigantic increase in energy consumption, a key challenge to the use of massive MIMO, is brought by scaling up the number of the antennas. An effective way to decrease the energy cost is to equip the BS with low-resolution analog-to-digital converters (ADCs) [6] and DACs [7].

A. Related Works

The performance of quantized massive MU-MIMO systems has been extensively evaluated in terms of different performance metrics, such as (coded or uncoded) bit error rate (BER) [8]–[10], coverage probability [11], [12], and sum rate [13]–[15]. For the simplest case that the massive MU-MIMO are deployed with 1-bit ADCs or DACs, it has been shown in recent study [16] that, compared with massive MIMO systems with ideal DACs, the sum rate loss in 1-bit massive MU-MIMO systems can be compensated by disposing approximately 2.5 times more antennas at the BS. Meanwhile, the authors in [17] presented a quantized constant envelope (QCE) precoder, which can significantly reduce the power consumption of large-size communication systems. For frequency-selective MU-MIMO systems, the achievable performance with 1-bit ADCs or DACs is satisfactory when the number of BS antennas is large [18]–[20]. The performance gap between 1-bit massive MU-MIMO systems and the ideal ones that have access to unquantized data, can be further narrowed by utilizing 3–4 bits ADCs or DACs [21]. However, coarsely quantized MU-MIMO systems have to pay a high price for performance loss in terms of BER, especially when the communication systems are applied with high-order modulations.

As novel alternatives, quantized nonlinear precoders can significantly outperform linear ones with additional computational complexity. A novel nonlinear precoder that can support the 1-bit MU-MIMO downlink system with high-order modulations is firstly proposed in [22], which enables 1-bit MU-MIMO system to work well not only with the QPSK signaling but also with high-order modulations. Then, the semidefinite relaxation based precoder, which has sound theoretical guarantees and can achieve a performance close to the optimal nonlinear precoder, has been proposed in [23]. Nevertheless, its high computational complexity becomes an obstacle to hinder its application in massive MIMO systems. Several more efficient precoders have been developed recently, such as the squared-infinity norm Douglas-Rachford splitting based precoder [24], the maximum safety margin precoder [24], the C1PO and C2PO precoders [25], the finite-alphabet precoder [26] and the alternating direction method of multipliers based precoder [27].

However, it is worthy of noting that though the use of low-resolution ADCs and DACs could greatly reduce the power consumption and maintain performance loss within acceptable

Lei Chu is with Department of Electrical Engineering, Research Center for Big Data Engineering Technology, Shanghai Jiaotong University, Shanghai 200240, China. (leocha@sjtu.edu.cn)

Fei Wen is with Department of Electronic Engineering, Shanghai Jiaotong University, Shanghai 200240, China. (feiwenn@sjtu.edu.cn)

Dr. Qiu and Lily Li is with the Department of Electrical and Computer Engineering Tennessee Technological University, Cookeville, TN 38505 USA.

Dr. Wu is also with Department of Electrical Engineering, Research Center for Big Data Engineering Technology, Shanghai Jiaotong University, Shanghai 200240, China. (e-mail: rqiu@tntech.edu).

Dr. Qiu’s work is partially supported by N.S.F. of China under Grant No.61571296 and N.S.F. of US under Grant No. CNS-1247778, No. CNS-1619250. Dr. Wen’s work is partially supported by N.S.F. of China under Grant No.61871265.

Lei Chu, Student Member, IEEE, Robert Qiu, Fellow, IEEE, Fei Wen Member, IEEE, and Lily Li
margin, the quantized precoders are very sensitive to the channel estimation error \([23, 26, 27]\). To the best of our knowledge, the precoding scheme for quantized massive MU-MIMO systems in the case of imperfect CSI has not been reported yet.

For the downlink of coarsely quantized massive MU-MIMO systems with imperfect CSI, existing analysis and results on imperfect CSI, e.g., \([28−34]\), are not directly applicable due to the nonlinearity introduced by low-resolution DACs. In this paper, based on the block random matrix theory and the Bussgang theorem \([35]\), we develop, for the first time, an EI precoder to address the quantized precoding problem in the case of imperfect CSI. The main contributions are as follows.

**B. Contributions**

Firstly, we provide some theoretical analysis for the channel matrix based on block random matrix theory \([36−38]\), which are the theoretical basis of the proposed precoder. Specially, the rectangular channel matrix is augmented into a block symmetric channel matrix (BSCA) by employing Girkos Hermitization trick, which in principle converts the spectral analysis of rectangular matrices into the spectral analysis of Hermitian matrices. Then, we derive the empirical distribution of the eigenvalues of the BSCA based on the block random matrix theory. Meanwhile, we establish the limiting spectra distribution connection between the desired BSCA and the noisy BSCA.Besides, to facilitate the analysis of the complicated precoding problem, we decompose the basic quantized precoding problem with imperfect CSI into three subproblems by using the well-established Bussgang theorem. The decomposed subproblems are then tackled one by one.

Secondly, based on these derived theoretical results, we propose an EI-based moments matching method to estimate the unknown CSI-related noise level, and further develop a rotation invariant estimation method to reconstruct the CSI from its noisy observation. Based on the reconstructed (refined) CSI, the precoding problem is solved via the Lagrangian multiplier method.

Finally, we have evaluated the new precoder in various conditions with imperfect CSI. The results show that, with the proposed precoding scheme, significant improvement in robustness against imperfect CSI can be achieved compared with existing precoders.

**C. Paper Outline and Notations**

The remainder of this paper is structured as follows. Section **II** introduces the system models and outlines the quantized problem for coarsely quantized massive systems with channel estimation errors. Section **III** presents the basic analysis for the addressed issue and shows, in detail, the proposed EI precoder. In Section **IV** numerical studies are provided to evaluate the effectiveness of the proposed EI precoder. The conclusion of this paper is given in Section **V**. For the sake of brevity, the derivations of the technical results are deferred to the Appendices.

**Notations:** Throughout this paper, vectors and matrices are given in lower and uppercase boldface letters, e.g., \(x\) and \(X\), respectively. We use \([X]_{kl}\) to denote the element at the \(k\)th row and \(l\)th column. The symbols \(E\{X\}\), \(\text{tr}(X)\), \(\text{diag}(X)\), \(\|X\|_p\), and \(X^\dagger\) denote the expectation operator, the trace operator, the diagonal operator, the Frobenius norm, the conjugate transpose of \(X\), respectively. \(\Re\{x\}\), \(\Im\{x\}\), and \(|x|\) represent the real part, the imaginary part and \(\ell_2\)-norm of vector \(x\). \(\partial f(\cdot)\) stands for the subdifferential of the function \(f\). \(I\) and \(O\) are respectively referred to an identity matrix and a zeros matrix with proper size.

**II. SYSTEM MODEL AND QUANTIZED PRECODING PROBLEMS**

**A. Quantized MU-MIMO with Perfect CSI**

We consider a single cell massive MU-MIMO downlink system in which the BS is equipped with low-resolution DACs. The \(M\) single antenna UEs communicate with the BS by \(N\) antennas simultaneously in the same time-frequency resource. Assuming perfect synchronization, the received signal at the UEs can be written as

\[
y = Hx + n,
\]

where \(n \in \mathbb{C}^{M \times 1}\) is additive white Gaussian noise (AWGN) with elements distributed as \(CN(0, \sigma^2)\), the transmitted streams (precoded signals) \(x\) obey the total power constraint \(E\{x^H x\} \leq P\). \(H\) denotes the \(M \times N\) channel matrix modeling the fast fading between the UEs and the BS. The entries of \(H\) are identically independent zero-mean complex Gaussian random variables with variance \(1/N\).

Let \(x = Ps\) be the precoded vector of the unquantized system. For the quantized MU-MIMO downlink system, each precoded signal component \(x_i\), \(i = 1, \cdots, N\), is quantized separately into a finite set of prescribed labels by a \(B\)-bit symmetric uniform quantizer \(Q\). It is assumed that the real and imaginary parts of precoded signals are quantized separately.

The resulting quantized signals read

\[
z = Q(x) = Q(Ps).
\]

Then, the input-output of a quantized MU-MIMO downlink system can be expressed as follows:

\[
y = Hz + n = HQ(Ps) + n.
\]

For the massive MU-MIMO downlink system with low-resolution DACs, the quantized precoding problem (QPP) can be formulated as \([23]\)

\[
\begin{align*}
\text{minimize} \quad & E\{\|s - \beta HQ(Ps)\|^2_2 + \beta^2 M \sigma^2\} \\
\text{s.t.} \quad & E\{xx^H\} \leq P.
\end{align*}
\]

Solving the QPP in \(\mathbb{R}\) is challenging as one needs to handle the nonlinear quantization function. Fortunately, such a difficulty has been overcome recently by applying the Bussgang theorem, with which one can decompose the nonlinear quantized signal output into a linear function of the input to the quantizers and an uncorrelated distortion term \(d\) \([35]\).

Specially, linearizing \([2]\) gives

\[
x = Q(Ps) = FP + d,
\]
where the coefficients matrix $\mathbf{F}$ has been shown to be a diagonal matrix. Substituting (5) into (4) and using a simple strategy in (4), one can obtain the precoding matrix and precoding factor. Though efficient, the above solutions degrade significantly in the case of imperfect CSI. To achieve a robust performance in the case of imperfect CSI, an eigen-inference precoder in this paper is proposed based on the block random matrix theory.

B. Quantized MU-MIMO with Imperfect CSI

In this paper, we follow [28], [29] for the model of imperfect CSI through a Gauss-Markov uncertainty of the form

$$
\tilde{\mathbf{H}} = \sqrt{1 - \eta} \mathbf{H} + \sqrt{\eta} \mathbf{E},
$$

(6)

where $\tilde{\mathbf{H}}$ is the imperfect observation of channel available to the BS and $\mathbf{E}$ is AWGN. The CSI-related parameter $\eta$ characterizes the partial CSI. Specifically, $\eta = 0$ means perfect CSI, the values of $0 < \eta < 1$ correspond to partial CSI and $\eta = 1$ accounts for no CSI. In this paper, the parameter $\eta$ is reasonably restricted to $0 \leq \eta < 1$. The proposed algorithm in this paper takes advantage of the Bussgang theorem, and is applicable for any linear precoding scheme, e.g., MRT, ZF, WF, and WFQ [39]. In what follows, we examine, in detail, the special case of the WF precoder.

In the case of imperfect CSI, the quantized precoding problem can be formulated as

$$
\min_{\mathbf{P} \in \mathbb{C}^{N \times N}, \beta \in \mathbb{R}^+} \mathbb{E} \left\| \mathbf{s} - \beta \mathbf{HQ} \left( \tilde{\mathbf{P}} \right) \right\|^2 + \beta^2 M \sigma^2
$$

(7)

subject to

$$
\mathbb{E} \{ \mathbf{x} \mathbf{x}^H \} \leq P
$$

where $\mathbf{P}$ is the noisy observation

$$
\tilde{\mathbf{P}} = \tilde{\mathbf{H}}^H (\tilde{\mathbf{H}} \tilde{\mathbf{H}}^H + M \sigma^2 \mathbf{I})^{-1},
$$

(8)

$$
\tilde{\beta} = \frac{1}{\sqrt{P}} \text{tr} \left( \tilde{\mathbf{P}} \tilde{\mathbf{P}}^H \right)^{-1/2}.
$$

The problem (7) is challenging since, in practice, we have no prior knowledge of the CSI-related parameter $\eta$. Besides, compared with the ideal CSI case, the precoding matrix $\mathbf{P}$ is corrupted by the environmental noise. In what follows, an eigen-inference precoder is proposed to ameliorate this difficulty.

III. THE PROPOSED SOLUTION: THE EIGEN-INFEERENCE PRECODER

We first provide a sketch of the algorithm development in the following. Firstly, by taking advantage of the well-established Bussgang theorem, the original quantized precoding problem (7) is decomposed into three subproblems:

- **Problem I**: Estimation of the unknown CSI-related parameter $\eta$;
- **Problem II**: Reconstruction of the CSI matrix $\mathbf{H}$ from the noisy observation $\tilde{\mathbf{H}}$;
- **Problem III**: Estimation of the coefficients matrix $\mathbf{F}$ in the case of imperfect CSI.

Secondly, we present some theoretical basics from random matrix theory that will be needed in the ensuing analysis. The main metrics we employ are introduced and some favorable properties are shown in detail. To estimate the CSI-related parameter $\eta$, an EI-based moments matching method is then proposed. Furthermore, following the strategies in [40], [41], the asymptotic overlap between the non-perturbed eigenvalues/eigenvectors with the perturbed ones is built. Accordingly, an EI-based rotation invariant estimator for constructing the estimation of cleaned CSI $\tilde{\mathbf{H}}$ is developed. Lastly, with the refined CSI, the desired precoding matrix, the cleaned precoding factor, and the estimate of the coefficient matrix $\tilde{\mathbf{F}}$ are accordingly obtained.

A. Problem Decomposition

It has been mentioned in Section I and Section II-B that there is no available algorithm directly applies to solve the quantized precoding problem in the case of imperfect CSI. In this subsection, the original quantized precoding problem (7) is decomposed into three subproblems. The detailed analysis is presented in the following.

By taking the advantage of the well-known Bussgang theorem, one can obtain

$$
\mathbf{x} = \mathbf{Q} \left( \tilde{\mathbf{P}} \right) = \tilde{\mathbf{F}} \tilde{\mathbf{P}} \mathbf{s} + \mathbf{d}.
$$

(9)

Plugging (9) into (7) and applying the Lagrangian multiplier method [41], in the case of imperfect CSI, one can get the quantized WF precoding matrix as:

$$
\mathbf{P}^{WF} = \tilde{\beta}^{WF} \tilde{\mathbf{F}} \tilde{\mathbf{P}},
$$

(10)

$$
\tilde{\mathbf{P}} = \tilde{\mathbf{H}}^H (\tilde{\mathbf{H}} \tilde{\mathbf{H}}^H + M \theta)^{-1},
$$

$$
\tilde{\beta}^{WF} = \frac{1}{\sqrt{P}} \text{tr} \left( \tilde{\mathbf{P}} \tilde{\mathbf{P}}^H \right)^{-1/2},
$$

$$
\theta = (\sigma^2 \mathbf{I} + \sigma^2_d),
$$

$$
\sigma^2_d = \left( 1 - \text{diag} \left( \tilde{\mathbf{F}} \right) \right) \left( M \sigma^2 + 1 \right).
$$

With results in (10), the original precoding problem now can be decomposed into three subproblems: a) Estimation of the CSI-related parameter $\eta$; b) Reconstruction of CSI $\mathbf{H}$ from the noisy observation $\tilde{\mathbf{H}}$; c) Estimation of the coefficients matrix $\mathbf{F}$ under imperfect CSI. The solutions for all the three subproblems are fundamentally based on the block random matrix theory [36], which is presented in details in the following.

B. Random Matrix Basics and Main Metrics

We start by presenting some basics from random matrix theory that will be needed in the following algorithm development. For a square random matrix $\mathbf{A} \in \mathbb{C}^{M \times M}$, the resolvent of $\mathbf{A}$ is defined as

$$
\mathbf{G}_\mathbf{A} (z) = (z \mathbf{I}_M - \mathbf{A})^{-1}.
$$

(11)

The normalized trace of (11) gives

$$
\mathbf{g}_\mathbf{A} (z) = \frac{1}{M} \text{tr} \left[ \mathbf{G}_\mathbf{A} (z) \right].
$$
In the limit of large dimension, one has \( g^N_\lambda (z) \to g_\lambda (z) \), where \( g_\lambda (z) \) is the Stieltjes transform of \( A \). The asymptotic empirical distribution of eigenvalues \( F_A(\lambda) \) can be described in terms of its Stieltjes transform \([22]\), defined by

\[
g_A(\lambda) = \int \frac{dF_A(\lambda)}{\lambda - z} = \int \frac{\rho(\lambda)}{\lambda - z} d\lambda, \quad \text{as} \quad M \to \infty.
\]

Besides, the \( R \) transform, a handy transform which enables the characterization of the limiting eigen-spectra of a sum of free random matrices from their individual limiting eigen-spectra, can be defined as

\[
R_A(z) = g_A^{-1}(z) - \frac{1}{z}.
\]

Furthermore, the \( R \) transform can be expanded as \([42]\):

\[
R_A(z) = \sum_{l=1}^{\infty} \kappa_l(A) z^{l-1},
\]

where \( \kappa_l(A) \) denotes the so-called free cumulant which can be expressed as a function of moments of \( A \). Specially, given the \( n \)th moments of \( A \) and using the so-called cumulant formula \([22]\), one can have

\[
\varphi(A^n) = \sum_{l=1}^{m} \kappa_l(A)^m \sum_{m_1,\ldots,m_l} \varphi(A^{m_1-1}) \cdots \varphi(A^{m_l-1}),
\]

where, for \( k = 1, \ldots, l, \ m_k \) is a non-negative integer and satisfies \( m_1 + \ldots + m_l = m \). For completeness, the first three cumulants are given by:

\[
\kappa_1 = \varphi_1, \ \kappa_2 = \varphi_2 - \varphi_1^2, \ \kappa_3 = \varphi_3 - 3\varphi_2\varphi_1 + 2\varphi_1^3.
\]

It is noted that, in our case, \( H \in \mathbb{C}^{M \times N} \) is a rectangular random matrix, having complex-valued eigenvalues and eigenvectors. The well-known Hermitian random matrix theory can not be applied unless a proper strategy is used. In this paper, our analysis is based on BSCA and the noisy observation of SCAM, of forms

\[
B = \begin{bmatrix} 0 & H \\ H^H & 0 \end{bmatrix},
\]

and

\[
\tilde{B} = \begin{bmatrix} 0 & \tilde{H} \\ \tilde{H}^H & 0 \end{bmatrix} = B + \alpha \begin{bmatrix} 0 & E \\ E^H & 0 \end{bmatrix} = B + \tilde{E},
\]

where \( \alpha = \frac{1}{\sqrt{1 + q}} \). The observations in \([17]\) and \([18]\) are known as the Girko’s Hermitization trick \([43], [44]\), which has been widely used in the case of theoretical analysis of Hermitian random matrix. In this paper, the rectangular random matrix is considered. Thanks to the Girko’s Hermitization trick, the theoretical analysis developed in this paper is based on the spectral theory of block Hermitian random matrices, instead of the much involved spectral theory of rectangular random matrices.

We, in particular, are aiming to build the connection between the limiting spectra distribution (LSD) of \( B \) and \( \tilde{B} \). Let us define two more auxiliary matrices:

\[
D = BB^H = \begin{bmatrix} H H^H & 0 \\ 0 & H H^H \end{bmatrix},
\]

and

\[
\tilde{D} = \tilde{B} \tilde{B}^H = \begin{bmatrix} \tilde{H} \tilde{H}^H & 0 \\ 0 & \tilde{H} \tilde{H}^H \end{bmatrix}.
\]

With notations in \([17]\)-\([20]\), one can find that a direct relationship among eigenvalues of \( B \) (or \( B \)), \( D \) (or \( \tilde{D} \)) and \( HH^H \) (or \( \tilde{H} \tilde{H}^H \)), which can be expressed as

\[
\lambda^D = \begin{cases} \lambda^H \tilde{H}^H, & \lambda^D \neq 0, \\
0, & \lambda^B = \pm \sqrt{\lambda^D},
\end{cases}
\]

and

\[
\lambda^{\tilde{D}} = \begin{cases} \lambda^H \tilde{H}^H, & \lambda^{\tilde{D}} \neq 0, \\
0, & \lambda^{\tilde{B}} = \pm \sqrt{\lambda^{\tilde{D}}},
\end{cases}
\]

Besides, the limiting spectra connection between the noisy SCAM and the true SCAM can be expressed as in the following results.

**Lemma III.1.** For the BSCA \( B \) defined in \([17]\), the Stieltjes transform of the LSD of \( HH^H \) can be represented by the Stieltjes transform of the LSD of \( B \) as:

\[
g_{HH^H}(z) = \frac{q + 1}{2q} g_B(z) - \frac{q - 1}{2q} \frac{1}{z}.
\]

**Lemma III.2.** Let \( q = M/N < 1 \). For the SCAM \( B \) and the sample covariance of the SCAM \( D \), the Stieltjes transforms of the LSDs of \( B \) and \( D \) satisfy

\[
g_B(z) = z g_D(z^2).
\]

**Lemma III.3.** Let the random matrices \( D \) and \( B \) be defined as in Lemma III.2. The \( S \) transforms of the LSDs of \( B \) and \( D \) meet

\[
[S_B(z)]^2 = \frac{1 + z}{z} S_D(z).
\]

**Lemma III.2** and Lemma III.3 state direct LSD linkage between \( B \) and \( D \). The proof of the above lemmas are respectively deferred to Appendix A, Appendix B and Appendix C. With the technical results shown in Lemma III.1, Lemma III.2 and Lemma III.3, the statistical properties (i.e., the asymptotic empirical distribution of eigenvalues) of \( B \) can be obtained by the well-established Hermitian random matrix theory. Besides, combined with the addition law \([42], [45]\),

\[
R_{X+Y}(z) = R_X(z) + R_Y(z),
\]

one can reveal the limiting spectra connection between the noisy SCAM and the true SCAM.

All technical results shown above act as fundamental roles which enable us to tackle the quantized precoding problem introduced in Section II-B. The detailed analysis is given in the following.
C. Determination of the Unknown Parameter $\eta$: An EI-based Moments Matching Algorithm

In this part, we propose a moments matching algorithm to determine the unknown parameter $\eta$ in (5). Our analysis is based on the free probability which is a powerful tool for analyzing the eigenspectra of large random matrices. The proposed moments matching algorithm can provide a robust and flexible estimation of $\eta$ without specifying an explicit relationship between the number of UEs and the number of antennas.

The estimation of $\eta$ is enabled by random matrix theory which offers a collection of useful results on the asymptotic behavior of the BSCA $B$ and its noisy observation $\tilde{B}$, which are shown in the following theoretical results.

**Theorem III.4.** For the BSCA $B$ defined in (17), the empirical distribution of the eigenvalues of $B$ converges almost surely, as $M, N \to \infty$ with $M/N \to q$, to a nonrandom distribution whose density function is

$$
\rho_B(x) = \frac{\sqrt{b^2 - a^2}}{(q + 1) \pi |x|} + \frac{1 - q}{1 + q} \delta(x),
$$

where $a \leq |x| \leq b$ and $a = 1 - \sqrt{q}, b = 1 + \sqrt{q}$.  

The proof of Theorem III.4 is deferred to Appendix E. As an illustration of the result obtained in Theorem III.4 Fig. 1 shows the empirical distribution of the nonzero eigenvalues of the BSCA. It can be seen that the empirical density function (27) agrees well with the simulation result.

Combined with the favorable results given in Theorem III.4 Lemma III.2, Lemma III.3 and Lemma III.1, one can obtain the following result for the sample covariance of the noisy channel matrix.

**Theorem III.5.** For the noisy channel matrix $\tilde{H} \in \mathbb{C}^{M \times N}$, the empirical distribution of eigenvalues of its sample covariance matrix $\tilde{H}\tilde{H}^H$ converges almost surely to a limit distribution whose Stieltjes transform, denoted by $G = \rho_{\tilde{H}\tilde{H}^H}(z)$, satisfies

$$
-2Gz + G^2 \left(1 + k^2\right) qz + h(G) = 0,
$$
as $M, N \to \infty$ with the ratio $q = M/N \leq 1$ fixed, and

$$
h(G) = \sqrt{1 + G^2} \left(4q - (2G^2q + G^2q^2)\right) + \sqrt{1 + G^2} \alpha^2q(4q - (2G^2q^2 + G^2q^3)).
$$

The proof of Theorem III.5 is deferred to Appendix E. Theoretical result in Theorem III.5 with the relation in (13), we can obtain the $R$ transform of $\tilde{H}\tilde{H}^H$. Accordingly, we can infer the parameters of the underlying noisy channel matrix $\tilde{H}$ from a realization of its covariance matrix. Taking Taylor’s expansion for the obtained $R$ transform, the first three moments of the sample covariance matrix can be analytically parameterized by the unknown parameters $\eta$ of form:

$$
\varphi \left[ \tilde{H}\tilde{H}^H \right] = \frac{1}{1 - \eta},
$$

$$
\varphi \left[ \tilde{H}\tilde{H}^H \right] = \frac{2(1 - \eta)(1 - q) + q}{(1 - \eta)^2},
$$

$$
\varphi \left[ \tilde{H}\tilde{H}^H \right] = \frac{q(3 - \eta)(1 - q) + q}{(1 - \eta)^2}.
$$

Given an observation $\tilde{H}$, we can compute estimates of the first three moments of its sample covariance matrix as

$$
\hat{\alpha} = \arg\min_{\alpha > 0} \left\| \sum_{k=1}^3 \varphi \left[ \tilde{H}\tilde{H}^H \right] - \varphi \left[ \tilde{H}\tilde{H}^H \right] \right\|^2.
$$

(33)

Though the EI-based moments matching algorithm described above is theoretically exact when $M, N \to \infty$, we verify through various simulations that the proposed method is sufficiently accurate for even small values of $N$ and $M$ (i.e., $M = 8, N = 32$). See experimental results in Section IV for more details.

D. CSI Cleaning Based on Random Matrix Theory: EI-based Rotation Invariance Estimation

We now attempt to construct an estimator $\hat{H}$ of the true channel matrix $H$ that relies on the imperfect observation $\tilde{H}$. Let $\lambda_1 \geq \cdots \geq \lambda_{M+N}$ and $u_1, \cdots, u_{M+N}$ be the eigenvalues and eigenvectors of a true matrix (denoted by $\mathbf{A}$), respectively. For the noisy matrix (denoted by $\mathbf{A}$), its eigenvalues and eigenvectors are respectively denoted by $\omega_1 \geq \cdots \geq \omega_{M+N}$ and $v_1, \cdots, v_{M+N}$.

Here we only employ the first three moments by considering the trade-off between the computational complexity and estimation accuracy. See more details in Section IV-A.
With the Hermitian expressions in (17) and (18), we can tackle the CSI reconstruction problem by taking advantage of the rotation invariant (RI) estimation method introduced in (40), (41), which is shown in the following.

To find the optimal estimate of $A$, a natural way to tackle the problem is to solve the following minimizing problem

$$
\hat{A} = \arg \min_{\hat{A} \in \text{RI } \Omega (A)} \left\| \Omega (A) - \hat{A} \right\|_F^2,
$$

(34)

where $\text{RI } \Omega (A)$ denotes the set of all possible RI estimators. Any rotation invariant estimator $\Omega (A)$ obtained by solving (34) has the same eigenvectors as the perturbed matrix $\hat{A}$. In other words, we can have

$$
\Omega (A) = \sum_{k=1}^{M+N} \xi_k v_k v_k^H.
$$

(35)

Substituting (35) into (34), the optimal solution of (34) is given by

$$
\hat{A} = \sum_{k=1}^{M+N} \xi_k v_k v_k^H,
$$

(36)

and

$$
\hat{\xi}_k = \sum_{j=1}^{M+N} \lambda_j O (\omega_k, \lambda_j),
$$

(37)

where

$$
O (\omega_k, \lambda_j) = (M + N) \mathbb{E} [u_j^H v_j].
$$

(38)

The expression $O (\omega_k, \lambda_j)$ in (38) is so-called miracle overlap which creates deep connection between the true eigenvalues $\lambda_j$ and perturbed eigenvalues $\omega_k$. By the miracle overlap, in large dimensional regime, one can estimate $\lambda_j$ only from the measurements of LSD of $A$. In particular, let $\hat{A} = \hat{H}^H \hat{H}$, we can obtain the estimates of the true eigenvalues of $A = \hat{H}^H \hat{H}$ by [41]

$$
\hat{\lambda}_k = \phi_1 (\omega_k) \phi_2 (\omega_k) + \phi_3 (\omega_k),
$$

(39)

where

$$
\begin{align*}
\phi_1 (\omega_k) &= 1 - qa^2 h_k, \\
\phi_2 (\omega_k) &= \omega_k - \alpha^2 (1 - q) - 2qa^2 h_k, \\
\phi_3 (\omega_k) &= h_k (\omega_k - \alpha^2 (1 - q) + qa^2 \omega_k (\rho_k - h_k)), \\
\rho_k &= \lim_{\epsilon_k \to 0^+} \Re [\mathbf{g}_A (\omega_k + i \epsilon_k)], \\
h_k &= \lim_{\epsilon_k \to 0^+} \Im [\mathbf{g}_A (\omega_k + i \epsilon_k)].
\end{align*}
$$

The results in (39) have been proven by means of the Replica method [40], [41] in the case of the Hermitian random matrix. With (39) and the direct relation shown in (21) and (22), one can get the estimates of the eigenvalues of $D$ and $B$. With (36), we can obtain the estimates of $D$ and $B$. Finally, using (17), the estimate of the channel matrix from the noisy observation is achieved.

The analysis performed above is based on the theoretical work of [40], [41]. However, the results shown in [40], [41]

are restricted to the Hermitian random matrix. We extend them to the rectangular random matrix. Besides, we provide an EI-based moments matching method to estimate the CSI-related parameter $\eta$, which is actually unknown in practice. Further, we derive the estimator of the Stieltjes transform of $\hat{A}$ in Theorem III.5 which is useful in (39).

E. Determination of the Coefficient Matrix $F$

With cleaned CSI, we are now ready to determine $F$. It has been widely assumed that the elements of the uniform quantization errors $d$ converge in distribution to a zero-mean Gaussian random variable, whose variance can be characterized in closed form. Indeed, such an assumption is essentially valid when the errors from individual channels are asymptotically pairwise independent, each uniformly distributed within the quantization thresholds [46]. Denote the estimated preceding matrix by $\hat{P}$. With Gaussian assumption of $d$ and based the Bussgang theorem, one can estimate $F$ by minimizing the square error

$$
\mathbb{E} [d]_F^2 = \mathbb{E} \| x - \hat{F} z \|_2^2,
$$

(40)

where $z = \hat{P} s$. Then

$$
\hat{F} = \arg \min_F \mathbb{E} [x - \hat{F} z]_2^2 = \arg \min_F \text{tr} \left( R_{xx} - R_{xz} F^H (R_{zz} + F R_{zz} F^H) \right)
$$

(41)

$F$ can be obtained. A solution of (41) has to satisfy

$$
\frac{\partial f (F)}{\partial F} = -2 R_{xz} + 2 F R_{zz},
$$

(42)

which yields

$$
\hat{F} = R_{xz} R_{zz}^{-1},
$$

(43)

where

$$
R_{zz} = \mathbb{E}_s (z z^H) = \hat{P} \mathbb{E}_s (s s^H) \hat{P}^H = \hat{P} \hat{P}^H,
$$

(44)

and

$$
R_{xz} = \mathbb{E}_s (z x^H) = \mathbb{E}_s (Q (z) x^H). \tag{45}
$$

It is assumed that $F$ is a diagonal matrix\footnote{This assumption, despite its simplicity, is a common assumption and proven to be accurate in previous works [16], [23], [25], [27] and verified in our experimental results in Sec. IV} then we can obtain its diagonal elements

$$
\left[ \begin{array}{c}
\hat{F} \\
\end{array} \right]_{m,m} = \frac{\mathbb{E}_s [Q (z) x^H]}{\sigma_m^2}, \tag{46}
$$

where $m = 1, 2, \ldots, M$ and $\sigma_m^2 = [R_{zz}]_{m,m}$. Let

$$
l_b = \Delta \left( b - \frac{2^B - 1}{2} \right), b = 1, \ldots, 2^B - 1,
$$

and

$$
\tau_b = \Delta \left( b - 2^B - 1 \right), b = 2, \ldots, 2^B,
$$

The set $\Omega (A)$ is rotation invariant and satisfies $\Phi \Omega (\hat{A}) \Phi^H = \Omega (\Phi \hat{A} \Phi^H)$, for any rotation matrix $\Phi$.\footnote{The set $\Omega (A)$ is rotation invariant and satisfies $\Phi \Omega (\hat{A}) \Phi^H = \Omega (\Phi \hat{A} \Phi^H)$, for any rotation matrix $\Phi$.}
respectively be the quantization labels and quantization thresholds of a $B$ bit uniform quantizer, whose output can be expressed as

$$Q(z) = \frac{\Delta}{2} (1 - 2^B) + \Delta \sum_{l=1}^{2^B-1} \delta_{\Delta(l-2^B-1), \infty}(z). \quad (47)$$

Substituting (47) into (46) gives

$$\left[ \hat{F} \right]_{m,m} = \frac{\Delta}{\sigma_u^{2B/2}} \sum_{l=1}^{2^B-1} \int_{1-2^{B-1}}^{\infty} x \exp \left( -\frac{x^2}{\sigma_u^2} \right) dx,
= \frac{\Delta}{\sqrt{\pi} \sigma_u} \sum_{l=1}^{2^B-1} \exp \left( -\frac{\Delta^2}{\sigma_u} \right) (1 - 2^B)^2 \quad (48)$$

So far, the quantized precoding problem has been tackled.

The proposed EI precoder is summarized in Algorithm 1.

**Algorithm 1** The Proposed EI Precoding Algorithm

1. **Input**: Imperfect SCM of CSI: $\hat{H}$.
2. Construct noisy BSCA $\tilde{B}$ using (18).
3. Calculate the estimates of the first three moments of $\tilde{H}\hat{H}^H$ using (59) and compute the related moments using (19).
4. Substituting the obtained moments $\hat{\varphi} \left[ \left( \tilde{H}\hat{H}^H \right) \right]$ into (58) to estimate the CSI-related parameter $\hat{\eta}$.
5. Performing eigenvalue decompositions of $\tilde{B}$ yields $[U, S] = \text{eig} \left( \tilde{B} \right)$.
6. Combined with the CSI-related parameter $\hat{\eta}$, calculating the estimates of eigenvalues of $\tilde{B}$ using (59) and (22) gives $\hat{S}$.
7. Reconstructing $\tilde{B}$ as $\tilde{B} = U\hat{S}U^H$.

The estimated CSI, denoted by $\hat{H}$, is the upper triangular part of $\tilde{B}$ and $\hat{H} = \tilde{B}_1 \cdot M \cdot \tilde{B}_{M+1} \cdot M \cdot N$.
8. Plugging the estimated CSI $\hat{H}$ into (10) yields the estimates of the precoding matrix and precoding factor, which are denoted by $\hat{P}$ and $\hat{\beta}$, respectively.
9. Calculating the estimate of the coefficients matrix $\hat{F}$ using (44) to obtain $\hat{F}$.
10. **Output**: $\hat{P}$, $\hat{F}$, and $\hat{\beta}$.

### IV. Numerical Studies

This section evaluates the performance of the new precoder via numerical simulations in comparison with five representative linear quantized precoders: MRT, WF, WFQ, ZF and QCE. Each provided result is an average over 1000 independent runs.

#### A. The Estimation Accuracy of the Unknown Parameter $\eta$

In the first experiment, we investigate the estimation accuracy of the unknown CSI-relates parameter $\eta$ in three different settings, in terms of the cumulative distribution function (CDF) of estimation error. The results shown in the following are the CDFs of the absolute estimation error, which is denoted by

$$\Delta \eta = |\eta - \hat{\eta}|,$$

where $\hat{\eta}$ is the estimate of $\eta$.

1) **The Effect of the Order of Moments**: The result shown in Fig. 2 is obtained by numerically solving (58). The selected root of (58) satisfies $0 < \eta < 1$. The true value of $\eta$ is fixed at 0.5. Fig. 2 shows the CDFs of the absolute estimation error $\Delta \eta$ in the case of different order of the moments estimates. In the case of a small-size MU-MIMO system (with $M=32$, $N=8$), FMS, SMS, and TMS are respectively referred to first-order, second-order, and third-order of moments estimates ($k = 1$ or $2$ or $3$ in (58)). Similarly, FML, SML, and TML mean first-order, second-order, and third-order of moments estimates in a large-scale MU-MIMO system ($M=256$, $N=30$), respectively. From Fig. 2 we can see that the estimation error decreases as the order of the moments increases. For the large-size MU-MIMO system, the maximum of absolute estimation error is less than 0.05, which is only 10% of $\eta$, indicating the robustness of the proposed method. Besides, compared to the small-size MIMO system, better estimation performance can be obtained in MU-MIMO systems with large-size transmit antennas. However, it is noted that employing higher order of moments pays the price of higher computational complexity. In what follows, we respectively use first-order and third-order moments estimates to determine the unknown parameter $\eta$ in small-size ($M \leq 32$) and large-size ($M \geq 128$) systems.

2) **The Effect of the System Size**: We now investigate the effect of the system size on the estimation of $\eta$. We consider an $M$-antenna MU-MIMO system with QPSK signaling. Fig. 3 plots the CDFs of the absolute estimation error for different numbers of transmit antennas, whilst Fig. 4 shows the results for different numbers of UEs. In Fig. 3, the number of UEs is $N = 20$. The numbers of antennas are 20, 32, 64, 128 and 256, respectively. It is shown in Fig. 3 that the estimation performance improves as the increase of the number of transmit antennas. In Fig. 4 the numbers of UEs are 5,
Fig. 3. The CDFs of the absolute estimation error in the case of different number of transmit antennas.

10, 15, 20, 20, and 30, respectively. The number of transmit antennas is $M = 256$. The results in Fig. 3 imply that one can obtain accurate estimates of $\eta$ with high probability for various numbers of UEs. Besides, the estimation performance improves as the increase of the number of UEs. Interestingly, the estimation performance does not improve distinctly when the number of UEs is above a certain threshold (e.g., $N \geq 20$). We will revisit such a phenomenon with a theoretical analysis in our future work.

Fig. 4. The CDFs of the absolute estimation error in the case of different number of UEs.

3) Robustness in Different Error Conditions: Fig. 5 presents the performance of the proposed EI-based moments matching method for a various values of $\eta$. In Fig. 5 we consider a massive MIMO system with 256 transmit antennas and 30 UEs. The results in Fig. 5 indicate that accurate estimation of the noise parameter can be achieved with high probability in a wide range of $\eta$, which demonstrates the robustness of the proposed algorithm.

Fig. 5. The CDFs of the absolute estimation error in various values of $\eta$.

B. The Accuracy of Channel Matrix Reconstruction

We consider the case with $N = 20$ UEs and $M = 20, 32, 64, 128, 256$ transmit antennas. In this case, the CSI is reconstructed by the proposed EI-based rotation invariant estimation method as shown in Algorithm 1. Let $\hat{H}$ be the estimate of $H$. Fig. 6 illustrates the the mean square error (MSE) of channel matrix reconstruction, which is denoted by

$$\text{MSE} = \frac{1}{MN} \| H - \hat{H} \|_F^2.$$

It is shown in Fig. 6 that the reconstruction error decreases as the increase in the number of transmit antennas. Besides, the proposed method can provide accurate estimates of CSI from noisy CSI in both small perturbation condition (e.g., $\eta = 0.1$) and large perturbation condition (e.g., $\eta = 0.9$), indicating the robustness of the new method.

Fig. 6. MSE of channel matrix reconstruction with different number of transmit antennas.

C. BER Performance

In this experiment, we investigate the BER performance of coarsely quantized MU-MIMO systems with imperfect CSI in
three cases. The BS is equipped with 256 transmit antennas and serves 30 UEs. For notation succinctness, we use EI to denote the proposed eigen-inference precoder. WF and WF0 are respectively referred to the WF precoding scheme implemented in the case of perfect CSI and imperfect CSI (without reconstruction). We have similar definition for other compared precoding schemes.

1) The Effect of Quantization Levels: Fig. 7 shows the BER of the WF variants versus signal-to-noise ratios (SNRs) for different quantization levels (B-bit DACs). The CSI-related parameter is fixed at $\eta = 0.3$. From Fig. 7, we observe that the performance of the WF precoder (with perfect CSI) and the proposed WF-IE precoder (with imperfect CSI) improve with the increase of the quantization level. The results indicate that the proposed precoder can provide reliable transmission of QPSK signaling in the case of imperfect CSI. Specially, for $\{2, 3, 4\}$-bit quantization, the performance gap to ideal BER (perfect CSI) with that of WF-IE is only about 3 dB for a target BER of $10^{-3}$.

2) The Effect of the SNR: In Fig. 8, we consider a massive MIMO system with 4-bit DACs. It is shown that the WF-IE precoder has similar performance with the ZF-IE precoder and outperforms the MRT-IE precoder. Moreover, the performance gap between the WF precoder with perfect CSI and the proposed WF-IE precoder (with imperfect CSI) is remarkably smaller than that between the WF precoder and the WF0 precoder. For example, for the same BER target of $10^{-3}$, the former is about 3 dB while the latter is more than 15 dB. The result implies that the implementation of robust precoding algorithms for coarsely quantized massive MIMO systems with imperfect CSI is helpful and possible by employing proper signal processing techniques (e.g., the proposed EI precoder).

3) The Effect of Different Values of $\eta$: In Fig. 9 we show the BER performance of all the precoders with different values of $\eta$. The SNR is fixed at 5 dB, a low-to-moderate (typical) SNR. In the case of perfect CSI, the WF precoder outperforms the WFQ precoder, which is in line with the result in [23]. In the presence of imperfect CSI, it can be seen from Fig. 9 that each precoder degrades as the increase of $\eta$. However, they have different sensitivity, e.g., WFQ0, WF0 and QCE0 are more sensitive to errors in CSI than their counterparts using the proposed CSI reconstruction method. As a consequence, WFQ-IE, WF-IE and QCE-IE significantly outperform WFQ0, WF0 and QCE0, and the advantage increases as $\eta$ increases. For example, WFQ0 and WF0 can approach a target BER of $10^{-3}$ only when $\eta < 0.15$, while WFQ-IE and WF-IE can achieve a target BER of $10^{-3}$ for any $\eta < 0.45$. Even in the limit case of $\eta = 0.9$, the proposed precoder can still achieve a BER of about $10^{-2.5}$. On the whole, the results imply that the employment of simplified hardware (low-resolution DACs) could be possible without severe BER performance loss in the case of imperfect CSI.
V. CONCLUSIONS

In the paper, we developed a novel precoding scheme for coarsely quantized massive MU-MIMO systems in the presence of imperfect CSI. Firstly, we provided some analysis using the block random matrix theory, based on which the limiting spectra distribution connection between the true BSCA and noisy BSCA has been established. Then, with the obtained theoretical results, we proposed an EI-based moments matching method to estimate the CSI-related noise level and a rotation invariant estimation method to reconstruct the CSI. Finally, experimental results have demonstrated that the proposed EI precoding scheme can significantly mitigate the deterioration caused by imperfect CSI in coarsely quantized massive MU-MIMO systems.

VI. ACKNOWLEDGEMENTS

The authors would like to thank for technical experts: Tian Pan, Guangyi Yang, Rui Gong and Yingzhe Li, all from Huawei Technologies for their fruitful discussions.

APPENDIX A

PROOF OF LEMMA III.1

Proof: Following [42], one can expand the Stieltjes transform of a random matrix $A$ as

$$ \mathcal{g}_A(z) = -\frac{1}{z} \sum_{k=0}^{\infty} \mathbb{E} \left[ A^k \right] z^{-k} \quad (A.1) $$

If $A$ is a square random matrix, i.e., $A \in \mathbb{C}^{N \times N}$, then one can have, in large dimensional regime,

$$ \mathbb{E} \left[ A^k \right] \equiv \mathbb{E} \left[ \frac{1}{N} \operatorname{tr} \left\{ A^k \right\} \right] \quad (A.2) $$

The Stieltjes transform of $D$ can be written as

$$ \mathcal{g}_D(z) = -\frac{1}{z} \sum_{k=0}^{\infty} \mathbb{E} \left[ B^k \right] z^{-k} \equiv \mathbb{E} \left[ \operatorname{tr} \left\{ (zI - D)^{-1} \right\} \right] \quad (A.3) $$

From the definition of $D$ in (19):

$$ D = BB^H = \begin{bmatrix} HH^H & 0 \\ 0 & H^HH \end{bmatrix}, $$

one can have, using the inverse of the block matrix,

$$ \mathcal{g}_D(z) = \frac{1}{M + N} \operatorname{tr} \left[ (zI - HH^H)^{-1} \right] = \frac{1}{M + N} \operatorname{tr} \left[ (zI - HH^H)^{-1} + 1 \right] = \frac{M}{M + N} \mathcal{g}_{HH^H}(z) + \frac{N}{M + N} \mathcal{g}_{H^HH}(z). \quad (A.4) $$

From Lemma 3.1 in [43], we have

$$ \mathcal{g}_{HH^H}(z) = \frac{M}{N} \mathcal{g}_{HH^H}(z) + \frac{M - N}{z} \quad (A.5) $$

Let $q = M/N$, plugging (A.5) into (A.3) gives

$$ \mathcal{g}_D(z) = \frac{2q}{q + 1} \mathcal{g}_{HH^H}(z) + \frac{q - 1}{q + 1} \frac{1}{z}, \quad (A.6) $$

which is the result shown in Lemma III.1.

APPENDIX B

PROOF OF LEMMA III.2

From (19), we have $D = B^2$. Let $F_B(x)$ and $F_D(x)$ be the empirical eigenvalue distribution of $B$ and $D$, respectively. Since the eigenvalues of $D$ are the radicals of the eigenvalues of $D$, let $\left\{ -\sqrt{b} - \sqrt{b} \right\} \cup \left\{ \sqrt{a}, \sqrt{b} \right\}$ be the support of the eigenvalues of $B$ and recall the definition of the Stieltjes transform in (12), then we have

$$ \mathcal{g}_B(z) = \int_{-\sqrt{a}}^{\sqrt{a}} \frac{1}{z - x} dF_B(x) + \int_{-\sqrt{b}}^{\sqrt{b}} \frac{1}{z - x} dF_B(x) = \int_{-\sqrt{a}}^{\sqrt{a}} \frac{2z}{z^2 - 1} dF_B(x) = \int_{a}^{b} \frac{z}{z^2 - y} dF_B((y) = z \mathcal{g}_D(z^2), \quad (B.1) $$

which completes the proof of Lemma III.2.

APPENDIX C

PROOF OF LEMMA III.3

We start with the definition of the $M$ transform [42] of a random matrix $A$, which is denoted by

$$ M_A(z) = \sum_{k=1}^{\infty} \mathbb{E} \left[ A^k \right] z^{-k}, \quad (C.1) $$

or

$$ M_A(z) = z^{-1} \mathcal{g}_A(z^{-1}) + 1. \quad (C.2) $$

Besides, the $S$ transform can be expressed with the $M$ transform as

$$ S_A(z) = \frac{1}{z} M_A^{-1}(z). \quad (C.3) $$

From (24) or (B.1), is follows that

$$ M_B(z) = z^{-1} \mathcal{g}_B(z^{-1}) + 1 = z^{-2} \mathcal{g}_D(z^2) + 1 = M_D(z^2) \quad (C.4) $$

Substituting (C.4) into (C.3) yields

$$ S_B(z) = \frac{1 + \frac{1}{z}}{z} M_B^{-1}(z) = \frac{1 + \frac{1}{z}}{z} \left[ M_D(z) \right]^{1/2} = \frac{1 + \frac{1}{z}}{z} \left[ \frac{1}{1 + z} S_B(z) \right]^{1/2} $$

or equivalently,

$$ [S_B(z)]^2 = \frac{1 + \frac{1}{z}}{z} S_B(z), \quad (C.5) $$

which completes the proof of Lemma III.3.
APPENDIX D
PROOF OF THEOREM III.4
With the technical results in Lemma III.1, Lemma III.2 and Lemma III.3 one can easily obtain Theorem III.4 by means of the transforms of random matrices.

From [42], the Stieltjes transform of $\HH^H$ is

$$\begin{align*}
g_{\HH^H}(z) &= \frac{1 - q - z - \sqrt{z^2 - 2(1+q)z + (q-1)^2}}{2qz}.
\end{align*}$$

Plugging (D.1) into (23) (or (A.6)), we have

$$\begin{align*}
g_B(z) &= \frac{2q}{q+1} g_{\HH^H}(z) + \frac{q-1}{q+1} \\
&= -z - \sqrt{z^2 - 2(1+q)z + (q-1)^2} \\
&\quad \quad \quad \quad \quad \frac{(q+1)z}{(q+1)}.
\end{align*}$$

With Lemma III.2 one can obtain the Stieltjes transform of $B$, which is denoted by

$$\begin{align*}
g_B(z) &= -z^2 - \sqrt{z^4 - 2(1+q)z^2 + (q-1)^2} \\
&\quad \quad \quad \quad \quad \frac{(q+1)z}{(q+1)}.
\end{align*}$$

Given $g_B(z)$, the inversion formula [42], [45] that yields the limiting probability density function of the eigenvalues of $B$ is

$$\begin{align*}
g_B(z) &= \lim_{y \to 0^+} \frac{1}{\pi} \Im [g_A(x+i\gamma)] \\
&= \frac{\sqrt{(b^2-x^2)(x^2-a^2)}}{(q+1)\pi |x|} + \frac{1-q}{1+q} \delta(x).
\end{align*}$$

where $a \leq |x| \leq b$ and $a = 1 - \sqrt{q}$, $b = 1 + \sqrt{q}$. This completes the proof of Theorem III.4.

APPENDIX E
PROOF OF THEOREM III.5
For the BSCA $\tilde{B}$, from (D.3), we have

$$\begin{align*}
g_B(z) &= -z^2 - \sqrt{z^4 - 2(1+q)z^2 + (q-1)^2} \\
&\quad \quad \quad \quad \quad \frac{(q+1)z}{(q+1)} \\
&= \frac{2q}{q+1} g_B(z) + \frac{q-1}{2qz}.
\end{align*}$$

where $B$ is an auxiliary matrix whose Stieltjes transform satisfies (E.1). With Lemma III.1, Lemma III.2, we can have

$$g_B(z) = z g_{\HH^H}(z^2).$$

Using (E.2) and Lemma III.3 one can obtain the $S$ transform of $B$, which can be denoted by

$$\begin{align*}
S_B(z) &= \left(1 + \frac{z}{z} S_{\HH^H}(z) \right)^{1/2}.
\end{align*}$$

The relation between the $R$ transform and the $S$ transform has to satisfy [45], [49]

$$\begin{align*}
\frac{1}{R(z)} &= S(R(z)).
\end{align*}$$

Plugging (E.3) into (E.4) gives

$$\begin{align*}
\left[ \frac{1}{R_B(z)} \right]^2 &= \left[ S_B(z R_B(z)) \right] \frac{1}{2} \\
&= \frac{1 + z}{z} S_{\HH^H}(z R_B(z)).
\end{align*}$$

Since $S_{\HH^H}(z) = \frac{1}{1+qz}$, (E.5), one can obtain the $R$ transform of $B$ by solving the equation (E.5), which yields

$$\begin{align*}
R_B(z) &= -1 + qz^2 + \sqrt{1 + 4z^2 - 2q^2z^2 + q^2z^4} \\
&\quad \quad \quad \quad \quad \frac{2}{2z}.
\end{align*}$$

For the noisy BSCA $\tilde{B}$, one can similarly define an auxiliary matrix $B$ that

$$g_B(z) = z g_{\HH^H}(z^2),$$

Combined with the addition law [42] shown in [26] and Lemma 4.1 in [45], the obtained $R$ transform of $B$ reads

$$\begin{align*}
R_B(z) &= \frac{1 + qz^2 + h(z)}{2z},
\end{align*}$$

where

$$h(z) = \sqrt{1 + qz^2} (4 + q(-2 + q z^2))$$

+ $1 + \alpha^2 z^2 (4 + q(-2 + z^2 \alpha^2 q))$.

From (13), we can have

$$\begin{align*}
R_B(g_B(z)) &= z - \frac{1}{R_B(z)}.
\end{align*}$$

Finally, plugging (E.7) and (E.8) into (E.9) and solving (E.9) yields the technical result shown in Theorem III.5 which completes the proof.

REFERENCES

[1] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, “An overview of massive mimo: Benefits and challenges,” IEEE Journal of Selected Topics in Signal Processing, vol. 8, no. 5, pp. 742–758, 2014.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, “Massive mimo for next generation wireless systems,” IEEE Communications Magazine, vol. 52, no. 2, pp. 186–195, 2014.

[3] F. Rusek, D. Persson, B. K. Lau, and E. G. Larsson, “Scaling up mimo: Opportunities and challenges with very large arrays,” IEEE Signal Processing Magazine, vol. 30, no. 1, pp. 40–60, 2012.

[4] M. Joham, W. Utschick, and J. A. Nossek, “Linear transmit processing in mimo communications systems,” IEEE Transactions on Signal Processing, vol. 53, no. 8, pp. 2700–2712, 2005.

[5] N. Jindal and A. Goldsmith, “Dirty-paper coding versus t dma for mimo broadcast channels,” IEEE Transactions on Information Theory, vol. 51, no. 5, pp. 1763–1794, 2005.

[6] J. Zhang, L. Dai, X. Li, Y. Liu, and L. Hanzo, “On low-resolution adcs in practical 5g millimeter-wave massive mimo systems,” IEEE Transactions on Signal Processing, vol. PP, no. 99, pp. 1–1, 2017.

[7] A. Gokceoglu, E. Bjornson, E. G. Larsson, and M. Valkama, “Spatio-temporal waveform design for multi-user massive mimo downlink with 1-bit receivers,” IEEE Journal of Selected Topics in Signal Processing, vol. PP, no. 99, pp. 1–1, 2017.

[8] J. Guerreiro, R. Dinis, and P. Montezuma, “Use of 1-bit digital-to-analogue converters in massive mimo systems,” Electronics Letters, vol. 52, no. 9, pp. 778–779, 2016.

[9] Z. Zhang, X. Cai, C. Li, C. Zhong, and H. Dai, “One-bit quantized massive mimo detection based on variational approximate message passing,” IEEE Transactions on Signal Processing, vol. PP, no. 99, pp. 1–1, 2017.

[10] P. Dong, H. Zhang, W. Xu, and X. You, “Efficient low-resolution adc relaying for multiuser massive mimo system,” IEEE Transactions on Vehicular Technology, vol. PP, no. 99, pp. 1–1, 2017.
[11] O. B. Usman, J. A. Nossek, C. A. Hofmann, and A. Knopp, “Joint mmse precoder and equalizer for massive mimo using 1-bit quantization,” in *IEEE International Conference on Communications*, 2017, pp. 1–6.

[12] H. Yang, G. Geraci, T. Q. S. Quek, and J. Andrews, “Cell-edge-aware precoding for downlink massive mimo cellular networks,” *IEEE Transactions on Signal Processing*, vol. PP, no. 99, pp. 1–1, 2017.

[13] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, “On one-bit quantized zi precoding for the multiuser massive mimo downlink,” in *Sensor Array and Multichannel Signal Processing Workshop*, 2016.

[14] J. Mo, A. Alkhateeb, S. Abu-Surra, and R. W. Heath, “Hybrid architectures with few-bit adc receivers: Achievable rates and energy-rate tradeoffs,” *IEEE Transactions on Wireless Communications*, vol. 16, no. 4, pp. 2274–2287, 2017.

[15] C. Kong, A. Mezghani, C. Zhong, A. L. Swindlehurst, and Z. Zhang, “Multipair massive mimo relaying systems with one-bit adcs and dacs,” *IEEE Transactions on Signal Processing*, vol. PP, no. 99, pp. 1–1, 2017.

[16] Y. Li, C. Tao, A. L. Swindlehurst, A. Mezghani, and L. Liu, “Downlink achievable rate analysis in massive mimo systems with 1-bit dacs,” *IEEE Communications Letters*, vol. 21, no. 7, pp. 1669–1672, 2017.

[17] H. Jedda and J. A. Nossek, “On the statistical properties of constant envelope quantizers,” *IEEE Wireless Communications Letters*, vol. PP, no. 99, pp. 1–1, 2018.

[18] S. Jacobsson, G. Durisi, M. Coldrey, and C. Studer, “Massive mu-mimo-ofdm downlink with one-bit dacs and linear precoding,” in *GLOBECOM 2017* - 2017 IEEE Global Communications Conference, 2017, pp. 1–6.

[19] C. Studer and G. Durisi, “Quantized massive mu-mimo-ofdm uplink,” *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2387–2399, 2016.

[20] C. Mollen, J. Choi, E. G. Larsson, and R. W. Heath, “Uplink performance of wideband massive mimo with one-bit adcs,” *IEEE Transactions on Wireless Communications*, vol. PP, no. 99, pp. 1–1, 2016.

[21] C. K. Wen, C. J. Wang, S. Jin, K. K. Wong, and P. Ting, “Bayes-optimal joint channel-and-data estimation for massive mimo with low-precision adcs,” *IEEE Transactions on Signal Processing*, vol. 60, no. 10, pp. 2541–2556, 2016.

[22] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, “Nonlinear 1-bit precoding for massive mu-mimo with higher-order modulation,” in *2016 50th Asilomar Conference on Signals, Systems and Computers*, 2016, pp. 763–767.

[23] ——, “Quantized precoding for massive mu-mimo,” in *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4670 – 4684, 2017.

[24] H. Jedda, J. A. Nossek, and A. Mezghani, “Minimum ber precoding in 1-bit massive mimo-systems,” in *Sensor Array and Multichannel Signal Processing Workshop*, 2016.

[25] O. Castaneda, S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, “1-bit massive mimo precoding in vlsi,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 7, no. 4, pp. 508 – 522, 2017.

[26] C. J. Wang, C. K. Wen, S. Jin, and S. H. L. Tsai, “Finite-alphabet precoding for massive mu-mimo with low-resolution dacs,” *IEEE Transactions on Wireless Communications*, vol. PP, no. 99, pp. 1–1, 2018.

[27] L. Chu, F. Wen, L. Li, and R. Qiu, “Efficient nonlinear precoding for massive mu-mimo downlink systems with 1-bit dacs,” http://cn.arxiv.org/abs/1804.08839, 2018.

[28] I. C. Wong and B. L. Evans, “Optimal resource allocation in the ofdma downlink with imperfect channel knowledge,” *IEEE Transactions on Communications*, vol. 57, no. 1, pp. 232–241, 2009.

[29] P. Marsch and G. Fettweis, “On downlink network mimo under a constrained backhaul and imperfect channel knowledge,” in *IEEE International Conference on Communications*, 2009, pp. 1–6.

[30] D. Ding and S. D. Blostein, “Maximum mutual information design for mimo systems with imperfect channel knowledge,” *IEEE Transactions on Information Theory*, vol. 56, no. 10, pp. 4793–4801, 2010.

[31] A. Mukherjee and A. L. Swindlehurst, “Robust beamforming for security in mimo wiretap channels with imperfect csi,” *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 351–361, 2010.

[32] M. Rezk and B. Friedlander, “On high performance mimo communications with imperfect channel knowledge,” *IEEE Transactions on Wireless Communications*, vol. 10, no. 2, pp. 602–613, 2011.

[33] N. Lee, O. Simeone, and J. Kang, “The effect of imperfect channel knowledge on a mimo system with interference,” *IEEE Transactions on Communications*, vol. 60, no. 8, pp. 2221–2229, 2012.

[34] M. H. Al-Ali and K. C. Ho, “Transmit precoding in underlay mimo cognitive radio with unavailable or unpakable knowledge of primary interference channel,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 8, pp. 5143–5155, 2016.