Analysis of One-Bit Quantized Precoding for the Multiuser Massive MIMO Downlink

A Kant Saxena, Inbar Fijalkow, A Lee Swindlehurst

Analysis of One-Bit Quantized Precoding for the Multiuser Massive MIMO Downlink

Amodh Kant Saxena, Inbar Fijalkow, *Senior Member, IEEE*, and A. Lee Swindlehurst, *Fellow, IEEE*

Abstract - We present a mathematical analysis of linear precoders for downlink massive MIMO multiuser systems that employ one-bit digital-to-analog converters at the basestation in order to reduce complexity and mitigate power usage. The analysis is based on the Bussgang theorem, and applies generally to any linear precoding scheme. We examine in detail the special case of the quantized zero-forcing (ZF) precoder, and derive a simple asymptotic expression for the resulting symbol error rate at each terminal. Our analysis illustrates that the performance of the quantized ZF precoder depends primarily on the ratio of the number of antennas to the number of users, and our simulations show that it achieves performance similar to a more complicated nonlinear least-squares encoder for low-to-moderate signal to noise ratios, where massive MIMO systems are presumed to operate. We also use the Bussgang theorem to derive a new linear precoder optimized for the case of one-bit quantization, and illustrate its improved performance.

I. Introduction

MASSIVE MIMO involves the use of many, perhaps hundreds, of antennas at the base station (BS) of a wireless network, and can potentially provide large increases in capacity via spatial multiplexing [1]. In a multi-user (MU) scenario, the massive MIMO BS typically serves a number of users much smaller than the number of antennas, and hence a large number of degrees-of-freedom can be offered to each user. This can in turn lead to improved robustness and correspondingly high data rates [1]–[3].

Under favorable propagation conditions, the user channels become asymptotically orthogonal as the number of antennas grows, and simple linear precoding at the BS can be used to invert the channel without noise enhancement. Many studies consider Maximal Ratio Transmission (MRT) [4] or Zero Forcing (ZF) precoders [5], and asymptotic results from random matrix theory show how an increasing number of antennas can result in a dramatic increase in downlink capacity [6] even for these simple precoding schemes.

While the benefits of massive MIMO at the BS grow with the number of antennas, so do the resulting power consumption and hardware costs. While one can scale down the transmit power with an increase in the number of antennas in order to maintain a certain level of performance (e.g., due to beamforming gain), there are certain sources of fixed power consumption at the circuit level that cannot be reduced and these sources will lead to an increase in power as the number of antennas is increased [7]. More important than this is the issue of energy efficiency; a standard RF implementation requires highly linear amplifiers that must as a result be operated with considerable power back-off, which severely limits the overall energy efficiency of the system. The more RF chains, the less and less efficient the system is.

One approach to addressing this problem in the massive MIMO downlink is the use of hybrid analog and digital RF front ends, which employ fewer RF chains in favor of an analog beamforming (precoding) network that is deployed after the digital-to-analog converters (DACs) [8]–[10]. However, this approach does not scale well for wideband systems, as one must either use the same RF beamforming network for the entire band (which is clearly suboptimal), or one must add complexity to the RF analog domain, in the form of either additional phase-shift networks for different frequency bands, or some type of analog tapped-delay line. Instead, we focus on another approach that has gained attention recently, namely the use of low-resolution DACs for each antenna and RF chain; in particular, we will investigate the simplest possible case involving one-bit DACs. Using one-bit ADCs/DACs considerably reduces power consumption, which grows linearly with increases in bandwidth and sampling rate, and exponentially in the number of quantization bits [11]–[13]. Unlike hybrid beamforming schemes, extending one-bit systems to the wideband case does not require further complicating the RF analog design; in particular (and more importantly for the downlink), it considerably simplifies the RF architecture by eliminating the need for highly linear amplifiers and back-off operation, which further reduces circuit complexity and dramatically improves energy efficiency. As we will show in this paper, the severe distortion caused by the one-bit DACs can be mitigated by proper signal processing, and the impact is not too significant in the low- to mid-SNR ranges where massive MIMO systems will likely operate.

Most of the work on one-bit quantization for wireless communication systems has focused on the uplink, where the BS employs one-bit analog-to-digital converters (ADCs).

Manuscript submitted for review on October 17, 2016. First revision submitted on April 14, 2017. Manuscript was accepted on May 24, 2017.

Amodh Kant Saxena and A. Lee Swindlehurst are with the Center for Pervasive Communications and Computing, University of California Irvine, Irvine, CA 92697, USA (e-mail: {aksaxena, swindle}@uci.edu)

Inbar Fijalkow is with ETIS, UMR 8051 / ENSEA, Université Cergy-Pontoise, CNRS, F-95000 Cergy, France (e-mail: inbar.fijalkow@ensea.fr)

The research was supported by the National Science Foundation under Grant ECCS-1547155, and by the Institute for Advanced Study at the Technische Universität München, funded by the German Excellence Initiative and the European Union Seventh Framework Programme under grant agreement No. 291763, and by the European Union under the Marie Curie COFUND Program, and by the French CNRS.

Single-antenna studies of the impact of one-bit ADCs can be found in [14]–[16]. More recently, their use in MIMO systems has been considered, and the resulting work has focused primarily on channel estimation and information theoretic rate analyses [17]–[22].While there has been considerable research on downlink precoding for massive MIMO (see e.g., [23]-[24]) very little has been reported on the impact of low-resolution DACs on transmit processing. In [25], transmit optimization for the case of flat fading MIMO systems with low resolution DACs is addressed. The mean squared error between the received symbol and the symbol vector input to the transmitter is minimized to find optimum quantizer levels, transmit matrix and scalar receiver. In [26], a precoding technique is introduced which aims to minimize the inter-user interference and quantization noise introduced by using a look-up table for all possible transmit sequence combinations. This paper also introduced a novel minimum Bit Error Rate (BER) performance metric. In [27], a two stage precoder is proposed, which comprises a digital precoder, and an analog precoder implemented after the quantizer. The precoders are optimized by minimizing the mean square error between the transmit vector and the receive vector. An iterative algorithm is utilized in the optimization problem.

In this paper, we study the impact of one-bit DACs on linear precoding for the massive MIMO downlink. We presented a preliminary analysis of this problem in [28] using a different approach. To focus on the performance degradation due to quantization, we assume that the BS has perfect channel state information, although this additional error source would have to be accounted for in a full analysis. Using the Bussgang theorem [29] to model the second-order statistics of the quantization noise introduced by the DACs, we provide a closed-form expression for the signal to quantization, interference and noise ratio (SQINR), which we use to deduce the symbol error rate for each terminal in the network. We then focus on the special case of the zero-forcing (ZF) precoder and use asymptotic arguments to obtain an even simpler expression. Our analysis illustrates that the performance of the quantized ZF precoder depends primarily on the ratio of the number of antennas to the number of user terminals, with asymptotically improving performance as this ratio increases. Our simulations indicate that it has performance similar to a more complicated nonlinear least squares (NLS) encoder for low-to-moderate signal to noise ratios, at least for small system dimensions where it is feasible to implement NLS. Finally, using our insights from the asymptotic analysis of the ZF precoder, we design a modified precoder that attempts to achieve the benefits of the ZF precoder even in the non-asymptotic case.

The paper is organized as follows. Section II introduces the one-bit downlink model, and describes both direct NLS precoding and the simpler quantized linear precoding approach. The SQINR performance of a general one-bit quantized linear precoder is then analyzed in Section III, and the approximate Symbol Error Rate (SER) for each user is derived. Section IV focuses on the special case of ZF precoding in the asymptotic regime where the number of BS antennas M and user terminals K become large, leading to a simpler and more insightful expression. In Section V, we introduce the Bussgang adapted precoding algorithm, which attempts to remove the interuser interference for non-asymptotic values of M and K. Simulation results comparing the various algorithms are presented throughout the paper.

II. One-Bit Downlink System Model

A. Mathematical Notation and Assumptions

In what follows, uppercase boldface letters, $\mathbf{A}$, indicate a matrix, with $[\mathbf{A}]_{kl}$ and a_{kl} interchangeably denoting the element at the k^{th} row and l^{th} column. Lower boldface letters, $\mathbf{a}$, indicate a column vector, with a_k denoting the k^{th} element of the column vector. The symbols $(.)^*$, $(.)^T$ and $(.)^H$ denote the complex conjugate, matrix transpose and the transpose-conjugate of the argument respectively. We will use diag$(\mathbf{C})$ to denote the square matrix whose main-diagonal elements are equal to those of the square matrix $\mathbf{C}$, and whose other entries are all zero. With a vector argument, Diag$(\mathbf{c})$ denotes the diagonal matrix whose main diagonal is composed of the elements of vector $\mathbf{c}$.

We assume a flat-fading downlink scenario in which an M-antenna BS is attempting to send QPSK symbols s_k to $k = 1, \cdots, K$ single-antenna users over the $K \times M$ channel $\mathbf{H}$. The BS transmits an $M \times 1$ vector $\sqrt{\rho}\mathbf{x}$, where $\sqrt{\rho}$ is a fixed gain and the elements of $\mathbf{x}$ are constrained to be equal to $\pm 1 \pm j$ due to the use of one-bit quantization of the in-phase and quadrature components of the signal at the BS. Let r_k be the signal received by user k, and define $\mathbf{r} = [r_1 \ \cdots \ r_K]^T$ so that we can write the overall system model as

$$\mathbf{r} = \sqrt{\rho}\mathbf{H}\mathbf{x} + \mathbf{n} \ , \tag{1}$$

where the $K \times 1$ vector $\mathbf{n}$ represents a vector of independent Gaussian noise terms of variance σ_n^2 at each user. For the downlink, the BS designs the vector $\mathbf{x}$ such that the elements of $\mathbf{r}$ can be correctly decoded as the appropriate QPSK symbols in the vector $\mathbf{s} = [s_1 \ \cdots \ s_K]^T$. The QPSK symbols for different users are assumed to be zero-mean and independent with power σ_s^2: $E(\mathbf{s}\mathbf{s}^H) = \sigma_s^2\mathbf{I}_K$. The assumption that the BS transmits QPSK symbols can be interpreted to mean that the individual users also employ one-bit ADCs, although this is not strictly necessary.

The k^{th} row of channel matrix $\mathbf{H}$ is denoted as $\mathbf{h}_k$, and represents the channel to user k. For our analysis, we will assume that the channel matrix is given by

$$\mathbf{H} = \text{Diag}(\sigma_1, \sigma_2, \ldots, \sigma_K)\tilde{\mathbf{H}} \ , \tag{2}$$

where the elements of $\tilde{\mathbf{H}}$ are complex Gaussian random variables, whose real and imaginary parts are both iid Gaussian random variables with zero-mean and unit-variance, and the parameters $\sigma_1, \sigma_2, \ldots, \sigma_K$ represent the individual channel gains for each user. As mentioned above, we will assume that the channel $\mathbf{H}$ is known at the BS in order to isolate the impact of the quantization and noise.

B. Nonlinear LS Precoding

Ideally, in the absence of noise, one might attempt to design a general non-linear precoder providing $\mathbf{x}$ such that $\mathbf{H}\mathbf{x} = \mathbf{s}$.

However, due to the finite alphabet constraint imposed by the one-bit DACs, one would have to find such an $\mathbf{x}$ with QPSK entries, which in general will prevent achieving $\mathbf{Hx} = \mathbf{s}$ with equality. Instead, one might choose to implement a nonlinear least squares (NLS) precoder that attempts to solve [30]–[32]

$$\mathbf{x} = \arg \min_{\delta>0, \mathbf{v}\in\mathcal{S}^M} ||\mathbf{s} - \delta\mathbf{Hv}||^2 , \tag{3}$$

where $\mathcal{S} = \{1+j, 1-j, -1+j, -1-j\}$ is the set of QPSK constellation points with $j = \sqrt{-1}$, and δ is a positive real-valued scaling. However, in general, (3) requires on exhaustive search of order $O(4^M)$, which is prohibitively expensive even for relatively small values of M, let alone in the massive MIMO case. Even a sphere-based encoder [33] would be too complex for large values of M, and in this case would require extra care since the matrix $\mathbf{H}$ has many more columns than rows. In such cases, one should transform (3) to

$$\mathbf{x} = \arg \min_{\delta>0, \mathbf{v}\in\mathcal{S}^M} ||\mathbf{D}(\mathbf{z} - \delta\mathbf{v})||^2 , \tag{4}$$

where $\mathbf{D}$ is the upper triangular matrix obtained by the Cholesky factorization of $\mathbf{G} = \mathbf{H}^H\mathbf{H} + \alpha\mathbf{I}_M$, $\mathbf{z} = \mathbf{G}^{-1}\mathbf{H}^H\mathbf{s}$ and α is a small regularization parameter as explained in [34]. Though less complex than direct NLS encoding, the generalized sphere encoder still has a complexity exponential in $M - K$, which is again costly in the massive MIMO case. Approaches that approximate the solution to (3) using convex semi-definite relaxation have been proposed in [32], although the resulting optimization, while considerably simpler, is still somewhat complex and must be done on a per symbol basis.

We note here that, for the case where the elements of the desired vector $\mathbf{s}$ are themselves drawn from a finite alphabet (QPSK here), the NLS encoder over-constrains the problem by attempting to force $\delta\mathbf{Hx}$ to be close to $\mathbf{s}$, when in fact all that is necessary is that its elements lie within the correct decision regions so that the users can properly decode them as the desired constellation points s_k. The noise-free received data $\delta\mathbf{Hx}$ can in principle be far away from $\mathbf{s}$ and still be decoded correctly; in fact, often the farther $\delta\mathbf{Hx}$ is away from $\mathbf{s}$, the farther the received signal is from the decision boundaries and hence the more resilient to noise. So we might expect that NLS encoding may not give optimal performance in this case, and in fact we will demonstrate this fact later in the paper.

C. One-bit Quantized Linear Precoding

As an alternative to NLS encoding, we will study the performance of a very simple approach in which the output of a standard linear precoder is quantized by one-bit DACs prior to transmission. In particular, assuming a linear precoding matrix $\mathbf{P}$, the transmitted signal is $\mathbf{x} = \mathbb{Q}(\mathbf{Ps})$, where the one-bit quantization operation is defined as

$$\mathbb{Q}(\mathbf{a}) = \text{sign}(\Re(\mathbf{a})) + j\text{sign}(\Im(\mathbf{a})), \tag{5}$$

with $\Re(\cdot)$ representing the real part, $\Im(\cdot)$ the imaginary part, and sign$(\cdot)$ the sign of their arguments. Figure 1 gives a graphical view of the assumed system, whose output is thus given by

$$\mathbf{r} = \sqrt{\rho}\mathbf{H}\mathbb{Q}(\mathbf{Ps}) + \mathbf{n} . \tag{6}$$

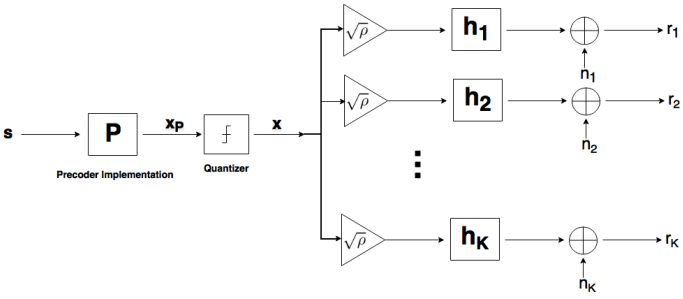

Fig. 1: System Model

In what follows, we will define $2\rho_0$ to be the total transmit power, which implies that $\rho = \frac{\rho_0}{M}$.

III. Bussgang Analysis of One-Bit Quantized Precoding

Let $\mathbf{x}_P = \mathbf{Ps}$ represent the precoded vector before quantization. In this section, we use the Bussgang decomposition to analyze the impact of the quantization on the signal of interest and to quantify the level of quantization noise. This will allow us in the sequel to approximate the SQINR.

A. Bussgang Decomposition

The one-bit quantization operation on the precoded vector $\mathbf{x}_P$ is modeled here using the Bussgang theorem [29]. We assume that the vector of QPSK symbols $\mathbf{s}$ is random with zero mean and covariance matrix $\sigma_s^2\mathbf{I}_K$. Although this implies that $\mathbf{x}_P$ is not strictly Gaussian, each element of $\mathbf{x}_P$ is formed as a result of a linear mixture of the K $i.i.d.$ elements of the vector $\mathbf{s}$, the Gaussian assumption is fulfilled for large enough K. We thus apply the Bussgang theorem to write

$$\mathbf{x} := \mathbb{Q}(\mathbf{x}_P) = \mathbf{F}\mathbf{x}_P + \mathbf{q} , \tag{7}$$

where $\mathbf{F}$ is chosen to satisfy $\mathbf{R}_{\mathbf{xp q}} = E(\mathbf{x}_P\mathbf{q}^H) = \mathbf{0}$. A similar property is demonstrated for transmit-side nonlinearities in [35]. The Bussgang theorem provides a linear representation of the quantization that is statistically equivalent up to the second moments of the data. To define the decomposition, we have

$$\begin{aligned}\mathbf{R}_{\mathbf{xx_P}} = E(\mathbf{x}\mathbf{x}_P^H) &= E(\{\mathbf{Fx_P} + \mathbf{q}\}\mathbf{x}_P^H) \\ &= \mathbf{F}\mathbf{R}_{\mathbf{x_P x_P}} ,\end{aligned} \tag{8}$$

where

$$\mathbf{R}_{\mathbf{x_P x_P}} = \sigma_s^2\mathbf{P}\mathbf{P}^H. \tag{9}$$

Note that the $M \times M$ matrix in (9) is rank deficient, and thus $\mathbf{F}$ cannot be solved for directly as in [16]. We will see shortly that a unique expression for $\mathbf{F}$ is unnecessary, and that (8) is sufficient.

Under the mutual Gaussian assumption between the components of $\mathbf{x}_P$, the inter-correlation between the one-bit quantized x_k and unquantized $x_{P,l}$ is equal to the normalized inter-correlation of the unquantized signals reduced by a factor of $\sqrt{2/\pi}$, as in [36]:

$$E(x_k x_{P,l}^*) = \sqrt{\frac{2}{\pi}}\frac{E(x_{P,k}x_{P,l}^*)}{\sigma_{x_{P,k}}} ,$$

where $\sigma_{x_{P,k}} = \{E(x_{P,k}x_{P,k}^*)\}^{\frac{1}{2}}$. In matrix form this yields

$$\begin{aligned}\mathbf{R}_{\mathbf{x}\mathbf{x}_\mathbf{P}} =&\sqrt{\frac{2}{\pi}}\left\{\text{diag}(E(\mathbf{x}_P\mathbf{x}_P^H))\right\}^{-\frac{1}{2}}E(\mathbf{x}_P\mathbf{x}_P^H)\\ =&\sqrt{\frac{2}{\pi}}\sigma_s\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\mathbf{P}\mathbf{P}^H.\end{aligned} \tag{10}$$

Thus, from (8),

$$\mathbf{F}\mathbf{P}\mathbf{P}^H = \frac{1}{\sigma_s}\sqrt{\frac{2}{\pi}}\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\mathbf{P}\mathbf{P}^H\ , \tag{11}$$

and since $\mathbf{P}$ is full column rank, we have

$$\mathbf{F}\mathbf{P} = \frac{1}{\sigma_s}\sqrt{\frac{2}{\pi}}\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\mathbf{P}. \tag{12}$$

Note that since $\mathbf{P}\mathbf{P}^H$ is not invertible, $\mathbf{F}$ can not be uniquely defined. However, we will see in Section III-B that this poses no problem since an expression for $\mathbf{F}\mathbf{P}$ will be sufficient for our analysis.

It is also useful to derive here the covariances of the quantization noise $\mathbf{q}$ and the data vector $\mathbf{x}$ after quantization. Using the arcsin law, for a hard limiting one-bit quantizer, we have [36]

$$\begin{aligned}E(x_k x_l^*) =&\frac{2}{\pi}\text{arcsin}\left(\Re\left(\frac{E(x_{P,k}x_{P,l}^*)}{\sigma_{x_{P,k}}\sigma_{x_{P,l}}}\right)\right)\\ &+j\frac{2}{\pi}\text{arcsin}\left(\Im\left(\frac{E(x_{P,k}x_{P,l}^*)}{\sigma_{x_{P,k}}\sigma_{x_{P,l}}}\right)\right),\end{aligned} \tag{13}$$

which implies

$$\begin{aligned}\mathbf{R}_{\mathbf{x}\mathbf{x}} =&\frac{2}{\pi}\text{arcsin}\{\{\text{diag}(\mathbf{R}_{\mathbf{x}_\mathbf{P}\mathbf{x}_\mathbf{P}})\}^{-\frac{1}{2}}\Re\left(\mathbf{R}_{\mathbf{x}_\mathbf{P}\mathbf{x}_\mathbf{P}}\right)\\ &\times\{\text{diag}(\mathbf{R}_{\mathbf{x}_\mathbf{P}\mathbf{x}_\mathbf{P}})\}^{-\frac{1}{2}}\}\\ &+j\frac{2}{\pi}\text{arcsin}\{\{\text{diag}(\mathbf{R}_{\mathbf{x}_\mathbf{P}\mathbf{x}_\mathbf{P}})\}^{-\frac{1}{2}}\Im\left(\mathbf{R}_{\mathbf{x}_\mathbf{P}\mathbf{x}_\mathbf{P}}\right)\\ &\times\{\text{diag}(\mathbf{R}_{\mathbf{x}_\mathbf{P}\mathbf{x}_\mathbf{P}})\}^{-\frac{1}{2}}\}.\end{aligned} \tag{14}$$

For the quantization noise vector $\mathbf{q}$, we have

$$\begin{aligned}\mathbf{R}_{\mathbf{q}\mathbf{q}} =&\mathbf{R}_{\mathbf{x}\mathbf{x}} - \mathbf{F}\mathbf{R}_{\mathbf{x}_\mathbf{P}\mathbf{x}}\\ &+\mathbf{F}\mathbf{R}_{\mathbf{x}_\mathbf{P}\mathbf{x}_\mathbf{P}}\mathbf{F}^H - \mathbf{R}_{\mathbf{x}\mathbf{x}_\mathbf{P}}\mathbf{F}^H\\ =&\mathbf{R}_{\mathbf{x}\mathbf{x}} - \mathbf{F}\mathbf{P}\mathbf{P}^H\mathbf{F}^H\sigma_s^2\\ =&\frac{2}{\pi}\{\text{arcsin}\{\{\text{diag}(\mathbf{P}\mathbf{P}^H)\}^{-\frac{1}{2}}\Re\left(\mathbf{P}\mathbf{P}^H\right)\\ &\times\{\text{diag}(\mathbf{P}\mathbf{P}^H)\}^{-\frac{1}{2}}\}\\ &+j\ \text{arcsin}\{\{\text{diag}(\mathbf{P}\mathbf{P}^H)\}^{-\frac{1}{2}}\Im\left(\mathbf{P}\mathbf{P}^H\right)\\ &\times\{\text{diag}(\mathbf{P}\mathbf{P}^H)\}^{-\frac{1}{2}}\}\\ &-\{\text{diag}(\mathbf{P}\mathbf{P}^H)\}^{-\frac{1}{2}}\mathbf{P}\mathbf{P}^H\{\text{diag}(\mathbf{P}\mathbf{P}^H)\}^{-\frac{1}{2}}\}.\end{aligned} \tag{15}$$

B. Impact on the Signal of Interest

Let $\tilde{\mathbf{s}}$ be the noiseless received signal vector

$$\tilde{\mathbf{s}} = \sqrt{\rho}\mathbf{H}\mathbf{x} = \frac{\sqrt{\rho_0}}{\sqrt{M}}\mathbf{H}\mathbf{x}. \tag{16}$$

The cross-correlation between the received $\tilde{\mathbf{s}}$ and desired $\mathbf{s}$ is

$$\begin{aligned}\mathbf{R}_{\tilde{\mathbf{s}}\mathbf{s}} =&\frac{\sqrt{\rho_0}}{\sqrt{M}}E(\mathbf{H}\mathbf{x}\mathbf{s}^H)\\ =&\frac{\sqrt{\rho_0}}{\sqrt{M}}\mathbf{H}E\{(\mathbf{F}\mathbf{P}\mathbf{s}+\mathbf{q})\mathbf{s}^H\}\\ =&\frac{\sqrt{\rho_0}\sigma_s^2}{\sqrt{M}}\mathbf{H}\mathbf{F}\mathbf{P}\ ,\end{aligned} \tag{17}$$

where (17) results because

$$E(\mathbf{x}_\mathbf{P}\mathbf{q}^H) = \mathbf{P}E(\mathbf{s}\mathbf{q}^H) = \mathbf{0}\ . \tag{18}$$

Since $\mathbf{P}$ is full column rank, we have

$$E(\mathbf{s}\mathbf{q}^H) = \mathbf{0} \tag{19}$$

In the sequel, we denote $\mathbf{G} = \mathbf{H}\mathbf{F}\mathbf{P}$. Using (12), $\mathbf{G} = \sqrt{\frac{2}{\pi}}\frac{1}{\sigma_s}\mathbf{H}\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\mathbf{P}$ and

$$\mathbf{R}_{\tilde{\mathbf{s}}\mathbf{s}} = \sqrt{\frac{2}{\pi}}\frac{\sqrt{\rho_0}\sigma_s}{\sqrt{M}}\mathbf{H}\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\mathbf{P}. \tag{20}$$

Equation (20) shows that, for any full rank precoder, the impact of the one-bit quantization on the signal of interest is the diagonal matrix $\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}$ and a scalar factor $\sqrt{\frac{2}{\pi}}$.

C. SQINR and Probability of Error

Using the Bussgang decomposition, the received vector after quantization can be represented as

$$\begin{aligned}\mathbf{r} =&\frac{\sqrt{\rho_0}}{\sqrt{M}}\mathbf{H}(\mathbf{F}\mathbf{P}\mathbf{s}+\mathbf{q}) + \mathbf{n}\\ =&\frac{\sqrt{\rho_0}}{\sqrt{M}}\mathbf{G}\mathbf{s} + \frac{\sqrt{\rho_0}}{\sqrt{M}}\mathbf{H}\mathbf{q} + \mathbf{n}\ .\end{aligned} \tag{21}$$

Letting $\mathbf{d} = \mathbf{H}\mathbf{q}$, we denote the covariance matrix of the received quantized noise as

$$\mathbf{R}_{\mathbf{d}\mathbf{d}} = \mathbf{H}\mathbf{R}_{\mathbf{q}\mathbf{q}}\mathbf{H}^H. \tag{22}$$

With these definitions, the SQINR experienced by user k for an arbitrary linear precoder $\mathbf{P}$ whose output is one-bit quantized can be expressed as

$$SQINR_k = \frac{\rho_0\frac{|g_{kk}|^2\sigma_s^2}{M}}{\rho_0\sum_{l=1,l\neq k}^{K}\frac{|g_{kl}|^2\sigma_s^2}{M} + \rho_0\frac{[\mathbf{R}_{\mathbf{d}\mathbf{d}}]_{kk}}{M} + \sigma_n^2}\ , \tag{23}$$

where $\rho_0\sum_{l=1,l\neq k}^{K}\frac{|g_{kl}|^2\sigma_s^2}{M}$ accounts for multi-user interference and $\rho_0\frac{[\mathbf{R}_{\mathbf{d}\mathbf{d}}]_{kk}}{M}$ for quantization noise. With the assumption of equally likely Gray-mapped QPSK signaling, using the nearest neighbour approximation, we can calculate the probability of a decoding error for user k as

$$\begin{aligned}P_e =&Pr(\mathbb{Q}(r_k)\neq s_k) \simeq 2Q(\sqrt{SQINR_k})\\ =&2Q\left(\sqrt{\frac{\rho_0\frac{|g_{kk}|^2\sigma_s^2}{M}}{\rho_0\sum_{l=1,l\neq k}^{K}\frac{|g_{kl}|^2\sigma_s^2}{M} + \rho_0\frac{[\mathbf{R}_{\mathbf{d}\mathbf{d}}]_{kk}}{M} + \sigma_n^2}}\right).\end{aligned} \tag{24}$$

IV. Asymptotic Performance of the One-bit Quantized Zero-Forcing Precoder

The previous section provides a closed-form expression for the SQINR for any one-bit quantized linear precoder $\mathbf{P}$. To get additional insight into the impact of the one-bit DACs, here we focus on the special case of the zero-forcing (ZF) precoder defined by

$$\mathbf{P} = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-1} \, . \tag{25}$$

In addition, we will further simplify the resulting expressions by adopting a massive MIMO assumption and letting both M and K be large [3]. In the following analysis, the mathematical results assume M and K to be approaching infinity while the ratio $\gamma = \frac{M}{K} > 1$ is finite. In a physical sense of course it is not possible for the number of antennas to grow to infinity, as a certain separation between the antennas is crucial for maintaining independence between the channel vectors of distinct antennas [37]. However, as we show in our analysis and simulation results, the approximations are very tight even for small and realistic values of the system dimensions and hence of practical value.

A. Approximations for the Asymptotic Case

In our asymptotic analysis, we let M and K grow large while maintaining a finite value for the ratio $\gamma > 1$. In what follows, we recall and extend some results on the asymptotic behaviour of the matrix $(\mathbf{H}\mathbf{H}^H)^{-1}$ needed for analyzing the ZF precoder. As mentioned earlier, the channel matrix is assumed to be described as

$$\mathbf{H} = \mathbf{\Sigma}\tilde{\mathbf{H}} \, , \tag{26}$$

where $\mathbf{\Sigma} = \text{Diag}(\sigma_1, ..., \sigma_K)$ denotes the individual channel gains, and we assume that the elements of $\tilde{\mathbf{H}}$ are $i.i.d.$ circularly symmetric Gaussian random variables with $\Re(\tilde{h}_{kl}) \sim \mathcal{N}(0,1)$ independent of $\Im(\tilde{h}_{kl}) \sim \mathcal{N}(0,1), \quad \forall k = 1, 2, \ldots, K$ and $l = 1, 2, \ldots, M$.

It is shown in [38] that $\mathbf{Z} = 2\mathbf{H}\mathbf{H}^H$ is a complex Wishart matrix (see [39]) with distribution

$$\mathbf{W}_K(M, \mathbf{\Sigma}^2; \mathbf{Z}) = \frac{|\mathbf{Z}|^{\frac{M-K-1}{2}} \exp\left(-\frac{1}{2}\text{trace}(\mathbf{\Sigma^{-2}Z})\right)}{2^{\frac{MK}{2}} \Gamma_K(\frac{M}{2}) |\mathbf{\Sigma^2}|^{\frac{M}{2}}} \tag{27}$$

where

$$\Gamma_K(M) = \pi^{\frac{K(K-1)}{4}} \prod_{l=1}^{K} \Gamma(M + \frac{1-l}{2}) \, ,$$

is the Gamma function. In our case, $\mathbf{\Sigma}$ is diagonal so that

$$(\det 2(\mathbf{\Sigma}))^{2M} = 2^{2KM} \prod_{k=1}^{K} \sigma_k^{2M} \, .$$

The variance of the elements of $\mathbf{Y} = \mathbf{Z}^{-1}$, due to the property of the Wishart distribution [39], is given by

$$Var(y_{kl}) = \frac{1}{(2M-K)(2M-K-1)(2M-K-3)\sigma_k^2\sigma_l^2}$$

for $k \neq l$ and

$$Var(y_{kk}) = \frac{2}{(2M-K-1)^2(2M-K-3)\sigma_k^4}$$

otherwise. Note that, for large values of $M > K$, the variance goes to zero proportionally to $\frac{1}{M^3}$, for all elements of $(\mathbf{H}\mathbf{H}^H)^{-1}$.

In the asymptotic case, the rows of $\mathbf{H}$ become quasi-orthogonal and the diagonal terms of $(\mathbf{H}\mathbf{H}^H)^{-1}$ have been studied in [38] (or [40] when $\mathbf{\Sigma} = \mathbf{I}_K$). Given the Wishart distribution (27), we have

$$(\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H)^{-1} \xrightarrow[M\to\infty]{} \frac{1}{2K(\gamma-1)}\mathbf{I}_K \tag{28}$$

and

$$(\mathbf{H}\mathbf{H}^H)^{-1} \xrightarrow[M\to\infty]{} \frac{1}{2K(\gamma-1)}\mathbf{\Sigma}^{-2} \, , \tag{29}$$

where $\xrightarrow[M\to\infty]{}$ denotes almost sure convergence (*i.e.,* convergence with probability one). Because the variance of the elements of these matrices scale as $1/M^3$, convergence to these mean values is very tight, even though they are asymptotically proportional to $1/M$. We will see in the simulations that the resulting approximations are very accurate even for practical values of M.

From (25) and (29),

$$\mathbf{P} = \tilde{\mathbf{H}}^H\mathbf{\Sigma}\{\mathbf{\Sigma}\tilde{\mathbf{H}}\tilde{\mathbf{H}}^H\mathbf{\Sigma}\}^{-1} \xrightarrow[M\to\infty]{} \frac{1}{2K(\gamma-1)}\tilde{\mathbf{H}}^H\mathbf{\Sigma}^{-1} \, . \tag{30}$$

From (30),

$$\mathbf{P}\mathbf{P}^H = \mathbf{H}^H(\mathbf{H}\mathbf{H}^H)^{-2}\mathbf{H} \xrightarrow[M\to\infty]{} \frac{1}{4K^2(\gamma-1)^2}\tilde{\mathbf{H}}^H\mathbf{\Sigma}^{-2}\tilde{\mathbf{H}}. \tag{31}$$

In what follows, we use these asymptotic approximations to analyze the one-bit quantized ZF precoder.

B. Asymptotic Received Downlink Signal

Using the results of the previous section, we have

$$\begin{aligned}\mathbf{F}\mathbf{P} = & \frac{1}{\sigma_s}\sqrt{\frac{2}{\pi}}\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\mathbf{P} \\ \xrightarrow[M\to\infty]{} & \frac{1}{2K\sigma_s(\gamma-1)}\sqrt{\frac{2}{\pi}}\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\tilde{\mathbf{H}}^H\mathbf{\Sigma}^{-1},\end{aligned} \tag{32}$$

$$\begin{aligned}\mathbf{G} = & \ \mathbf{H}\mathbf{F}\mathbf{P} \\ \xrightarrow[M\to\infty]{} & \frac{1}{2K\sigma_s(\gamma-1)}\sqrt{\frac{2}{\pi}}\mathbf{\Sigma}\tilde{\mathbf{H}}\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\tilde{\mathbf{H}}^H\mathbf{\Sigma}^{-1}.\end{aligned} \tag{33}$$

Although the expression for $\mathbf{P}\mathbf{P}^H$ in (31) is rank deficient for any finite value of K, in the limit as both M and K go to infinity (with $\gamma = M/K$ being finite), $\mathbf{P}\mathbf{P}^H$ converges to the following full-rank diagonal matrix assuming that there is not a significant variation among the channel gains [41]:

$$\mathbf{P}\mathbf{P}^H \xrightarrow[K,M\to\infty]{} \frac{\varsigma}{2K(\gamma-1)^2}\mathbf{I}_M \, , \tag{34}$$

where we have defined $\varsigma = \frac{1}{K}\sum_{i=1}^{K}\sigma_i^{-2}$ to be the average of the inverse squared channel gains. Thus, using (34) and (28) in (33),

$$\begin{aligned}\boldsymbol{\Sigma}\tilde{\mathbf{H}}\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\tilde{\mathbf{H}}^H\boldsymbol{\Sigma}^{-1} \\ \xrightarrow[K,M\to\infty]{} \frac{2K\sqrt{2K}(\gamma-1)^2}{\sqrt{\varsigma}}\mathbf{I}_K,\end{aligned} \tag{35}$$

so that

$$\mathbf{G} \xrightarrow[K,M\to\infty]{} \frac{2\sqrt{K}(\gamma-1)}{\sigma_s\sqrt{\pi\varsigma}}\mathbf{I}_K. \tag{36}$$

Using (20) and (32)-(36), the cross-covariance matrix of $\tilde{\mathbf{s}}$ and $\mathbf{s}$ can be expressed as

$$\begin{aligned}\mathbf{R}_{\tilde{\mathbf{s}}\mathbf{s}} &= \sqrt{\frac{2}{\pi}}\frac{\sqrt{\rho_0}\sigma_s}{\sqrt{M}}\mathbf{H}\left\{\text{diag}(\mathbf{P}\mathbf{P}^H)\right\}^{-\frac{1}{2}}\mathbf{P} \\ &= \frac{\sqrt{\rho_0}\sigma_s^2}{\sqrt{M}}\mathbf{G} \\ &\xrightarrow[K,M\to\infty]{} \frac{2\sigma_s\sqrt{\rho_0}(\gamma-1)}{\sqrt{\pi\varsigma\gamma}}\mathbf{I}_K = \beta\mathbf{I}_K\ .\end{aligned} \tag{37}$$

When the channel gains are all equal to σ, this expression becomes

$$\mathbf{R}_{\tilde{\mathbf{s}}\mathbf{s}} \xrightarrow[K,M\to\infty]{} \frac{2\sigma_s\sigma\sqrt{\rho_0}(\gamma-1)}{\sqrt{\pi\gamma}}\mathbf{I}_K\ . \tag{38}$$

We observe that for large M and K, $\mathbf{R}_{\tilde{\mathbf{s}}\mathbf{s}}$ is diagonal with positive real diagonal entries. This indicates that $\tilde{s}_k$ and s_l are uncorrelated for $k \neq l$, and hence the multi-user interference disappears, and also that the received signal constellation is not rotated. In addition, we see that the signal of interest is received with a gain of $\frac{\beta}{\sigma_s^2}$, which grows as $\sqrt{\gamma}$. Hence, the larger the value of $\gamma = M/K$, the deeper the received signal will be pushed into the correct decision region, and hence the lower the probability of a decoding error in the presence of noise at the receiver.

In Fig. 2, we plot the scaling factor, $\frac{\beta}{\sigma_s^2}$ found by simulation when averaging over all $K = 20$ users and over 10^4 channel realizations for a case with $\sigma_s = \sqrt{2}$, $\rho_0 = 1$ and $\sigma_i = \sigma = 1, \forall i$. The simulation curve is compared with the asymptotic formula in (38), and shows very good accuracy for $\gamma \geq 10$.

C. Asymptotic SQINR and Probability of Error

As described above, the matrix $\mathbf{P}\mathbf{P}^H$ asymptotically becomes a scaled identity matrix in (34). Using this result together with (15), we have

$$\begin{aligned}\mathbf{R}_{\mathbf{dd}} = \mathbf{H}\mathbf{R}_{\mathbf{qq}}\mathbf{H}^H &= \boldsymbol{\Sigma}\tilde{\mathbf{H}}\mathbf{R}_{\mathbf{qq}}\tilde{\mathbf{H}}^H\boldsymbol{\Sigma} \\ &\xrightarrow[K,M\to\infty]{} 2\left(1-\frac{2}{\pi}\right)K(\gamma-1)\boldsymbol{\Sigma}^2,\end{aligned} \tag{39}$$

Substituting Eq. (39) and (36) into the general SQINR expression (23) yields

$$SQINR_k \xrightarrow[K,M\to\infty]{} \frac{\frac{\rho_0}{\sigma_n^2}\frac{4}{\pi\varsigma}\left(\gamma+\frac{1}{\gamma}-2\right)}{\frac{\rho_0}{\sigma_n^2}2\left(1-\frac{2}{\pi}\right)\left(1-\frac{1}{\gamma}\right)\sigma_k^2+1}\ . \tag{40}$$

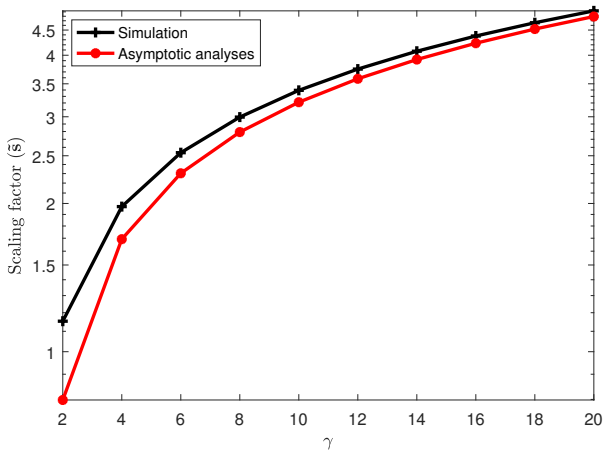

Fig. 2: Asymptotic and simulated average scaling factor, $\frac{\beta}{\sigma_s^2}$ with respect to γ.

As before, assuming equally likely Gray-mapped QPSK symbols, we have the probability of error for the k^{th} user as

$$\begin{aligned}P_e =& Pr(\mathbb{Q}(r_k) \neq s_k) \simeq 2Q(\sqrt{SQINR_k}) \\ \xrightarrow[K,M\to\infty]{}\ & 2Q\left(\sqrt{\frac{\frac{\rho_0}{\sigma_n^2}\frac{4}{\pi\varsigma}\left(\gamma+\frac{1}{\gamma}-2\right)}{\frac{\rho_0}{\sigma_n^2}2\left(1-\frac{2}{\pi}\right)\left(1-\frac{1}{\gamma}\right)\sigma_k^2+1}}\right).\end{aligned} \tag{41}$$

In the case of equal channel gains,

$$SQINR_k \xrightarrow[K,M\to\infty]{} \frac{\frac{\rho_0}{\sigma_n^2}\frac{4\sigma^2}{\pi}\left(\gamma+\frac{1}{\gamma}-2\right)}{\frac{\rho_0}{\sigma_n^2}2\sigma^2\left(1-\frac{2}{\pi}\right)\left(1-\frac{1}{\gamma}\right)+1}, \tag{42}$$

so that

$$P_e \simeq 2Q\left(\sqrt{\frac{\frac{\rho_0}{\sigma_n^2}\frac{4\sigma^2}{\pi}\left(\gamma+\frac{1}{\gamma}-2\right)}{\frac{\rho_0}{\sigma_n^2}2\sigma^2\left(1-\frac{2}{\pi}\right)\left(1-\frac{1}{\gamma}\right)+1}}\right). \tag{43}$$

For high SNR scenarios, the $SQINR_k$ can be approximated by the signal to quantization and interference ratio ($SQIR_k$)

$$SQIR_k \simeq \frac{\frac{2}{\pi}(\gamma-1)}{\left(1-\frac{2}{\pi}\right)\sigma_k^2\varsigma}, \tag{44}$$

and thus the probability of error will experience the following error floor

$$\begin{aligned}P_e \simeq& 2Q(\sqrt{SQIR_k}) \\ \simeq& 2Q\left(\sqrt{\frac{\frac{2}{\pi}}{\left(1-\frac{2}{\pi}\right)}}\sqrt{\frac{\gamma-1}{\sigma_k^2\varsigma}}\right).\end{aligned} \tag{45}$$

For equal channel gains, the high-SNR $SQIR_k$ is given by

$$SQIR_k \simeq \frac{\frac{2}{\pi}}{1-\frac{2}{\pi}}(\gamma-1), \tag{46}$$

so that the error floor becomes

$$P_e \simeq 2Q\left(\sqrt{\frac{\frac{2}{\pi}}{1-\frac{2}{\pi}}(\gamma-1)}\right). \tag{47}$$

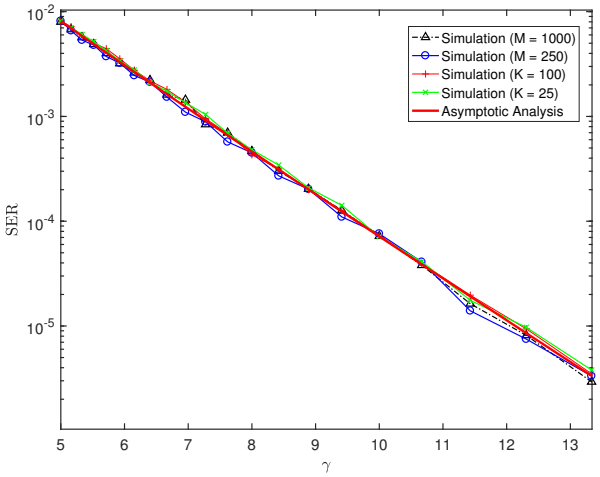

Fig. 3: Variation of one-bit ZF precoding SER with the ratio γ in the noiseless case.

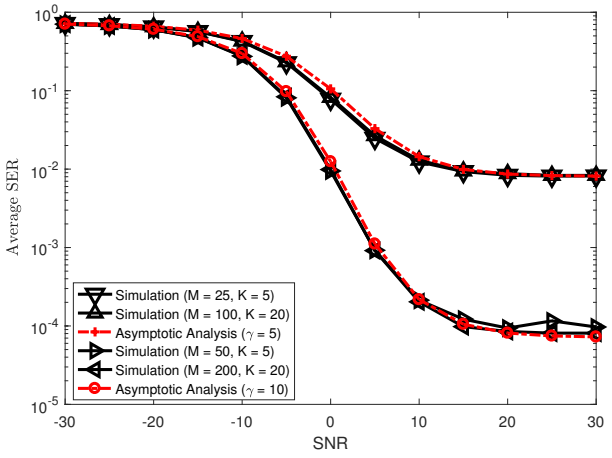

Fig. 4: Variation of SER with SNR, for varying number of users, K and BS antennas, M.

In all cases we note the critical dependence of the SQINR and probability of error on the quantity $\gamma = M/K$; in particular, the SQINR increases approximately linearly with γ. In Fig. 3, we have plotted the symbol error rate (SER) for the case of no additive noise as a function of γ for various choices of M and K averaged over 10^6 channel realizations. We see that the simulations match the analysis very well, and illustrate the importance of the ratio γ on performance. Massive MIMO systems are typically envisioned to operate with loading factors on the order of $\gamma \simeq 10$, and we see that for this case the SER due to the one-bit quantization alone is below 10^{-4}, which bodes well for the use of the quantized ZF precoder in practical scenarios.

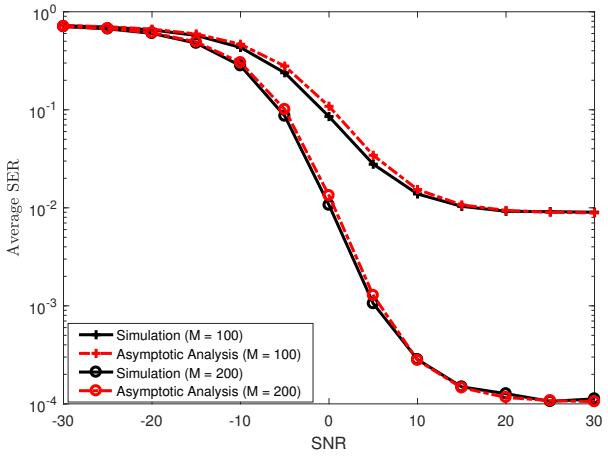

Fig. 5: Variation of SER with SNR for $K = 20$ users with unequal channel gains and varying values of BS antennas M.

D. Simulations

In Fig. 4, we have plotted both the predicted and simulated SER[1] for one-bit quantized ZF precoding at the BS for varied number of users as a function of the SNR, $\frac{2\rho_0}{\sigma_n^2}$, as described in Section II-C. Again, 10^6 channel realizations were used to generate the results, assuming equal channel gains, $\sigma_i = 1, \forall i = 1, \ldots, K$. We note the excellent match between the simulations and analytical approximation in (43), which validates our analysis. As expected, we observe that the SER approaches an error floor at high SNR; for example, with $M \sim 10K$, the SER floor is of the order of 10^{-4}. To see how the analysis holds for non-asymptotic values of K and M, we have also performed the simulation for the case of $K = 5$ users. We observe that our asymptotic analysis is accurate even for this non-asymptotic value. The results of simulation are similar for $K = 5$ and $K = 20$ for both $M/K = 5, 10$. This result reinforces the observation that performance is governed by the ratio $\gamma = M/K$, independent of their specific values.

In Fig 5, we plot the average SER for the quantized ZF precoding scenario, again with $K = 20$. Unlike the previous example, we have assumed here that the users have unequal channel gains; in this simulation, the square of the gains were chosen as independent log-normal random variables, such that $\ln(\sigma_i^2) \sim \mathcal{N}(\mu, v^2)$, $\forall\, i = 1, \ldots, K$ with parameters $v = 0.125$ and $\mu = -v^2/2$, so that for large enough K, $\frac{1}{K}\sum_{k=1}^{K} \sigma_k^2 \to e^{\mu + \frac{v^2}{2}} = 1$. The results were generated using 10^5 channel realizations. For this case also, we can see that the SER approaches an error floor of the order of 10^{-4} when $M \sim 10K$ for high SNR. The simulations agree very well with the asymptotic analysis.

In Fig. 6 we compare the NLS encoding approach (3) with the quantized ZF precoder. Due to the complexity of the NLS encoder, we can only perform the simulation for the relatively

[1]Assuming a Grey-encoded QPSK constellation, the bit-error rate (BER) will typically equal $1/2$ the SER except for cases with extremely low SNR where performance is not particularly of interest. We observed this correspondence in our simulation results; in the several cases we tested, there was virtually no difference between the BER and $1/2$ the SER.

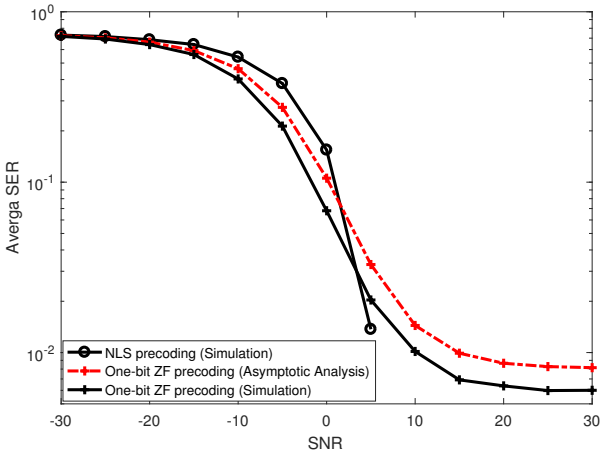

Fig. 6: Variation of SER versus SNR for $K = 2$ users and $M = 10$ BS antennas for one-bit ZF precoding and NLS encoding for equal channel gains.

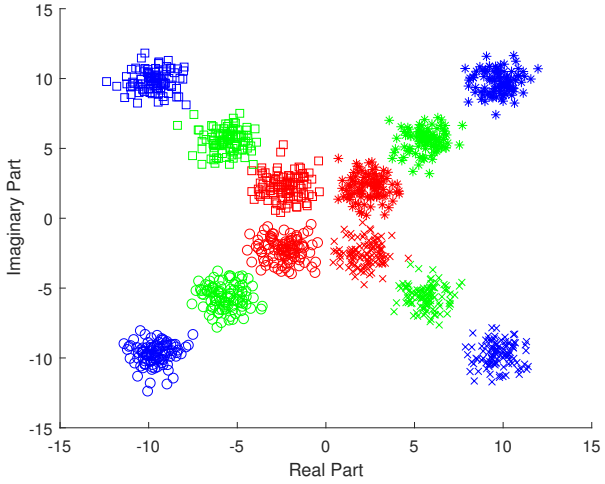

Fig. 7: Plot of the received vector in a noiseless scenario, for $K = 4$, and different values of M, $M = 20$ (red), 100 (green), 300 (blue). Different components of the received vector are shown as different symbols.

small values $M = 10$ and $K = 2$. The gains of the channel rows are assumed to be equal: $\sigma_1 = \sigma_2 = 1$, and 10^4 channel realizations were generated to compute the results. The value of δ for the NLS approach was chosen to be $\delta = 1/\sqrt{2\sigma^2 M}$, which makes the average norm of $\|\mathbf{Hx}\| \approx \sqrt{2\sigma^2 KM}$ equal to that of $\|\mathbf{s}\| = \sqrt{2K}$ for QPSK signals.[2] While the NLS encoder is superior at high SNR, there is a broad range of low- to medium-SNRs where the simple quantized ZF precoder provides the same or slightly better performance, although the complexity of the NLS approach prevents us from establishing this in general for anything but systems with small M and K. Still, the observed performance bodes well for the simple quantized ZF precoder, since the low- to medium-SNR range is of particular interest for massive MIMO implementations.

For the next example, we take the special case of $K = 4$ and assume that each user is being sent a different symbol. In particular, the desired user symbols are chosen to be $\mathbf{s} = [1 + j, 1 - j, -1 + j, -1 - j]^T$. In Fig. 7, we plot the signals received by each of the four users (different symbol for each user) and for $M = 20, 100, 300$ (different color for each M), assuming no receiver noise (the only source of error here is the quantization at the transmitter). The simulations have been performed over 10^2 independent channel realizations. We clearly see that as the value of M increases, the average distance of the symbols from the decision boundary increases, which indicates an increased robustness with respect to additive noise with increasing γ.

V. Bussgang Adapted One-Bit ZF Precoder

In this section, we use insights gained from our analysis of the quantized ZF precoder to improve its performance. We have seen that the cross-covariance matrix of $\tilde{\mathbf{s}}$ and intended vector $\mathbf{s}$ (37) reduces asymptotically with M and K to a diagonal matrix with positive diagonal elements when using the ZF precoder. This is a desirable property since it implies that multiuser interference is eliminated, and the received signal constellation is not rotated. Our objective in the approach we present here is to enforce this diagonal structure of $\mathbf{H}\left\{\text{diag}(\mathbf{PP}^H)\right\}^{-1/2}\mathbf{P}$ for any value of M and K, by an improved choice of $\mathbf{P}$. We will refer to this method as "Bussgang-adapted" since the expression for $\mathbf{R}_{\tilde{\mathbf{s}}\mathbf{s}}$ we wish to diagonalize was derived using the Bussgang theorem.

A. Principle

For $\mathbf{H}\left\{\text{diag}(\mathbf{PP}^H)\right\}^{-1/2}\mathbf{P}$ to be diagonal, we must choose $\mathbf{P}$ so that

$$\left\{\text{diag}(\mathbf{PP}^H)\right\}^{-1/2}\mathbf{P} = \mathbf{H}^\dagger\mathbf{D}, \tag{48}$$

where $\mathbf{H}^\dagger = \mathbf{H}^H(\mathbf{HH}^H)^{-1}$ and $\mathbf{D}$ represents some real-valued positive diagonal matrix. For this to be true, we can see that $\mathbf{P}$ must satisfy,

$$\mathbf{P} = \mathbf{\Lambda}\mathbf{H}^\dagger\mathbf{D}, \tag{49}$$

with $\mathbf{\Lambda} = \left\{\text{diag}(\mathbf{PP}^H)\right\}^{1/2}$. Using (49) in (48), the following condition must then hold:

$$\mathbf{\Lambda}^{-1/2}\{\text{diag}(\mathbf{H}^\dagger\mathbf{D}^2(\mathbf{H}^\dagger)^H)\}^{-1/2}\mathbf{\Lambda}^{1/2}\mathbf{H}^\dagger = \mathbf{H}^\dagger,$$

which simplifies to

$$\text{diag}(\mathbf{H}^\dagger\mathbf{D}^2(\mathbf{H}^\dagger)^H) = \mathbf{I}_M\ . \tag{50}$$

Note that $\mathbf{\Lambda}$ has vanished from (50) indicating that it does not affect the signal of interest, and thus we take $\mathbf{\Lambda} = \mathbf{I}_M$ for simplicity.

[2]Note that better performance for NLS can be obtained by using a value of δ that is optimized to change at the symbol rate with each $\mathbf{x}$ and $\mathbf{s}$, as in [32]. However, allowing an infinite precision (unquantized) gain that varies at the symbol rate can be viewed as being at odds with the assumption of simple one-bit DACs.

Denoting $\mathbf{T} = \mathbf{H}^\dagger$, (50) becomes $\text{diag}(\mathbf{T}\mathbf{D}^2\mathbf{T}^H) = \mathbf{I}_M$. It can be written with respect to the diagonal entries of $\mathbf{D} = \text{Diag}(d_1, d_2, \ldots, d_K)$ as

$$\begin{bmatrix} |t_{11}|^2 \ldots |t_{1K}|^2 \\ |t_{21}|^2 \ldots |t_{2K}|^2 \\ \ldots \\ |t_{M1}|^2 \ldots |t_{MK}|^2 \end{bmatrix} \begin{bmatrix} d_1^2 \\ d_2^2 \\ \ldots \\ d_K^2 \end{bmatrix} = \mathbf{1}_M \ , \tag{51}$$

where $\mathbf{1}_M$ is an $M \times 1$ vector of ones. Now define the $M \times K$ matrix

$$\tilde{\mathbf{T}} = \begin{bmatrix} |t_{11}|^2 \ldots |t_{1K}|^2 \\ |t_{21}|^2 \ldots |t_{2K}|^2 \\ \ldots \\ |t_{M1}|^2 \ldots |t_{MK}|^2 \end{bmatrix}$$

so that

$$\begin{bmatrix} d_1^2 \\ d_2^2 \\ \ldots \\ d_K^2 \end{bmatrix} = (\tilde{\mathbf{T}}^H \tilde{\mathbf{T}})^{-1} \tilde{\mathbf{T}}^H \mathbf{1}_M. \tag{52}$$

Solving (52), we can immediately deduce $\mathbf{D}$.

For the precoder we obtain solving (48), the cross-covariance matrix between $\tilde{\mathbf{s}}$ and $\mathbf{s}$ is

$$\mathbf{R}_{\tilde{\mathbf{s}}\mathbf{s}} = \frac{\sqrt{\rho_0}\mathbf{H}\mathbf{F}\mathbf{P}}{\sqrt{M}} \mathbf{R}_{\mathbf{s}\mathbf{s}} = \sqrt{\frac{2\rho_0}{M\pi}} \sigma_s \mathbf{D} \ , \tag{53}$$

indicating that all multi-user interference has been canceled. The received signal can thus be described as

$$\mathbf{r} = \sqrt{\frac{2\rho_0}{M\pi}} \frac{1}{\sigma_s} \mathbf{D}\mathbf{s} + \sqrt{\frac{\rho_0}{M}} \mathbf{H}\mathbf{q} + \mathbf{n}. \tag{54}$$

The $SQINR$ for the k^{th} received symbol will be

$$SQINR_k = \frac{\frac{2\rho_0}{M\pi} d_k^2}{\frac{\rho_0}{M} [\mathbf{H}\mathbf{R}_{\mathbf{q}\mathbf{q}}\mathbf{H}^H]_{kk} + \sigma_n^2} \ , \tag{55}$$

and the SER for the k^{th} user for QPSK modulation and equally likely signaling is

$$P_e \simeq 2Q \left(\sqrt{\frac{\frac{2\rho_0}{M\pi} d_k^2}{\frac{\rho_0}{M} [\mathbf{H}\mathbf{R}_{\mathbf{q}\mathbf{q}}\mathbf{H}^H]_{kk} + \sigma_n^2}} \right) \tag{56}$$

B. Proposed Algorithm

Since we know $\mathbf{H}$, it is easy to check whether or not the Bussgang-adapted precoded vector induces fewer errors at the receiver than the ZF precoder. Therefore, we propose the algorithm shown in the table labeled "Algorithm 1" below. For a given vector $\mathbf{s}$ and channel matrix, $\mathbf{H}$, the initial ZF precoding matrix $\mathbf{P} = \mathbf{H}^\dagger$ is estimated. It is then checked whether the channel output of the precoded matrix, when subjected to quantization, gives an output that is equal to the vector $\mathbf{s}$. If it is, then the precoding matrix is taken to be $\mathbf{P} = \mathbf{H}^\dagger$ without further computation. If not, we compute the Bussgang-adapted precoder, $\tilde{\mathbf{P}} = \mathbf{H}^\dagger\mathbf{D}$, where $\mathbf{D}$ is calculated from (52), and we compare the number of symbol errors produced at the receivers by both precoders after quantization and multiplication by the channel matrix $\mathbf{H}$. If the Bussgang-adapted precoder $\tilde{\mathbf{P}}$ produces fewer errors than the quantized ZF precoder $\mathbf{P}$, then we reset $\mathbf{P}$ to be $\tilde{\mathbf{P}}$. Mathematically, we can write this condition as

$$\|\mathbf{1}_{\mathbb{Q}(\mathbf{H}\mathbb{Q}(\mathbf{P}\mathbf{s}))\neq\mathbf{s}}\|_1 > \|\mathbf{1}_{\mathbb{Q}(\mathbf{H}\mathbb{Q}(\tilde{\mathbf{P}}\mathbf{s}))\neq\mathbf{s}}\|_1 \ ,$$

where we define $\mathbf{1}_{\mathbf{a}\neq\mathbf{b}}$ to be a vector whose k-th element is 1 when $a_k \neq b_k$ and is 0 otherwise.

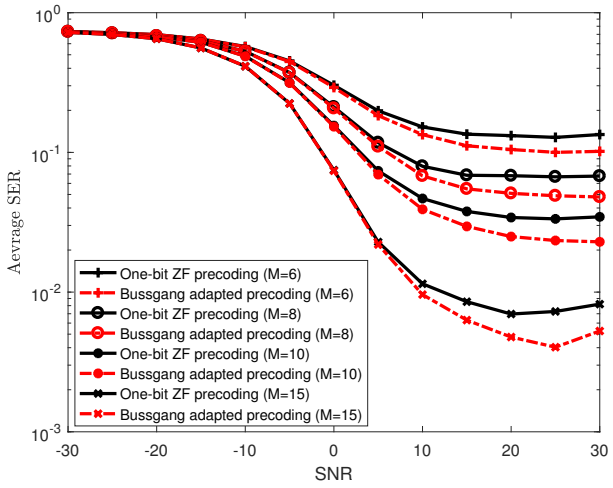

Fig. 8: Variation of SER for one-bit quantized Bussgang adapted and ZF precoding versus SNR for $K = 3$ users and varying values of M.

Algorithm 1: Bussgang adapted precoding algorithm

Input: $\mathbf{s}, \mathbf{H}$
Let $\mathbf{P} = \mathbf{H}^\dagger$
if $\mathbb{Q}(\mathbf{H}\mathbb{Q}(\mathbf{P}\mathbf{s})) \neq \mathbf{s}$ **then**
  Let $\tilde{\mathbf{P}} = \mathbf{H}^\dagger\mathbf{D}$, where $\mathbf{D}$ is found using eq. (52)
  if $\|\mathbf{1}_{\mathbb{Q}(\mathbf{H}\mathbb{Q}(\mathbf{P}\mathbf{s}))\neq\mathbf{s}}\|_1 > \|\mathbf{1}_{\mathbb{Q}(\mathbf{H}\mathbb{Q}(\tilde{\mathbf{P}}\mathbf{s}))\neq\mathbf{s}}\|_1$ **then**
    $\mathbf{P} = \tilde{\mathbf{P}}$
  end
end
Output: Resulting precoded vector is $\mathbf{x}_P = \mathbf{P}\mathbf{s}$

In Figs. 8 and 9, we compare the SER performance of the Bussgang adapted precoder with quantized ZF precoding as a function of the SNR for a case with $K = 3$ and $K = 10$ users respectively. We see that the new algorithm achieves a lower error floor in all cases compared with ZF precoding. Also, we can observe that for a fixed number of users, K, the improvement with respect to the ZF precoding increases slightly with an increasing number of BS antennas, M.

VI. Conclusion

We have studied the use of quantized linear precoding for the massive MIMO downlink scenario with one-bit DACs. We derived closed form expressions for the SQINR and SER for any linear precoder using the Bussgang decomposition. We provided an analysis to show that asymptotically in the number of antennas M and the number of users K, the algorithm yields signals at the user terminals that are scaled versions of

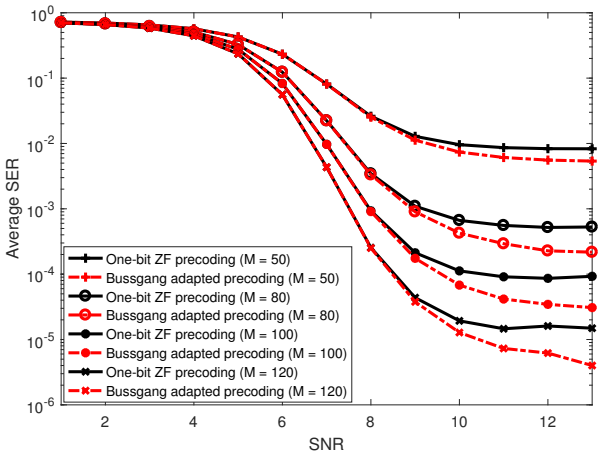

Fig. 9: Variation of SER for one-bit quantized Bussgang adapted and ZF precoding versus SNR for $K = 10$ users and varying values of M.

the desired symbols, with the scaling increasing proportionally to $\sqrt{M/K}$. Simulations suggest that, at least for small system dimensions, the algorithm can achieve slightly better performance than the NLS encoder for low to moderate SNRs, at a fraction of the computational cost. We also presented a modified version of the quantized ZF precoder that yields lower SERs at high SNR.

REFERENCES

[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, pp. 186–195, February 2014.

[2] 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 Processing Magazine*, vol. 30, pp. 40–60, Jan 2013.

[3] 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, pp. 742–758, Oct 2014.

[4] T. K. Y. Lo, "Maximum ratio transmission," in *IEEE International Conference on Communications*, vol. 2, pp. 1310–1314, 1999.

[5] A. Wiesel, Y. C. Eldar, and S. Shamai, "Zero-forcing precoding and generalized inverses," *IEEE Transactions on Signal Processing*, vol. 56, pp. 4409–4418, Sept 2008.

[6] H. Yang and T. L. Marzetta, "Performance of conjugate and zero-forcing beamforming in large-scale antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, pp. 172–179, February 2013.

[7] D. Ha, K. Lee, and J. Kang, "Energy efficiency analysis with circuit power consumption in massive MIMO systems," in *2013 IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pp. 938–942, Sept 2013.

[8] S. Han, C. l. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Communications Magazine*, vol. 53, pp. 186–194, January 2015.

[9] W. Roh, J. Y. Seol, J. Park, B. Lee, J. Lee, Y. Kim, J. Cho, K. Cheun, and F. Aryanfar, "Millimeter-wave beamforming as an enabling technology for 5G cellular communications: Theoretical feasibility and prototype results," *IEEE Communications Magazine*, vol. 52, pp. 106–113, February 2014.

[10] A. Alkhateeb, O. E. Ayach, G. Leus, and R. W. Heath, "Hybrid precoding for millimeter wave cellular systems with partial channel knowledge," in *Information Theory and Applications Workshop (ITA), 2013*, Feb 2013.

[11] R. H. Walden, "Analog-to-digital converter survey and analysis," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 539–550, Apr 1999.

[12] C. Svensson, S. Andersson, and P. Bogner, "On the power consumption of analog to digital converters," in *2006 NORCHIP*, pp. 49–52, Nov 2006.

[13] J. Singh, S. Ponnuru, and U. Madhow, "Multi-gigabit communication: The ADC bottleneck," in *2009 IEEE International Conference on Ultra-Wideband*, pp. 22–27, Sept 2009.

[14] A. Mezghani and J. A. Nossek, "Analysis of Rayleigh-fading channels with 1-bit quantized output," in *2008 IEEE International Symposium on Information Theory*, pp. 260–264, July 2008.

[15] A. Mezghani, F. Antreich, and J. A. Nossek, "Multiple parameter estimation with quantized channel output," in *Smart Antennas (WSA), 2010 International ITG Workshop on*, pp. 143–150, Feb 2010.

[16] F. Wendler, M. Stein, A. Mezghani, and J. A. Nossek, "Quantization-loss reduction for 1-bit BOC positioning," in *Proc. ION Int'l Technical Mtg.*, pp. 509–518, Jan 2013.

[17] N. Liang and W. Zhang, "Mixed-ADC Massive MIMO," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 983–997, 2016.

[18] S. Jacobsson, G. Durisi, M. Coldrey, U. Gustavsson, and C. Studer, "One-bit massive MIMO: Channel estimation and high-order modulations," in *2015 IEEE International Conference on Communication Workshop (ICCW)*, pp. 1304–1309, June 2015.

[19] Y. Li, C. Tao, L. Liu, G. Seco-Granados, and A. Swindlehurst, "Channel estimation and uplink achievable rates in one-bit massive MIMO systems," in *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM) (SAM 2016)*, (Rio de Janeiro, Brazil), July 2016.

[20] J. Mo and R. W. Heath, "Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information," *IEEE Transactions on Signal Processing*, vol. 63, pp. 5498–5512, Oct 2015.

[21] A. Mezghani and J. A. Nossek, "On ultra-wideband MIMO systems with 1-bit quantized outputs: Performance analysis and input optimization," in *2007 IEEE International Symposium on Information Theory*, pp. 1286–1289, June 2007.

[22] A. Mezghani and J. A. Nossek, "Capacity lower bound of MIMO channels with output quantization and correlated noise," in *IEEE International Symposium on Information Theory Proceedings (ISIT)*, 2012.

[23] L. Zhao, K. Zheng, H. Long, and H. Zhao, "Performance analysis for downlink massive MIMO system with ZF precoding," *Transactions on Emerging Telecommunications Technologies*, vol. 25, no. 12, pp. 1219–1230, 2014.

[24] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?," *IEEE Journal on Selected Areas in Communications*, vol. 31, pp. 160–171, February 2013.

[25] A. Mezghani, R. Ghiat, and J. A. Nossek, "Transmit processing with low resolution D/A-converters," in *Electronics, Circuits, and Systems, 2009. ICECS 2009. 16th IEEE International Conference on*, pp. 683–686, Dec 2009.

[26] H. Jedda, J. A. Nossek, and A. Mezghani, "Minimum BER precoding in 1-bit massive MIMO systems," in *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM) (SAM 2016)*, (Rio de Janeiro, Brazil), July 2016.

[27] O. B. Usman, H. Jedda, A. Mezghani, and J. A. Nossek, "MMSE precoder for massive MIMO using 1-bit quantization," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 3381–3385, March 2016.

[28] A. K. Saxena, I. Fijalkow, and L. Swindlehurst, "On One-bit quantized ZF precoding for the multiuser massive MIMO downlink," in *2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM) (SAM 2016)*, (Rio de Janeiro, Brazil), July 2016.

[29] J. Bussgang and M. I. of Technology. Research Laboratory of Electronics, *Crosscorrelation functions of amplitude-distorted Gaussian signals*. Technical report (Massachusetts Institute of Technology. Research Laboratory of Electronics), Research Laboratory of Electronics, Massachusetts Institute of Technology, 1952.

[30] C. B. Peel, B. M. Hochwald, and A. L. Swindlehurst, "A vector-perturbation technique for near-capacity multiantenna multiuser communication - Part I: Channel inversion and regularization," *IEEE Transactions on Communications*, vol. 53, pp. 195–202, Jan 2005.

[31] 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*, pp. 763–767, Nov 2016.

[32] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized precoding for massive MU-MIMO," *arXiv preprint arXiv:1610.07564*, 2016.

[33] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Mathematics of computation*, vol. 44, no. 170, pp. 463–471, 1985.
[34] T. Cui and C. Tellambura, "An efficient generalized sphere decoder for rank-deficient MIMO systems," *IEEE Communications Letters*, vol. 9, pp. 423–425, May 2005.
[35] W. Zhang, "A general framework for transmission with transceiver distortion and some applications," *IEEE Transactions on Communications*, vol. 60, no. 2, pp. 384–399, 2012.
[36] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*, pp. 393–397. New York: McGraw-Hil, 2002.
[37] M. T. Ivrlac and J. A. Nossek, "On physical limits of massive miso systems," in *Smart Antennas (WSA 2016); Proceedings of the 20th International ITG Workshop on*, VDE, 2016.
[38] D. A. Gore, R. W. Heath, and A. J. Paulraj, "Transmit selection in spatial multiplexing systems," *IEEE Communications Letters*, vol. 6, pp. 491–493, Nov 2002.
[39] R. J. Muirhead, *Aspects of multivariate statistical theory*, vol. 197. John Wiley & Sons, 2009.
[40] B. Hochwald and S. Vishwanath, "Space-time multiple access: Linear growth in the sum rate," in *Proc. 40th Annual Allerton Conf. Communications, Control and Computing*, 2002.
[41] R. Couillet and M. Debbah, *Random matrix methods for wireless communications*. Cambridge University Press, 2011.