Analysis of one-bit quantized ZF precoding for the multiuser massive MIMO downlink — Source link

Amodh Kant Saxena, Inbar Fijalkow, Amine Mezghani, A. Lee Swindlehurst

Institutions: University of California, Irvine, Cergy-Pontoise University

Related papers:

- On one-bit quantized ZF precoding for the multiuser massive MIMO downlink

- Quantized Precoding for Massive MU-MIMO

- Linear Precoding of Data and Artificial Noise in Secure Massive MIMO Systems

- Analog Versus Hybrid Precoding for Multiuser Massive MIMO With Quantized CSI Feedback

- Low-Complexity Hybrid Precoding in Massive Multiuser MIMO Systems

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

A Saxena, Inbar Fijalkow, A Mezghani, A Swindlehurst

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

Amodh Kant Saxena[1], Inbar Fijalkow[2], Amine Mezghani[1], A. Lee Swindlehurst[1]

[1] Center for Pervasive Communications and Computing, University of California Irvine, Irvine, CA 92697, USA,
{aksaxena, amine.mezghani, swindle}@uci.edu

[2] ETIS, UMR 8051 / ENSEA, Université Cergy-Pontoise, CNRS, F-95000 Cergy, France, inbar.fijalkow@ensea.fr

Abstract - We present a mathematical analysis of linear precoders for downlink multiuser massive MIMO. In the scenario we consider, one-bit digital-to-analog converters are employed at the basestation antennas for mitigating power consumption and hardware complexity. Using the Bussgang theorem, a probability-of-error analysis for a general precoder is presented. The special case of zero-forcing (ZF) precoders is considered and simplified error expressions are derived for asymptotic scenarios. Our analysis illustrates that the probability of error for the quantized ZF precoder depends primarily on the ratio of the number of antennas to the number of users. Our simulations also show that the quantized ZF precoder outperforms the high complexity ML encoder for low- to medium-SNRs

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]. 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 [2], [3].

However, the resulting power consumption and hardware costs of massive MIMO increase with the number of antennas [4]. In order to mitigate these costs, we focus on an approach that has gained attention recently, namely the use of low-resolution DACs for each antenna; in particular we will explore 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 [5]–[7].

In this paper, we study the impact of one-bit DACs on linear precoding for the massive MIMO downlink. To focus on the performance degradation due to quantization, we assume that the BS has perfect channel state information. Using the Bussgang theorem [8] 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, and our simulations show that it can outperform the much more complicated maximum likelihood encoder for low-to-moderate signal to noise ratios, where massive MIMO systems are presumed to operate.

The paper is organized as follows. Section II introduces the one-bit downlink model, and describes both direct non-linear maximum likelihood 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.

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 kth row and l^{th} column. Lower boldface letters, $\mathbf{a}$, indicate a column vector, with a_k denoting the kth 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 which are zero-mean and independent with power σ_s^2, 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

$$\mathbf{r} = \sqrt{\rho}\mathbf{H}\mathbf{x} + \mathbf{n} \, , \qquad (1)$$

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 kth 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 elements of the channel matrix $\mathbf{H}$ are complex Gaussian random variables, whose real and imaginary parts are both iid Gaussian random variables with zero-mean and variance σ^2. As mentioned above, we will assume that the channel $\mathbf{H}$ is known at the BS.

B. ML Precoding

Assuming Gaussian noise, one might choose to implement a maximum likelihood (ML) encoder, which attempts to solve [9]

$$\mathbf{x} = \arg \min_{\mathbf{v} \in \mathcal{S}^M} \|\mathbf{s} - \mathbf{H}\mathbf{v}\|^2 \, , \qquad (2)$$

where $\mathcal{S} = \{1 + j, 1 - j, -1 + j, -1 - j\}$ is the set of QPSK constellation points. However, in general, (2) 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.

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 ML encoder over-constrains the problem by attempting to force $\mathbf{H}\mathbf{x}$ 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 data $\mathbf{H}\mathbf{x}$ can in principle be far away from $\mathbf{s}$ and still be decoded correctly; in fact, often the farther $\mathbf{H}\mathbf{x}$ 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 ML 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 ML 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 $\mathbf{x} = \mathbb{Q}(\mathbf{P}\mathbf{s})$, where the one-bit quantization operation is defined as

$$\mathbb{Q}(\mathbf{a}) = \text{sign}(\Re(\mathbf{a})) + j.\text{sign}(\Im(\mathbf{a})), \qquad (3)$$

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

$$\mathbf{r} = \sqrt{\rho}\mathbf{H}\mathbb{Q}(\mathbf{P}\mathbf{s}) + \mathbf{n} \, . \qquad (4)$$

In what follows, we will assume that $\rho = \frac{\rho_0}{M}$, where, ρ_0 is defined to be the transmit SNR.
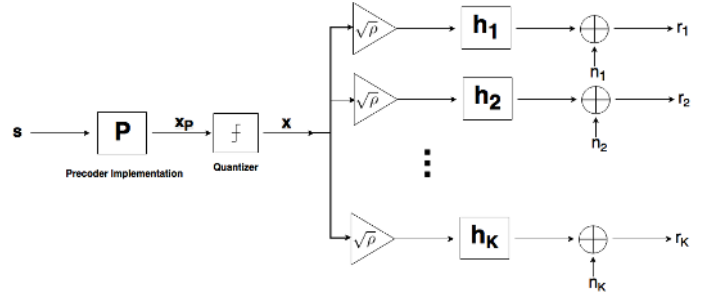

Fig. 1: System Model

III. BUSSGANG ANALYSIS OF ONE-BIT QUANTIZED PRECODING

Let $\mathbf{x}_P = \mathbf{P}\mathbf{s}$ 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.

A. Bussgang Decomposition

The one-bit quantization operation on the precoded vector $\mathbf{x}_P$ is modeled here using the Bussgang theorem [8]. 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 $\mathbf{s}$, hence the Gaussian assumption is fulfilled for large enough K. We thus apply the Bussgang theorem to write

$$\mathbf{x} := \mathbb{Q}(\mathbf{x}_P) = \mathbf{F}\mathbf{x}_P + \mathbf{q} \, , \qquad (5)$$

where $\mathbf{F}$ is chosen to satisfy $\mathbf{R}_{\mathbf{x}_P \mathbf{q}} = E(\mathbf{x}_P \mathbf{q}^H) = \mathbf{0}$. 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

$$\mathbf{R}_{\mathbf{x}\mathbf{x}_P} = E(\mathbf{x}\mathbf{x}_P^H) = E(\{\mathbf{F}\mathbf{x}_P + \mathbf{q}\}\mathbf{x}_P^H)$$
$$= \mathbf{F}\mathbf{R}_{\mathbf{x}_P\mathbf{x}_P} \, , \qquad (6)$$

where

$$\mathbf{R}_{\mathbf{x}_P\mathbf{x}_P} = \sigma_s^2 \mathbf{P}\mathbf{P}^H. \qquad (7)$$

Note that the $M \times M$ matrix in (7) is rank deficient, and thus $\mathbf{F}$ cannot be solved for directly as in [10]. We will see shortly that a unique expression for $\mathbf{F}$ is unnecessary, and that (6) is sufficient. Under the mutual Gaussian assumption between components of $\mathbf{x}_P$, and since $\mathbf{P}$ is column rank, we have (for details, see [11]),

$$\mathbf{F}\mathbf{P} = \frac{1}{\sigma_s}\sqrt{\frac{2}{\pi}} \left\{ \text{diag}(\mathbf{P}\mathbf{P}^H) \right\}^{-\frac{1}{2}} \mathbf{P}. \qquad (8)$$

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 [11], [12],

$$\mathbf{R}_{\mathbf{x}\mathbf{x}} = \frac{2}{\pi}\arcsin[\{\text{diag}(\mathbf{R}_{\mathbf{x}_P\mathbf{x}_P})\}^{-\frac{1}{2}}$$
$$\mathbf{R}_{\mathbf{x}_P\mathbf{x}_P}\{\text{diag}(\mathbf{R}_{\mathbf{x}_P\mathbf{x}_P})\}^{-\frac{1}{2}}]. \qquad (9)$$

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

$$\mathbf{R_{qq}} = \frac{2}{\pi}[\arcsin\{\{\text{diag}(\mathbf{PP}^H)\}^{-\frac{1}{2}}\mathbf{PP}^H\{\text{diag}(\mathbf{PP}^H)\}^{-\frac{1}{2}}\} \\ - \{\text{diag}(\mathbf{PP}^H)\}^{-\frac{1}{2}}\mathbf{PP}^H\{\text{diag}(\mathbf{PP}^H)\}^{-\frac{1}{2}}]. \tag{10}$$

B. Impact on the Signal of Interest

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

$$\tilde{\mathbf{s}} = \sqrt{\rho}\mathbf{Hx} = \frac{\sqrt{\rho_0}}{\sqrt{M}}\mathbf{Hx}. \tag{11}$$

The cross-correlation between the received $\tilde{\mathbf{s}}$ and desired $\mathbf{s}$ is (see [11]),

$$\mathbf{R_{\tilde{s}s}} = \frac{\sqrt{\rho_0}\sigma_s^2}{\sqrt{M}}\mathbf{HFP} \\ = \sqrt{\frac{2}{\pi}}\frac{\sqrt{\rho_0}\sigma_s}{\sqrt{M}}\mathbf{H}\left\{\text{diag}(\mathbf{PP}^H)\right\}^{-\frac{1}{2}}\mathbf{P} , \tag{12}$$

Equation (12) shows that, for any full rank precoder, the impact of the one-bit quantization on the signal of interest is the diagonal matrix $\left\{\text{diag}(\mathbf{PP}^H)\right\}^{-\frac{1}{2}}$ 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

$$\mathbf{r} = \frac{\sqrt{\rho_0}}{\sqrt{M}}\mathbf{H}(\mathbf{FPs} + \mathbf{q}) + \mathbf{n} \\ = \frac{\sqrt{\rho_0}}{\sqrt{M}}\mathbf{Gs} + \frac{\sqrt{\rho_0}}{\sqrt{M}}\mathbf{Hq} + \mathbf{n} . \tag{13}$$

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

$$\mathbf{R_{dd}} = \mathbf{HR_{qq}H}^H. \tag{14}$$

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

$$SQINR_k = \frac{\rho_0\frac{|g_{kk}|^2\sigma_s^2}{M}}{\rho_0\sum_{l=1,l\neq k}^K\frac{|g_{kl}|^2\sigma_s^2}{M} + \rho_0\frac{[\mathbf{R_{dd}}]_{kk}}{M} + \sigma_n^2} , \tag{15}$$

where $\rho_0\sum_{l=1,l\neq k}^K\frac{|g_{kl}|^2\sigma_s^2}{M}$ accounts for multi-user interference and $\rho_0\frac{[\mathbf{R_{dd}}]_{kk}}{M}$ 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

$$P_e = Pr(\mathbb{Q}(r_k) \neq s_k) \simeq 2Q(\sqrt{SQINR_k}) \\ = 2Q\left(\sqrt{\frac{\rho_0\frac{|g_{kk}|^2\sigma_s^2}{M}}{\rho_0\sum_{l=1,l\neq k}^K\frac{|g_{kl}|^2\sigma_s^2}{M} + \rho_0\frac{[\mathbf{R_{dd}}]_{kk}}{M} + \sigma_n^2}}\right). \tag{16}$$

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

$$\mathbf{P} = \mathbf{H}^H(\mathbf{HH}^H)^{-1} . \tag{17}$$

In addition, we will further simplify the resulting expressions by adopting a massive MIMO assumption and letting both M and K be large [13].

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 $\frac{M}{K}$ that is greater than 1. In what follows, we recall and extend some results on the asymptotic behaviour of the matrix $(\mathbf{HH}^H)^{-1}$ needed for analyzing the ZF precoder. As mentioned earlier, we assume that the elements of $\mathbf{H}$ are $i.i.d.$ circularly symmetric Gaussian random variables with $\Re(h_{ij}) \sim \mathcal{N}(0,\sigma^2)$ independent of $\Im(h_{ij}) \sim \mathcal{N}(0,\sigma^2)$, $\forall i = 1,2,\ldots,K$ and $j = 1,2,\ldots,M$.

It is shown in [14] that $\mathbf{Z} = 2\mathbf{HH}^H$ is a complex Wishart matrix (see [15]) with distribution

$$\mathbf{W}_K(M,\sigma^2\mathbf{I}_K;\mathbf{Z}) = \frac{|\mathbf{Z}|^{\frac{M-K-1}{2}}\exp\left(-\frac{1}{2}\text{trace}(\sigma^{-2}\mathbf{Z})\right)}{2^{\frac{MK}{2}}\Gamma_K(\frac{M}{2})|\sigma^2\mathbf{I}_K|^{\frac{M}{2}}} \tag{18}$$

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 what follows, we use the properties of the Wishart distribution to analyze the asymptotic behavior of one-bit quantized ZF precoder (for details, see [11]).

B. Asymptotic Received Downlink Signal

Using the results of the previous section (for details, see [11]), we have

$$\mathbf{FP} = \frac{1}{\sigma_s}\sqrt{\frac{2}{\pi}}\left\{\text{diag}(\mathbf{PP}^H)\right\}^{-\frac{1}{2}}\mathbf{P} \\ \xrightarrow[M\to\infty]{}\sqrt{\frac{2}{\pi}}\frac{1}{2\sigma_s(M-K)\sigma^2}\left\{\text{diag}(\mathbf{PP}^H)\right\}^{-\frac{1}{2}}\mathbf{H}^H, \tag{19}$$

Also, for $\mathbf{G}$, we have (using (19)),

$$\mathbf{G} = \mathbf{HFP} \\ \xrightarrow[M\to\infty]{}\sqrt{\frac{2}{\pi}}\frac{1}{2\sigma_s(M-K)\sigma^2}\mathbf{H}\left\{\text{diag}(\mathbf{PP}^H)\right\}^{-\frac{1}{2}}\mathbf{H}^H. \tag{20}$$

Since the columns of $\mathbf{H}$ become quasi-orthogonal as K becomes large, the $M \times M$ rank K matrix, $\mathbf{PP}^H$, asymptotically becomes a diagonal matrix with K non-zero elements that can be approximated as:

$$[\mathbf{PP}^H]_{kk} \xrightarrow[K\to\infty]{} \frac{K}{2(M-K)^2\sigma^2}. \tag{21}$$

Moreover, the non-zero diagonal values correspond to eigenvectors that lie in the size K subspace spanned by $\mathbf{H}$. Thus, using (21) and (20) (see [11]),

$$\mathbf{G} \xrightarrow[K\to\infty,M\to\infty]{} \frac{2(M-K)\sigma}{\sigma_s\sqrt{\pi K}}\mathbf{I}_K. \qquad (22)$$

The cross-covariance matrix of $\tilde{\mathbf{s}}$ and $\mathbf{s}$ becomes

$$\mathbf{R}_{\tilde{\mathbf{s}}\mathbf{s}} \simeq \sigma_s^2 \frac{\sqrt{\rho_0}\mathbf{H}\mathbf{H}^H}{\sqrt{M\pi K}\sigma}$$

$$\qquad (23)$$

$$\xrightarrow[K\to\infty,M\to\infty]{} \frac{2\sigma_s\sigma\sqrt{\rho_0}}{\sqrt{\pi}}\left(\frac{M}{K}-1\right)\sqrt{\frac{K}{M}}\mathbf{I}_K$$

$$=\beta\mathbf{I}_K \,,$$

We observe that for large M and K, $\tilde{s}_k$ and s_l are uncorrelated for $k \neq l$, and hence the multi-user interference disappears. In addition, we see that the signal of interest is received with a gain of $\frac{\sqrt{2}}{\sigma_s}\beta$ which grows as $\sqrt{\frac{M}{K}}$. Hence, the larger the value of 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.

C. Asymptotic SQINR and Probability of Error

As described in Section IV-B, the ZF precoder $\mathbf{P}\mathbf{P}^H$ asymptotically becomes a diagonal matrix with K non-zero diagonal elements which correspond to eigenvectors lying on the subspace spanned by $\mathbf{H}$. Using (21) and the Law of Large Numbers [16], we have

$$\mathbf{R}_{\mathbf{dd}} \xrightarrow[K\to\infty,M\to\infty]{} 2\left(1-\frac{2}{\pi}\right)(M-K)\sigma^2\mathbf{I}_K, \qquad (24)$$

The SQINR then becomes

$$SQINR_k \xrightarrow[K\to\infty,M\to\infty]{} \frac{\rho_0\frac{4(M-K)^2\sigma^2}{M\pi K}}{\rho_0\frac{2}{M}\left(1-\frac{2}{\pi}\right)(M-K)\sigma^2+\sigma_n^2}, \qquad (25)$$

As before, assuming equally likely Gray-mapped QPSK symbols, we have the probability of error for the kth user as

$$P_e = Pr(\mathbb{Q}(r_k) \neq s_k) \simeq 2Q(\sqrt{SQINR_k})$$

$$\xrightarrow[K\to\infty,M\to\infty]{} 2Q\left(\sqrt{\frac{\rho_0\frac{4(M-K)^2\sigma^2}{M\pi K}}{\rho_0\frac{2}{M}\left(1-\frac{2}{\pi}\right)(M-K)\sigma^2+\sigma_n^2}}\right).$$

$$\qquad (26)$$

For a high SNR, $SQINR_k$ is given by

$$SQINR_k \simeq \frac{\frac{2}{\pi}}{1-\frac{2}{\pi}}\left(\frac{M}{K}-1\right), \qquad (27)$$

so that the error floor becomes

$$P_e \simeq 2Q\left(\sqrt{\frac{\frac{2}{\pi}}{1-\frac{2}{\pi}}\left(\frac{M}{K}-1\right)}\right). \qquad (28)$$

In all cases above (25-28), we note the critical dependence of the SQINR and probability of error on the quantity M/K; in particular for a high SNR (see (27)), the SQINR increases approximately linearly with M/K. In Fig. 2, we have plotted
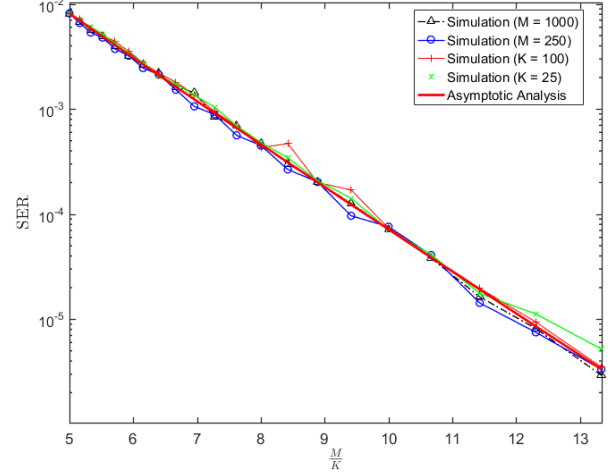

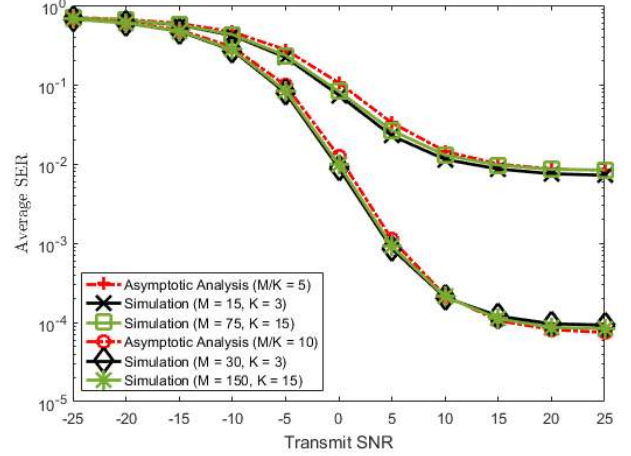Fig. 2: Variation of SER with the ratio $\frac{M}{K}$ in the noiseless case.

Fig. 3: Variation of SER with transmit SNR, for varying number of users, K and BS antennas, M.

the SER for the case of no additive noise as a function of M/K 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 M/K on performance. Massive MIMO systems are typically envisioned to operate with loading factors on the order of $M/K \simeq 10$, and we see that for this case the SER is below 10^{-4}, which bodes well for the use of the quantized ZF precoder in practical scenarios.

D. Simulations

In Fig. 3, we have plotted both the predicted and simulated SER for one-bit quantized ZF precoding at the BS for varied number of users as a function of the transmit SNR, ρ_0, as described in Section II-C. The variance of the real and imaginary parts of the channel coefficients are independently assumed to be $\sigma = 1$. We note the excellent match between the simulations and analytical approximation in (26), which
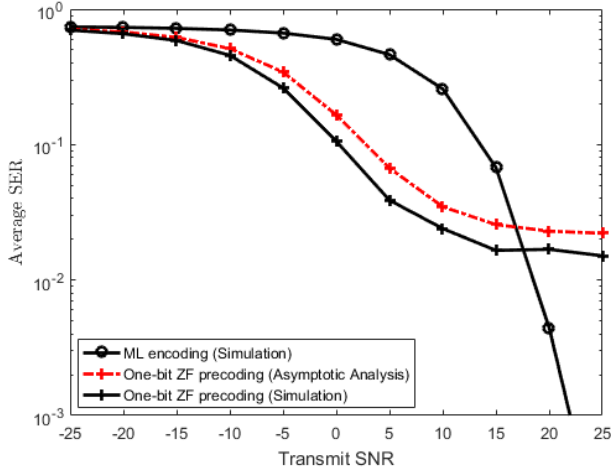
Fig. 4: Variation of SER versus transmit SNR for $K = 2$ users and $M = 8$ BS antennas for one-bit ZF precoding and ML encoding

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 $K = 3$ users. We observe that our asymptotic analysis is accurate even for this non-asymptotic value. The results of simulation are similar for $K = 3$ and $K = 15$ for both $M/K = 5, 10$. This result reinforces the observation that performance is governed by the ratio M/K, independent of their specific values.

In Fig. 4 we compare the ML encoding approach (2) with the quantized ZF precoder. Due to the complexity of the ML encoder, we perform the simulation for the relatively small values $M = 8$ and $K = 2$. Again, $\sigma = 1$ is assumed for the channel. While the ML encoder is superior at high SNR, there is a broad range of low- to medium-SNRs where the simple quantized ZF precoder provides significantly better performance. The low- to medium-SNR range is of particular interest for massive MIMO implementations, and thus the quantized ZF precoder is an attractive simple approach for such scenarios.

V. CONCLUSION

We have studied the use of quantized linear precoding for the massive MIMO downlink 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 quantized ZF precoder yields signals at the user terminals that are scaled versions of the desired symbols, with the scaling increasing proportionally to $\sqrt{M/K}$. We also derived the asymptotic error performance of the quantized ZF precoder and showed that it primarily depends on the ratio M/K. Simulations show that the algorithm outperforms the ML encoder for low to moderate SNRs for the scenarios of interest.

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] T. K. Y. Lo, "Maximum ratio transmission," in *Communications, 1999. ICC '99. 1999 IEEE International Conference on*, vol. 2, pp. 1310–1314 vol.2, 1999.

[3] 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.

[4] 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.

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

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

[7] 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.

[8] 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.

[9] 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.

[10] F. Wendler, M. Stein, A. Mezghani, and J. A. Nossek, "Quantization-loss reduction for 1-bit BOC positioning," in *Proceedings of the ION International Technical Meeting*, Jan 2013.

[11] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive mimo downlink," *arXiv preprint arXiv:1610.06659*, 2016.

[12] I. Miller, "Probability, random variables, and stochastic processes," *Technometrics*, vol. 8, no. 2, pp. 378–380, 1966.

[13] 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.

[14] 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.

[15] R. J. Muirhead, *Aspects of multivariate statistical theory*, vol. 197. John Wiley & Sons, 2009.

[16] J. A. Gubner, *Probability and Random Processes for Electrical and Computer Engineers*. New York, NY, USA: Cambridge University Press, 2006.