This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

Linear CE and 1-bit quantized precoding with optimized dithering

Amodh Kant Saxena, Amine Mezghani, *Member, IEEE*, Robert W Heath Jr., *Fellow, IEEE*

High power amplifiers (HPA), used at transmission, add nonlinear impairments to the output signals. Through Constant envelope (CE) transmission, distortion in the signal can be avoided without wasting power on PA linearization. A more restricted form of CE transmission, 1-bit quantized transmission, further simplifies the RF chain and reduces the DAC power consumption. In this paper, for CE transmission and 1-bit quantized transmission at the BS antennas, we analyze downlink transmission for low complexity linear precoding. We observe that for small numbers of users in the downlink, correlation among the quantization error components across BS antennas is high, deteriorating the performance rapidly as number of users become smaller. To improve performance for smaller numbers of downlink users, we propose the addition of correlated Gaussian dither to the precoded signal before quantization and subsequent transmission. We observe that the receive SQINR peaks for finite non-trivial dither power. For given value of transmit power, number of BS antennas and number of users, SQINR is maximized analytically by the transmitter, to find the optimum dither power, using the Bussgang decomposition. We observe that with the implementation of optimized dithering, the error floor in the coded BER at high transmit power, for CE and 1-bit quantized transmissions, is pushed down significantly. We also observe that optimum dither power increases monotonically with transmit power, with rate of increase decreasing with increasing transmit power. Further, the optimum dither power strictly increases with number of BS antennas.

***Index Terms*—1-bit quantized transmission, massive MIMO, downlink precoding, constant envelope transmission, 1-bit DACs, Bussgang decomposition.**

I. Introduction

Massive MIMO is a technology for increasing receive SNR, supporting multiple users on the spatial channel, and improving rates in cellular networks [3]–[5]. Achieving the best performance, however, requires implementation of higher order modulations [6], which results in a signal with high peak to average power ratio. This introduces in-band distortion, out-of-band distortion and hardware impairments in the transmitted signal due to usage of high power amplifier with nonlinear characteristics [7]. Power back-off can be used to linearize the behavior of the HPA, but it reduces energy efficiency. Digital predistortion (DPD) of the input signal can help linearize the behavior of PA in the saturation region [8], [9]. DPD systems, however, cannot be directly applied to the 5G systems, as using dedicated DPD circuitry at each element of a massive MIMO transmitter would result in an infeasible amount of power consumption as number of antennas becomes large. For 5G communication systems operating at bandwidths of 100s of MHz, the DPD system would also need ADCs with huge sampling rate and power consumption, at each antenna of the massive MIMO antenna array [10], increasing total power consumption significantly.

CE signals have low PAPR and thus are not affected by signal distortion when operating in the saturation region of the HPA, removing the need for power back-off and predistortion. In this paper, therefore, we focus on eliminating the need for linearization by using CE transmission at the BS antennas, and developing a downlink precoding algorithm assuming BS antennas with CE constraint. We also focus on a more restricted form of transmission, a 1-bit DAC implementation, at the BS, which further simplifies the RF frontend architecture, and reduces power consumption at the DAC for mmWave frequencies [11].

Some part of this work has been published as part of proceedings of IEEE SPAWC, held in July 2019 [1], and Asilomar CSSC, held in November 2019 [2].

A. K. Saxena, and R. W. Heath Jr are with Wireless Networking and Communications Group, UT Austin, Austin, TX-78712, USA (email:{amodh, rheath}@utexas.edu).

A. Mezghani is with Department of ECE, University of Manitoba, Winnipeg, MB, R3T 5V6, Canada (email: amine.mezghani@utexas.edu).

The work was supported in part by AT&T Labs, Austin, TX, USA. This material is based upon work supported in part by the National Science Foundation under Grant No. NSF-CCF-1514275, ECCS-1711702, and CNS-1731658.

A. Prior Work

Downlink transmission with CE precoding has been investigated extensively in prior work. In [12], a precoding algorithm based on gradient descent was proposed to minimize error norm of the received signal. A nonlinear least square (NLS) based algorithm has been proposed to precode the CE signal for transmission in [13], which has further been improved upon by researchers in [14], implementing cross-entropy optimization (CEO) instead of gradient descent for the NLS algorithm. Cross-entropy optimization is used to maximize the minimum value of Contructive interference metric to perform CE precoding in [15]. Minimization of symbol error rate at receiver, using projected gradient descent algorithm has been proposed for CE precoding in [16] and [17]. Precoding algorithms for CE transmitted signals with quantized phases have been discussed in [18]–[22]. A special case of 1-bit quantized precoding has also been investigated extensively. In [29], a precoder to maximize the distance of the constructive region from the decision boundary was designed using relaxed non-convex optimization. In [30], the work of [29] was improved upon through a partial-branch and bound algorithm. Many other algorithms using relaxed non-convex optimizations have also been proposed, [17], [23]–[28], [31], [32]. These precoding algorithms, however, are nonlinear iterative algorithms, with polynomial complexity with respect to number of BS antennas, making them infeasible for massive MIMO scenarios which may have an order of magnitude more BS antennas relative to number of users. Linear precoding has a complexity which is linear with respect to the number of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

BS antennas making it highly suitable for downlink massive MIMO transmission. In this paper, therefore, we focus on linear precoding solutions for the downlink massive MIMO transmission.

While papers like [11], [24], [33], [34] have investigated CE and 1-bit quantized downlink linear precoding using the Bussgang decomposition, they have used it only up to the first-order approximation, which implicitly assumes quantization error elements to be uncorrelated across the BS antennas. Correlation among quantization error elements is negligible for large number of users. For small number of users and high number of BS antennas, however, it has substantial effects on performance [1]. According to [1], which used Bussgang decomposition up to the third-order of approximation for analyzing 1-bit quantized transmission at the BS, downlink receive performance deteriorates rapidly with reduction in number of users, in those scenarios. There is a need, therefore, to address this shortcoming arising in downlink linear transmission for CE and 1-bit quantized transmission system due to small number of users.

B. Contribution

In this paper, we use the Bussgang decomposition up to the third-order approximation to investigate the downlink performance of CE and 1-bit quantized linear transmission for MU-massive MIMO. We use a Taylor series expansion of the input-output correlation's relationship provided by the Price's theorem [35], up to third-order harmonics, thereby taking into account the correlation between quantization noise elements across the BS antennas. Our contributions can be summarized as follows.

- Through the use of the Bussgang decomposition, we aim to develop a closed form expression for the signal-to-quantization, interference and noise ratio (SQINR) as a function of number of users, BS antennas, and transmit power for sufficiently large number of users and BS antennas.
- We aim to examine the deterioration experienced by SQINR for small number of users for CE as well as 1-bit quantized transmission at the BS antennas. Deterioration experienced for small number of users in the downlink has been investigated in [1] for 1-bit quantized transmission at the BS.
- We improve receive per-user SQINR for small number of users by reducing cross-correlation among quantization error vector elements across the BS antennas, through the introduction of dither. Addition of dither to a signal before quantization, or dithering, can be used to modify characteristics of the quantization error [36]–[38]. Correlated Gaussian dither is added to the precoded signal, before quantization and subsequent transmission, by the transmitter, for both CE and 1-bit quantized transmission scenarios. We use Bussgang decomposition up to the third-order approximation to develop an analytical expression for the downlink system's SQINR as a function of dither power, transmit power, number of BS antennas, and number of users. We observe that with other system parameters being fixed, the SQINR peaks for finite non-trivial dither power.
- We maximize the SQINR with respect to dither power to get the optimal dither power as a function of system parameters such as number of BS antennas, number of users, and the total transmit power. Using the closed form expression of optimum dither as a function of number of users, BS antennas, and transmit power, we analyze the received per-user SQINR and BER at the users, and measure its improvement as a function of transmit power, number of users and number of BS antennas. In [2], we have presented a part of our algorithm, analyzing the improvement in received per-user SQINR with addition of correlated Gaussian dither and estimating the optimal dither power, for 1-bit quantized downlink transmission.

We observe that for systems incorporating the addition of optimal dither power, or optimized dithering, deterioration in the SQINR for small number of users is prevented, and the SQINR monotonically increases with decrease in number of users. We also observe that, while the increase in SQINR is sub-linear with respect to number of BS antennas for the original scenario, with optimized dithering SQINR increases linearly with number of BS antennas. Optimized dithering also improves BER performance of the CE and 1-bit quantized transmission systems, reducing the error floor experienced by quantized linear transmission systems significantly. We also examine the variation of optimum dither power with system parameters. We observe that optimum dither power increases monotonically with BS transmit power, approaching saturation at high transmit power. We also observe that with transmit power and number of users kept fixed, optimum dither power strictly decreases with number of BS antennas.

The remainder of this paper is structured as follows. In Section II, we introduce the downlink CE and 1-bit quantized MU massive MIMO systems with linear precoding at the BS. In Section III, we use the Bussgang decomposition to derive linear models for CE and 1-bit quantized transmission which are valid up to third-order harmonics of the output signal. Using the models introduced in Section III, we analyze the received signal at the users deriving closed-form expression of the received SQINR for asymptotic scenarios with ZF precoding at the BS in Section IV. In Section V, we measure the improvement in per-user SQINR with the addition of correlated Gaussian dither at the transmitter and derive optimal dither power, or optimized dithering, to achieve maximum improvement in SQINR as a function of the number of BS antennas, number of users and BS transmit power. We numerically analyze impact of dither on BER performance of the received signal and receive per-user SQINR, along with properties of optimal dither power, in Section VI. In Section VII, we present conclusions derived from our work, and directions for future work.

Notation: Throughout the paper, matrix $\mathrm{diag}(\mathbf{A})$, for a square matrix $\mathbf{A}$, equals a purely diagonal matrix with same diagonal entries as $\mathbf{A}$. As a counterpart, matrix $\mathrm{nondiag}(\mathbf{B}) = \mathbf{B} - \mathrm{diag}(\mathbf{B})$. Matrix operations $\mathbb{R}(\cdot), \mathbb{I}(\cdot)$ and $(\cdot)^*$ result in real part, imaginary part, and conjugate transpose. The matrix operation $\mathbf{C} \circ \mathbf{D}$ is an element-wise operation, such

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

that $[\mathbf{C} \circ \mathbf{D}]_{i,k} = \mathbb{R}[C]_{i,k}\mathbb{R}[D]_{i,k} + \mathrm{j}\ \mathbb{I}[C]_{i,k}\mathbb{I}[D]_{i,k}$, while $\mathbf{C}^{\circ n}$ results in entry-wise n^{th} power of matrix, such that $[\mathbf{C}^{\circ n}]_{i,k} = \mathbb{R}[C]^n_{i,k} + \mathrm{j}\ \mathbb{I}[C]^n_{i,k}$. Some other important notations used in the paper are provided in TABLE I.

TABLE I: Table of Notations

$\mathbf{s}$	Symbol vector
$\mathbf{H}$	Channel matrix
$\mathbf{x}$	Transmitted vector
$\mathbf{P}$	Linear precoder
$\mathbf{P}_{\text{d}}$	$\text{diag}(\mathbf{PP}^*)$
$\mathbf{P}_{\text{nd}}$	$\text{nondiag}(\mathbf{PP}^*)$
$\mathbf{q}_{\text{CE}}$	Quantization error vector for CE transmission at BS
$\mathbf{q}_{\text{b}}$	Quantization error vector for 1-bit transmission at BS
$\mathbf{C}_{\mathbf{q},\text{CE}}$	Covariance matrix for $\mathbf{q}_{\text{CE}}$
$\mathbf{C}_{\mathbf{q},\text{b}}$	Covariance matrix for $\mathbf{q}_{\text{b}}$
SQINR_{CE}	Per-user receive SQINR for CE scenario at user 1
SQINR_{b}	Per-user receive SQINR for 1-bit scenario at user 1
$\mathbf{R}_{\mathbf{H}}$	$\mathbf{I} - \mathbf{H}^*(\mathbf{HH}^*)^{-1}\mathbf{H}$
σ^2	Dither power
$\sigma^2_{\text{opt}}(N_{\text{t}}, U, \rho)$	Optimum dither power as a function of N_{t}, U and ρ

II. Downlink system model

We assume a narrowband, downlink communication model. A BS with N_{t} antennas communicates with U single-antenna users in the downlink. Complex symbol vector $\mathbf{s} = [s_1, \ldots, s_U]^{\text{T}}$ is sent to the U users, such that $\mathbb{E}[\mathbf{ss}^*] = \mathbf{I}$. An $N_{\text{t}} \times 1$ vector $\mathbf{x}$ is transmitted from the BS through the narrowband channel $\mathbf{H}$. The system model is given in Fig. 1. We assume that the BS has access to complete channel state information (CSI), and $\mathbf{x}$ is a precoded signal estimated using the values of $\mathbf{H}$ and $\mathbf{s}$. Assuming a total transmit power of ρ, and $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_U)$ as additive white Gaussian noise at the receivers, the $U \times 1$ received symbol vector is

$$\mathbf{r} = \sqrt{\frac{\rho}{N_{\text{t}}}}\mathbf{Hx} + \mathbf{n}. \tag{1}$$

Transmitted signal undergoes CE quantization before transmission and thus follows constant amplitude constraint at transmission. For CE signals, therefore, $x_\ell \in \{e^{\mathrm{j}\phi} : \phi \in [0, 2\pi)\}\ \forall \ell = 1, \ldots, N_{\text{t}}$. For a vector $\mathbf{a}$, with ϕ_k being phase of the k^{th} entry, the CE quantization operation on $\mathbf{a}$ gives the output

$$[\mathcal{Q}(\mathbf{a})]_k = e^{\mathrm{j}\phi_k}, \forall k = 1, \ldots, N_{\text{t}}. \tag{2}$$

We assume quantized linear transmission at the BS. With an $N_{\text{t}} \times U$ linear precoder $\mathbf{P}$ at the BS with transmitted symbol vector $\mathbf{x} = \mathcal{Q}(\mathbf{Ps})$, the received signal is

$$\mathbf{r} = \sqrt{\frac{\rho}{N_{\text{t}}}}\mathbf{H}\mathcal{Q}(\mathbf{Ps}) + \mathbf{n}. \tag{3}$$

Without loss of generality, the signal at the first user, with $\mathbf{h}_1^{\text{T}}$ being the $N_{\text{t}} \times 1$ channel observed at this user, is

$$r_1 = \sqrt{\frac{\rho}{N_{\text{t}}}}\mathbf{h}_1^{\text{T}}\mathcal{Q}(\mathbf{Ps}) + n_1. \tag{4}$$

We also consider the scenario of 1-bit quantized transmission, which assigns 1-bit to both the inphase and quadrature branches of the signal before transmission. With 1-bit quantized transmission at the BS antennas, $x_\ell \in \{\pm\frac{1}{\sqrt{2}} \pm \frac{\mathrm{j}}{\sqrt{2}}\}\ \forall \ell = 1, \ldots, N_{\text{t}}$. For a linear precoder $\mathbf{P}$, assuming 1-bit quantization, expression for the received signal (3) holds with the quantization operation given by

$$\mathcal{Q}(\mathbf{Ps}) = \frac{1}{\sqrt{2}}\left(\text{sign}(\mathbb{R}(\mathbf{Ps})) + \mathrm{j} \cdot \text{sign}(\mathbb{I}(\mathbf{Ps}))\right). \tag{5}$$

For CE transmission (2), as well as 1-bit quantized transmission (5), the factor of $\sqrt{\frac{1}{N_{\text{t}}}}$ in (1) ensures that the transmitted symbol vector has a total transmit power of ρ. We characterize the receive performance of the downlink transmission model given in (3) with per-user receive SQINR. We introduce the Bussgang decomposition to develop linear model of the nonlinear quantization operations of (2) and (5) in Section III. We use the linear model to develop a closed form expression of the per-user receive SQINR.

III. Modeling downlink transmission using Bussgang Decomposition

In this section, we obtain a linear model of the nonlinear quantization operation using the Bussgang decomposition. The Bussgang decomposition is used for stochastic analysis of nonlinear operations with Gaussian inputs. For a nonlinear operation, it states that, with C depending on the nonlinear operation considered, the cross-correlation of a Gaussian input signal X and the output signal Y is related to the auto-correlation of the input signal as [39]

$$R_{\text{XY}}(\tau) = C\ R_{\text{X}}(\tau). \tag{6}$$

A phase-only nonlinear operation $\mathbf{z} = \text{f}(\mathbf{y})$ for a Gaussian vector $\mathbf{y}$, denoting $\mathbf{F}$ and $\mathbf{q}$ as the coefficient matrix and the quantization error, can be written according to the Bussgang decomposition in a linear form as [39]

$$\mathbf{z} = \mathbf{Fy} + \mathbf{q}. \tag{7}$$

By means of the Bussgang decomposition, with an appropriately chosen $\mathbf{F}$, one can enforce that,

$$\mathbb{E}[\mathbf{qy}^*] = \mathbf{0}. \tag{8}$$

The derivation of $\mathbf{F}$ is based on (6), and is discussed in the following. For some value of ξ which depend upon nonlinear phase-only operation $\text{f}(\cdot)$ considered, the covariance of the input and output of the operation are related through Price's theorem (a generalization of the Bussgang property (6)) [35],

$$\mathbb{E}[z_k y_l^*] = \xi \frac{\mathbb{E}[y_k y_l^*]}{\sqrt{\mathbb{E}[y_k y_k^*]}}. \tag{9}$$

In matrix form

$$\mathbb{E}[\mathbf{zy}^*] = \xi\ \text{diag}(\mathbb{E}[\mathbf{yy}^*])^{-\frac{1}{2}}\mathbb{E}[\mathbf{yy}^*]. \tag{10}$$

Using (8) and (10)

$$\mathbf{F}\ \mathbb{E}[\mathbf{yy}^*] = \xi\ \text{diag}(\mathbb{E}[\mathbf{yy}^*])^{-\frac{1}{2}}\mathbb{E}[\mathbf{yy}^*]. \tag{11}$$

If $\mathbb{E}[\mathbf{yy}^*]$ is invertible, we get,

$$\mathbf{F} = \xi\ \text{diag}(\mathbb{E}[\mathbf{yy}^*])^{-\frac{1}{2}}. \tag{12}$$

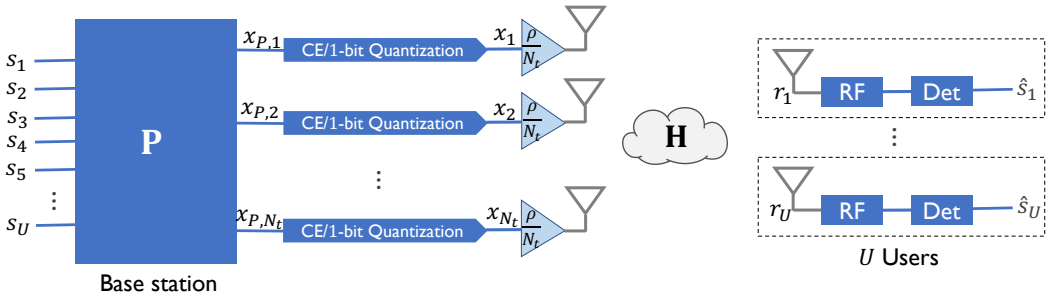


Fig. 1: System model for linear CE and 1-bit quantized transmission in the downlink.

Equation (10) and (11) give the intuition that the cross-correlation of elements of input Gaussian signal and elements of output signal is directly proportional to normalized cross-correlation of the corresponding elements of the Gaussian input signal vector, in accordance to the Bussgang property (6). Through the Central limit theorem [35], for an IID channel $\mathbf{H}$ and $\mathbf{P} = \mathbf{H}^*(\mathbf{H}\mathbf{H}^*)^{-1}$, vector $\mathbf{Ps}$ has entries which have a distribution sufficiently close to the Gaussian distribution. In Section III-A and Section III-B, we use the equation derived in this section to develop a linear model for the CE operation, and 1-bit quantization operation.

A. Downlink communication model for CE transmission

For $\mathbf{x}_\text{P} = \mathbf{Ps}$, and CE operation, the Bussgang decomposition yields

$$\mathbf{x} = \mathbf{F}\mathbf{x}_\text{P} + \mathbf{q}_\text{CE}, \tag{13}$$

with

$$\mathbb{E}[\mathbf{q}_\text{CE}\mathbf{x}_\text{P}^*] = \mathbf{0}. \tag{14}$$

For the CE operation, according to Price's theorem, $\xi = \sqrt{\frac{\pi}{4}}$ [35]. Since,

$$\mathbb{E}[\mathbf{x}_\text{P}\mathbf{x}_\text{P}^*] = \mathbf{P}\mathbf{P}^*, \tag{15}$$

equation (11) for the coefficient matrix $\mathbf{F}$, can be modified for CE as

$$\mathbf{F}\mathbf{P}\mathbf{P}^* = \sqrt{\frac{\pi}{4}}\text{diag}(\mathbf{P}\mathbf{P}^*)^{-\frac{1}{2}}\mathbf{P}\mathbf{P}^*. \tag{16}$$

Since $\mathbf{P}$ has full column rank, using (16), we have,

$$\mathbf{F}\mathbf{P} = \sqrt{\frac{\pi}{4}}\text{diag}(\mathbf{P}\mathbf{P}^*)^{-\frac{1}{2}}\mathbf{P}. \tag{17}$$

The cross-covariance of the elements of $\mathbf{x} = \mathcal{Q}(\mathbf{x}_\text{P})$ as a function of the elements of $\mathbf{x}_\text{P}$ for the CE operation, using the arcsin law is [35] is

$$\begin{aligned}&\mathbb{E}[x_k, x_\ell^*]\\ &= \frac{1}{4}\int_0^{2\pi} e^{-\text{j}\phi}\arcsin\left(\mathbb{R}\left(e^{\text{j}\phi}\mathbb{E}[x_{k,P}, x_{\ell,P}^*]\right)\right)\text{d}\phi,\\ &\forall \ell \neq k.\end{aligned} \tag{18}$$

Using the Taylor series expansion of the arcsin function and expanding for $\ell \neq k$,

$$\begin{aligned}&\mathbb{E}[x_k, x_\ell^*]\\ &= \frac{1}{4}\int_0^{2\pi} e^{-\text{j}\phi}\left(\mathbb{R}\left(e^{\text{j}\phi}\mathbb{E}[x_{k,P}, x_{\ell,P}^*]\right)\right.\\ &\quad\left.+\frac{\mathbb{R}\left(\exp(\text{j}\phi)\mathbb{E}[x_{k,P}, x_{\ell,P}^*]\right)^3}{6} + \mathcal{O}(\mathbb{E}[x_{k,P}, x_{\ell,P}^*]^5)\right)\text{d}\phi\end{aligned} \tag{19}$$

$$\begin{aligned}&= \frac{\pi}{4}[\mathbf{P}\mathbf{P}^*]_{k,\ell} + \frac{\pi}{32}([\mathbf{P}\mathbf{P}^*]_{k,\ell}^{\circ 3} + \mathbb{I}([\mathbf{P}\mathbf{P}^*]_{k,\ell})^2\mathbb{R}([\mathbf{P}\mathbf{P}^*]_{k,\ell})\\ &\quad + \text{j}\,\mathbb{I}([\mathbf{P}\mathbf{P}^*]_{k,\ell})\mathbb{R}([\mathbf{P}\mathbf{P}^*]_{k,\ell})^2) + \mathcal{O}([\mathbf{P}\mathbf{P}^*]_{k,\ell}^{\circ 5}).\end{aligned} \tag{20}$$

Assuming $\mathbf{P}_\text{d} = \text{diag}(\mathbf{P}\mathbf{P}^*)$ and $\mathbf{P}_\text{nd} = \text{nondiag}(\mathbf{P}\mathbf{P}^*)$, the covariance matrix for $\mathbf{x}$, using (20), is

$$\begin{aligned}&\mathbb{E}[\mathbf{x}\mathbf{x}^*]\\ &= \mathbf{I} + \frac{\pi}{4}\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbf{P}_\text{nd}\mathbf{P}_\text{d}^{-\frac{1}{2}} + \frac{\pi}{32}((\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbf{P}_\text{nd}\mathbf{P}_\text{d}^{-\frac{1}{2}})^{\circ 3} + (\mathbf{P}_\text{d}^{-\frac{1}{2}}\\ &\quad\mathbb{I}(\mathbf{P}\mathbf{P}^*)\mathbf{P}_\text{d}^{-\frac{1}{2}})^{\circ 2}\circ(\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbb{R}(\mathbf{P}\mathbf{P}^*)\mathbf{P}_\text{d}^{-\frac{1}{2}}) + \text{j}(\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbb{R}(\mathbf{P}\mathbf{P}^*)\\ &\quad\mathbf{P}_\text{d}^{-\frac{1}{2}})^{\circ 2}\circ(\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbb{I}(\mathbf{P}\mathbf{P}^*)\mathbf{P}_\text{d}^{-\frac{1}{2}})) + \mathcal{O}((\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbf{P}_\text{nd}\mathbf{P}_\text{d}^{-\frac{1}{2}})^{\circ 5}).\end{aligned} \tag{21}$$

Using the linear expansion of $\mathbf{x}$ in terms of $\mathbf{F}$, $\mathbf{x}_\text{P}$ and $\mathbf{q}_\text{CE}$ (13), and the expression of the covariance matrix of $\mathbf{x}$ (21), the covariance matrix of the quantization error $\mathbf{q}_\text{CE}$ is

$$\begin{aligned}&\mathbf{C}_{\mathbf{q},\text{CE}}\\ &= \mathbb{E}[\mathbf{q}_\text{CE}\mathbf{q}_\text{CE}^*]\\ &= \mathbb{E}[\mathbf{x}\mathbf{x}^*] - \mathbf{F}\mathbf{P}\mathbf{P}^*\mathbf{F}^*\end{aligned} \tag{22}$$

$$\begin{aligned}&= (1-\frac{\pi}{4})\mathbf{I} + \frac{\pi}{32}((\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbf{P}_\text{nd}\mathbf{P}_\text{d}^{-\frac{1}{2}})^{\circ 3} + (\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbb{I}(\mathbf{P}\mathbf{P}^*)\\ &\quad\mathbf{P}_\text{d}^{-\frac{1}{2}})^{\circ 2}\circ(\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbb{R}(\mathbf{P}\mathbf{P}^*)\mathbf{P}_\text{d}^{-\frac{1}{2}}) + \text{j}(\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbb{R}(\mathbf{P}\mathbf{P}^*)\mathbf{P}_\text{d}^{-\frac{1}{2}})^{\circ 2}\\ &\quad(\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbb{I}(\mathbf{P}\mathbf{P}^*)\mathbf{P}_\text{d}^{-\frac{1}{2}})) + \mathcal{O}((\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbf{P}_\text{nd}\mathbf{P}_\text{d}^{-\frac{1}{2}})^{\circ 5}).\end{aligned} \tag{23}$$

Using the value of $\mathbf{F}$ from (17) and the linear model given in (13), the downlink communication system as given in (1) can be expanded as

$$\mathbf{r} = \sqrt{\frac{\pi\rho}{4N_\text{t}}}\mathbf{H}\mathbf{P}_\text{d}^{-\frac{1}{2}}\mathbf{Ps} + \sqrt{\frac{\rho}{N_\text{t}}}\mathbf{H}\mathbf{q}_\text{CE} + \mathbf{n}. \tag{24}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

The SQINR at the first user, using (24), with CE quantization at the BS antennas is

$$\mathrm{SQINR}_{\mathrm{CE}} = \frac{\frac{\rho}{N_\mathrm{t}}\frac{\pi}{4}\left|[\mathbf{H}\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbf{P}]_{1,1}\right|^2}{\frac{\rho}{N_\mathrm{t}}\frac{\pi}{4}\sum_{u=2}^{u=U}\left|[\mathbf{H}\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbf{P}]_{1,u}\right|^2 + \frac{\rho}{N_\mathrm{t}}[\mathbf{H}\mathbf{C}_{\mathbf{q},\mathrm{CE}}\mathbf{H}^*]_{1,1} + 1}. \tag{25}$$

Through equation (25), we can see that CE operation leads to the normalization of rows of $\mathbf{P}$ with its corresponding diagonal entries increasing the MUI, along with the addition of quantization noise power in the denominator with the AWGN. In Section IV, we perform asymptotic analysis for the downlink system with quantized ZF precoding at the BS, and estimate a closed-form expression for the $\mathrm{SQINR}_{\mathrm{CE}}$ for sufficiently large numbers of BS antennas, and users.

B. Downlink communication model for 1-bit quantized transmission

For the 1-bit quantization operation, the matrix form of the Price's theorem (10) holds for $\xi = \sqrt{\frac{2}{\pi}}$ [35]. Following a similar analysis as in (13)-(17) for 1-bit quantization, denoting $\mathbf{F}$ as the coefficient matrix and $\mathbf{q}_\mathrm{b}$ as the quantization error uncorrelated with the input signal, the quantized vector $\mathbf{x}$ can be decomposed as

$$\mathbf{x} = \mathbf{F}\mathbf{x}_\mathrm{P} + \mathbf{q}_\mathrm{b}, \tag{26}$$

with matrix $\mathbf{F}$ related to precoding matrix $\mathbf{P}$ as

$$\mathbf{F}\mathbf{P} = \sqrt{\frac{2}{\pi}}\mathrm{diag}(\mathbf{P}\mathbf{P}^*)^{-\frac{1}{2}}\mathbf{P}. \tag{27}$$

Similar to how covariance matrix for $\mathbf{x}$ was estimated in (21), we use the arcsin law to obtain second-order statistics of the 1-bit quantization [35] as

$$\begin{aligned}\mathbb{E}[\mathbf{x}\mathbf{x}^*] &= \frac{2}{\pi}\Big(\arcsin(\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbb{R}(\mathbf{P}\mathbf{P}^*)\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}) \\ &\quad + \mathrm{j}\cdot\arcsin(\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbb{I}(\mathbf{P}\mathbf{P}^*)\mathbf{P}_\mathrm{d}^{-\frac{1}{2}})\Big) \qquad (28)\\ &= \frac{2}{\pi}\left(\frac{\pi}{2}\mathbf{I} + \mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbf{P}_\mathrm{nd}\mathbf{P}_\mathrm{d}^{-\frac{1}{2}} + \frac{1}{6}\mathbf{P}_\mathrm{d}^{-\frac{3}{2}}\mathbf{P}_\mathrm{nd}^{\circ 3}\mathbf{P}_\mathrm{d}^{-\frac{3}{2}}\right) \\ &\quad + \mathcal{O}((\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbf{P}_\mathrm{nd}\mathbf{P}_\mathrm{d}^{-\frac{1}{2}})^{\circ 5}). \qquad (29)\end{aligned}$$

We use the arcsin law to get (28) and Taylor series expansion of the $\arcsin(\cdot)$ up to the third order to get (29). Using the linear model in (26), expression for $\mathbf{F}$ from (27), and covariance matrix for $\mathbf{x}$ from (29), covariance matrix of the quantization error vector is

$$\begin{aligned}\mathbf{C}_{\mathbf{q},\mathrm{b}} &= \mathbb{E}[\mathbf{q}_\mathrm{b}\mathbf{q}_\mathrm{b}^*] \\ &= \mathbb{E}[\mathbf{x}_\mathrm{b}\mathbf{x}_\mathrm{b}^*] - \mathbf{F}_\mathrm{b}\mathbf{P}\mathbf{P}^*\mathbf{F}_\mathrm{b}^* \qquad (30)\\ &= \left(1 - \frac{2}{\pi}\right)\mathbf{I} + \frac{1}{3\pi}\mathbf{P}_\mathrm{d}^{-\frac{3}{2}}\mathrm{nondiag}(\mathbf{P}\mathbf{P}^*)^{\circ 3}\mathbf{P}_\mathrm{d}^{-\frac{3}{2}} \\ &\quad + \mathcal{O}((\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbf{P}_\mathrm{nd}\mathbf{P}_\mathrm{d}^{-\frac{1}{2}})^{\circ 5}). \qquad (31)\end{aligned}$$

Using downlink transmission model (1), the linear 1-bit quantization model from (26), and expression for $\mathbf{F}$ (27), the received signal is

$$\mathbf{r}_\mathrm{b} = \sqrt{\frac{2}{\pi}\frac{\rho}{N_\mathrm{t}}}\mathbf{H}\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbf{P}\mathbf{s} + \sqrt{\frac{\rho}{N_\mathrm{t}}}\mathbf{H}\mathbf{q}_\mathrm{b} + \mathbf{n}. \tag{32}$$

The SQINR at the first user, with 1-bit quantization at the BS antennas, is

$$\mathrm{SQINR}_\mathrm{b} = \frac{\frac{\rho}{N_\mathrm{t}}\frac{2}{\pi}\left|[\mathbf{H}\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbf{P}]_{1,1}\right|^2}{\frac{\rho}{N_\mathrm{t}}\frac{2}{\pi}\sum_{u=2}^{u=U}\left|[\mathbf{H}\mathbf{P}_\mathrm{d}^{-\frac{1}{2}}\mathbf{P}]_{1,u}\right|^2 + \frac{\rho}{N_\mathrm{t}}[\mathbf{H}\mathbf{C}_{\mathbf{q},\mathrm{b}}\mathbf{H}^*]_{1,1} + 1}. \tag{33}$$

Through equation (33), we can see that, similar to SQINR expression for CE transmission scenario (25), 1-bit quantization operation leads to normalization of rows of $\mathbf{P}$ with its corresponding diagonal entries, increasing the MUI, along with the addition of quantization noise power in the denominator. The signal power component and the MUI component are multiplied by factors of $\pi/4 \approx 0.785$ and $2/\pi \approx 0.637$ in the SQINR expression for CE and 1-bit transmissions, as shown in (25) and (33). Factors of $\pi/16$ and $1/2\pi$ are multiplied by the quantization noise component of the SQINR expression for CE and 1-bit transmissions. Intuitively, this would make the SQINR for CE transmission higher than SQINR for 1-bit transmission, as is confirmed by the simulations in the next section. In Section IV, we analyze the received signal for BS with 1-bit quantized ZF precoding, and estimate a closed form expression for $\mathrm{SQINR}_\mathrm{b}$ for asymptotic scenarios.

IV. Performance Analysis of Downlink Transmission with Quantized ZF Precoding

In this section, we mathematically analyze downlink performance, in terms of per-user SQINR, of MU-MIMO transmission with CE and 1-bit quantized transmission at the BS, assuming sufficiently large numbers of BS antennas and users. We also assume that ratio of number of BS antennas to number of users is finite, that is $\left|\frac{N_\mathrm{t}}{U}\right| < \infty$, for asymptotic analysis and simulations in the remaining part of the paper. While the asymptotic assumption on numbers of users and antennas makes the mathematical calculations feasible, through simulations we will observe that the analysis holds well for small numbers of users and antennas as well. We assume ZF precoding at the BS, with $\mathbf{P} = \mathbf{H}^*(\mathbf{H}\mathbf{H}^*)^{-1}$. The channel $\mathbf{H}$ is assumed to be IID Rayleigh with elements having unit variance for mathematical analysis, and simulations in this section, Section V and Section VI. Since $\mathbf{H}\mathbf{H}^*$ is a Wishart matrix, the expected value of the trace of $(\mathbf{H}\mathbf{H}^*)^{-1}$ in terms of number of users U, and number of BS antennas N_t, is [40]

$$\mathbb{E}[\mathrm{tr}((\mathbf{H}\mathbf{H}^*)^{-1})] = \frac{U}{N_\mathrm{t} - U}. \tag{34}$$

The variance of the diagonal entries of $(\mathbf{H}\mathbf{H}^*)^{-1}$ is [41]

$$\mathrm{Var}([(\mathbf{H}\mathbf{H}^*)^{-1}]_{k,k})$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

$$= \frac{8}{(2N_{\mathrm{t}}-U)(2N_{\mathrm{t}}-U-1)(2N_{\mathrm{t}}-U-3)}, \forall k = 1, \ldots, U, \tag{35}$$

making the variance of the trace bounded as

$$\begin{aligned} &\mathrm{Var}(\mathrm{tr}((\mathbf{H}\mathbf{H}^*)^{-1})) \\ &\leqslant \frac{8U}{(2N_{\mathrm{t}}-U)(2N_{\mathrm{t}}-U-1)(2N_{\mathrm{t}}-U-3)}. \end{aligned} \tag{36}$$

Since the variance, as given in (36), goes to zero as N_{t} becomes large, using (34) and the law of large numbers [35] asymptotically, the power of the precoded symbols is given by

$$[\mathbf{P}_{\mathrm{d}}]_{\ell,\ell} \xrightarrow[|\frac{N_{\mathrm{t}}}{U}|<\infty]{N_{\mathrm{t}}\to\infty, U\to\infty} \frac{\mathrm{tr}(\mathbf{P}\mathbf{P}^*)}{N_{\mathrm{t}}} \tag{37}$$

$$= \frac{\mathrm{tr}\left((\mathbf{H}\mathbf{H}^*)^{-1}\right)}{N_{\mathrm{t}}} \tag{38}$$

$$\xrightarrow[|\frac{N_{\mathrm{t}}}{U}|<\infty]{N_{\mathrm{t}}\to\infty, U\to\infty} \frac{U}{(N_{\mathrm{t}}-U)N_{\mathrm{t}}}, \quad \forall \ell = 1, \ldots, N_{\mathrm{t}}. \tag{39}$$

Using (39) we observe that for asymptotic scenarios, since diagonal entries of $\mathbf{P}\mathbf{P}^*$ are same, multi-user interference component of the $\mathrm{SQINR}_{\mathrm{CE}}$ as given in (25) vanishes. The channel response of the quantization error vector at the first user is

$$\begin{aligned} &\mathbf{h}_1^* \mathbf{C}_{\mathbf{q},\mathrm{CE}} \mathbf{h}_1 \\ &= \left(1-\frac{\pi}{4}\right) \mathbf{h}_1^* \mathbf{h}_1 + \frac{(N_{\mathrm{t}}-U)^3 {N_{\mathrm{t}}}^3}{U^3} \frac{\pi}{32} \mathbf{h}_1^* \left(\mathbf{P}_{\mathrm{nd}}^{\circ 3} + \mathbb{I}(\mathbf{P}\mathbf{P}^*)^{\circ 2} \right. \\ &\quad \left. \circ \mathbb{R}(\mathbf{P}\mathbf{P}^*) + \mathrm{j} \cdot \mathbb{R}(\mathbf{P}\mathbf{P}^*)^{\circ 2} \circ \mathbb{I}(\mathbf{P}\mathbf{P}^*)\right) \mathbf{h}_1 \\ &\quad + \mathcal{O}\left(\frac{{N_{\mathrm{t}}}^{10}}{U^5} \mathbf{h}_1^* \mathbf{P}_{\mathrm{nd}}^{\circ 5} \mathbf{h}_1\right) \end{aligned} \tag{40}$$

$$= \left(1-\frac{\pi}{4}\right) N_{\mathrm{t}} + \frac{\pi}{16} \frac{N_{\mathrm{t}}(N_{\mathrm{t}}-2U)}{U^2} + \mathcal{O}\left(\frac{{N_{\mathrm{t}}}^2}{U^3}\right) \tag{41}$$

$$\xrightarrow[|\frac{N_{\mathrm{t}}}{U}|<\infty]{N_{\mathrm{t}}\to\infty, U\to\infty} \left(1-\frac{\pi}{4}\right) N_{\mathrm{t}} + \frac{\pi}{16} \frac{N_{\mathrm{t}}(N_{\mathrm{t}}-2U)}{U^2}, \tag{42}$$

where the expression of $\mathbf{C}_{\mathbf{q},\mathrm{CE}}$ from (23), and the asymptotic approximation of $\mathbf{P}_{\mathrm{d}}$ (39), are used for the result in (40), and the asymptotic value of the second expression of (40), calculated at (95) in Appendix B is used for (41). Using $\mathrm{SQINR}_{\mathrm{CE}}$ expression (25), asymptotic expression for entries of $\mathbf{P}_{\mathrm{d}}$ (39), and asymptotic channel response to quantization error vector (42), the $\mathrm{SQINR}_{\mathrm{CE}}$ for sufficiently large values of N_{t} and U is

$$\mathrm{SQINR}_{\mathrm{CE}} \approx \frac{\rho \frac{\pi}{4} \frac{N_{\mathrm{t}}-U}{U}}{\rho\left(1-\frac{\pi}{4}\right) + \rho \frac{\pi}{16} \frac{N_{\mathrm{t}}-2U}{U^2} + 1}. \tag{43}$$

For a sufficiently large number of antennas and users, MUI at the users vanishes, quantization error along with AWGN remains the only distortion affecting the received signal and the received SQINR is attenuated by a factor of $\frac{\pi}{4}$ with respect to the unquantized case [42].

When using the Bussgang decomposition with the arcsin law up to the first order of approximation, along with the law of large numbers [35], the channel response to the quantization noise is

$$\mathbf{h}_1^* \mathbf{C}_{\mathbf{q},\mathrm{CE}} \mathbf{h}_1 = \left(1-\frac{\pi}{4}\right) \mathbf{h}_1^* \mathbf{h}_1 + \mathcal{O}\left(\frac{{N_{\mathrm{t}}}^6}{U^3} \mathbf{h}_1^* \mathbf{P}_{\mathrm{nd}}^{\circ 3} \mathbf{h}_1\right) \tag{44}$$

$$= \left(1-\frac{\pi}{4}\right) N_{\mathrm{t}} + \mathcal{O}\left(\frac{{N_{\mathrm{t}}}^2}{U^2}\right), \tag{45}$$

which uses derivation similar the scenario for 1-bit quantized transmission in [11]. The $\mathrm{SQINR}_{\mathrm{CE}}$, for sufficiently large values of N_{t} and U with quantization error component calculated through (45) is

$$\mathrm{SQINR}_{\mathrm{CE}} \approx \frac{\rho \frac{\pi}{4} \frac{N_{\mathrm{t}}-U}{U}}{\rho\left(1-\frac{\pi}{4}\right) + 1}. \tag{46}$$

From the quantization noise term in (46), derived from (44), we can see that the first order approximation of the covariance matrix $\mathbb{E}[\mathbf{x}\mathbf{x}^*]$ assumes the diagonal elements of $\mathbf{P}\mathbf{P}^*$ and subsequently of $\mathbf{C}_{\mathbf{q},\mathrm{CE}}$ to be zero. For the first order approximation to be accurate, therefore, cross-correlation among elements of quantization error $\mathbf{q}$ should be negligible relative to the variance of elements of $\mathbf{q}$.

In Fig. 2 (a), we plot simulated and asymptotic values of $\mathrm{SQINR}_{\mathrm{CE}}$ as a function of number of users U, for number of BS antennas $N_{\mathrm{t}} = 100$, and transmit power $\rho = 0$ dB. For unquantized ZF precoding at the BS and fixed transmit power, the receive SNR should consistently increase, in proportion to $\frac{1}{U}$ as U is decreased, when complete CSI is available at the BS [40]. With decreasing U, however, the $\mathrm{SQINR}_{\mathrm{CE}}$ increases to a maximum, and then starts decreasing with further decrease in U. The value of U, as a function of ρ and N_{t}, which maximizes the $\mathrm{SQINR}_{\mathrm{CE}}$ as given in (43) is

$$U_{\text{opt,CE}} = \frac{N_{\mathrm{t}}}{1+\sqrt{2N_{\mathrm{t}}\left(\frac{8}{\pi\rho}+\frac{8}{\pi}-2\right)-1}}. \tag{47}$$

The behavior of $\mathrm{SQINR}_{\mathrm{CE}}$ captured by analysis using Bussgang decomposition up to the third order approximation is very close to the simulated behavior, as shown in Fig. 2 (a). This behavior is different from what we obtain from the analysis using Bussgang decomposition up to the first order approximation, wherein the $\mathrm{SQINR}_{\mathrm{CE}}$ monotonically increases with decrease in U, for all values of U. The difference and deterioration in $\mathrm{SQINR}_{\mathrm{CE}}$ is due to high cross-correlation among the elements of $\mathbf{q}$ for small U, which is not captured by the first order analysis, but by the third-order term in the channel response of quantization error, as given in (40). Due to high correlation among the quantization error components, quantization error elements across the BS antennas combine coherently at the users, enhancing the quantization noise in the received signal.

For 1-bit quantized transmission at the BS, the channel response of the quantization noise is

$$\begin{aligned} &\mathbf{h}_1^* \mathbf{C}_{\mathbf{q},\mathrm{b}} \mathbf{h}_1 \\ &= \left(1-\frac{2}{\pi}\right) \mathbf{h}_1^* \mathbf{h}_1 + \frac{1}{3\pi} \frac{(N_{\mathrm{t}}-U)^3 {N_{\mathrm{t}}}^3}{U^3} \mathbf{h}_1^* \mathbf{P}_{\mathrm{nd}}^{\circ 3} \mathbf{h}_1 \\ &\quad + \mathcal{O}\left(\frac{{N_{\mathrm{t}}}^{10}}{U^5} \mathbf{h}_1^* \mathbf{P}_{\mathrm{nd}}^{\circ 5} \mathbf{h}_1\right) \end{aligned} \tag{48}$$

$$= \left(1-\frac{2}{\pi}\right) N_{\mathrm{t}} + \frac{1}{2\pi} \frac{N_{\mathrm{t}}(N_{\mathrm{t}}-2U)}{U^2} + \mathcal{O}\left(\frac{{N_{\mathrm{t}}}^2}{U^3}\right) \tag{49}$$

$$\xrightarrow[|\frac{N_{\mathrm{t}}}{U}|<\infty]{N_{\mathrm{t}}\to\infty, U\to\infty} \left(1-\frac{2}{\pi}\right) N_{\mathrm{t}} + \frac{1}{2\pi} \frac{N_{\mathrm{t}}(N_{\mathrm{t}}-2U)}{U^2}, \tag{50}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

where we used the asymptotic expression for $\mathbf{C}_{\mathbf{q},\mathrm{b}}$ from (31), and the asymptotic expression for $\mathbf{P}_{\mathrm{nd}}$ from (39) to get (48). For (49), we used the asymptotic result for $\mathbf{h}_1^*\mathbf{P}_{\mathrm{nd}}^{\circ 3}\mathbf{h}_1$ from (93) in Appendix A. Using the SQINR expression from (33), and asymptotic channel response to quantization error vector given in (50), for sufficiently large numbers of BS antennas, N_{t} and users, U the

$$\mathrm{SQINR}_{\mathrm{b}} \approx \frac{\rho \frac{2}{\pi}\frac{(N_{\mathrm{t}}-U)}{U}}{\rho\left(1-\frac{2}{\pi}+\frac{1}{2\pi}\frac{N_{\mathrm{t}}-2U}{U^2}\right)+1}. \tag{51}$$

For the case of 1-bit quantized transmission in (51), for sufficiently large numbers of antennas and numbers of users, MUI at the users vanishes, and the received SQINR is attenuated by a factor of $\frac{2}{\pi}$ in comparison to the unquantized case. We derived this result as a part of our work in [1].

When using the Bussgang decomposition up to the first order approximation, channel response of the quantization noise for 1-bit quantization at the BS, as calculated in [11] is

$$\mathbf{h}_1^*\mathbf{C}_{\mathbf{q},\mathrm{b}}\mathbf{h}_1 = \left(1-\frac{2}{\pi}\right)\mathbf{h}_1^*\mathbf{h}_1 + \mathcal{O}\left(\frac{{N_{\mathrm{t}}}^6}{U^3}\mathbf{h}_1^*\mathbf{P}_{\mathrm{nd}}^{\circ 3}\mathbf{h}_1\right) \tag{52}$$

$$= \left(1-\frac{2}{\pi}\right)N_{\mathrm{t}} + \mathcal{O}\left(\frac{{N_{\mathrm{t}}}^2}{U^2}\right). \tag{53}$$

For sufficiently large N_{t} and U, using (53), and the Bussgang decomposition up to the first order approximation the

$$\mathrm{SQINR}_{\mathrm{b}} \approx \frac{\rho \frac{2}{\pi}\frac{(N_{\mathrm{t}}-U)}{U}}{\rho\left(1-\frac{2}{\pi}\right)+1}. \tag{54}$$

Similar to (45), the first order analysis, conducted in [11] for the 1-bit quantized transmission, assumes that the diagonal elements of $\mathbf{P}\mathbf{P}^*$ are much larger than the non-diagonal elements. The cross-correlation of elements of quantization error vector $\mathbf{q}_{\mathrm{b}}$ across the BS antennas is assumed to be negligible when compared to the variance of its elements.

In Fig. 2 (b) [1], we plot the simulated and asymptotic closed-form $\mathrm{SQINR}_{\mathrm{b}}$ as a function of number of users U. We observe that the $\mathrm{SQINR}_{\mathrm{b}}$ does not increase monotonically as U is decreased, for small U. It instead increases to a peak, and then deteriorates rapidly with further decrease in U, while $\mathrm{SQINR}_{\mathrm{b}}$ given by first order asymptotic analysis increases monotonically for all values of U. Similar to equation (47), $\mathrm{SQINR}_{\mathrm{b}}$ attains a peak for the case of 1-bit quantized transmission at the BS at [1]

$$U_{\mathrm{opt,b}} = \frac{N_{\mathrm{t}}}{1+\sqrt{2N_{\mathrm{t}}\left(\frac{\pi}{\rho}+\pi-2\right)-1}}. \tag{55}$$

Since, the SQINR given by first order analysis (54) assumes non-diagonal entries of $\mathbf{P}\mathbf{P}^*$ and $\mathbf{C}_{\mathbf{q},\mathrm{b}}$ to be negligible relative to their diagonal entries, we infer that the deterioration of downlink performance for small U is caused due to high correlation among components of $\mathbf{q}_{\mathrm{b}}$. The deterioration in the $\mathrm{SQINR}_{\mathrm{b}}$, and the cross-correlation among elements of quantization error is captured by the third-order term of channel response to quantization error in (48). Decreasing cross-correlation among the quantization error vector elements across the BS antennas should, therefore, improve the receive per-user SQINR for both the CE and 1-bit quantized transmission at the BS, for small U. We, therefore, propose to add dither to the precoded signal before quantization, and subsequent transmission, in order to decrease cross-correlation among elements of quantization noise. In the next section, we explore the idea of adding dither to the precoded signal before quantization, and optimizing its power to get the maximum SQINR for given system parameters.

V. Improving downlink performance for quantized linear transmission by adding correlated dither

For small U, downlink linear transmission with CE and 1-bit quantized transmission at the BS antennas shows rapid deterioration in per-user receive SQINR, with decreasing U. At low power, SU transmission along the dominant eigenvector is optimal for a MIMO system. With CE and 1-bit quantized transmission, however, downlink massive MIMO based SU transmission would experience significant attenuation in SQINR. As established in the previous section, while the receive per-user SQINR calculated by the asymptotic first-order analysis, with the assumption of negligible correlation in the quantization noise components across BS antennas, increases monotonically with decrease in U, the actual SQINR along with the third-order approximation experiences monotonic decrease with decreasing U for small values of U. We predict, therefore, that reducing correlation among the quantization error components should help in reducing the deterioration in SQINR experienced for small U.

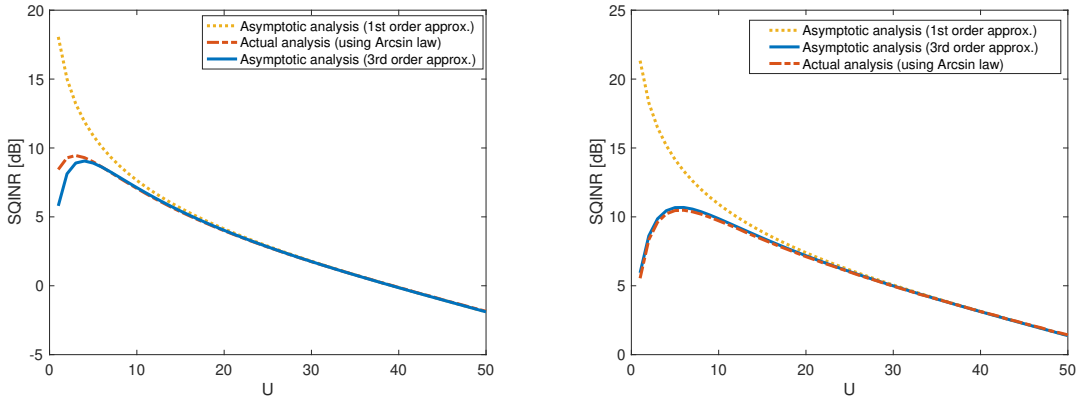
To reduce correlation among quantization noise components for small values of U, we propose the addition of correlated dither to the precoded signal before quantization, by the transmitter. Researchers in the past have used dither to whiten the quantization noise [36], and enhance effective bit resolution of the quantization operation [37]. Previous work has also explored the idea of experimentally enhancing the performance of quantized downlink communication through the addition of dither by the transmitter [38]. In this paper, we use the Bussgang decomposition to maximize the received SQINR with respect to the dither power added, and derive the optimal dither power as a function of the transmit power ρ, number of users U, and number of BS antennas N_{t}. IID Gaussian dither of power σ^2, $\mathbf{w} \sim \mathcal{CN}(\mathbf{0}, \sigma^2\mathbf{I})$ is projected on the null space of the channel $\mathbf{H}$ to remove it from the users' subspace, as done in work [38], and reduce the detrimental effect of dither on the downlink transmission through the channel. We, therefore, get correlated dither

$$\mathbf{v} = (\mathbf{I} - \mathbf{H}^*(\mathbf{H}\mathbf{H}^*)^{-1}\mathbf{H})\mathbf{w}. \tag{56}$$

Considering the general quantization case, correlated dither obtained through (56) is added to the precoded signal before quantization to give

$$\mathbf{x} = \mathcal{Q}(\mathbf{P}\mathbf{s} + \mathbf{v}). \tag{57}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing



(a) For CE transmission

(b) For 1-bit quantized transmission

Fig. 2: Comparison of the per-user SQINR with $N_{\mathrm{t}} = 100$ and different values of U, for CE and 1-bit quantized transmissions at the BS antennas. Transmit power is $\rho = 0$ dB. We can observe that for small values of U, SQINR rapidly deteriorates as U is decreased. The Bussgang decomposition up to the third-order approximation accurately depicts the SQINR for small values of U, unlike the first-order approximation. A variation of the plot (b) has been published as part of [1].

Using the Bussgang decomposition, assuming the inputs to be Gaussian, with the coefficient matrix $\mathbf{F}$, and the quantization noise vector $\mathbf{q}$, $\mathbf{x}$ can be decomposed as

$$\mathbf{x} = \mathbf{F}(\mathbf{P}\mathbf{s} + \mathbf{v}) + \mathbf{q}. \tag{58}$$

Letting $\mathbf{x_v} = \mathbf{P}\mathbf{s} + \mathbf{v}$ and $\mathbf{R_H} = \mathbf{I} - \mathbf{H}^*(\mathbf{H}\mathbf{H}^*)^{-1}\mathbf{H}$, the covariance matrix of the input to the quantizer is

$$\mathbb{E}[\mathbf{x_v}\mathbf{x_v}^*] = \mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H}. \tag{59}$$

As with (11), we have

$$\begin{aligned}&\mathbf{F}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H})\\ &= \xi\left(\mathbf{P}_{\mathrm{d}}^{-\frac{1}{2}} + \frac{1}{\sigma}\mathrm{diag}(\mathbf{R_H})^{-\frac{1}{2}}\right)(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H}),\end{aligned} \tag{60}$$

which results in

$$\mathbf{F} = \xi\mathrm{diag}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H})^{-\frac{1}{2}}. \tag{61}$$

For CE transmission, the covariance matrix of the quantization error vector $\mathbf{q}_{\mathrm{CE}}$, using the expression of $\mathbf{C}_{\mathbf{q},\mathrm{CE}}$ from (22), and $\mathbf{F}$ from (61), with $\xi = \sqrt{\frac{\pi}{4}}$, is

$$\begin{aligned}&\mathbf{C}_{\mathbf{q},\mathrm{CE}}\\ &= \mathbb{E}[\mathbf{x}\mathbf{x}^*] - \mathbf{F}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H})\mathbf{F}^* &(62)\\ &= \frac{\pi}{32}\mathrm{diag}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H})^{-\frac{3}{2}}(\mathrm{nondiag}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H})^{\circ 3}\\ &+ \mathbb{I}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H})^{\circ 2} \circ \mathbb{R}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H}) + \mathrm{j}\cdot\mathbb{R}(\mathbf{P}\mathbf{P}^*\\ &+ \sigma^2\mathbf{R_H})^{\circ 2} \circ \mathbb{I}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H}))\mathrm{diag}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H})^{-\frac{3}{2}}\\ &+ \mathcal{O}(\mathrm{nondiag}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H})^{\circ 5}). &(63)\end{aligned}$$

With the received signal at the users given by

$$\mathbf{r} = \sqrt{\frac{\rho}{N_{\mathrm{t}}}}\mathbf{H}\mathbf{F}\mathbf{P}\mathbf{s} + \sqrt{\frac{\rho}{N_{\mathrm{t}}}}\mathbf{H}\mathbf{F}\mathbf{v} + \sqrt{\frac{\rho}{N_{\mathrm{t}}}}\mathbf{H}\mathbf{q}_{\mathrm{CE}} + \mathbf{n}, \tag{64}$$

at the first user,

$$\begin{aligned}&\mathrm{SQINR}_{\mathrm{CE}}\\ &= \frac{\frac{\rho}{N_{\mathrm{t}}}\left|[\mathbf{h}_1^{\mathrm{T}}\mathbf{F}\mathbf{P}]_1\right|^2}{\frac{\rho}{N_{\mathrm{t}}}(\sum_{u=2}^{u=U}\left|[\mathbf{h}_1^{\mathrm{T}}\mathbf{F}\mathbf{P}]_u\right|^2 + \sigma^2\mathbf{h}_1^*(\mathbf{F}\mathbf{R_H}\mathbf{F}^* + \mathbf{C}_{\mathbf{q},\mathrm{CE}})\mathbf{h}_1) + 1}.\end{aligned} \tag{65}$$

Along with the quantization noise, MUI and AWGN, an additional term due to addition of correlated dither to the precoded signal is added to the distortion experienced by the received signal.

We now proceed to estimate $\mathrm{SQINR}_{\mathrm{CE}}$ for asymptotic scenarios with sufficiently large numbers of BS antennas and users. For sufficiently large N_{t} and U, using the law of large numbers [35], and the expression for $\mathrm{tr}((\mathbf{H}\mathbf{H}^*)^{-1})$ from (39), power of the symbols before quantization is,

$$\begin{aligned}&[\mathrm{diag}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R_H})]_{\ell,\ell}\\ &= [\mathrm{diag}(\mathbf{P}\mathbf{P}^*)]_{\ell,\ell} + [\mathrm{diag}(\sigma^2\mathbf{R_H})]_{\ell,\ell} &(66)\\ &\xrightarrow[|\frac{N_{\mathrm{t}}}{U}|<\infty]{N_{\mathrm{t}}\to\infty, U\to\infty} \frac{1}{N_{\mathrm{t}}}(\mathrm{tr}(\mathbf{P}\mathbf{P}^*) + \sigma^2\mathrm{tr}(\mathbf{R_H})) &(67)\\ &= \frac{1}{N_{\mathrm{t}}}(\mathrm{tr}((\mathbf{H}\mathbf{H}^*)^{-1}) + \sigma^2(\mathrm{N_t} - \mathrm{U})) &(68)\\ &= \frac{U + \sigma^2(N_{\mathrm{t}} - U)^2}{(N_{\mathrm{t}} - U)N_{\mathrm{t}}} + \mathcal{O}\left(\frac{\sqrt{U}}{{N_{\mathrm{t}}}^{5/2}}\right) &(69)\\ &\xrightarrow[|\frac{N_{\mathrm{t}}}{U}|<\infty]{N_{\mathrm{t}}\to\infty, U\to\infty} \frac{U + \sigma^2(N_{\mathrm{t}} - U)^2}{(N_{\mathrm{t}} - U)N_{\mathrm{t}}}, \quad \forall \ell = 1, \ldots, N_{\mathrm{t}}. &(70)\end{aligned}$$

We can see that in (70) the weight of dither power relative to the power of precoded symbol grows quadratically with N_{t}. Using the asymptotic expression for the diagonal values of $\mathbb{E}[\mathbf{x_v}\mathbf{x_v}^*]$ in (70), and the expression for $\mathbf{F}$ in (61), we can see that for sufficiently large values of N_{t} and U, MUI and channel response to dither has negligible effect on the SQINR, as given in (65). In Section VI through simulations, we show that (70) is highly accurate for small values of U. Using ideas from estimation of channel response to quantization error at users for CE transmission without the addition of dither (42), and asymptotic result (98) from Appendix C, the channel response

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

of the quantization error for CE transmission after addition of correlated dither at the users is

$$\begin{aligned}&\mathbf{h}_1^*\mathbf{C}_{\mathbf{q},\text{CE}}\mathbf{h}_1\\&=\left(1-\frac{\pi}{4}\right)\mathbf{h}_1^*\mathbf{h}_1+\frac{\pi}{32}\frac{{N_\text{t}}^3(N_\text{t}-U)^3}{(U+\sigma^2(N_\text{t}-U)^2)^3}\mathbf{h}_1^*\\&\quad(\text{nondiag}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})^{\circ 3}+\mathbb{I}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})^{\circ 2}\\&\quad\circ\mathbb{R}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})+\text{j}\mathbb{R}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})^{\circ 2}\circ\mathbb{I}(\mathbf{P}\mathbf{P}^*\\&\quad+\sigma^2\mathbf{R}_\mathbf{H}))\mathbf{h}_1+\mathcal{O}(\mathbf{h}_1^*(\text{nondiag}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})^{\circ 5})\mathbf{h}_1)\end{aligned}\tag{71}$$

$$\begin{aligned}&=\left(1-\frac{\pi}{4}\right)N_\text{t}+\frac{\pi}{16}UN_\text{t}(N_\text{t}-2U-\sigma^2(N_\text{t}-U)^2)\\&\quad\frac{(1-\sigma^2(N_\text{t}-U))^2}{(U+\sigma^2(N_\text{t}-U)^2)^3}+\mathcal{O}\left(\frac{U^2}{{N_\text{t}}^5}\right)\end{aligned}\tag{72}$$

$$\begin{aligned}&\xrightarrow[|\frac{N_\text{t}}{U}|<\infty]{N_\text{t}\to\infty,U\to\infty}\left(1-\frac{\pi}{4}\right)N_\text{t}+\frac{\pi}{16}UN_\text{t}(N_\text{t}-2U\\&\qquad-\sigma^2(N_\text{t}-U)^2)\frac{(1-\sigma^2(N_\text{t}-U))^2}{(U+\sigma^2(N_\text{t}-U)^2)^3},\end{aligned}\tag{73}$$

for sufficiently large N_t and U, where expression for covariance matrix of quantization error derived at (63), and asymptotic result (70) are used to get (71), and asymptotic result for (71) from (104) in Appendix D are used to get (73). For sufficiently large values of N_t and U, using $\mathbf{F}$ derived from (61), expression for SQINR_CE calculated in (64), and asymptotic channel response to the quantization error from (73), the

$$\begin{aligned}&\text{SQINR}_\text{CE}\\&\approx\frac{\rho\frac{\pi(N_\text{t}-U)}{4(U+\sigma^2(N_\text{t}-U)^2)}}{\rho\left(1-\frac{\pi}{4}+\frac{\pi}{16}\frac{U(N_\text{t}-2U-\sigma^2(N_\text{t}-U)^2)(1-\sigma^2(N_\text{t}-U))^2}{(U+\sigma^2(N_\text{t}-U)^2)^3}\right)+1}.\end{aligned}\tag{74}$$

From (74), we can see that dither affects both the signal power component of the received signal and the quantization noise component of the received signal. From (73), we can see that effect of dither power on the quantization error is depicted in the third-order term of the channel response of quantization error in the denominator, which represents cross-correlation among quantization error elements. Dither power, therefore, directly impacts the cross-correlation among elements of quantization error.

As dither power σ^2 is increased from 0, its effect on the quantization operation increases, and quantization noise component decreases due to decrease in correlation among elements of quantization error, as given by the expression $\frac{U(N_\text{t}-2U-\sigma^2(N_\text{t}-U)^2)(1-\sigma^2(N_\text{t}-U))^2}{(U+\sigma^2(N_\text{t}-U)^2)^3}$. With increasing σ^2, the signal power component, $\frac{\pi(N_\text{t}-U)}{4(U+\sigma^2(N_\text{t}-U)^2)}$, also decreases. For sufficiently large σ^2, quantization noise power saturates as it is $\mathcal{O}(1)$ with respect to σ^2. We therefore predict that SQINR_CE increases as σ^2 is increased from 0, reaching a maximum for finite non-trivial value of σ^2. In Fig. 3, we plot the variation of SQINR_CE with σ, validating our prediction. By differentiating (74) with respect to σ^2 and equating it to zero, we calculate the σ^2_opt which maximizes SQINR for given values of system parameters: transmit power ρ, number of BS antennas N_t, and number of users U. Differentiating the SQINR_CE with respect to σ^2 and equating to zero, we get the following equation

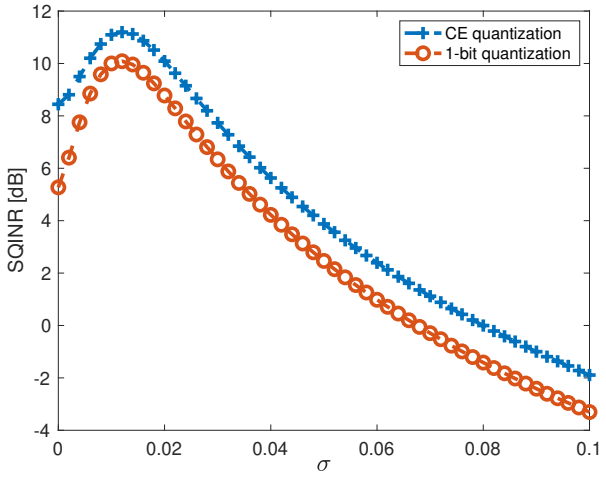


Fig. 3: Variation of the per-user SQINR (calculated using the Arcsin law) for transmit power $\rho = 0$ dB, number of BS antennas $N_\text{t} = 100$, and number of users $U = 1$ with σ. It can be seen that as dither power increases from 0, SQINR increases, reaching a maximum for finite non-trivial value of σ^2, and decreases with further increase in dither power.

$$\begin{aligned}0=&-(N_\text{t}-U)\left(\left(\rho\left(1-\frac{\pi}{4}\right)+1\right)(U+\sigma^2(N_\text{t}-U)^2)^3\right.\\&+\frac{\rho\pi}{16}U(N_\text{t}-2U)(1-\sigma^2(N_\text{t}-U))^3\Big)\\&+\frac{3\rho\pi}{16}UN_\text{t}(N_\text{t}-2U)(1-\sigma^2(N_\text{t}-U))^2.\end{aligned}\tag{75}$$

Equation (75) is cubic in σ^2. Analyzing the discriminant of the equation, we deduce that the equation has one real root and two complex roots for all suitable values of arguments. We use the real root as optimum dither power needed to maximize the SQINR_CE. The theoretical expression of optimum dither power, $\sigma^2_\text{opt}(N_\text{t},U,\rho)$, is a function of N_t, U, ρ and being a very large expression cannot be accommodated in this paper. It, however, gives a closed-form value of optimal dither power for given system parameters, which can be calculated by the BS for downlink transmission.

For 1-bit quantization, covariance matrix for quantization error vector $\mathbf{q}_\text{b}$, when correlated dither is added to the signal before quantization, using similar $\mathbf{C}_{\mathbf{q},\text{b}}$ expression from (31) and expression for $\mathbf{F}$ from (61) with $\xi=\sqrt{\frac{2}{\pi}}$, is

$$\begin{aligned}&\mathbf{C}_{\mathbf{q},\text{b}}\\&=\mathbb{E}[\mathbf{x}\mathbf{x}^*]-\mathbf{F}_\text{b}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})\mathbf{F}_\text{b}^*\end{aligned}\tag{76}$$

$$\begin{aligned}&=\left(1-\frac{2}{\pi}\right)\mathbf{I}+\frac{1}{3\pi}\text{diag}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})^{-\frac{3}{2}}\text{nondiag}(\mathbf{P}\mathbf{P}^*\\&\quad+\sigma^2\mathbf{R}_\mathbf{H})^{\circ 3}\text{diag}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})^{-\frac{3}{2}}\\&\quad+\mathcal{O}(\text{nondiag}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})^{\circ 5}).\end{aligned}\tag{77}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

Asymptotic channel response of the quantization error vector at the first user is, therefore,

$$\begin{aligned}&\mathbf{h}_1^*\mathbf{C}_{\mathbf{q},\mathrm{b}}\mathbf{h}_1\\&=\left(1-\frac{2}{\pi}\right)\mathbf{h}_1^*\mathbf{h}_1+\frac{1}{3\pi}\frac{{N_\mathrm{t}}^3(N_\mathrm{t}-U)^3}{(U+\sigma^2(N_\mathrm{t}-U)^2)^3}\mathbf{h}_1^*\mathrm{nondiag}(\mathbf{P}\mathbf{P}^*\\&\quad+\sigma^2\mathbf{R}_\mathbf{H})^{\circ 3}\mathbf{h}_1+\mathcal{O}(\mathbf{h}_1^*(\mathrm{nondiag}(\mathbf{P}\mathbf{P}^*+\sigma^2\mathbf{R}_\mathbf{H})^{\circ 5})\mathbf{h}_1)\end{aligned}\tag{78}$$

$$\begin{aligned}&=\left(1-\frac{2}{\pi}\right)N_\mathrm{t}+\frac{1}{2\pi}UN_\mathrm{t}(N_\mathrm{t}-2U-\sigma^2(N_\mathrm{t}-U)^2)\\&\quad\frac{(1-\sigma^2(N_\mathrm{t}-U))^2}{(U+\sigma^2(N_\mathrm{t}-U)^2)^3}+\mathcal{O}\left(\frac{U^2}{{N_\mathrm{t}}^5}\right)\end{aligned}\tag{79}$$

$$\begin{aligned}&\xrightarrow[|\frac{N_\mathrm{t}}{U}|<\infty]{N_\mathrm{t}\to\infty,U\to\infty}\left(1-\frac{2}{\pi}\right)N_\mathrm{t}+\frac{1}{2\pi}UN_\mathrm{t}(N_\mathrm{t}-2U\\&\qquad-\sigma^2(N_\mathrm{t}-U)^2)\frac{(1-\sigma^2(N_\mathrm{t}-U))^2}{(U+\sigma^2(N_\mathrm{t}-U)^2)^3},\end{aligned}\tag{80}$$

where asymptotic result (70) is used to get (78), and asymptotic result for second term of (78) from Appendix C is used to get (79). The $\mathrm{SQINR_b}$ at the first user, for sufficiently large values of N_t and U, is given as,

$$\begin{aligned}&\mathrm{SQINR_b}\\&\approx\frac{\rho\frac{2(N_\mathrm{t}-U)}{\pi(U+\sigma^2(N_\mathrm{t}-U)^2)}}{\rho\left(1-\frac{2}{\pi}\right)+\rho\frac{1}{2\pi}\frac{U(N_\mathrm{t}-2U-\sigma^2(N_\mathrm{t}-U)^2)(1-\sigma^2(N_\mathrm{t}-U))^2}{(U+\sigma^2(N_\mathrm{t}-U)^2)^3}+1}.\end{aligned}\tag{81}$$

From (81), we can see that dither affects both the signal power, and quantization noise component of the received signal. From (80), we can see that dither directly impacts the third-order term of the channel response of quantization error, modifying the cross-correlation among elements of quantization error.

Similar to what we observed for (74), for (81), we can see that with an increase in dither power σ^2 from zero both the quantization noise power and signal power components of the $\mathrm{SQINR_b}$ decrease, with quantization noise power reaching saturation for sufficiently high σ^2. Thus, the $\mathrm{SQINR_b}$ maximizes at finite non-trivial value of σ^2. We plot the variation of $\mathrm{SQINR_b}$ with σ, in Fig. 3. We differentiate the SQINR with respect to σ^2 to get the optimum dither power, $\sigma^2_{\mathrm{opt}}(N_\mathrm{t},U,\rho)$, required to maximize the SQINR. For the downlink system with 1-bit quantized transmission, differentiating the $\mathrm{SQINR_b}$ from (81) with respect to σ^2 and equating it to zero, we get following equation in terms of σ^2, N_t, U and ρ,

$$\begin{aligned}0=&-(N_\mathrm{t}-U)\left(\left(\rho\left(1-\frac{2}{\pi}\right)+1\right)(U+\sigma^2(N_\mathrm{t}-U)^2)^3\right.\\&\left.+\frac{\rho}{2\pi}U(N_\mathrm{t}-2U)(1-\sigma^2(N_\mathrm{t}-U))^3\right)\\&+\frac{3\rho}{2\pi}UN_\mathrm{t}(N_\mathrm{t}-2U)(1-\sigma^2(N_\mathrm{t}-U))^2.\end{aligned}\tag{82}$$

Similar to equation (75), equation (82) is cubic in σ^2. We have derived this result as a part of our work in [2]. For all suitable values of arguments, this equation has two complex roots and one real root which we use as optimum dither power needed to maximize the SQINR. The expression can be calculated by hand or using tools like Mathematica, to give a closed form expression of optimal dither power for given system parameters. The BS can, therefore, calculate the power of optimum dither required, as a function of system parameters for downlink transmission.

VI. Numerical analysis of linear transmission with optimized dithering

In this section, we analyze impact of dither, and optimized dither calculated from the methods developed in previous sections, on the SQINR and the BER of the received signal, with CE and 1-bit quantized transmissions at the BS antennas.

In Fig. 4, we plot the per user SQINR for different values of N_t and U, with and without optimized dithering, with ZF precoding at the BS. Simulations are performed over 1000 IID Rayleigh channel instances for each setting. The transmit power at the BS is $\rho = 0$ dB. For high-resolution downlink MU-MIMO system with ZF precoding, the received SNR should increase proportionally to $\frac{1}{U}$ [40], when N_t is kept constant. Received SNR should also increase linearly with N_t for constant U, for high resolution systems. We can see that for both CE and 1-bit quantized transmissions, the SQINR deteriorates rapidly as U is reduced, for small values of U when dither is not added at the transmitter. When N_t is increased for a constant U, the SQINR increases sub-linearly, not benefitting from the array gain of downlink massive MIMO transmission. The addition of dither with optimal power $\sigma^2_{\mathrm{opt}}(N_\mathrm{t},U,\rho)$, reduces correlation among quantization error vector components, maximizing the SQINR for given values of N_t, U and ρ. Thus, the SQINR for downlink communication with optimized dithering increases monotonically as U decreases, for all values of U. It also increases almost linearly with N_t exploiting the advantage of array gain.

In Fig. 5, we plot uncoded BER performance of the downlink transmission system incorporating dithering at the transmitter, as a function of σ for the CE and 1-bit quantized downlink massive MIMO systems, with $N_\mathrm{t} = 100$ BS antennas and $U = 1$ user. Signal modulation of 16-QAM is assumed and ZF precoding is implemented at the BS. We simulate over 1000 IID Rayleigh channel instances, with 1000 bits transmitted over each channel instance. We can see that as dither power is increased from zero, BER decreases, as the decrease in cross-correlation among elements of quantization error increases with increase in dither power. The received signal power also reduces with the increase in dither power, leading to a minimum in the BER for a finite non-trivial value of σ. As the reduction in signal power starts to dominate the reduction in quantization noise, BER starts increasing. Dither power corresponding to the minimum value of BER gives the $\sigma^2_{\mathrm{opt}}(N_\mathrm{t},U,\rho)$ for the system model considered.

In Fig. 6, we plot coded BER as a function of transmit power ρ, for 16-QAM symbol modulation, with and without optimized dithering. The transmitted symbols are coded using a rate-$1/2$ LDPC code, for number of BS antennas, $N_\mathrm{t} = 100$, and number of users, $U = 2$. Simulations are performed over 1000 IID Rayleigh channel instances with 1000 bits transmitted in each channel instance. The quantization noise

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

component of the SQINR increases with transmit power ρ, as shown in (74) and (81). Dithering improves SQINR by optimally reducing quantization error, at the receiver. It, therefore, offers small improvement for low transmit power at the BS. For medium to high transmit power it offers significant improvement in the receive SQINR. The error floor experienced at high ρ, which is a major limitation of quantized linear precoded systems [11], is pushed down significantly when optimized dithering is used.

In Fig. 7, we plot the optimal dither power σ^2_{opt} as a function of transmit power, ρ, for different values of N_{t} and U. We can see that σ^2_{opt} increases monotonically with ρ. As mentioned in Section V, the signal power in SQINR decreases with σ^2. Quantization noise power in SQINR decreases as σ^2 is increased from 0, saturating for high σ^2. Signal power and quantization noise power are linear with respect to transmit power ρ in the SQINR expression given in (74), and (81). Optimal dither power, which maximizes the SQINR, thus, strictly increases with ρ for low values of ρ. For high values of ρ, AWGN becomes small and SQINR, as given in (74), and (81), does not depend on ρ. Optimal dither power, thus, saturates at high values of ρ.

In Fig. 8, we plot the variation of σ^2_{opt} with number of BS antennas, N_{t}, for $U = 1$ and transmit power $\rho = 0$ dB. We can observe that σ^2_{opt} strictly decreases with rise in number of BS antennas, for both cases of CE and 1-bit quantized transmissions. Due to ZF precoding, the power of the precoded symbols asymptotically falls quadratically with N_{t} as given in (39). Therefore, the relative power of dither with respect to power of the precoded symbols in the quantizer input increases quadratically with N_{t}. The rate of decrease in signal power component of SQINR as a function of σ^2, thus, increases with N_{t}. A faster decrease in signal power component decreases the σ^2 needed to maximize the SQINR, leading to this variation of σ^2_{opt} with N_{t}.

VII. Conclusion

In this work, we analyzed the per-user receive SQINR of linear precoding for downlink MU-massive MIMO with CE transmission and 1-bit quantized transmission assumed at the BS antennas. ZF precoding is assumed at the BS. We used the Bussgang decomposition up to the third-order approximation to model the quantization operation. For a small number of users, we observed that as U is decreased, receive SQINR at the users deteriorated rapidly, due to high correlation among the quantization error components across BS antennas, for both CE and 1-bit quantized transmissions. We inferred that the deterioration, and cross-correlation among elements of quantization error for small U was depicted by the third-order term of the channel response to quantization error at the receivers, as given in (42), and (48) for CE and 1-bit quantized transmissions. To improve the performance for small number of users, we propose the addition of correlated Gaussian dither to the precoded signal before quantization, to improve per user SQINR. Using the Bussgang decomposition, we estimate closed form expression for per-user received SQINR as a function of dither power, number of BS antennas, number of users and total transmit power. We observed that dither directly impacted the third-order term of the channel response to quantization error at the receivers as given in (73) and (80), thereby modifying the cross-correlation among elements of quantization error, along with affecting the signal power component of the SQINR. SQINR achieved maximum for finite non-trivial value of dither power. We estimate the optimum dither power by maximizing the SQINR for a given number of BS antennas, number of users and transmit power.

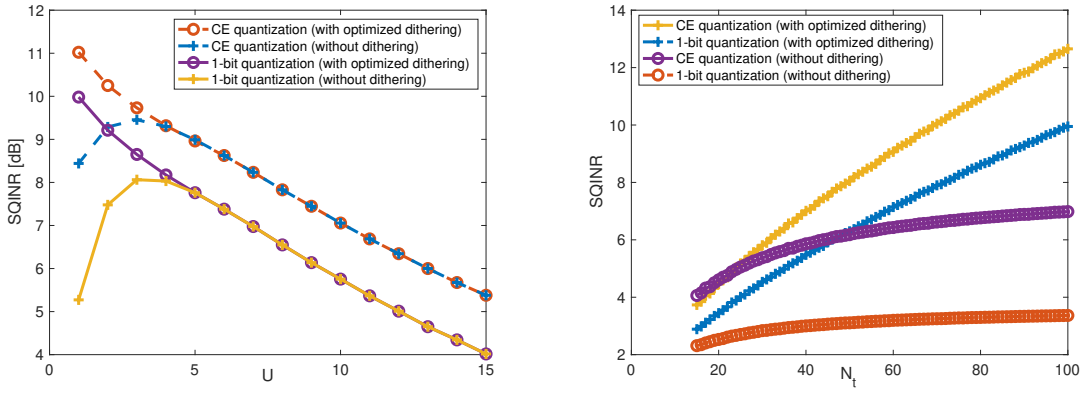
Addition of optimized dither at the transmitter dramatically improves the SQINR for small values of U, which shows monotonic increase with decreasing values of U. The downlink system is also able to show a linear increase in SQINR with N_{t}, given that other parameters are kept constant, when optimized dithering is implemented. We also observe that with optimized dithering, the error floor observed for BER at high BS transmit power, which is a significant limitation of linear precoded CE and 1-bit quantized transmissions, is pushed down significantly. Optimum dither also decreases monotonically with increase in number of BS antennas, as the rate of decrease in signal power component of SQINR as a function of σ^2, increases with N_{t}. Addition of dither, can, therefore, help improve downlink performance for small number of users without any significant increase in computational complexity, or hardware complexity. For a given BS, optimal Gaussian dither can be computed by the transmitter based on the system parameters of the downlink communication, and applied for per-user receive performance enhancement. Similar research can be undertaken to explore the effect on downlink communication with low resolution quantization at the transmitter when dither other than Gaussian dither, like triangular dither and exponential dither, are added.

Appendix A

We calculate the asymptotic third-order harmonics for channel response of quantization noise with 1-bit quantized transmission at the BS, given at (48), as follows. The derivation follows similar reasoning to the derivation in [43]. With (84) and (85) holding for large U and, approximating the matrix $\mathbf{P}^{\circ 2}_{\text{nd}}$ by its mean in (86), we have

$$\begin{aligned}
&\mathbf{h}_1^* \mathbf{P}^{\circ 3}_{\text{nd}} \mathbf{h}_1 \\
&\xrightarrow[|\frac{N_{\text{t}}}{U}|<\infty]{N_{\text{t}}\to\infty} \frac{\text{tr}(\mathbf{H}\,\mathbf{P}^{\circ 3}_{\text{nd}}\mathbf{H}^*)}{U} && (83)\\
&\approx \frac{3}{U}\sum_k \text{tr}\big(\mathbf{H}\big(\text{nondiag}(\mathbf{p}_k\mathbf{p}_k^*) \\
&\quad \circ \text{nondiag}\big(\sum_{k'\neq k}\mathbf{p}_{k'}\mathbf{p}_{k'}^*\big)^{\circ 2}\big)\mathbf{H}^*\big) && (84)\\
&\approx \frac{3}{U}\sum_k \text{tr}\big(\mathbf{H}\big(\text{nondiag}(\mathbf{p}_k\mathbf{p}_k^*)\circ \mathbf{P}^{\circ 2}_{\text{nd}}\big)\mathbf{H}^*\big) && (85)\\
&\approx \frac{3}{U}\sum_k \text{tr}\big(\mathbf{H}\big(\text{nondiag}(\mathbf{p}_k\mathbf{p}_k^*)\circ \mathbb{E}[\mathbf{P}^{\circ 2}_{\text{nd}}]\big)\mathbf{H}^*\big). && (86)
\end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing



(a) SQINR as a function of U, for $N_{\mathrm{t}} = 100$ (in dB).

(b) SQINR as a function of N_{t}, for $U = 1$.

Fig. 4: Plot of SQINR, for different values of N_{t} and U, for CE and 1-bit quantized transmission at the BS antennas. With optimized dithering, the SQINR strictly increases with decreasing value of U, for constant N_{t}. It also increases linearly with N_{t}, for constant U, capturing the advantage of array gain provided by downlink massive MIMO communication. A part of plot (a), for 1-bit quantized transmission, has been published as a part of [2].

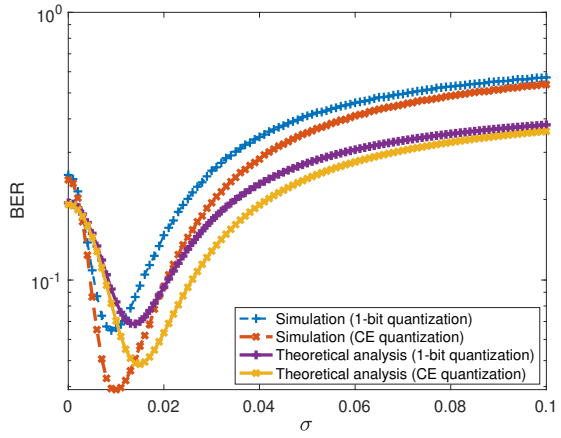


Fig. 5: BER plot for $N_t = 100$ and $U = 1$ for varying dither power, σ^2. Minimum BER is achieved at non-trivial finite optimal dither power, σ^2_{opt}, beyond which the performance starts deteriorating. The gap observed between experimental and theoretical plots is a result of the usage of 16-QAM constellation instead of Gaussian signaling.

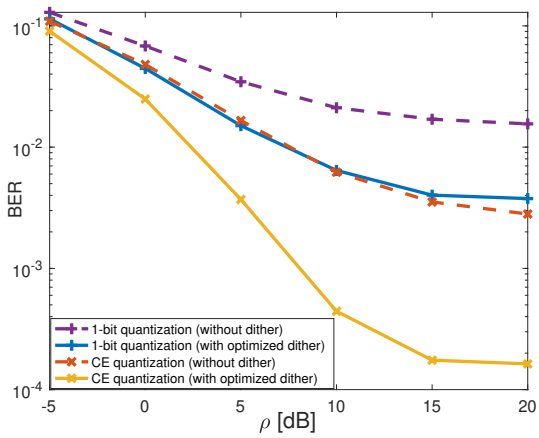


Fig. 6: Comparison of the error performance with varying transmit power ρ, for $N_{\mathrm{t}} = 100$ and $U = 2$ for 16-QAM transmission, with and without the addition of optimized dither. A rate-1/2 LDPC code is used for coding the transmitted symbols. The improvement in BER performance increases with BS transmit power, and error floor experienced by the BER for high BS transmit power is lowered significantly, due to optimized dithering.

Using the singular value decomposition $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$, and the fact that these matrices are independent if $\mathbf{H}$ is IID Gaussian, we have,

$$\mathbb{E}[\mathbf{P}_{\mathrm{nd}}^{\circ 2}] = \mathbb{E}[\mathrm{nondiag}(\mathbf{H}^* \left(\mathbf{H}\mathbf{H}^*\right)^{-1} \mathbf{H})^{\circ 2}] \tag{87}$$

$$= \mathbb{E}[\mathrm{nondiag}(\mathbf{V}\mathbf{\Sigma}^{-2}\mathbf{V}^*)^{\circ 2}]. \tag{88}$$

Then, we approximate the distribution of the off-diagonal elements of $\mathbf{V}\mathbf{\Sigma}^{-2}\mathbf{V}^*$ conditioned on $\mathbf{\Sigma}$ by a complex Gaussian distribution with standard deviation $\mathrm{tr}(\mathbf{\Sigma}^{-2})/\sqrt{U}/N_{\mathrm{t}}$, which asymptotically holds for large number of users. Therefore, using the fact that $\mathrm{tr}(\mathbf{\Sigma}^{-2}) \to U/(N_{\mathrm{t}} - U)$ using (39), we obtain

$$\mathbb{E}[\mathbf{P}_{\mathrm{nd}}{}^{\circ 2}] \approx \frac{U(1+\mathrm{j})}{2(N_{\mathrm{t}} - U)^2 N_{\mathrm{t}}{}^2} \mathbf{1}\mathbf{1}^T. \tag{89}$$

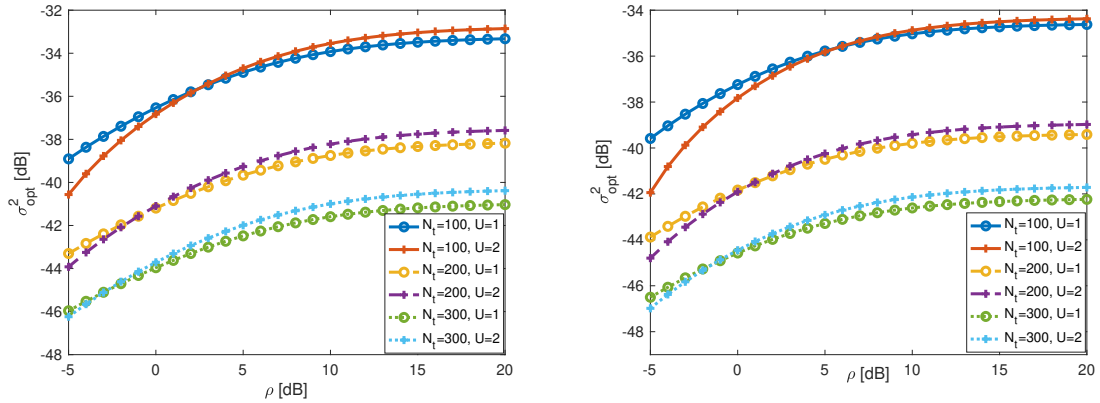
As a result, (86) becomes

$$\mathrm{tr}(\mathbf{H}\,\mathbf{P}_{\mathrm{nd}}{}^{\circ 3}\mathbf{H}^*)/U$$

$$\approx \frac{3}{2(N_{\mathrm{t}} - U)^2 N_{\mathrm{t}}{}^2} \mathrm{tr}(\mathbf{H}\,\mathbf{P}_{\mathrm{nd}}\mathbf{H}^*) \tag{90}$$

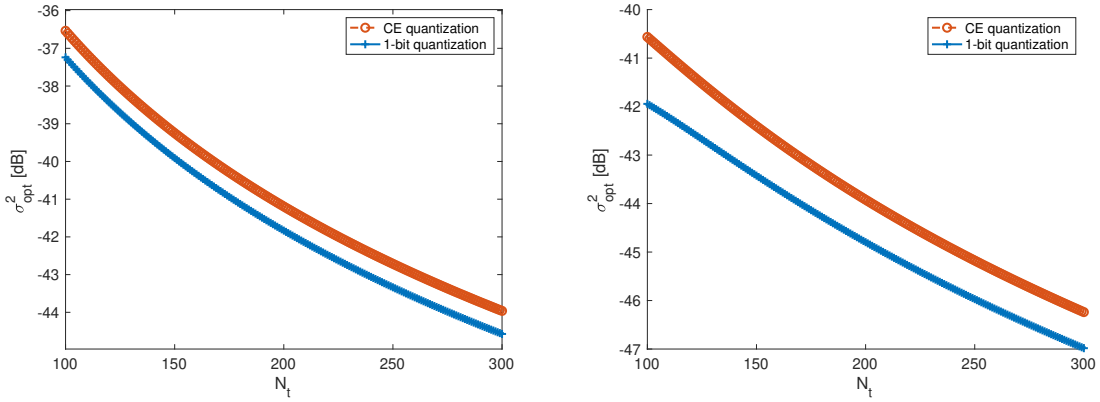
$$= \frac{3}{2(N_{\mathrm{t}} - U)^2 N_{\mathrm{t}}{}^2} (U - \mathrm{tr}(\mathbf{H}\,\mathbf{P}_{\mathrm{d}}\mathbf{H}^*)) \tag{91}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing



(a) For CE transmission (b) For 1-bit quantized transmission

Fig. 7: Plot of optimal dither power, σ^2_{opt} as a function of transmit power ρ, for different values of N_{t} and U, for CE and 1-bit quantized transmission at the BS antennas. The optimal dither power strictly increases for low transmit power, but saturates to a constant value at high transmit power. A variation of plot (b) has been published as a part of [2].



(a) For $\rho = 0$ dB, U $= 1$ (b) For $\rho = -5$ dB, U $= 2$

Fig. 8: Plot of optimal dither power, σ^2_{opt} as a function of number of BS antennas N_{t}. Optimal dither power decreases monotonically with N_{t}. Since the rate of decrease of the signal power component of SQINR as a function of dither power increases with N_{t}, the SQINR attains maximum at a lower dither power for larger N_{t}.

$$\overset{(39)}{\rightarrow} \frac{3U}{2(N_{\text{t}}-U)^2{N_{\text{t}}}^2}\left(1-\frac{U}{N_{\text{t}}-U}\right) \tag{92}$$

$$= \frac{3U(N_{\text{t}}-2U)}{2(N_{\text{t}}-U)^3{N_{\text{t}}}^2}. \tag{93}$$

We have used this calculation in our previous paper to calculate the channel response to the quantization noise at the receivers for downlink 1-bit quantized transmission [1].

Appendix B

We calculate the asymptotic third-order harmonics for channel response of quantization noise with CE transmission at the BS, given at (40), as follows. Using (86) and (89) we have,

$$\mathbf{h}_1^*\left(\mathbb{I}(\mathbf{P}\mathbf{P}^*)^{\circ 2}\circ\mathbb{R}(\mathbf{P}\mathbf{P}^*)+\mathrm{j}\cdot\mathbb{R}(\mathbf{P}\mathbf{P}^*)^{\circ 2}\circ\mathbb{I}(\mathbf{P}\mathbf{P}^*)\right)\mathbf{h}_1$$

$$\xrightarrow[|\frac{N_{\text{t}}}{U}|<\infty]{N_{\text{t}}\to\infty} \frac{\mathrm{tr}\mathbf{H}(\mathbb{I}(\mathbf{P}\mathbf{P}^*)^{\circ 2}\circ\mathbb{R}(\mathbf{P}\mathbf{P}^*)}{U}$$

$$+\mathrm{j}\frac{\mathbb{R}(\mathbf{P}\mathbf{P}^*)^{\circ 2}\circ\mathbb{I}(\mathbf{P}\mathbf{P}^*))\mathbf{H}^*}{U}$$

$$\approx \frac{1}{U}\sum_k \mathrm{tr}(\mathbf{H}\big(\mathrm{nondiag}(\mathbf{p}_k\mathbf{p}_k^*)\circ\mathbb{E}[\mathbf{P}_{\text{nd}}^{\circ 2}]\big)\mathbf{H}^*) \tag{94}$$

$$\approx \frac{U(N_{\text{t}}-2U)}{2(N_{\text{t}}-U)^3{N_{\text{t}}}^2}. \tag{95}$$

Appendix C

We calculate the asymptotic third-order harmonics for channel response of quantization noise with correlated dither added before quantization, and 1-bit quantized transmission at the BS, given at (78), as follows. Using the singular value decomposition $\mathbf{H} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^*$, we have

$$\mathbb{E}[(\mathbf{P}_{\text{nd}}+\mathrm{nondiag}(\sigma^2\mathbf{R}_{\mathbf{H}}))^{\circ 2}]$$

$$= \mathbb{E}[(\mathbf{P}_{\text{nd}}-\mathrm{nondiag}(\sigma^2\mathbf{H}^*(\mathbf{H}\mathbf{H}^*)^{-1}\mathbf{H}))^{\circ 2}] \tag{96}$$

$$= \mathbb{E}[\mathrm{nondiag}(\mathbf{V}(\mathbf{\Sigma}^{-2}-\sigma^2\mathbf{I})\mathbf{V}^*)^{\circ 2}]. \tag{97}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/OJSP.2020.3040590, IEEE Open Journal of Signal Processing

Using the fact that asymptotically $\mathrm{tr}(\mathbf{\Sigma}^{-2} - \sigma^2\mathbf{I}) \rightarrow \left(\frac{U}{N_\mathrm{t}-U} - \sigma^2 U\right)$ using (39), with computation similar to (89) we have

$$\mathbb{E}[(\mathbf{P}_{\text{nd}} + \text{nondiag}(\sigma^2\mathbf{R}_{\mathbf{H}}))^{\circ 2}] \approx \frac{U(1-\sigma^2(N_\mathrm{t}-U))^2(1+\mathrm{j})}{2(N_\mathrm{t}-U)^2{N_\mathrm{t}}^2}\mathbf{1}\mathbf{1}^T. \tag{98}$$

Since $\mathbf{P}\mathbf{P}^*$ and $\mathbf{R}_{\mathbf{H}}$ are positive semi-definite matrices, $\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}}$ is positive semi-definite and can be expressed as product of a matrix and its complex-transpose, $\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}} = \tilde{\mathbf{P}}\tilde{\mathbf{P}}^*$. Assuming $\tilde{\mathbf{p}}_k$ as k^{th} column of $\tilde{\mathbf{P}}$, using computation similar to (86), and result from (98) we get,

$$\mathbf{h}_1^*(\mathbf{P}_{\text{nd}} + \text{nondiag}(\sigma^2\mathbf{R}_{\mathbf{H}}))^{\circ 3}\mathbf{h}_1 \xrightarrow[|\frac{N_\mathrm{t}}{U}|<\infty]{N_\mathrm{t}\rightarrow\infty} \mathrm{tr}(\mathbf{H}\,(\mathbf{P}_{\text{nd}} + \text{nondiag}(\sigma^2\mathbf{R}_{\mathbf{H}}))^{\circ 3}\mathbf{H}^*)/\mathrm{U} \tag{99}$$

$$\approx \frac{3}{U}\sum_k \mathrm{tr}(\mathbf{H}\big(\text{nondiag}(\tilde{\mathbf{p}}_k\tilde{\mathbf{p}}_k^*) \circ \mathbb{E}[(\mathbf{P}_{\text{nd}} + \text{nondiag}(\sigma^2\mathbf{R}_{\mathbf{H}}))^{\circ 2}]\big)\mathbf{H}^*) \tag{100}$$

$$\approx \frac{3(1-\sigma^2(N_\mathrm{t}-U))^2}{2(N_\mathrm{t}-U)^2{N_\mathrm{t}}^2}\mathrm{tr}(\mathbf{H}(\mathbf{P}_{\text{nd}} + \text{nondiag}(\sigma^2\mathbf{R}_{\mathbf{H}}))\mathbf{H}^*) \overset{(39)}{\rightarrow} \frac{3U(1-\sigma^2(N_\mathrm{t}-U))^2(N_\mathrm{t}-2U-\sigma^2(N_\mathrm{t}-U)^2)}{2(N_\mathrm{t}-U)^3{N_\mathrm{t}}^2}. \tag{101}$$

We have derived this result as a part of our work in [2].

APPENDIX D

We calculate the asymptotic third-order harmonics for channel response of quantization noise with correlated dither added before quantization, and CE transmission at the BS, given at (71), as follows. Using results from (94), (98) and (100) we have,

$$\mathbf{h}_1^*\left(\mathbb{I}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}})^{\circ 2} \circ \mathbb{R}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}}) + \mathrm{j}\cdot\mathbb{R}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}})^{\circ 2} \circ \mathbb{I}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}})\right)\mathbf{h}_1 \xrightarrow[|\frac{N_\mathrm{t}}{U}|<\infty]{N_\mathrm{t}\rightarrow\infty} \frac{\mathrm{tr}(\mathbf{H}\,\mathbb{I}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}})^{\circ 2} \circ \mathbb{R}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}})\mathbf{H}^*)}{U} + \mathrm{j}\frac{\mathrm{tr}(\mathbf{H}\,\mathbb{R}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}})^{\circ 2} \circ \mathbb{I}(\mathbf{P}\mathbf{P}^* + \sigma^2\mathbf{R}_{\mathbf{H}})\mathbf{H}^*)}{U}$$

$$\approx \frac{1}{U}\sum_k \mathrm{tr}(\mathbf{H}\big(\text{nondiag}(\tilde{\mathbf{p}}_k\tilde{\mathbf{p}}_k^*) \circ \mathbb{E}[(\mathbf{P}_{\text{nd}} + \text{nondiag}(\sigma^2\mathbf{R}_{\mathbf{H}}))^{\circ 2}]\big)\mathbf{H}^*) \tag{102}$$

$$\approx \frac{(1-\sigma^2(N_\mathrm{t}-U))^2}{2(N_\mathrm{t}-U)^2{N_\mathrm{t}}^2}\mathrm{tr}(\mathbf{H}(\mathbf{P}_{\text{nd}} + \text{nondiag}(\sigma^2\mathbf{R}_{\mathbf{H}}))\mathbf{H}^*) \tag{103}$$

$$\overset{(39)}{\rightarrow} \frac{U(1-\sigma^2(N_\mathrm{t}-U))^2(N_\mathrm{t}-2U-\sigma^2(N_\mathrm{t}-U)^2)}{2(N_\mathrm{t}-U)^3{N_\mathrm{t}}^2}. \tag{104}$$

REFERENCES

[1] A. K. Saxena, A. Mezghani, R. Bendlin, R. W. Heath, J. G. Andrews, N. Sairamesh, and A. Chopra, "Asymptotic performance of downlink massive MIMO with 1-bit quantized zero-forcing precoding," in *Proceedings of the 19th International Workshop on Signal Processing Advances in Wireless Communications*, July 2019, pp. 1–5.

[2] A. K. Saxena, A. Mezghani, R. W. Heath, and J. G. Andrews, "Linear transmit precoding with optimized dithering," in *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, 2019, pp. 838–842.

[3] 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, February 2014.

[4] 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, Oct 2014.

[5] 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, no. 1, pp. 40–60, Jan 2013.

[6] A. Goldsmith, "Adaptive modulation and coding for fading channels," in *Proceedings of the IEEE Information Theory and Communications Workshop*, June 1999, pp. 24–26.

[7] T. Lee and H. Ochiai, "A new trellis shaping design for peak power reduction of SC-FDMA signals with high-order QAM," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 6, pp. 5030–5042, June 2017.

[8] R. N. Braithwaite, "Fixed point considerations for digital predistortion of a RF power amplifier using recursive least square (RLS) estimation," in *Proceedings of IEEE Topical Conference on RF/Microwave Power Amplifiers for Radio and Wireless Applications*, Jan 2019, pp. 1–3.

[9] P. Jaraut, M. Rawat, and F. M. Ghannouchi, "Harmonically related concurrent tri-band behavioral modeling and digital predistortion," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 66, no. 6, pp. 1073–1077, June 2019.

[10] Y. Beltagy, P. Mitran, and S. Boumaiza, "Direct learning algorithm for digital predistortion training using sub-nyquist intermediate frequency feedback signal," *IEEE Transactions on Microwave Theory and Techniques*, vol. 67, no. 1, pp. 267–277, Jan 2019.

[11] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, "Analysis of one-bit quantized precoding for the multiuser massive MIMO downlink," *IEEE Transactions on Signal Processing*, vol. 65, no. 17, pp. 4624–4634, Sept 2017.

[12] S. K. Mohammed and E. G. Larsson, "Single-user beamforming in large-scale MISO systems with per-antenna constant-envelope constraints: The doughnut channel," *IEEE Transactions on Wireless Communications*, vol. 11, no. 11, pp. 3992–4005, November 2012.

[13] ——, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Transactions on Communications*, vol. 61, no. 3, pp. 1059–1071, March 2013.

[14] J. Chen, C. Wen, and K. Wong, "Improved constant envelope multiuser precoding for massive MIMO systems," *IEEE Communications Letters*, vol. 18, no. 8, pp. 1311–1314, Aug 2014.

[15] P. V. Amadori and C. Masouros, "Constant envelope precoding by interference exploitation in phase shift keying-modulated multiuser transmission," *IEEE Transactions on Wireless Communications*, vol. 16, no. 1, pp. 538–550, Jan 2017.

[16] M. Shao, Q. Li, W. Ma, and A. M. So, "Minimum symbol error rate-based constant envelope precoding for multiuser massive MISO downlink," in *Proceedings of the IEEE Statistical Signal Processing Workshop*, June 2018, pp. 727–731.

[17] ——, "A framework for one-bit and constant-envelope precoding over multiuser massive MISO channels," *IEEE Transactions on Signal Processing*, vol. 67, no. 20, pp. 5309–5324, Oct 2019.

[18] M. Kazemi, H. Aghaeinia, and T. M. Duman, "Discrete-phase constant envelope precoding for massive MIMO systems," *IEEE Transactions on Communications*, vol. 65, no. 5, pp. 2011–2021, May 2017.

[19] H. Jedda, A. Mezghani, A. L. Swindlehurst, and J. A. Nossek, "Quantized constant envelope precoding with PSK and QAM signaling," *IEEE Transactions on Wireless Communications*, vol. 17, no. 12, pp. 8022–8034, Dec 2018.

[20] S. Jacobsson, O. Castañeda, C. Jeon, G. Durisi, and C. Studer, "Nonlinear precoding for phase-quantized constant-envelope massive MU-MIMO-OFDM," in *Proceedings of the International Conference on Telecommunications*, June 2018, pp. 367–372.

[21] M. Shao, Q. Li, and W. Ma, "Discrete constant envelope transceiver design for multiuser massive MIMO downlink," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, May 2019, pp. 4684–4688.

[22] J. Chen, "Efficient constant envelope precoding with quantized phases for massive MU-MIMO downlink systems," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 4, pp. 4059–4063, April 2019.

[23] H. Jedda, J. A. Nossek, and A. Mezghani, "Minimum BER precoding in 1-bit massive MIMO systems," in *Proc. of the IEEE SAM*, July 2016, pp. 1–5.

[24] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized precoding for massive MU-MIMO," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4670–4684, Nov 2017.

[25] O. Castañeda, S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "1-bit massive MU-MIMO precoding in VLSI," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 7, no. 4, pp. 508–522, Dec 2017.

[26] J. Chen, "Alternating minimization algorithms for one-bit precoding in massive multiuser MIMO systems," *IEEE Transactions on Vehicular Technology*, vol. 67, no. 8, pp. 7394–7406, Aug 2018.

[27] Y. Yapici, S. J. Maeng, I. Guvenc, H. Dai, and A. Bhuyan, "SLNR based precoding for one-bit quantized massive MIMO in mmwave communications," in *Proceedings of the IEEE International Conference on Communications Workshops*, May 2019, pp. 1–6.

[28] A. Balatsoukas-Stimming, O. Castañeda, S. Jacobsson, G. Durisi, and C. Studer, "Neural-network optimized 1-bit precoding for massive MU-MIMO," in *Proceedings of the International Workshop on Signal Processing Advances in Wireless Communications*, July 2019, pp. 1–5.

[29] A. Li, C. Masouros, F. Liu, and A. L. Swindlehurst, "Massive MIMO 1-bit DAC transmission: A low-complexity symbol scaling approach," *IEEE Transactions on Wireless Communications*, vol. 17, no. 11, pp. 7559–7575, Nov 2018.

[30] A. Li, F. Liu, C. Masouros, Y. Li, and B. Vucetic, "Interference exploitation 1-bit massive MIMO precoding: A partial branch-and-bound solution with near-optimal performance," *IEEE Transactions on Wireless Communications*, vol. 19, no. 5, pp. 3474–3489, 2020.

[31] H. Jedda, A. Mezghani, J. A. Nossek, and A. L. Swindlehurst, "Massive MIMO downlink 1-bit precoding with linear programming for PSK signaling," in *Proceedings of the International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2017, pp. 1–5.

[32] A. Li, C. Masouros, and A. L. Swindlehurst, "1-bit massive MIMO downlink based on constructive interference," in *Proceedings of the 26th European Signal Processing Conference (EUSIPCO)*, 2018, pp. 927–931.

[33] F. Wiffen, M. Z. Bocus, A. Doufexi, and A. Nix, "Phase-only OFDM communication for downlink massive MIMO systems," in *Proceedings of the IEEE Vehicular Technology Conference*, June 2018, pp. 1–5.

[34] J. Xu, W. Xu, F. Gong, H. Zhang, and X. You, "Optimal multiuser loading in quantized massive MIMO under spatially correlated channels," *IEEE Transactions on Vehicular Technology*, vol. 68, no. 2, pp. 1459–1471, Feb 2019.

[35] A. Papoulis and S. U. Pillai, *Probability, Random Variables, and Stochastic Processes*, 4th ed. McGraw Hill, 2002.

[36] L. Schuchman, "Dither signals and their effect on quantization noise," *IEEE Transactions on Communication Technology*, vol. 12, pp. 162 – 165, 01 1965.

[37] P. Carbone and D. Petri, "Performance of stochastic and deterministic sed quantizers," *IEEE Transactions on Instrumentation and Measurement*, vol. 49, no. 2, pp. 337–340, April 2000.

[38] U. Gustavsson, C. Sanchéz-Perez, T. Eriksson, F. Athley, G. Durisi, P. Landin, K. Hausmair, C. Fager, and L. Svensson, "On the impact of hardware impairments on massive MIMO," in *Proceedings of the IEEE Globecom Workshops*, Dec 2014, pp. 294–300.

[39] J. J. Bussgang, *Crosscorrelation functions of amplitude-distorted Gaussian signals*. Research Laboratory of Electronics, Massachusetts Institute of Technology, 1952.

[40] R. W. Heath Jr. and A. Lozano, *Foundations of MIMO Communication*. Cambridge University Press, 2018.

[41] R. J. Muirhead, *Aspects of Multivariate Statistical Theory*. Wiley-Interscience, 2005.

[42] 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, no. 2, pp. 160–171, 2013.

[43] A. Mezghani and A. L. Swindlehurst, "mmWave massive MIMO with simple RF and appropriate DSP," in *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, Oct 2017, pp. 277–284.