Title

Quantized Constant Envelope Precoding With PSK and QAM Signaling

Journal

IEEE Transactions on Wireless Communications, 17(12)

ISSN

1536-1276

Authors

Jedda, Hela
Mezghani, Amine
Swindlehurst, A Lee
et al.

Publication Date

2018-12-01

DOI

10.1109/twc.2018.2873386

Quantized Constant Envelope Precoding with PSK and QAM Signaling

Hela Jedda[1], Amine Mezghani[2], A. Lee Swindlehurst[3], and Josef A. Nossek[1,4]

[1]Associate Professorship of Signal Processing, Technical University of Munich, 80290 Munich, Germany
[2]Wireless Networking and Communications Group, University of Texas at Austin, Austin, TX 78712, USA
[3]Center for Pervasive Communications and Computing, University of California, Irvine, Irvine, CA 92697, USA
[4]Department of Teleinformatics Engineering, Federal University of Ceará, Fortaleza, Brazil
Email: {hela.jedda, josef.a.nossek}@tum.de, amine.mezghani@utexas.edu, swindle@uci.edu

***Abstract*—Coarsely quantized massive Multiple-Input Multiple-Output (MIMO) systems are gaining more interest due to their power efficiency. We present a new precoding technique to mitigate the Multi-User Interference (MUI) and the quantization distortions in a downlink Multi-User (MU) MIMO system with coarsely Quantized Constant Envelope (QCE) signals at the transmitter. The transmit signal vector is optimized for every desired received vector taking into account the QCE constraint. The optimization is based on maximizing the safety margin to the decision thresholds of the receiver constellation modulation. Simulation results show a significant gain in terms of the uncoded Bit Error Ratio (BER) compared to the existing linear precoding techniques.**

***Index Terms*—Constant Envelope, Coarse Quantization, Constructive Interference, Downlink Massive Multi-User MIMO, Precoding.**

I. Introduction

THE next generation of mobile communication aims at increasing 1000-fold the network capacity, 10-100-fold the number of connected devices and decreasing 5-fold the latency time and the power consumption compared to 4G networks [1]. To achieve these challenging requirements, the following technologies are the subject of current research:

- massive Multiple-Input Multiple-Output (MIMO) systems, where the base stations (BSs) are equipped with a very large number of antennas (100 or more) that can simultaneously serve many users [2]–[6],
- millimeter-Wave (mmW) communication, i.e. frequencies ranging between 30 GHz and 300 GHz, where the spectrum is less crowded and greater bandwidth is available [7]–[9] and
- smaller cells with ranges on the order of 10-200 m, i.e. pico- and femtocells.

First, massive MIMO systems lead to a drastic increase in the number of Radio Frequency (RF) chains at the BS and hence in the number of the wireless front-end hardware components. Second, mmW communication implies that the wireless front-end hardware components are operated at much higher frequencies. Third, reducing the cell size means that the number of cells per unit area is increased and thus results in a much more dense wireless network. Combining the three technologies means a dramatic increase in the number of RF hardware elements operating at very high frequencies per unit area. Hence, the RF power consumption per unit area alarmingly increases. While the above technologies are foreseen as key technologies for future communication systems, the increase in power consumption represents a crucial concern.

The term green communication refers to the idea of minimizing power consumption while guaranteeing a certain quality of service. The most critical front-end elements in terms of power consumption, depending on whether the large number of antennas is situated at the transmitter or at the receiver, are the Analog-to-Digital Converters (ADCs) in the uplink scenario, and mainly the Power Amplifiers (PAs) and secondarily the Digital-to-Analog Converters (DACs) in the downlink scenario, which is the focus of this contribution. The PA is considered as the most power hungry device at the transmitter side [10], [11]. When the PA is run in the saturation region, i.e. the highly non-linear region, high power efficiency is achieved and hence less power is consumed [12]. However, in the saturation region strong non-linear distortions are introduced to the signals. To avoid the PA distortions when run in the saturation region, we resort to PA input signals of Constant Envelope (CE). CE signals have the property of constant magnitude leading to a unit Peak-to-Average-Power Ratio (PAPR). Thus, the information is carried by the signal phase.

To this end, polar (phase-based) DACs at the transmitter are designed to convert the time-discrete and value-discrete base-band signals into time-continuous but value-discrete, i.e. phase-discrete, CE signals. The number of possible discrete phases is determined by the resolution of the DACs. The larger the resolution is, the more accurate the phase information at the DACs' outputs, but the larger their power consumption [13]. To further reduce the hardware power consumption, the DACs' resolution can be reduced. The use of coarsely quantized DACs is also beneficial in terms of reduced cost and circuit area and can further simplify the surrounding RF circuitry due to the relaxed linearity constraint, leading to very efficient hardware implementations. In this way, the power consumption is reduced twofold: power efficient PAs due to the CE signals and less power consuming DACs due to the low resolution. However, this approach leads to non-linear distortions that degrade the system performance and have to

be mitigated by the precoder design in massive Multi-User (MU) MIMO downlink systems.

A. Related Works

The first precoding techniques in the context of CE transmit signals were introduced in [14]–[16]. The idea of CE transmit signals was further exploited for massive MIMO systems in [17]–[20], where the Multi-User Interference (MUI) is minimized subject to the CE constraint, whereas recent works in [21] and [22] exploit the constructive part of the MUI to design the CE precoder. The authors in [23] design a CE precoder to maximize the Signal-to-Leakage-plus-Noise Ratio (SLNR). In the above contributions, the DACs are assumed to have infinite resolution.

The contribution in [24] is an early work that addressed the precoding task with low resolution DACs at the transmitter. A linear Minimum Mean Squared Error (MMSE) precoder is designed, while quantization distortion is taken into account. This precoding design is not in the context of coarsely Quantized Constant Envelope (QCE) signals since the DACs are not polar but cartesian (inphase- and quadrature- based). However, the extreme case of 1-bit DACs in [24] represents a special case of coarsely QCE signals. Many contributions in the literature have studied this special case. They can be categorized in two groups: linear and non-linear precoders. In addition to the linear precoder in [24], we introduced in [25] another linear precoder, where the second-order statistics of the 1-bit DAC signals are computed based on Price's theorem [26]. Non-linear precoding techniques in this context were introduced in [27]–[32]. The non-linear methods can be classified with respect to two design criteria: symbol-wise Minimum Squared Error (MSE) and symbol-wise Maximum Safety Margin (MSM) exploiting the idea of constructive interference. In the context of the symbol-wise MSE, the authors in [28] presented a convex formulation of the problem and applied it to higher-order modulation scheme in [29]. The problem formulation is based on semidefinite relaxation and squared ℓ_∞-norm relaxation. The same optimization problem was solved more efficiently in [31] and [33].

In the context of symbol-wise MSM in [27], we presented a precoding technique based on a minimum-bit-error-ratio criterion and made use of the box norm (ℓ_∞) to relax the 1-bit constraint. Recently, the work in [32] proposed a method to significantly improve linear precoding solutions in conjunction with 1-bit quantization by properly perturbing the linearly precoded signal for each given input signal to favorably impact the probability of correct detection. In [30] the safety margin to the decision thresholds of the received Phase-Shift Keying (PSK) symbols is maximized subject to a relaxed 1-bit constraint using linear programming. The same optimization problem was solved by the Branch-and Bound algorithm in [34] for the particular QPSK case. To the best of our knowledge, the only works that considered the case of coarsely QCE transmit signals are [35], [36] and [37]. In [35], we propose a symbol-wise MSE precoders based on the gradient-descent method under the strict CE constraint or a relaxed polygon constraint. In [36], the authors extend the method in [28] to fit in the context of QCE transmit signals. In [37], the authors use the greedy approach for the precoder design while having the symbol-wise MSE as the design criterion, too. The contribution in [38] addresses the task of QCE precoding in the context of using a single common PA and separate digital phase shifters for the antenna front-ends. The optimization problem consists of designing the QCE precoder while minimizing the MUI, and the idea of constructive interference, [39], [40], is not exploited as in our work. The concept of QCE precoding and general constellations is studied in this contribution. It is worth mentioning that the QCE precoding can be combined with appropriate pulse shaping strategies as in [41], [42] to ensure an efficient spectral confinement. In [43], it was shown that CE precoding is still power efficient even when considering the time processing. The same investigation can be conducted for the case of QCE precoding. Here, we focus rather on the spatial design problem.

B. Main Contributions

The main contributions in this paper are summarized as follows

1) We propose a QCE precoding in the context of massive MIMO systems, where the transmit signals have constant magnitude and phases are drawn from a discrete set. The precoder design exploits the idea of constructive interference and adapts the design criterion in [30] to the coarsely QCE case. The optimization problem is solved using linear programming, which is very advantageous in terms of complexity as it is one of the most widely applied and studied optimization techniques.
2) We extend the proposed QCE precoder to the case of QAM signaling. In particular, this is a novel extension of the constructive interference idea to non-PSK signals.

C. Remainder and Notation

The remainder of this paper is organized as follows. In Section II, we present the system model. In Section III, the motivation behind formulating the precoding problem as a linear programming problem is explained. Sections IV and V present the corresponding optimization problems for PSK and Quadrature Amplitude Modulation (QAM) signals, respectively. The complexity of each optimization problem is discussed in Section VI. Simulation results are introduced in VII. Finally, Section VIII summarizes this work.

Notation: Bold lower case and upper case letters indicate vectors and matrices, non-bold letters express scalars. The operators $(.)^*$, $(.)^{\mathrm{T}}$ and $(.)^{\mathrm{H}}$ stand for complex conjugation, transposition and Hermitian transposition, respectively. The $n \times n$ identity (zero) matrix is denoted by $\mathbf{I}_n$ ($\mathbf{0}_{n,n}$). The n dimensional one (zero) vector is denoted by $\mathbf{1}_n$ ($\mathbf{0}_n$). The vector $\mathbf{e}_m$ represents a zero-vector with 1 at the m-th position. Additionally, $\mathrm{diag}\,(\mathbf{a})$ denotes a diagonal matrix containing the entries of the vector $\mathbf{a}$. Every vector $\mathbf{a}$ of dimension L is defined as $\mathbf{a} = \sum_{\ell=1}^{L} a_\ell \mathbf{e}_l$. The operator $\otimes$ denotes the Kronecker product. The operator $\leq$ in the context of vector inequalities applies element-wise to the vector entries.

II. System Model

The system model shown in Fig.1 consists of a single-cell massive MU-MIMO downlink scenario with coarsely QCE signals at the transmitter. The Base Station (BS) is equipped with N antennas and serves M single-antenna users simultaneously, where $N \gg M$. The input signal vector $\mathbf{s}$ contains the signals to be transmitted to each of the M users. Each user's signal is drawn from the set $\mathbb{S}$ that represents either an S-PSK or S-QAM constellation, where S denotes the number of constellation points. We assume that $\mathrm{E}[\mathbf{s}] = \mathbf{0}_M$ and $\mathrm{E}[\mathbf{s}\mathbf{s}^{\mathrm{H}}] = \sigma_{\mathrm{s}}^2 \mathbf{I}_M$. The signal vector $\mathbf{s}$ is precoded into the vector $\mathbf{x} \in \mathbb{X}^N$ prior to the DACs. The non-linear function $\mathcal{P}(\bullet)$ is a symbol-wise precoder to reduce the distortions caused by the coarse quantization and the MUI. The operator $Q_{\mathrm{CE}}(\bullet)$ models the non-linear behavior of the DACs combined with the power allocation at the PAs as

$$\mathbf{t} = Q_{\mathrm{CE}}(\mathbf{x}) = \sqrt{\frac{P_{\mathrm{tx}}}{N}}\, \mathrm{e}^{\mathrm{j}\, Q_\phi(\arg(\mathbf{x}))}, \tag{1}$$

where the total transmit power P_{tx} is allocated equally among the transmit antennas. The phase quantizer $Q_\phi(\bullet)$ is a symmetric uniform real-valued quantizer. It is characterized by its resolution q that defines the number of discrete output phases

$$Q = 2^q. \tag{2}$$

In other words, the 2π-phase range is divided into Q $\frac{2\pi}{Q}$-rotationally symmetric sectors. The input signal that belongs to the k-th sector is quantized (mapped) to $\mathrm{e}^{\mathrm{j}(2k-1)\frac{\pi}{Q}}$. This can be mathematically expressed as

$$Q_\phi(\arg(x)) = \left(\left\lfloor \frac{\arg(x)}{2\pi/Q} \right\rfloor + \frac{1}{2} \right) \frac{2\pi}{Q}. \tag{3}$$

Thus, the information after the CE quantizer lies only in the phase. Hence, the set $\mathbb{T}$ is defined as

$$\mathbb{T} = \left\{ \sqrt{\frac{P_{\mathrm{tx}}}{N}} \exp\left(\mathrm{j}\,(2i-1)\,\frac{\pi}{Q} \right) : i = 1, \cdots, Q \right\}. \tag{4}$$

The signal $\mathbf{t}$ is transmitted through a flat-fading channel that is modeled by the channel matrix $\mathbf{H}$. We assume that the (m,n)-th element h_{mn} is a zero-mean unit-variance channel tap between the n-th transmit antenna and the m-th user. At the M receive antennas, additive white Gaussian noise (AWGN), which is denoted by the vector $\boldsymbol{\eta} \sim \mathcal{CN}_{\mathbb{C}}\left(\mathbf{0}_M, \mathbf{C}_{\boldsymbol{\eta}} = \mathbf{I}_M\right)$, perturbs the received signals

$$\mathbf{r} = \mathbf{H}\mathbf{t} + \boldsymbol{\eta}. \tag{5}$$

Coherent data transmission with multiple BS antennas leads to an antenna gain, which depends on the channel realization. The entries of the received signal vector $\mathbf{r}$ do not belong to the nominal decision regions of $\mathbb{S}$ but to a scaled version of them. Therefore, rescaling the received signal at each receive antenna is required to make the signal belong to the nominal decision region. The rescaling operation is modeled by the diagonal real-valued matrix $\mathbf{G}$, as follows

$$\mathbf{u} = \mathbf{G}\,(\mathbf{H}\mathbf{t} + \boldsymbol{\eta}), \tag{6}$$

where

$$\mathbf{G} = \sum_{m=1}^{M} g_m \mathbf{e}_m \mathbf{e}_m^{\mathrm{T}}, \tag{7}$$

with $g_m > 0,\ m = 0, \cdots, M$. Note that no receive processing $\mathbf{G}$ is required if $\mathbb{S}$ represents the PSK constellation. Finally, based on the decision regions to which the entries of the signal $\mathbf{u}$ belong, the decision operation $\mathcal{D}(\bullet)$ produces the detected symbols $\hat{\mathbf{s}}$ at the users

$$\hat{\mathbf{s}} = \mathcal{D}\left(\mathbf{G}\left(\mathbf{H}\mathbf{t} + \boldsymbol{\eta}\right)\right). \tag{8}$$

III. Precoding Task

In this work, we make use of the idea of constructive interference optimization [39], [40]. When the downlink channel and all users' data are known at the transmitter, instantaneous constructive MUI can be exploited to move the received signals further from the decision thresholds [40]. In contrast to this, conventional precoding methods (MMSE, Zero-forcing) aim at minimizing the total MUI such that the received signals lie as close as possible to the nominal constellation points. Constructive interference optimization exploits the larger symbol decision regions and thus leads to a more relaxed optimization.

For every given input signal $\mathbf{s}$ and for each channel realization $\mathbf{H}$, the precoding task is to find

$$\mathbf{x} = \mathcal{P}\left(\mathbf{s}, \mathbf{H}\right). \tag{9}$$

The task consists in designing the transmit vector $\mathbf{x}$ such that $\hat{\mathbf{s}} = \mathbf{s}$ holds true with high probability to reduce the detection error probability. The symbol-wise precoder aims to mitigate all sources of distortion

- the quantization distortions
- the channel distortions, and
- the additive white Gaussian noise.

Our goal is to develop a problem formulation that jointly minimizes all three distortion sources.

First, it is obvious that the quantization distortions can be omitted if we design the quantizer input such that it belongs to $\mathbb{T}^N$, i.e. $\mathbb{X} = \mathbb{T}$. Consequently, we would get an undistorted signal

$$\mathbf{t} = \mathbf{x}, \text{ if } \mathbf{x} \in \mathbb{T}^N. \tag{10}$$

In what follows we enforce the QCE constraint in (10) to ensure the non-distorting behavior of the quantizer $Q_{\mathrm{CE}}(\bullet)$.

Second, to minimize the channel distortions and the noise, we look deeper at the constellation properties. As illustrated in Fig. 2 and Fig. 3, each constellation is defined by the decision thresholds that separate the distinct decision regions of the constellation points. In total, we have as many contiguous decision regions as constellation points. Each constellation symbol lies within a Symbol Region (SR) that is a downscaled version of the decision region. In contrast to the decision region, the SR has a safety margin denoted by δ that separates it from the decision thresholds. When each entry of the noiseless received signal vector $\mathbf{y}$ belongs to the correct SR and thus the correct decision region, the channel distortions

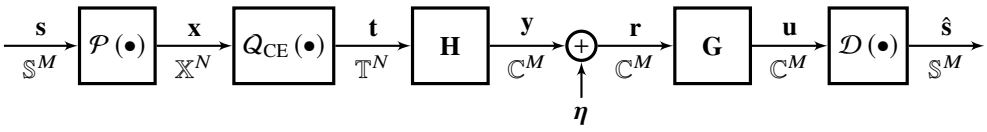

Figure 1. Downlink MU-MIMO system model.

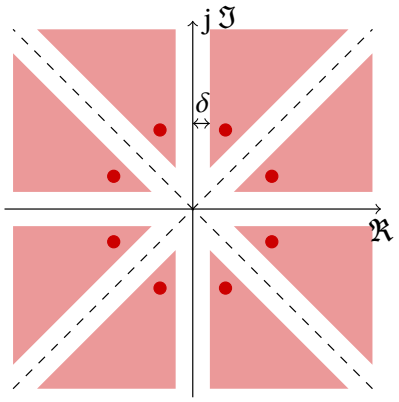

Figure 2. Decision and symbol regions (in red) for 8PSK symbols.

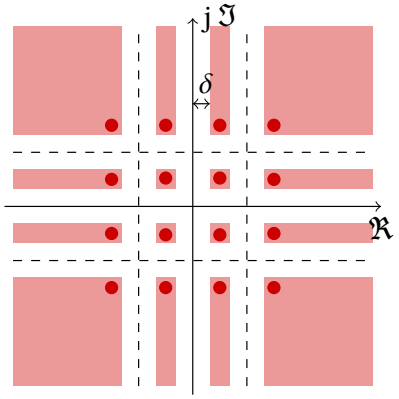

Figure 3. Decision and symbol regions (in red) for 16QAM symbols.

are mitigated. Additionally, the safety margin δ has to be large enough such that the received signals when perturbed by the additive noise do not jump to the neighboring unintended decision regions.

In summary, the problem formulation has to take into account the QCE constraint, the SR for each received signal and maximizing the safety margin. Thus, the optimization problem for the symbol-wise precoder can be written in general as follows

$$\max_{\mathbf{x}} \delta \tag{11}$$

$$\text{s.t. } y_m \in \text{SR}_m, \forall m \tag{12}$$

$$\text{and } \mathbf{x} \in \mathbb{T}^N. \tag{13}$$

This problem formulation depends on one hand on the input symbol vector $\mathbf{s}$, which determines the intended SR for each received signal, on the channel $\mathbf{H}$, and on the other hand on the transmit power P_{tx} that affects the noiseless received signal $\mathbf{y}$. Since the optimization variables depend linearly on the square-root of the transmit power $\sqrt{P_{\text{tx}}}$, it is sufficient to solve the optimization problem for one transmit power value. Therefore, we consider the following specific case for the subsequent derivations:

$$\mathbf{y}' = \mathbf{y}\,|_{P_{\text{tx}}=N,(10)} = \mathbf{H}\mathbf{x}. \tag{14}$$

This signal vector $\mathbf{y}'$ is equal to the noiseless received signal $\mathbf{y}$ for a transmit power $P_{\text{tx}} = N$ and when (10) is fulfilled. The optimization is based on this special case. However, the QCE constraint leads to a discrete optimization problem that cannot be solved efficiently. Therefore, the QCE constraint will be relaxed to a convex constraint as shown in Section IV-B. The constraint relaxation does not satisfy the equality in (10) and thus the quantization distortions are not fully omitted. However, they are reduced significantly as shown later.

IV. Problem Formulation for PSK Signaling

A. Symbol Region for PSK Signals

In this section, we assume that the input signals s_m, $m = 1, \cdots, M$, belong to the S-PSK constellation. The set $\mathbb{S}$ in this case is defined as

$$\mathbb{S} := \{\exp(\mathrm{j}(2i-1)\theta) : i = 1, \cdots, Q\}, \text{ where } \theta = \frac{\pi}{S}. \tag{15}$$

Each SR in the PSK constellation, as shown in Fig. 2, is a circular sector of infinite radius and angle 2θ. To find a mathematical expression for the SR, the original coordinate system is rotated by the phase of the symbol of interest s_m to get a modified coordinate system as illustrated in Fig. 4. The

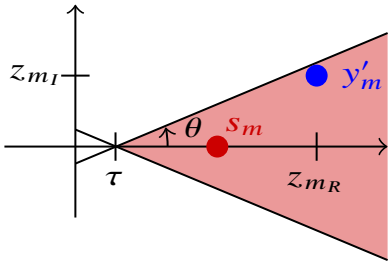

Figure 4. Illustration of the PSK symbol region in a modified coordinate system.

coordinates of the noiseless received signal y'_m in the modified coordinate system are given by

$$z_{m_R} = \Re\{y'_m s_m^*\}\frac{1}{|s_m|} \tag{16}$$

$$z_{m_I} = \Im\{y'_m s_m^*\}\frac{1}{|s_m|}. \tag{17}$$

Since PSK signals have unit magnitude, plugging (14) into the above equations gives

$$z_{m_R} = \Re\{\mathbf{e}_m^{\mathrm{T}}\mathbf{H}\mathbf{x}s_m^*\} = \Re\{\mathbf{e}_m^{\mathrm{T}}\tilde{\mathbf{H}}\mathbf{x}\} \tag{18}$$

$$z_{m_I} = \Im\{\mathbf{e}_m^{\mathrm{T}}\mathbf{H}\mathbf{x}s_m^*\} = \Im\{\mathbf{e}_m^{\mathrm{T}}\tilde{\mathbf{H}}\mathbf{x}\}, \tag{19}$$

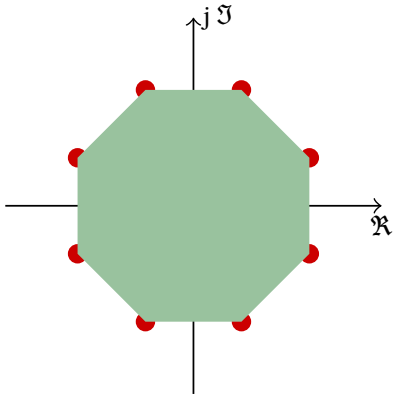

Figure 5. Illustration of the relaxed polygon constraint for Q=8.

where

$$\tilde{\mathbf{H}} = \text{diag}(\mathbf{s}^*)\mathbf{H}. \tag{20}$$

The m-th SR can be hence described by

$$z_{m_R} \geq \tau \tag{21}$$

$$|z_{m_I}| \leq \left(z_{m_R} - \tau\right) \tan\theta, \forall m, \tag{22}$$

where $\tau = \frac{\delta}{\sin\theta}$. Note that the inequality in (21) is already fulfilled if the inequality in (22) is satisfied. Plugging (18) and (19) into (22), the SRs for all M users can be defined by

$$|\Im\{\tilde{\mathbf{H}}\mathbf{x}\}| \leq \left(\Re\{\tilde{\mathbf{H}}\mathbf{x}\} - \tau\mathbf{1}_M\right) \tan\theta. \tag{23}$$

When using the following real-valued representation

$$\Re\{\tilde{\mathbf{H}}\mathbf{x}\} = \underbrace{\begin{bmatrix}\Re\{\tilde{\mathbf{H}}\} & -\Im\{\tilde{\mathbf{H}}\}\end{bmatrix}}_{=\mathbf{A}} \underbrace{\begin{bmatrix}\Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\}\end{bmatrix}}_{=\mathbf{x}'} = \mathbf{A}\mathbf{x}' \tag{24}$$

$$\Im\{\tilde{\mathbf{H}}\mathbf{x}\} = \underbrace{\begin{bmatrix}\Im\{\tilde{\mathbf{H}}\} & \Re\{\tilde{\mathbf{H}}\}\end{bmatrix}}_{=\mathbf{B}} \begin{bmatrix}\Re\{\mathbf{x}\} \\ \Im\{\mathbf{x}\}\end{bmatrix} = \mathbf{B}\mathbf{x}', \tag{25}$$

the constraint in (12) can be rewritten as

$$\begin{bmatrix}\mathbf{B} - \tan\theta\mathbf{A} & \frac{1}{\cos\theta}\mathbf{1}_M \\ -\mathbf{B} - \tan\theta\mathbf{A} & \frac{1}{\cos\theta}\mathbf{1}_M\end{bmatrix} \begin{bmatrix}\mathbf{x}' \\ \delta\end{bmatrix} \leq \mathbf{0}_{2M}. \tag{26}$$

B. Relaxed Polygon Constraint

The non-convex QCE constraint is relaxed to a convex constraint, which we call the polygon constraint, such that the entries of the vector $\mathbf{x}$ belong to the polygon built by the Q-PSK symbols, as shown in Fig. 5. For the case of $Q = 4$, we can describe the constraint as follows

$$\mathbf{x}' \leq \frac{1}{\sqrt{2}}\mathbf{1}_{2N} \text{ and } -\mathbf{x}' \leq \frac{1}{\sqrt{2}}\mathbf{1}_{2N}. \tag{27}$$

For q-bit DACs, i.e., where the transmitted data are constrained to be Q-PSK symbols, the polygon can be constructed by the intersection of $Q/4$ squares that have an angular shift of $2\pi/Q$. To this end, we define

$$\mathbf{T}_i = \begin{bmatrix}\cos\beta_i & \sin\beta_i \\ -\sin\beta_i & \cos\beta_i\end{bmatrix} \otimes \mathbf{I}_N, \; i = 1, ..., Q/4, \tag{28}$$

where $\beta_i = \frac{2\pi}{Q}(i-1)$. The system of inequalities that considers the feasible set (polygon constraint) and hence relaxes the constraint in (13) is given by

$$\begin{bmatrix}\mathbf{T}_1 \\ -\mathbf{T}_1 \\ \vdots \\ \mathbf{T}_{\frac{Q}{4}} \\ -\mathbf{T}_{\frac{Q}{4}}\end{bmatrix} \mathbf{x}' \leq \cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{NQ}. \tag{29}$$

Since $\mathbf{T}_1 = \mathbf{I}_{2N}$, the first $4N$ inequalities in (29) define the bounds of $\mathbf{x}'$. Hence, (29) can be rewritten as

$$-\cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{2N} \leq \mathbf{x}' \leq \cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{2N},$$

$$\text{and} \quad \underbrace{\begin{bmatrix}\mathbf{T}_2 \\ -\mathbf{T}_2 \\ \vdots \\ \mathbf{T}_{\frac{Q}{4}} \\ -\mathbf{T}_{\frac{Q}{4}}\end{bmatrix}}_{=\mathbf{E}} \mathbf{x}' \leq \cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{NQ}. \tag{30}$$

This reformulation leads to significant computational savings since the final optimization problem will be written as a linear program with bounded variables. It is beneficial in terms of computational complexity to have less number of inequalities as discussed in Section VI.

C. Optimization Problem with the Relaxed Polygon Constraint

Finally, the optimization problem for the symbol-wise precoder with PSK signaling is obtained by combining (11), (26) and (30) and is expressed as

$$\begin{aligned}&\max_{\mathbf{v}} \begin{bmatrix}\mathbf{0}_{2N}^{\mathrm{T}} & 1\end{bmatrix}\mathbf{v} \\ &\text{s.t.} \begin{bmatrix}\mathbf{B} - \tan\theta\mathbf{A} & \frac{1}{\cos\theta}\mathbf{1}_M \\ -\mathbf{B} - \tan\theta\mathbf{A} & \frac{1}{\cos\theta}\mathbf{1}_M \\ \mathbf{E} & \mathbf{0}_{N(Q-4)}\end{bmatrix}\mathbf{v} \leq \begin{bmatrix}\mathbf{0}_{2M} \\ \cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{N(Q-4)}\end{bmatrix}, \\ &\text{and} \begin{bmatrix}-\cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{2N} \\ 0\end{bmatrix} \leq \mathbf{v} \leq \begin{bmatrix}\cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{2N} \\ \infty\end{bmatrix},\end{aligned} \tag{31}$$

where $\mathbf{v}^{\mathrm{T}} = \begin{bmatrix}\mathbf{x}'^{\mathrm{T}} & \delta\end{bmatrix}$. The resulting optimization problem is a linear programming problem for which there exist very efficient solving methods [44].

When the optimization terminates, the optimal signal $\mathbf{x} \in \mathbb{X}^N$ is found. The signal $\mathbf{t}$ that goes through the channel is obtained as described in (1). In other words, each entry in $\mathbf{x}$ gets mapped to the corresponding CE point depending on the circular sector that it lies in.

V. Problem Formulation for QAM Signaling

A. The Need for an Additional Degree of Freedom α

In this section, we assume that the input signals s_m, $m = 1, \cdots, M$, belong to the S-QAM constellation, where S is assumed to be a power of 4. The QAM symbols are drawn from the set $\mathbb{S}$ defined as

$$\mathbb{S} := \left\{\pm(2i-1) \pm \mathrm{j}(2i-1) : i = 1, \cdots, \log_4(S)\right\}. \tag{32}$$

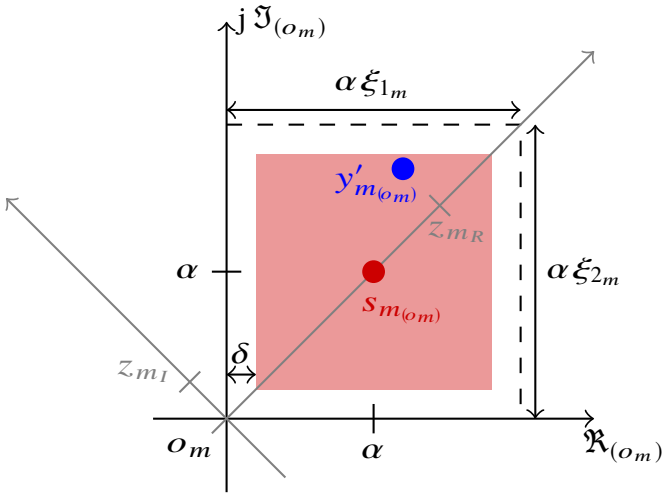

Figure 6. Illustration of the QAM receiver symbol region for $\Re\{s_m\} > 0$ and $\Im\{s_m\} > 0$ in the shifted coordinate system (in black) and in the shifted and rotated coordinate system (in gray): $\xi_{1/2_m} \in \{2, \infty\}$.

As explained in Section III, the safety margin δ has to be maximized such that the entries of the noiseless received signal $\mathbf{y}$ belong to the intended SRs. The SRs in turn are determined by the constellation set $\mathbb{S}$ and the safety margin δ. Hence, the safety margin δ cannot exceed 1,

$$\delta \leq 1. \tag{33}$$

Independently of the available transmit power, the entries of $\mathbf{y}$ cannot have a distance to the decision thresholds larger than 1. Hence, the available transmit power cannot be exploited to the fullest. This results already in a limitation of the problem formulation.

Thanks to the receive processing $\mathbf{G}$, we can introduce an additional degree of freedom α such that the entries of the received signal $\mathbf{y}$ do not have to belong to the SRs of the set $\mathbb{S}$ but rather to a scaled version of them; that is, the QAM constellation at each receiver gets scaled by α. Thus, the constraint in (33) is replaced by

$$\delta \leq \alpha, \tag{34}$$

where α has to be jointly optimized with δ. Note that maximizing δ results in turn to maximizing α, which leads to a maximal exploitation of the available transmit power. Thus, the entries of the signal vector $\mathbf{x}$ will get closer to the polygon corners, which decreases the variations between $\mathbf{t}$ and $\mathbf{x}$.

The ratio α denotes the expansion or shrinkage factor of the constellation at the receiver side depending on the available transmit power P_{tx}. As explained in Section III, the optimization problem is formulated for the specific case, i.e. $P_{\text{tx}} = N$.

B. Scaled Symbol Region for QAM Signals

To describe the SRs for QAM signaling after considering α, we define a new coordinate system, that is a shifted and rotated version of the original coordinate system. First, the original receiver constellation system is shifted by o_m

$$o_m = \alpha\Big((\Re\{s_m\} - \operatorname{sgn}(\Re\{s_m\})) + \mathrm{j}\,(\Im\{s_m\} - \operatorname{sgn}(\Im\{s_m\}))\Big). \tag{35}$$

We get the following expressions for the received and the desired signal in the new coordinate system depicted in Fig. 6

$$\begin{aligned} y'_{m_{(o_m)}} &= y'_m - o_m \\ &= \mathbf{e}_m^{\mathrm{T}}\mathbf{H}\mathbf{x} - o_m \end{aligned} \tag{36}$$

$$\begin{aligned} s_{m(o_m)} &= \alpha s_m - o_m \\ &= \alpha\,(\operatorname{sgn}(\Re\{s_m\}) + \mathrm{j}\operatorname{sgn}(\Im\{s_m\}))\,. \end{aligned} \tag{37}$$

Second, the intermediate coordinate system is rotated by the phase of the symbol of interest $s_{m(o_m)}$. So the received signal y'_m has the following coordinates in the shifted and rotated coordinate system

$$z_{m_R} = \frac{\Re\left\{y'_{m_{(o_m)}} s^*_{m_{(o_m)}}\right\}}{|s_{m(o_m)}|}, \tag{38}$$

and

$$z_{m_I} = \frac{\Im\left\{y'_{m_{(o_m)}} s^*_{m_{(o_m)}}\right\}}{|s_{m(o_m)}|}. \tag{39}$$

We get

$$\begin{aligned} \frac{y'_{m_{(o_m)}} s^*_{m_{(o_m)}}}{|s_{m(o_m)}|} &= \frac{1}{\sqrt{2}\alpha}\left(\mathbf{e}_m^{\mathrm{T}}\mathbf{H}\mathbf{x} - o_m\right) s^*_{m_{(o_m)}} \\ &= \frac{1}{\sqrt{2}\alpha}\left(s^*_{m_{(o_m)}}\mathbf{e}_m^{\mathrm{T}}\mathbf{H}\mathbf{x} - o_m s^*_{m_{(o_m)}}\right) \\ &= \frac{(\operatorname{sgn}(\Re\{s_m\}) - \mathrm{j}\operatorname{sgn}(\Im\{s_m\}))}{\sqrt{2}}\mathbf{e}_m^{\mathrm{T}}\mathbf{H}\mathbf{x} \\ &\qquad - \frac{1}{\sqrt{2}\alpha} o_m s^*_{m_{(o_m)}} \\ &= \mathbf{e}_m^{\mathrm{T}}\hat{\mathbf{H}}\mathbf{x} - \alpha c_m, \end{aligned} \tag{40}$$

where

$$\hat{\mathbf{H}} = \frac{1}{\sqrt{2}}\operatorname{diag}(\operatorname{sgn}(\Re\{\mathbf{s}\}) - \mathrm{j}\operatorname{sgn}(\Im\{\mathbf{s}\}))\,\mathbf{H}, \tag{41}$$

and

$$c_m = \frac{o_m s^*_{m_{(o_m)}}}{\sqrt{2}\alpha^2}. \tag{42}$$

Note that c_m does not actually depend on α as can be concluded from (35), (37) and (42). Plugging (40) into (38) and (39), we get

$$z_{m_R} = \mathbf{e}_m^{\mathrm{T}}\mathbf{V}\mathbf{x}' - \alpha\Re\{c_m\} \tag{43}$$

$$z_{m_I} = \mathbf{e}_m^{\mathrm{T}}\mathbf{W}\mathbf{x}' - \alpha\Im\{c_m\}, \tag{44}$$

where

$$\mathbf{V} = \begin{bmatrix}\Re\{\hat{\mathbf{H}}\} & -\Im\{\hat{\mathbf{H}}\}\end{bmatrix} \tag{45}$$

$$\mathbf{W} = \begin{bmatrix}\Im\{\hat{\mathbf{H}}\} & \Re\{\hat{\mathbf{H}}\}\end{bmatrix}. \tag{46}$$

The m-th SR, as shown in Fig. 6, can be hence described by

$$z_{m_R} \geq \sqrt{2}\delta \tag{47}$$

$$z_{m_R} \leq \sqrt{\left(\alpha\xi_{1_m} - \delta\right)^2 + \left(\alpha\xi_{2_m} - \delta\right)^2} \tag{48}$$

$$|z_{m_I}| \leq \left(z_{m_R} - \sqrt{2}\delta\right) \tag{49}$$

$$z_{m_I} \leq -z_{m_R} + \sqrt{2}\left(\alpha\xi_{2_m} - \delta\right) \tag{50}$$

$$z_{m_I} \geq z_{m_R} - \sqrt{2}\left(\alpha\xi_{1_m} - \delta\right). \tag{51}$$

Note that ξ_{1_m} and $\xi_{2_m} \in \{2, \infty\}$ depending on which constellation point the symbol of interest s_m corresponds to. If s_m is one of the outer constellation points, then at least ξ_{1_m} or ξ_{2_m} must be equal to ∞. Since (47) and (48) are inherently fulfilled by (49), (50) and (51), the constraint in (12) can be rewritten as

$$\begin{bmatrix} \mathbf{W}-\mathbf{V} & \mathbf{1}_M & \Re\{\mathbf{c}\} - \Im\{\mathbf{c}\} \\ -\mathbf{W}-\mathbf{V} & \mathbf{1}_M & \Re\{\mathbf{c}\} + \Im\{\mathbf{c}\} \\ \mathbf{W}+\mathbf{V} & \mathbf{1}_M & -\Re\{\mathbf{c}\} - \Im\{\mathbf{c}\} - \sqrt{2}\boldsymbol{\xi}_2 \\ -\mathbf{W}+\mathbf{V} & \mathbf{1}_M & -\Re\{\mathbf{c}\} + \Im\{\mathbf{c}\} - \sqrt{2}\boldsymbol{\xi}_1 \end{bmatrix} \begin{bmatrix} \mathbf{x}' \\ \sqrt{2}\delta \\ \alpha \end{bmatrix} \leq \mathbf{0}_{4M}. \tag{52}$$

C. Optimization Problem with the Relaxed Polygon Constraint

We are interested in maximizing the safety margin as presented in (11). In contrast to the PSK case, there is a constraint on δ in the QAM case, stated in (34), which is inherently fulfilled by (52). Combining (11) with the SR constraint in (52) and the relaxed polygon constraint in (30), we get a linear programming problem for the design of the symbol-wise precoder for QAM signaling. The optimization problem is given in (53), where $\mathbf{v}^{\mathrm{T}} = \begin{bmatrix} \mathbf{x}'^{\mathrm{T}} & \sqrt{2}\delta & \alpha \end{bmatrix}$.

Again the optimized vector $\mathbf{x} \in \mathbb{X}^N$ goes through the quantizer, as stated in (1), to obtain the transmit vector $\mathbf{t}$.

D. Receive Processing

The variables of the optimization problem are the transmit vector $\mathbf{x}$, the safety margin δ and the expansion factor α. The latter determines the receive processing $\mathbf{G}$. Note that the optimal value of α is determined on a symbol-by-symbol basis, and its value cannot be communicated to the receiver. However, due to the massive MIMO assumption and the induced hardening effect, the flactuations of α across the symbols are small (see also the discussion in the following subsections). Therefore an exact value of α is not required at the receiver. Only the positions of the decision thresholds are needed to rescale the receiver constellation points to the nominal constellation points, and these only depend on the mean value of α. An estimate of the mean of α can easily be computed by averaging over a block of received signals.

After multiplication with the receiver coefficient g_m, the scaled received signal should equal

$$u_m = g_m r_m = g_m \, \mathrm{e}_m^{\mathrm{T}} \, \mathbf{H}\mathbf{t} + g_m \eta_m = s_m + \eta'_m, \tag{54}$$

where η'_m denotes the deviation of u_m from the nominal point s_m due to the Additive White Gaussian Noise (AWGN) η_m, the SR constraint and the quantization applied on the relaxed optimized vector $\mathbf{x}$. Then, we can write

$$|\mathrm{Re}\{r_m\}| + |\mathrm{Im}\{r_m\}| = g_m^{-1}\left(|\mathrm{Re}\{s_m + \eta'_m\}| + |\mathrm{Im}\{s_m + \eta'_m\}|\right) \tag{55}$$

$$\overset{\text{w/ high prob. at SINR} \gg 1}{=} g_m^{-1}\left(|\mathrm{Re}\{s_m\}| + |\mathrm{Im}\{s_m\}|\right) + g_m^{-1}\left(\mathrm{Re}\{\eta'_m\} + \mathrm{Im}\{\eta'_m\}\right), \tag{56}$$

meaning that with zero-mean noise plus interference η'_m we have

$$\mathrm{E}[|\mathrm{Re}\{r_m\}| + |\mathrm{Im}\{r_m\}|] \approx g_m^{-1}\mathrm{E}[|\mathrm{Re}\{s_m\}| + |\mathrm{Im}\{s_m\}|]. \tag{57}$$

Based on (57), we propose a blind estimation method to obtain the scaling factor g_m for each user prior to the decision operation. The method does not require any feedback or training from the BS nor any knowledge of the noise plus interference power at the user terminal:

$$g_m = T \cdot \frac{\mathrm{E}\left[|\mathrm{Re}\{s\}| + |\mathrm{Im}\{s\}|\right]}{\sum_{t=1}^{T} |\mathrm{Re}\{r_m[t]\}| + |\mathrm{Im}\{r_m[t]\}|}, \tag{58}$$

where T is the length the received sequence.

E. Symbol-wise Processing vs. Block-wise Processing

One might ask why we opt for symbol-wise processing and not block-wise processing. The factor α cannot be communicated to the receiver and hence has to be estimated. The estimation is based on averaging over a block of T received signals. Thus, one expects that the design of α at the transmitter has to be computed for the same block length, i.e. $B = T$. However, fixing α for a certain block length not only increases the complexity of the problem but also reduces the degrees of freedom of the optimization problem at the transmitter and the vectors $\mathbf{x}$ have to be designed with a greater restriction on α. This leads to the entries of the vector $\mathbf{x}$ moving farther from the polygon corners, thus increasing the quantization distortions. This effect is illustrated in Fig. 7, where the entries of $\mathbf{e}_m^{\mathrm{T}}\mathbf{H}\mathbf{x}$, $\mathbf{e}_m^{\mathrm{T}}\mathbf{H}\mathbf{t}$ and $\frac{1}{\alpha}\mathbf{e}_m^{\mathrm{T}}\mathbf{H}\mathbf{t}$ of an arbitrary user m are obtained by transmitting 1024 16QAM signal vectors through an independent and identically distributed (i.i.d.) channel of $N = 64$, $M = 8$ and $Q = 4$. The optimization is computed for both symbol-wise processing, i.e. $B = 1$, and block-wise processing with $B = 4$. As can be deduced from the plots, the block-wise processing leads to a larger safety margin with the relaxed vector $\mathbf{x}$. However, after applying the quantization this gain is lost and the symbol-wise processing seems to be more robust against the quantization operation. This can be further explained by the results in Table I, which shows $\mathrm{E}\left[\frac{\|\mathbf{t}\neq\mathbf{x}\|_1}{N}\right]$, the percentange of entries of $\mathbf{x}$ that are distorted due to the quantization and the MSE between $\mathbf{t}$ and $\mathbf{x}$. We see that increasing B significantly increases the quantization distortion. Therefore, the symbol-wise processing is chosen in this contribution, i.e, an optimal value of α is designed for each vector $\mathbf{x}$.

F. One Joint α vs. M Distinct α's for M Users

The symbol-wise transmit processing followed by the block-wise receive processing is reliable only if the obtained values

$$\max_{\mathbf{v}} \begin{bmatrix} \mathbf{0}_{2N}^{\mathrm{T}} & 1 & 0 \end{bmatrix} \mathbf{v} \text{ s.t. } \begin{bmatrix} \mathbf{W}-\mathbf{V} & \mathbf{1}_M & \Re\{\mathbf{c}\}-\Im\{\mathbf{c}\} \\ -\mathbf{W}-\mathbf{V} & \mathbf{1}_M & \Re\{\mathbf{c}\}+\Im\{\mathbf{c}\} \\ \mathbf{W}+\mathbf{V} & \mathbf{1}_M & -\Re\{\mathbf{c}\}-\Im\{\mathbf{c}\}-\sqrt{2}\boldsymbol{\xi}_2 \\ -\mathbf{W}+\mathbf{V} & \mathbf{1}_M & -\Re\{\mathbf{c}\}+\Im\{\mathbf{c}\}-\sqrt{2}\boldsymbol{\xi}_1 \\ \mathbf{E} & \mathbf{0}_{N(Q-4)} & \mathbf{0}_{N(Q-4)} \end{bmatrix} \mathbf{v} \leq \begin{bmatrix} \mathbf{0}_{4M} \\ \cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{N(Q-4)} \end{bmatrix}$$

$$\text{and } \begin{bmatrix} -\cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{2N} \\ 0 \\ 0 \end{bmatrix} \leq \mathbf{v} \leq \begin{bmatrix} \cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{2N} \\ \infty \\ \infty \end{bmatrix}. \tag{53}$$

Table I
QUANTIZATION DISTORTION VS. B.

	$B=1$	$B=4$
$\mathrm{E}\left[\frac{\|\mathbf{t}\neq\mathbf{x}\|_1}{N}\right]$	0.2176	0.4432
$\mathrm{E}\left[\|\mathbf{t}-\mathbf{x}\|_2^2\right]$	2.5458	12.6429

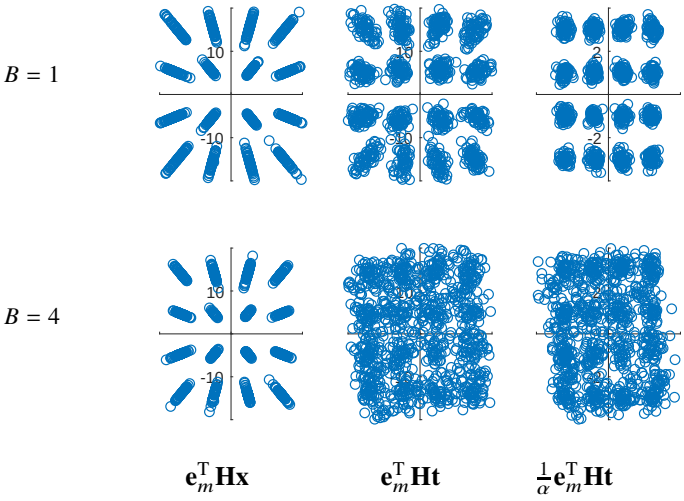

Figure 7. The noiseless received symbols at one arbitrary user m for an arbitrary i.i.d. channel realization with $N=64$, $M=8$ and $Q=4$.

of α do not vary much from one vector $\mathbf{x}$ to another. Otherwise, estimating the mean value of α at the receiver would not be sufficient for correct detection. This explains why we choose one joint α for all users. If a different value α_m per user is chosen, this would result in more degrees of freedom and the values $\alpha_m, m = 1, \cdots, M$, would fluctuate much more from one vector $\mathbf{x}$ to another, which worsens the estimation result at the receiver. For a large number of users, the jointly designed α will not vary much due to the channel hardening effect. This behavior can be better illustrated by looking at the relative range of α

$$\mathrm{E}_{\mathbf{H}}\left[\frac{\max\{\alpha\}-\min\{\alpha\}}{\mathrm{E}[\alpha]}\right] \tag{59}$$

and the maximal relative range of $\alpha_m,\ m = 1, \cdots, M$,

$$\max_m \mathrm{E}_{\mathbf{H}}\left[\frac{\max\{\alpha_m\}-\min\{\alpha_m\}}{\mathrm{E}[\alpha_m]}\right] \tag{60}$$

averaged over many channel realizations and for different values of N and M in Table II and Table III with 16QAM signaling and i.i.d. channel. As a consequence, exploiting the

Table II
RELATIVE RANGE OF α: $\mathrm{E}_{\mathbf{H}}\left[\frac{\max\{\alpha\}-\min\{\alpha\}}{\mathrm{E}[\alpha]}\right]$.

M \ N	64	200
2	1.52	1.44
8	0.78	0.71
14	0.61	0.50

Table III
MAXIMAL RELATIVE RANGE OF α_m: $\max_m \mathrm{E}_{\mathbf{H}}\left[\frac{\max\{\alpha_m\}-\min\{\alpha_m\}}{\mathrm{E}[\alpha_m]}\right]$.

M \ N	64	200
2	1.51	1.44
8	1.37	0.75
14	2	0.97

channel hardening effect by using only one single α for all users is crucial for the robustness of the method and to allow for adequate receiver processing for calculating the scaling factors.

VI. COMPUTATIONAL COMPLEXITY OF MSM

A. On the Computational Complexity of General Linear Programming Problems

In this section, we study the computational complexity of the simplex method for a general linear programming problem with bounded variables in inequality form:

$$\max_{\mathbf{x}} \mathbf{c}^{\mathrm{T}}\mathbf{x} \text{ s.t. } \mathbf{A}\mathbf{x} \leq \mathbf{b} \quad \text{and } \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \tag{61}$$

where $\mathbf{c}$, $\mathbf{x}$, $\mathbf{l}$ and $\mathbf{u} \in \mathbb{R}^n$, $\mathbf{A} \in \mathbb{R}^{m\times n}$ and $\mathbf{b} \in \mathbb{R}^m$.

First, we have to make sure that the entries of $\mathbf{b}$ are non-negative. To this end, we change the signs of the inequalities that correspond to negative entries in $\mathbf{b}$. So we get

$$\min_{\mathbf{x}} \mathbf{c}^{\mathrm{T}}\mathbf{x} \text{ s.t. } \tilde{\mathbf{A}}\mathbf{x} \gtreqless \tilde{\mathbf{b}} \quad \text{and } \mathbf{l} \leq \mathbf{x} \leq \mathbf{u}, \tag{62}$$

where $\tilde{\mathbf{b}} \in \mathbb{R}_+^m$ and some inequalities hold with the sign $\leq$ and others with the sign $\geq$.

Second, the linear programming problem is transformed to the canonical form by introducing m slack and surplus variables denoted by $\mathbf{x}_\mathrm{s}$. Additionally, a artificial variables denoted by $\mathbf{x}_\mathrm{a}$, with $0 \leq a \leq m$, are added to set up an initial feasible solution [45]. The equivalent enlarged problem reads as

$$\min_{\bar{\mathbf{x}}} \bar{\mathbf{c}}^\mathrm{T}\bar{\mathbf{x}} \text{ s.t. } \bar{\mathbf{A}}\bar{\mathbf{x}} = \tilde{\mathbf{b}} \quad \text{and } \bar{\mathbf{l}} \leq \bar{\mathbf{x}} \leq \bar{\mathbf{u}}, \tag{63}$$

where $\bar{\mathbf{A}} = \begin{bmatrix}\tilde{\mathbf{A}} & \mathbf{A}_\mathrm{s} & \mathbf{I}_a\end{bmatrix} \in \mathbb{R}^{m\times(n+m+a)}$, $\bar{\mathbf{x}}^\mathrm{T} = \begin{bmatrix}\mathbf{x}^\mathrm{T} & \mathbf{x}_\mathrm{s}^\mathrm{T} & \mathbf{x}_\mathrm{a}^\mathrm{T}\end{bmatrix} \in \mathbb{R}^{n+m+a}$, $\bar{\mathbf{l}}^\mathrm{T} = \begin{bmatrix}\mathbf{l}^\mathrm{T} & \mathbf{0}_{m+a}^\mathrm{T}\end{bmatrix}$ and $\bar{\mathbf{u}}^\mathrm{T} = \begin{bmatrix}\mathbf{u}^\mathrm{T} & \infty\mathbf{1}_{m+a}^\mathrm{T}\end{bmatrix}$. The matrix $\mathbf{A}_\mathrm{s}$ is a diagonal matrix with entries equal to 1 or -1 depending on whether the inequality sign in (62) is $\leq$ or $\geq$, respectively. The number a of artificial variables is defined by the number of negative entries in $\mathbf{A}_\mathrm{s}$, such that the concatenation of m columns from $\begin{bmatrix}\mathbf{A}_\mathrm{s} & \mathbf{I}_a\end{bmatrix}$ can construct the identity matrix $\mathbf{I}_m$. For the special case $\mathbf{b} = \tilde{\mathbf{b}}$, i.e. the entries of $\mathbf{b}$ are non-negative, $\mathbf{A}_\mathrm{s} = \mathbf{I}_m$. Hence, no artificial variables are needed, i.e. $a = 0$.

With the use of the simplex method to solve (63), the number of operations (multiplication and addition pairs) on each iteration is given by, [45, p.83],

$$3m \qquad \text{or} \qquad (m+1)(n+a+1)+2m, \tag{64}$$

depending on whether pivoting is required or not. According to [45, p.86], in most iterations no pivoting is required and hence less computation is needed.

B. Computational Complexity of MSM for PSK Signaling

As can be seen from (31), there are $m = 2M + N(Q-4)$ inequalities and $n = 2N + 1$ variables. The number a of artificial variables reduces to 0, since the vector $\mathbf{b}^\mathrm{T} = \begin{bmatrix}\mathbf{0}_{2M}^\mathrm{T} & \cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{N(Q-4)}^\mathrm{T}\end{bmatrix}$ has only non-negative entries. Thus, the number of operations (multiplication and addition pairs) on each iteration calculates in this case to

$$6M + 3(Q-4)N, \tag{65}$$

or

$$2N + 4MN + 8M + 2(Q-4)\left(2N^2 + 4N\right). \tag{66}$$

For the special case of 1-bit quantization, i.e. $Q = 4$, the complexity is linear in N and M.

C. Computational Complexity of MSM for QAM Signaling

From (53), we have $m = 4M + N(Q-4)$ inequalities and $n = 2N + 2$ variables. The number a of artificial variables reduces to 0, since the vector $\mathbf{b}^\mathrm{T} = \begin{bmatrix}\mathbf{0}_{4M+1}^\mathrm{T} & \cos\left(\frac{\pi}{Q}\right)\mathbf{1}_{N(Q-4)}^\mathrm{T}\end{bmatrix}$ has only non-negative entries. Thus, the number of operations (multiplication and addition pairs) on each iteration calculates in this case to

$$12M + 3(Q-4)N, \tag{67}$$

or

$$2N + 8MN + 20M + 3 + (Q-4)\left(2N^2 + 5N\right). \tag{68}$$

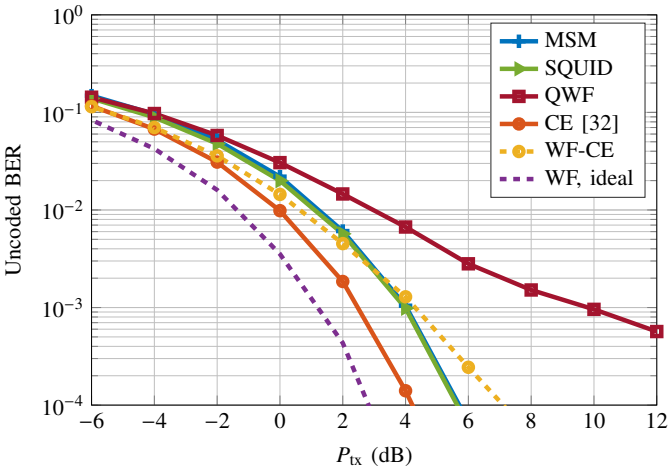

Figure 8. Uncoded BER performance for a MU-MIMO system with $N = 64$ and $M = 8$ with different precoding designs and QPSK signaling.

For the special case of 1-bit quantization, i.e. $Q = 4$, the complexity is linear in N and M. Note that the sparsity of $\mathbf{E}$ can be exploited by deploying the revised simplex method to reduce the number of required operations in the case of $Q > 4$ [45, p.89].

VII. Simulation Results

For the simulations, we assume a BS with $N = 64$ antennas serving $M = 8$ single-antenna users. The channel $\mathbf{H}$ is composed of i.i.d. Gaussian random variables with zero-mean and unit variance. The numerical results are obtained with Monte Carlo simulations of 100 independent channel realizations. The additive noise is also i.i.d. with variance one at each antenna. The performance metric is the uncoded Bit Error Ratio (BER) averaged over the single-antenna users. For the blind estimation of the coefficients g_m we use a block length of $T = 128$.

In the first simulation set, depicted in Fig. 8, we assume full Channel State Information (CSI), choose QPSK modulation and compare the uncoded BER as a function of the transmit power P_tx for the following precoders

- the proposed MSM method with $Q = 4$,
- the SQUID precoder presented in [28] with $Q = 4$,
- the quantized Wiener Filter (WF) precoder denoted by "QWF" from [24] with $Q = 4$,
- the CE precoder presented in [35] denoted by "CE [35]", with $Q = \infty$, where the precoding gain α is taken into account,
- the WF precoder followed by the CE quantizer with $Q = \infty$ denoted by "WF-CE", and
- the WF precoder in the ideal case denoted by "WF, ideal", where neither quantization nor the CE constraint is applied to the transmit signal.

It can be seen that the CE constraint leads to a loss, compared to the ideal case, of almost 2 dB at a BER of 10^{-2} compared to WF and a loss of less than 1.5 dB when using the symbol-wise precoder proposed in [35]. The 1-bit quantization, which represents the QCE case of $Q = 4$, leads to more losses that

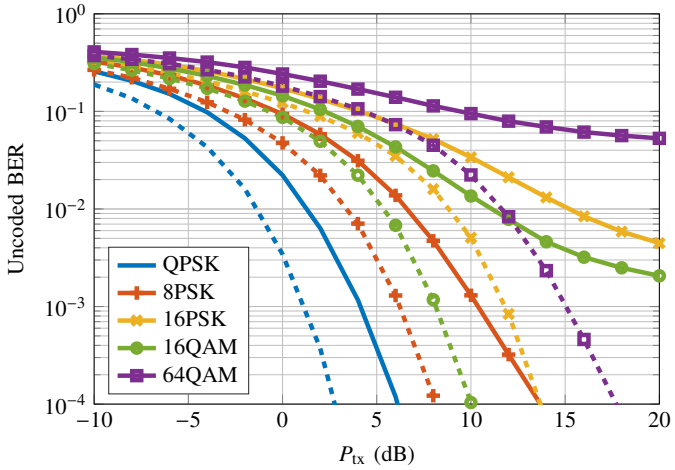

Figure 9. Uncoded BER performance for a MU-MIMO system with $N = 64$ and $M = 8$ for different modulation schemes and $Q = 4$: the dashed lines represent the uncoded BER results obtained in the case of WF, ideal.

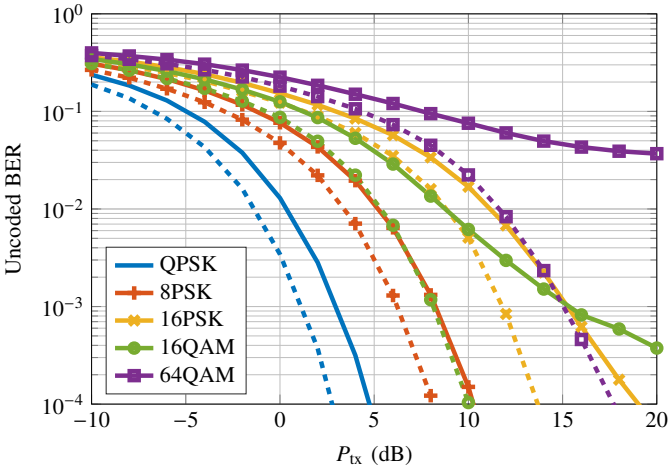

Figure 10. Uncoded BER performance for a MU-MIMO system with $N = 64$ and $M = 8$ for different modulation schemes and $Q = 8$: the dashed lines represent the uncoded BER results obtained in the case of WF, ideal.

depend on the precoder design. With the use of the linear precoder QWF a loss of more than 4 dB at a BER of 10^{-2} is noticed. However, the non-linear precoders MSM and SQUID improve the performance drastically and show a loss of slightly more than 2 dB compared to the ideal caseat the cost of higher computational complexity. Nevertheless, the proposed MSM method appears to be more efficient than SQUID as it is based on a pure linear programming formulation that has been intensively investigated in the literature.

In the second simulation set, depicted in Fig. 9 and Fig. 10, the uncoded BER is plotted as a function of the transmit power P_{tx} using the MSM precoder for different modulation schemes and two different values of Q: $Q = 4$ and $Q = 8$. Higher values of Q are omitted since the obtained results do not differ much from the case of $Q = 8$. In addition, it is beneficial in terms of computational complexity and power consumption to keep Q as small as possible. As expected, the higher the number of symbols in the modulation scheme, the higher the BER for a given P_{tx} value. However, the increase of the DAC resolution q and thus the increase of Q leads to a performance improvement, which depends on the modulation scheme. Interestingly, the 16QAM results outperform the 16PSK results with a gain of almost 4 dB at a BER of 10^{-2} for the case of $Q = 4$, whereas in the case of $Q = 8$ the gain reduces to 3 dB and the 16PSK modulation outperforms the 16QAM for transmit power values larger than 15 dB.

The third simulation set, depicted in Fig. 11, addresses the system performance in the presence of channel estimation errors. The estimated channel is defined as

$$\mathbf{H}_\nu = \sqrt{1-\nu}\mathbf{H} + \sqrt{\nu}\mathbf{\Gamma}, \tag{69}$$

where $\mathbf{\Gamma}$ is a random matrix with i.i.d. zero-mean and unit-variance entries. We can see that the performance of the proposed MSM precoder in the case of erroneous channel estimation is still better than the linear WF followed by the CE quantizer with $Q = \infty$.

In the last simulation set, we counted the average number of iterations required by the MSM precoder. The results are summarized in Table IV, where we observe that around 50 iterations are required for all different modulation schemes for $Q = 4$ and more than 100 iterations for $Q > 4$.

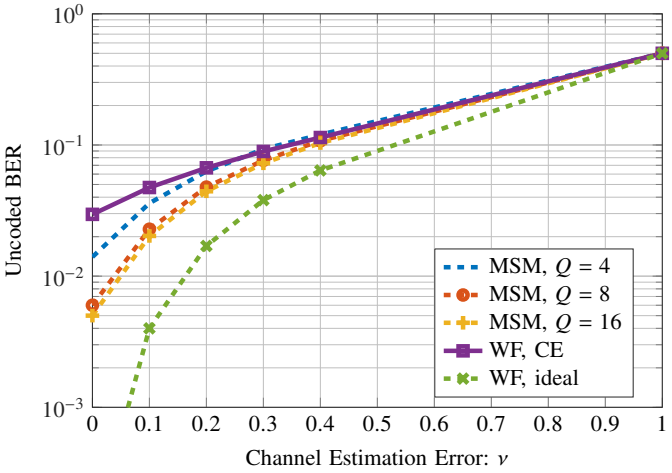

Figure 11. Uncoded BER performance as a function of the channel estimation error variance for 16QAM signaling and $P_{\text{tx}} = 10$ dB.

Table IV
AVERAGE NUMBER OF ITERATIONS OF THE MSM PRECODER.

Nb. of iter	$Q = 4$	$Q = 8$	$Q = 16$
QPSK	45.77	121.05	187.63
8PSK	50.15	123.91	191.55
16PSK	54.94	128.74	199.61
16QAM	43.25	120.42	187.32
64QAM	43.04	120.30	188.30

VIII. CONCLUSION

We proposed a symbol-wise precoder for a massive MU-MIMO downlink system with coarsely QCE signals at the transmit antennas. The CE constraint is motivated by the high PA power efficiency for CE input signals, and the coarse quantization provides further power savings due to the use

of the low-resolution DACs. The MSM precoder is based on maximizing the safety margin to the receiver decision thresholds taking the QCE into account. When relaxing the QCE constraint to a convex set, the optimization problem can be formulated as a linear programming problem, and thus can be efficiently solved via a number of methods. The proposed precoding method comprises both PSK and QAM modulation schemes.

REFERENCES

[1] A. Osseiran, F. Boccardi, V. Braun, K. Kusume, P. Marsch, M. Maternia, O. Queseth, M. Schellmann, H. Schotten, H. Taoka, H. Tullberg, M. A. Uusitalo, B. Timus, and M. Fallgren, "Scenarios for 5G Mobile and Wireless Communications: The vision of the METIS project," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 26–35, 2014.

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

[3] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. Marzetta, and F. Tufvesson, "Scaling Up MIMO: Opportunities and Challenges with Very Large Arrays," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 40–60, 2013.

[4] 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.

[5] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and Spectral Efficiency of Very Large Multiuser MIMO Systems," *IEEE Transactions on Communications*, vol. 61, no. 4, pp. 1436–1449, 2013.

[6] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An Overview of Massive MIMO: Benefits and Challenges," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 742–758, 2014.

[7] A. L. Swindlehurst, E. Ayanoglu, P. Heydari, and F. Capolino, "Millimeter-Wave Massive MIMO: The Next Wireless Revolution?" *IEEE Communications Magazine*, vol. 52, no. 9, pp. 56–62, 2014.

[8] Y. Niu, Y. Li, D. Jin, L. Su, and A. V. Vasilakos, "A Survey of Millimeter Wave Communications (mmWave) for 5G: Opportunities and Challenges," *Wireless Networks*, vol. 21, no. 8, pp. 2657–2676, 2015.

[9] T. S. Rappaport, *Millimeter Wave Wireless Communications*. Upper Saddle River N.J.: Prentice Hall, 2015.

[10] G. Auer, O. Blume, V. Giannini, I. Godor, M. A. Imran, Y. Jading, E. Katranaras, M. Olsson, D. Sabella, P. Skillermark, and W. Wajda, "Energy Efficiency Analysis of the Reference Systems, Areas of Improvements and Target Breakdown," in *EARTH Project*, 2012.

[11] O. Blume, D. Zeller, and U. Barth, "Approaches to Energy Efficient Wireless Access Networks," in *Proc. 4th International Symposium on Communications, Control and Signal Processing (ISCCSP)*, 2010.

[12] P. Varahram, S. Mohammady, B. M. Ali, and N. Sulaiman, *Power Efficiency in Broadband Wireless Communications*, 1st ed. Boca Raton: CRC Press, 2014.

[13] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-Constrained Modulation Optimization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, 2005.

[14] D. J. Love and R. W. Heath, "Equal Gain Transmission in Multiple-Input Multiple-Output Wireless Systems," *IEEE Transactions on Communications*, vol. 51, no. 7, pp. 1102–1110, 2003.

[15] X. Zheng, Y. Xie, J. Li, and P. Stoica, "MIMO Transmit Beamforming Under Uniform Elemental Power Constraint," *IEEE Transactions on Signal Processing*, vol. 55, no. 11, pp. 5395–5406, 2007.

[16] W. Yu and T. Lan, "Transmitter Optimization for the Multi-Antenna Downlink With Per-Antenna Power Constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646–2660, 2007.

[17] 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, 2012.

[18] ——, "Per-Antenna Constant Envelope Precoding for Large Multi-User MIMO Systems," *IEEE Transactions on Communications*, vol. 61, no. 3, pp. 1059–1071, 2013.

[19] ——, "Constant-Envelope Multi-User Precoding for Frequency-Selective Massive MIMO Systems," *IEEE Wireless Communications Letters*, vol. 2, no. 5, pp. 547–550, 2013.

[20] C. Mollen and E. G. Larsson, "Multiuser MIMO Precoding with Per-Antenna Continuous-Time Constant-Envelope Constraints," in *2015 IEEE 16th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*. IEEE, 2015, pp. 261–265.

[21] F. Liu, C. Masouros, P. V. Amadori, and H. Sun, "An Efficient Manifold Algorithm for Constructive Interference Based Constant Envelope Precoding," *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1542–1546, 2017.

[22] 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, 2017.

[23] H. Shen, W. Xu, A. Lee Swindlehurst, and C. Zhao, "Transmitter Optimization for Per-Antenna Power Constrained Multi-Antenna Downlinks: An SLNR Maximization Methodology," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2712–2725, 2016.

[24] A. Mezghani, R. Ghiat, and J. A. Nossek, "Transmit Processing with Low Resolution D/A-Converters," in *Proc. 16th IEEE International Conference on Electronics, Circuits and Systems - (ICECS)*, 2009, pp. 683–686.

[25] O. B. Usman, H. Jedda, A. Mezghani, and J. A. Nossek, "MMSE Precoder for Massive MIMO Using 1-Bit Quantization," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2016, pp. 3381–3385.

[26] R. Price, "A Useful Theorem for Nonlinear Devices Having Gaussian Inputs," *IEEE Transactions on Information Theory*, vol. 4, no. 2, pp. 69–72, 1958.

[27] H. Jedda, J. A. Nossek, and A. Mezghani, "Minimum BER Precoding in 1-Bit Massive MIMO Systems," in *Proc. IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2016.

[28] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Nonlinear 1-Bit Precoding for Massive MU-MIMO with Higher-Order Modulation," in *Proc. 50th Asilomar Conference on Signals, Systems and Computers*, 2016, pp. 763–767.

[29] ——, "Quantized Precoding for Massive MU-MIMO," *IEEE Transactions on Communications*, vol. 65, no. 11, pp. 4670–4684, 2017.

[30] H. Jedda, A. Mezghani, J. A. Nossek, and A. L. Swindlehurst, "Massive MIMO Downlink 1-Bit Precoding with Linear Programming for PSK Signaling," in *Proc. 18th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2017.

[31] O. Castañeda, T. Goldstein, and C. Studer, "POKEMON: A Non-Linear Beamforming Algorithm for 1-Bit Massive MIMO," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 3464–3468.

[32] A. Swindlehurst, A. Saxena, A. Mezghani, and I. Fijalkow, "Minimum Probability-of-Error Perturbation Precoding for the One-Bit Massive MIMO Downlink," in *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017, pp. 6483–6487.

[33] 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, 2017.

[34] L. T. N. Landau and R. C. de Lamare, "Branch-and-Bound Precoding for Multiuser MIMO Systems with 1-Bit Quantization," *IEEE Wireless Communications Letters*, 2017.

[35] A. Noll, H. Jedda, and J. A. Nossek, "PSK Precoding in Multi-User MISO Systems," in *Proc. 21th International ITG Workshop on Smart Antennas (WSA)*, 2017.

[36] S. Jacobsson, O. Castañeda, C. Jeon, G. Durisi, and C. Studer, "Non-linear Phase-Quantized Constant-Envelope Precoding for Massive MUMIMO-OFDM," *arXiv:1709.04846*, Oct. 2017.

[37] A. Nedelcu, F. Steiner, M. Staudacher, G. Kramer, W. Zirwas, R. S. Ganesan, P. Baracca, and S. Wesemann, "Quantized Precoding for Multi-Antenna Downlink Channels with MAGIQ," *arXiv:1712.08735*, Dec. 2017.

[38] 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, 2017.

[39] C. Masouros, T. Ratnarajah, M. Sellathurai, C. B. Papadias, and A. K. Shukla, "Known Interference in the Cellular Downlink: A Performance Limiting Factor or a Source of Green Signal Power?" *IEEE Communications Magazine*, vol. 51, no. 10, pp. 162–171, 2013.

[40] C. Masouros and G. Zheng, "Exploiting Known Interference as Green Signal Power for Downlink Beamforming Optimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3628–3640, 2015.

[41] H. Jedda, M. M. Ayub, J. Munir, A. Mezghani, and J. A. Nossek, "Power- and Spectral Efficient Communication System Design Using

1-Bit Quantization," in *Proc. International Symposium on Wireless Communication Systems (ISWCS)*, 2015, pp. 296–300.

[42] H. Jedda, A. Mezghani, and J. A. Nossek, "Spectral Shaping with Low Resolution Signals," in *Proc. 49th Asilomar Conference on Signals, Systems and Computers*, 2015, pp. 1437–1441.

[43] C. Mollen, E. G. Larsson, and T. Eriksson, "Waveforms for the Massive MIMO Downlink: Amplifier Efficiency, Distortion, and Performance," *IEEE Transactions on Communications*, vol. 64, no. 12, pp. 5050–5063, 2016.

[44] S. P. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge: Cambridge University Press, 2004.

[45] G. B. Dantzig and M. N. Thapa, *Linear Programming*, ser. Springer series in operations research. New York and London: Springer, 1997-.