Weighted Sum Rate Maximization for Multi-User MISO Systems with Low Resolution Digital to Analog Converters

Anastasios Kakkavas*†, Jawad Munir*, Amine Mezghani*, Hans Brunner* and Josef A. Nossek*
*Institute for Circuit Theory and Signal Processing
Technical University of Munich, 80290 Munich, Germany
Email: {tasos.kakkavas, jawad.munir, amine.mezghani, brunner, josef.a.nossek}@tum.de
†Huawei Technologies Duesseldorf GmbH, European Research Center, 80992 Munich, Germany

***Abstract*—We study the problem of downlink beamforming for the Weighted Sum Rate maximization (WSR) of Multi-User Multiple-Input-Single-Output systems with low-resolution Digital-to-Analog Converters (DACs) in a single-cell setup. The DACs, modeled as quantizers, are performing a nonlinear operation on the signals and are linearized using Bussgang decomposition and a linear approximation of the covariance of quantized signals. For the maximization of the WSR of the linearized system, we propose a gradient-based solution and a lower-complexity heuristic solution, based on the structure of the globally optimal solution. Through numerical simulations, we show that taking quantization into account in the filter design results in significant performance improvement when the number of transmit antennas is comparable to the number of users. When the number of transmit antennas becomes much larger than the number of users, it is found that the heuristic solution achieves near-optimal performance and that a quantization-aware design becomes less important.**

I. Introduction

Researchers both in academia and industry have concentrated their efforts to the development of the 5th Generation Wireless Systems (5G), which are expected to offer 1000 times higher mobile data volume per area and 10 to 100 times higher user data rate, at a similar cost and energy dissipation as today [1]. Two key technologies, compatible with and, maybe, complementary to each other, have received voluminous attention and are serious candidates for adoption in 5G. The first is the use of very large antenna arrays at the base station to serve a comparatively smaller number of users, a technique called massive MIMO [2] and the other is the use of the millimeter-wave (mmWave) frequency bands (30 to 300 GHz), where the vast amount of available spectrum will allow for higher data rates [3].

A major concern for the adoption of both technologies is the power dissipation in the Radio Frequency (RF)-chains. On the transmitter side, substantial portion of the power, especially in the case of short-range communications, is consumed by the Digital-to-Analog Converters (DACs). Moreover, the dissipated power in the DACs increases when the number of RF-chains increases (massive MIMO) and/or the sampling rates are increased (mmWave). The power consumed by a DAC has an exponential dependence on the bit resolution b of the converter [4]: $P_{\text{DAC}} \propto 2^b$.

In order to tackle the DAC problem, two approaches have been considered in the literature. The first approach is based on the deployment of hybrid precoding schemes, with both analog and digital processing blocks, which exploit the spatial structure of the channel [5], [6]. The other approach is the use of low resolution DACs. Systems with such DACs are usually referred to as coarsely quantized systems, as the converters are modeled as quantizers. In [7] and [8] modified linear and non-linear transmit Wiener filter designs were proposed, taking the low resolution DAC into account. In this work we take the latter approach, but, instead of the minimization of the mean square error, we focus on the maximization of the Weighted Sum Rate (WSR) of Multi-User Multiple-Input-Single-Output (MU-MISO) systems. Aiming to keep the complexity of the precoding filter as low as possible, we restrict our attention to the linear designs.

Linear downlink beamforming for WSR maximization under a total power constraint is a non-convex optimization problem [9], which has been extensively studied for the case of unquantized systems. Its globally optimal solution has been identified using the framework of monotonic optimization. In [10] the outer polyblock approximation (PA) algorithm was used, whereas in [11] the Brach-Reduce-and-Bound (BRB) algorithm was used, having a better scaling with the number of users than the PA algorithm. In [12] the system model is extended to include transmitter and receiver hardware imperfections.

The complexity of both the PA and the BRB algorithm increases exponentially with the number of users, which is prohibitive for use in practical scenarios. Hence, these algorithms can only be used as benchmarks for the evaluation of lower-complexity suboptimal methods. A popular suboptimal precoding strategy, considered in [12] and [13], is the weighted Minimum Mean Square Error (MMSE) or

Most of the work was conducted while A. Kakkavas was with the Technical University of Munich.

ISBN 978-3-8007-4177-9

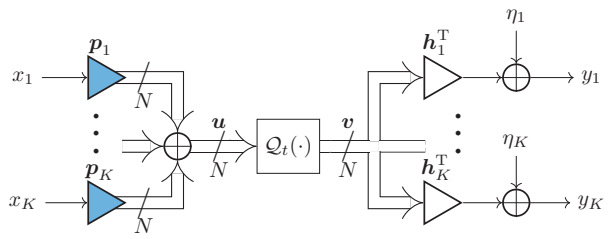

Fig. 1. Downlink beamforming for a MU-MISO system under low resolution DAC.

Signal-to-Leakage-and-Noise-Ratio (SLNR) solution, which balances the trade-off between Signal-to-Noise-Ratio (SNR) and unintended interference to other users. As shown in [12] and [13], for systems where the number of transmit antennas is much larger than the number of users, this beamforming strategy exhibits near-optimal performance. Motivated by these results we aimed to identify similar designs for systems with low-resolution DACs. Our main contribution is the derivation of a gradient-based and a low-complexity suboptimal heuristic solution for coarsely quantized MU-MISO systems.

The rest of the paper is organized as follows: The quantized system model is introduced in Section II, the design of an optimal uniform quantizer is presented in Section III and the linearized system model is derived in Sections IV and V. In Sections VI and VII the optimization problem is formulated and solved. Finally, simulation results are presented in Section VIII.

A. Notation

Scalars are denoted by italic letters, vectors by lower case bold italic letters and matrices by upper case bold italic letters. $(\bullet)^*, (\bullet)^{\mathrm{T}}, (\bullet)^{\mathrm{H}}, \mathbb{E}[\bullet], \mathrm{tr}(\bullet), \| \bullet \|_2, \| \bullet \|_{\mathrm{F}}, \Re\{\bullet\}$ and $\Im\{\bullet\}$ are used for the complex conjugate, transpose, conjugate transpose, expectation, trace, Euclidean norm, Frobenius norm, real and imaginary part. We use $\mathrm{diag}\,(\boldsymbol{A})$ to denote diagonal matrix containing only the diagonal elements of $\boldsymbol{A}$ and $\mathrm{nondiag}\,(\boldsymbol{A}) = \boldsymbol{A} - \mathrm{diag}\,(\boldsymbol{A})$.

II. Quantized system model

Figure 1 shows the channel model of the downlink of a single-cell scenario, where the base station (BS) has N antennas serving K single antenna users. The signal for each user $x_k \in \mathbb{C}$ is precoded with a beamforming vector $\boldsymbol{p}_k \in \mathbb{C}^N$. The precoded output vector $\boldsymbol{u} \in \mathbb{C}^N$ is given as

$$\boldsymbol{u} = \sum_{k=1}^{K} \boldsymbol{p}_k x_k = \boldsymbol{P}\boldsymbol{x}, \tag{1}$$

where $\boldsymbol{P} = \left[\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_K\right] \in \mathbb{C}^{N\times K}$ and $\boldsymbol{x} = \left[x_1, x_2, \ldots, x_K\right]^{\mathrm{T}} \in \mathbb{C}^K$. Without loss of generality, the variance of signal $\boldsymbol{x}$ is taken as $\mathbb{E}\left[\boldsymbol{x}\boldsymbol{x}^{\mathrm{H}}\right] = \boldsymbol{I}_K$. In our system, the real parts $u_{i,R}$ and the imaginary parts $u_{i,I}$ of the unquantized precoded output $u_i, 1 \le i \le N$ are each quantized by b-bit resolution quantizer. Thus, the resulting quantized signal is read as:

$$v_{i,c} = \mathcal{Q}\left(u_{i,c}\right) = u_{i,c} + q_{i,c}, c \in \{R, I\}\,,\ 1 \le i \le N, \tag{2}$$

where $\mathcal{Q}\left(\bullet\right)$ denotes the quantization operation. Now, let $u_i = u_{i,R} + \mathrm{j}u_{i,I}, q_i = q_{i,R} + \mathrm{j}q_{i,I}$ and $v_i = v_{i,R} + \mathrm{j}v_{i,I}$ be the complex input, the complex quantization error and the complex output, respectively, of the i-th antenna. Eq. 2 can be written in vector form as

$$\boldsymbol{v} = \boldsymbol{u} + \boldsymbol{q}, \tag{3}$$

with $\boldsymbol{v} = \left[u_1, u_2, \ldots, u_N\right]^{\mathrm{T}}, \boldsymbol{q} = \left[q_1, q_2, \ldots, q_N\right]^{\mathrm{T}}, \boldsymbol{v} = \left[v_1, v_2, \ldots, v_N\right]^{\mathrm{T}} \in \mathbb{C}^N$. The signal received by the k^{th} user takes the form

$$y_k = \boldsymbol{h}_k^T \boldsymbol{v} + \eta_k. \tag{4}$$

Here $\boldsymbol{h}_k \in \mathbb{C}^N$ is the channel vector between the k^{th} user and the BS, and η_k is the zero mean additive white Gaussian noise with variance $\sigma_{\eta_k}^2 = \mathbb{E}\left[|\eta_k|^2\right]$.

III. Optimal Quantizer

The quantizer design is based on the minimization of the mean square distortion between the input $u_{i,c}$ and the output $v_{i,c}$ of each quantizer, i.e.,

$$\Delta_{\mathrm{opt}_{i,c}} = \underset{\Delta_{i,c}}{\mathrm{argmin}}\, \mathbb{E}\left[\left(v_{i,c} - u_{i,c}\right)^2\right] = \underset{\Delta_{i,c}}{\mathrm{argmin}}\, \mathbb{E}\left[q_{i,c}^2\right]. \tag{5}$$

Each quantization process has a distortion factor $\rho_q^{(i,c)}$ to indicate the relative amount of quantization noise generated, which is defined as follows

$$\rho_q^{(i,c)} = \frac{\mathbb{E}\left[q_{i,c}^2\right]}{\sigma_{u_{i,c}}^2}, \tag{6}$$

where $\sigma_{u_{i,c}}^2 = \mathbb{E}\left[u_{i,c}^2\right]$. The distortion factor $\rho_q^{(i,c)}$ depends on the number of quantization bits b, the quantizer type (uniform or non-uniform) and the probability density function of $u_{i,c}$. Under this optimal design 5 of the scalar finite resolution quantizer, whether uniform or not, the following equations hold for all $0 \le i \le N, c \in \{R, I\}$ [14], [15]:

$$\mathbb{E}[q_{i,c}] = 0, \tag{7}$$

$$\mathbb{E}[v_{i,c}q_{i,c}] = 0, \tag{8}$$

$$\mathbb{E}[u_{i,c}q_{i,c}] = -\rho_q^{(i,c)}\sigma_{u_{i,c}}^2, \tag{9}$$

where (9) results from (6) and (8). For the uniform quantizer case, (7) holds only if the probability density function of $u_{i,c}$ is even.

For a large number of users, the quantizer input signals $u_{i,c}$ are approximately Gaussian distributed and thus, they undergo nearly the same distortion factor ρ_q, i.e., $\rho_q^{(i,c)} = \rho_q, \forall i, \forall c$. Under the assumption of uncorrelated real and imaginary part of u_i, we easily obtain:

$$\sigma_{q_i}^2 = \mathbb{E}[q_i q_i^*] = \rho_q \sigma_{u_i}^2, \tag{10}$$

$$r_{u_i q_i} = \mathbb{E}[u_i q_i^*] = -\rho_q \sigma_{u_i}^2. \tag{11}$$

ISBN 978-3-8007-4177-9

IV. Computation of Covariance Matrices

In order to derive the linearized system model with uncorrelated quantization noise in (4), we need the convariance matrices involving the unquantized input $\boldsymbol{u}$, the quantized outptut signal $\boldsymbol{v}$ and the distortion $\boldsymbol{q}$. Using (3), the covariance matrices can be written as

$$\begin{aligned}\boldsymbol{R_{vv}} &= \mathbb{E}\left[(\boldsymbol{u}+\boldsymbol{q})\left(\boldsymbol{u}^{\mathrm{H}}+\boldsymbol{q}^{\mathrm{H}}\right)\right] \\ &= \boldsymbol{R_{uu}}+\boldsymbol{R_{uq}}+\boldsymbol{R_{uq}}^{\mathrm{H}}+\boldsymbol{R_{qq}} \quad (12)\\ \boldsymbol{R_{uv}} &= \mathbb{E}\left[\boldsymbol{u}\left(\boldsymbol{u}^{\mathrm{H}}+\boldsymbol{q}^{\mathrm{H}}\right)\right]=\boldsymbol{R_{uu}}+\boldsymbol{R_{uq}}. \quad (13)\end{aligned}$$

Hence, $\boldsymbol{R_{uq}}$ and $\boldsymbol{R_{qq}}$ have to be computed. For $i \neq j$

$$\begin{aligned}[\boldsymbol{R_{uq}}]_{ij} &= r_{u_iq_j}=\mathbb{E}\left[u_iq_j^*\right] \\ &= \mathbb{E}_{u_j}\left[\mathbb{E}\left[u_iq_j^*|u_j\right]\right] \\ &\overset{(a)}{=} \mathbb{E}_{u_j}\left[\mathbb{E}[u_i|u_j]\,\mathbb{E}\left[q_j^*|u_j\right]\right] \\ &\overset{(b)}{\approx} \mathbb{E}_{u_j}\left[r_{u_iu_j}\sigma_{u_j}^{-2}u_j\mathbb{E}\left[q_j^*|u_j\right]\right] \\ &= r_{u_iu_j}\sigma_{u_j}^{-2}\mathbb{E}\left[u_jq_j^*\right] \\ &\overset{(c)}{=} -\rho_q r_{u_iu_j}, \quad (14)\end{aligned}$$

where (a) results from the fact that the quantization error q_j, conditioned on u_j, is statistically independent from all other random variables, (b) follows by approximating the Bayesian estimator with the linear estimator (which is accurate if $\boldsymbol{u}$ is jointly Gaussian distributed) and (c) follows from (9). Hence, from (9) and (14) we get

$$\boldsymbol{R_{uq}} \approx -\rho_q\boldsymbol{R_{uu}}. \quad (15)$$

Therefore

$$\boldsymbol{R_{uv}} \approx (1-\rho_q)\,\boldsymbol{R_{uu}}=\alpha_q\boldsymbol{R_{uu}}, \quad (16)$$

where $\alpha_q = 1-\rho_q$. Similarly, for $i \neq j$

$$\begin{aligned}[\boldsymbol{R_{qq}}]_{ij} &= r_{q_iq_j}=\mathbb{E}\left[q_iq_j^*\right] \\ &= \mathbb{E}_{u_j}\left[\mathbb{E}\left[q_iq_j^*|u_j\right]\right] \\ &= \mathbb{E}_{u_j}\left[\mathbb{E}[q_i|u_j]\,\mathbb{E}\left[q_j^*|u_j\right]\right] \\ &\overset{(a)}{\approx} \mathbb{E}_{u_j}\left[r_{q_iu_j}\sigma_{u_j}^{-2}u_j\mathbb{E}\left[q_j^*|u_j\right]\right] \\ &= r_{q_iu_j}\sigma_{u_j}^{-2}\mathbb{E}\left[u_jq_j^*\right] \\ &\overset{(b)}{=} -\rho_q r_{q_iu_j}=-\rho_q\mathbb{E}\left[q_iu_j^*\right]=-\rho_q\left(\mathbb{E}[u_jq_i^*]\right)^* \\ &\overset{(c)}{\approx} -\rho_q\left(-\rho_q r_{u_ju_i}\right)^*=\rho_q^2r_{u_ju_i}^*=\rho_q^2r_{u_iu_j}, \quad (17)\end{aligned}$$

where (a) follows from approximating the Bayesian estimator with the linear one, (b) follows from (11) and (c) follows from (14). So, from (6) and (17)

$$\begin{aligned}\boldsymbol{R_{qq}} &\approx \rho_q\,\mathrm{diag}\left(\boldsymbol{R_{uu}}\right)+\rho_q^2\,\mathrm{nondiag}\left(\boldsymbol{R_{uu}}\right) \\ &= \rho_q\boldsymbol{R_{uu}}-\left(1-\rho_q\right)\rho_q\,\mathrm{nondiag}\left(\boldsymbol{R_{uu}}\right) \\ &= \rho_q\boldsymbol{R_{uu}}-\alpha_q\rho_q\,\mathrm{nondiag}\left(\boldsymbol{R_{uu}}\right). \quad (18)\end{aligned}$$

From (12), (15) and (18) we obtain

$$\begin{aligned}\boldsymbol{R_{vv}} &\approx \alpha_q\left(\boldsymbol{R_{uu}}-\rho_q\,\mathrm{nondiag}\left(\boldsymbol{R_{uu}}\right)\right) \\ &= \alpha_q^2\boldsymbol{R_{uu}}+\alpha_q\rho_q\,\mathrm{diag}\left(\boldsymbol{R_{uu}}\right). \quad (19)\end{aligned}$$

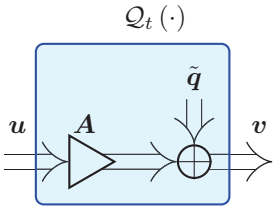

Fig. 2. Bussgang decomposition of the quantizer.

V. Linearized System Model Using Bussgang Decomposition

According to the Bussgang theorem [16], a nonlinear function with Gaussian input can be modeled as a linear function consisting of a linear transformation of the input signal and an additive distortion that is uncorrelated with the input. Hence, for the quantizer $\mathcal{Q}(\bullet)$ with input $\boldsymbol{u} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R_{uu}}) \in \mathbb{C}^N$ we can write

$$\begin{aligned}\boldsymbol{v} &= \mathcal{Q}(\boldsymbol{u}) \\ &= \boldsymbol{A}\boldsymbol{u}+\tilde{\boldsymbol{q}}. \quad (20)\end{aligned}$$

The Bussgang decomposition of the quantizer is depicted in Fig. 2. $\boldsymbol{A}$ can be computed from the requirement that the distortion is uncorrelated with the input:

$$\begin{aligned}\mathbb{E}\left[\tilde{\boldsymbol{q}}\boldsymbol{u}^{\mathrm{H}}\right] &= \mathbb{E}\left[(\boldsymbol{v}-\boldsymbol{A}\boldsymbol{u})\,\boldsymbol{u}^{\mathrm{H}}\right]=\boldsymbol{0}_{N\times N} \\ \Rightarrow \boldsymbol{A} &= \boldsymbol{R_{vu}}\boldsymbol{R_{uu}}^{-1}=\boldsymbol{R_{uv}}^{\mathrm{H}}\boldsymbol{R_{uu}}^{-1} \\ &\overset{(16)}{\approx} \alpha_q\boldsymbol{I}_N. \quad (21)\end{aligned}$$

The covariance of the distortion $\tilde{\boldsymbol{q}}$ reads as

$$\begin{aligned}\boldsymbol{R_{\tilde{q}\tilde{q}}} &= \mathbb{E}\left[(\boldsymbol{v}-\boldsymbol{A}\boldsymbol{u})\left(\boldsymbol{v}^{\mathrm{H}}-\boldsymbol{u}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\right)\right] \\ &= \boldsymbol{R_{vv}}-\boldsymbol{R_{uv}}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}-\boldsymbol{A}\boldsymbol{R_{uv}}+\boldsymbol{A}\boldsymbol{R_{uu}}\boldsymbol{A}^{\mathrm{H}} \\ &\overset{(16),(21)}{\approx} \boldsymbol{R_{vv}}-\alpha_q^2\boldsymbol{R_{uu}} \\ &\overset{(19)}{\approx} \alpha_q^2\boldsymbol{R_{uu}}+\alpha_q\rho_q\,\mathrm{diag}\left(\boldsymbol{R_{uu}}\right)-\alpha_q^2\boldsymbol{R_{uu}} \\ &= \alpha_q\rho_q\,\mathrm{diag}\left(\boldsymbol{R_{uu}}\right). \quad (22)\end{aligned}$$

Note that, although the covariance matrix of the distortion $\tilde{\boldsymbol{q}}$ is known, its distribution isn't. The linearized system model with uncorrelated quantization noise by using Bussgang decomposition yields

$$y_k = \boldsymbol{h}_k^{\mathrm{T}}\left(\alpha_q\sum_{i=1}^{K}\boldsymbol{p}_ix_i+\tilde{\boldsymbol{q}}\right)+\eta_k, \quad (23)$$

with

$$\begin{aligned}\boldsymbol{R_{\tilde{q}\tilde{q}}} &\overset{(22)}{\approx} \alpha_q\rho_q\,\mathrm{diag}\left(\boldsymbol{R_{uu}}\right) \\ &= \alpha_q\rho_q\,\mathrm{diag}\left(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\right), \quad (24)\end{aligned}$$

and it is depicted in Fig. 3.

ISBN 978-3-8007-4177-9

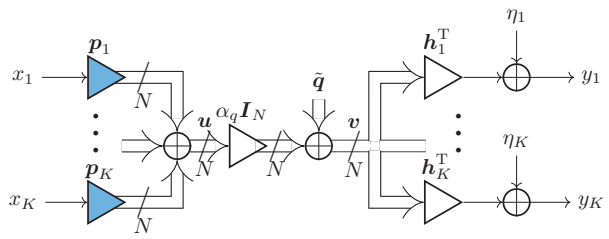

Fig. 3. Linearized model of a coarsely quantized MU-MISO system.

VI. Optimization Problem

Our aim is to find the beamforming vectors that maximize the WSR of the system under a transmit power constraint:

$$\max_{\{\boldsymbol{p}_k\}_{k=1}^K} \sum_{k=1}^K w_k I\left(x_k; y_k\right) \quad \text{s.t. } \mathbb{E}\left[\|\boldsymbol{v}\|_2^2\right] \leq P_{\text{tr}}. \tag{25}$$

We can't write an analytically tractable expression for the rate $I\left(x_k; y_k\right) = h(y_k) - h(y_k|x_k)$ of user k, but, assuming Gaussian input x_k, we can derive a lower bound for it. First, $h(y_k|x_k)$ can be upper bounded as follows:

$$\begin{aligned} h\left(y_k|x_k\right) &= h\left(\boldsymbol{h}_k^{\mathrm{T}}\left(\alpha_q \sum_{i=1}^K \boldsymbol{p}_i x_i + \tilde{\boldsymbol{q}}\right) + \eta_k \middle| x_k\right) \\ &= h\left(\boldsymbol{h}_k^{\mathrm{T}}\left(\alpha_q \sum_{i=1, i\neq k}^K \boldsymbol{p}_i x_i + \tilde{\boldsymbol{q}}\right) + \eta_k \middle| x_k\right) \\ &\leq h\left(\boldsymbol{h}_k^{\mathrm{T}}\left(\alpha_q \sum_{i=1, i\neq k}^K \boldsymbol{p}_i x_i + \tilde{\boldsymbol{q}}\right) + \eta_k\right) \\ &= h\left(\eta_k'\right), \end{aligned} \tag{26}$$

where equality holds if $\tilde{\boldsymbol{q}}$ and x_k are independent. Now, knowing that, under second moment constraints, the Gaussian distributed noise is the mutual information minimizing [17], we assume the quantization noise, and, hence, the effective noise η_k' to be Gaussian distributed to get

$$I\left(x_k; y_k\right) \geq \log_2\left(\pi \,\mathrm{e}\, \sigma_{y_k}^2\right) - \log_2\left(\pi \,\mathrm{e}\, \sigma_{\eta_k'}^2\right), \tag{27}$$

with

$$\sigma_{y_k}^2 = \alpha_q^2 \sum_{i=1}^K \left|\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_i\right|^2 + \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}} \boldsymbol{h}_k^* + \sigma_{\eta_k}^2 \tag{28}$$

$$\sigma_{\eta_k'}^2 = \alpha_q^2 \sum_{i=1, i\neq k}^K \left|\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_i\right|^2 + \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}} \boldsymbol{h}_k^* + \sigma_{\eta_k}^2. \tag{29}$$

Hence

$$I\left(x_k; y_k\right) \geq \log_2\left(1 + \text{SIDNR}_k\right), \tag{30}$$

where

$$\text{SIDNR}_k = \frac{\alpha_q^2 \left|\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_k\right|^2}{\alpha_q^2 \sum_{i=1, i\neq k}^K \left|\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_i\right|^2 + \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}} \boldsymbol{h}_k^* + \sigma_{\eta_k}^2}. \tag{31}$$

Multiplying (30) and summing over all $k = 1, \ldots, K$ we get

$$\sum_{k=1}^K w_k I\left(x_k; y_k\right) \geq \sum_{k=1}^K w_k \log_2\left(1 + \text{SIDNR}_k\right). \tag{32}$$

Finally, instead of maximizing the actual WSR, we maximize the weighted sum of the lower bounds, or else, the lower bound on the WSR:

$$\max_{\{\boldsymbol{p}_k\}_{k=1}^K} \sum_{k=1}^K w_k \log_2\left(1 + \text{SIDNR}_k\right) \quad \text{s.t. } \mathbb{E}\left[\|\boldsymbol{v}\|_2^2\right] \leq P_{\text{tr}}, \tag{33}$$

This problem is non-convex as it has a non-convex objective function. Note that the Karush-Kuhn-Tucker (KKT) conditions are necessary for the optimal solution.

VII. Solution to Optimization Problem

In this section, first the globally optimal solution of problem (33) is discussed and then two suboptimal solutions are presented.

A. Optimal Solution

The elements of the (diagonal) covariance matrix of the distortion can be expressed as

$$\left[\boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}}\right]_{nn} = \alpha_q \rho_q \left\|\boldsymbol{T}_n \boldsymbol{P}\right\|_{\mathrm{F}}^2, \quad n = 1, \ldots, N, \tag{34}$$

where $\boldsymbol{T}_n = \boldsymbol{e}_n \boldsymbol{e}_n^{\mathrm{T}}$ and $\boldsymbol{e}_n$ is a vector whose n-th entry is equal to 1 and the rest are equal to zero. The power of the transmit signal $\boldsymbol{v}$ can be written as

$$\begin{aligned} \mathbb{E}\left[\|\boldsymbol{v}\|_2^2\right] &= \mathrm{tr}\left(\alpha_q^2 \boldsymbol{R}_{\boldsymbol{u}\boldsymbol{u}} + \boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}}\right) \\ &= \alpha_q^2 \sum_{k=1}^K \|\boldsymbol{p}_k\|_2^2 + \alpha_q \rho_q \sum_{n=1}^N \left\|\boldsymbol{T}_n \boldsymbol{P}\right\|_{\mathrm{F}}^2. \end{aligned} \tag{35}$$

Introducing the auxiliary variables $\gamma_k, k = 1, \ldots, K$ and $t_n, n = 1, \ldots, N$ and observing that the phase of $\boldsymbol{v}_k$ can be selected arbitrarily, the optimization problem (33) can be reformulated as

$$\max_{\{\boldsymbol{p}_k, \gamma_k\}_{k=1}^K, \{t_n\}_{n=1}^N} \sum_{k=1}^K w_k \log_2\left(1 + \gamma_k\right) \tag{36a}$$

$$\text{s.t. } \sqrt{\alpha_q^2 \sum_{i=1}^K \left|\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_i\right|^2 + \sum_{n=1}^N t_n^2 \left|\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{e}_n\right|^2 + \sigma_{\eta_k}^2}$$

$$\leq \sqrt{\frac{\gamma_k + 1}{\gamma_k}} \alpha_q \Re\left\{\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_k\right\}, \forall k, \tag{36b}$$

$$\Im\left\{\boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_k\right\} = 0, \forall k, \tag{36c}$$

$$\alpha_q^2 \sum_{k=1}^K \|\boldsymbol{p}_k\|_2^2 + \alpha_q \rho_q \sum_{n=1}^N \left\|\boldsymbol{T}_n \boldsymbol{P}\right\|_{\mathrm{F}}^2 \leq P_{\text{tr}}, \tag{36d}$$

$$\sqrt{\alpha_q \rho_q} \left\|\boldsymbol{T}_n \boldsymbol{P}\right\|_{\mathrm{F}} \leq t_n, \forall n, \tag{36e}$$

where (36b), (36d) and (36e) are met with equality at the optimal point. As already mentioned, (36) is a non-convex monotonic optimization problem, that can be optimally solved

ISBN 978-3-8007-4177-9

using the BRB algorithm [12]. Getting into the details of this algorithm is outside the scope of this work, but it suffices to say that, exploiting the monotonicity of the objective function, it approximates the Pareto boundary around the optimal solution. The Pareto boundary is identified by solving at each iteration a series of quasi-convex optimization problems, whose constraints are identical to those of (36). The complexity of the algorithm scales exponentially with the number of users; hence, it is inapplicable to practical scenarios and can only be used as a benchmark. Therefore, suboptimal alternatives have to be considered.

B. Gradient-based solution

At first, we rewrite the term corresponding to the quantization distortion in the denominator of the SIDNR$_k$ as

$$\begin{aligned}\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}}\boldsymbol{h}_k^* &= \alpha_q\rho_q\boldsymbol{h}_k^{\mathrm{T}}\operatorname{diag}\left(\sum_{i=1}^{K}\boldsymbol{p}_i\boldsymbol{p}_i^{\mathrm{H}}\right)\boldsymbol{h}_k^* \\ &= \alpha_q\rho_q\operatorname{tr}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}\operatorname{diag}\left(\sum_{i=1}^{K}\boldsymbol{p}_i\boldsymbol{p}_i^{\mathrm{H}}\right)\right) \\ &= \alpha_q\rho_q\operatorname{tr}\left(\operatorname{diag}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}\right)\sum_{i=1}^{K}\boldsymbol{p}_i\boldsymbol{p}_i^{\mathrm{H}}\right) \\ &= \alpha_q\rho_q\sum_{i=1}^{K}\boldsymbol{p}_i^{\mathrm{H}}\operatorname{diag}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}\right)\boldsymbol{p}_i, \end{aligned} \tag{37}$$

where the trace identities $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA})$ and $\operatorname{tr}(\boldsymbol{A}\operatorname{diag}(\boldsymbol{B})) = \operatorname{tr}(\operatorname{diag}(\boldsymbol{A})\boldsymbol{B})$ were used. Now (28) and (29) can be rewritten as

$$\sigma_{y_k}^2 = \alpha_q\sum_{i=1}^{K}\boldsymbol{p}_i^{\mathrm{H}}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}} - \rho_q\operatorname{nondiag}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}\right)\right)\boldsymbol{p}_i + \sigma_{\eta_k}^2 \tag{38}$$

$$\sigma_{\eta_k'}^2 = \sigma_{y_k}^2 - \alpha_q^2\left|\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{p}_k\right|^2. \tag{39}$$

According to (27), the lower bound on the WSR can be expressed as

$$S = \sum_{k=1}^{K} w_k\left(\log_2\left(\pi\,\mathrm{e}\,\sigma_{y_k}^2\right) - \log_2\left(\pi\,\mathrm{e}\,\sigma_{\eta_k'}^2\right)\right) \tag{40}$$

and its gradient is found to be

$$\begin{aligned}\frac{\partial S}{\partial\boldsymbol{p}_k^*} = \frac{1}{\ln 2}\Bigg[&\frac{w_k\alpha_q^2}{\sigma_{\eta_k'}^2}\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}} + \sum_{i=1}^{K}\frac{w_i\alpha_q\left(\sigma_{\eta_i'}^2 - \sigma_{y_i}^2\right)}{\sigma_{y_i}^2\sigma_{\eta_i'}^2} \\ &\cdot\left(\boldsymbol{h}_i^*\boldsymbol{h}_i^{\mathrm{T}} - \rho_q\operatorname{nondiag}\left(\boldsymbol{h}_i^*\boldsymbol{h}_i^{\mathrm{T}}\right)\right)\Bigg]\boldsymbol{p}_k. \end{aligned} \tag{41}$$

We express the power of the transmit signal as

$$\begin{aligned}\mathbb{E}\left[\|\boldsymbol{v}\|_2^2\right] &= \operatorname{tr}\left(\alpha_q^2\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}} + \alpha_q\rho_q\operatorname{diag}\left(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\right)\right) \\ &= \operatorname{tr}\left(\alpha_q\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}} - \alpha_q\rho_q\operatorname{nondiag}\left(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\right)\right) \\ &= \alpha_q\operatorname{tr}\left(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\right). \end{aligned} \tag{42}$$

Input: $\mu > 0, \epsilon > 0, \boldsymbol{P}, \boldsymbol{P}_{\text{old}} : \|\boldsymbol{P} - \boldsymbol{P}_{\text{old}}\|_{\mathrm{F}} > \epsilon$
1: **while** $\|\boldsymbol{P} - \boldsymbol{P}_{\text{old}}\|_{\mathrm{F}} > \epsilon$ **do**
2: $\quad\boldsymbol{P}_{\text{old}} \leftarrow \boldsymbol{P}$
3: $\quad\boldsymbol{P} \leftarrow \boldsymbol{P} + \mu\frac{\partial S}{\partial\boldsymbol{P}^*}$
4: $\quad\zeta_n \leftarrow \sqrt{\frac{P_{tr}}{\alpha_q\operatorname{tr}\left(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\right)}}$
5: $\quad\boldsymbol{P} \leftarrow \zeta_n\boldsymbol{P}$
6: **end while**
Output: $\boldsymbol{P}$

Fig. 4. Gradient-projection algorithm for the computation of a locally optimal solution

Using (41) and (42), a locally optimal solution can be obtained through the gradient projection algorithm described in Fig. 4.

C. Heuristic solution

Considering again the optimization problem as posed in (33), with the power of the transmit signal expressed as in (42), the dual feasiblilty KKT condition of the problem is expressed as

$$\frac{\partial S}{\partial\boldsymbol{p}_k^*} - \mu\alpha_q\boldsymbol{p}_k \overset{!}{=} \boldsymbol{0}, \quad \mu \geq 0, \quad k = 1,\ldots,K. \tag{43}$$

Setting $\mu' = \mu\ln 2$

$$\begin{aligned}-\boldsymbol{p}_k + \Bigg[&\frac{w_k\alpha_q}{\mu'\sigma_{\eta_k'}^2}\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}} - \sum_{i=1}^{K}\frac{w_i\left(\sigma_{y_i}^2 - \sigma_{\eta_i'}^2\right)}{\mu'\sigma_{y_i}^2\sigma_{\eta_i'}^2} \\ &\cdot\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)\Bigg]\boldsymbol{p}_k \overset{!}{=} 0 \end{aligned} \tag{44}$$

where $\tilde{\boldsymbol{H}}_k = \boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}$. Now setting

$$c_k = \frac{w_k\alpha_q}{\mu'\sigma_{\eta_k'}^2}\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{p}_k \tag{45}$$

$$\lambda_i = \frac{w_i\sigma_{\eta_i}^2\left(\sigma_{y_i}^2 - \sigma_{\eta_i'}^2\right)}{\mu'\sigma_{y_i}^2\sigma_{\eta_i'}^2} \tag{46}$$

we get

$$\left[\boldsymbol{I}_N + \sum_{i=1}^{K}\frac{\lambda_i}{\sigma_{\eta_i}^2}\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)\right]\boldsymbol{p}_k = c_k\boldsymbol{h}_k^*. \tag{47}$$

Finally, setting

$$\sqrt{P_k} = c_k\left\|\left(\boldsymbol{I}_N + \sum_{i=1}^{K}\frac{\lambda_i\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{\sigma_{\eta_i}^2}\right)^{-1}\boldsymbol{h}_k^*\right\|_2 \tag{48}$$

$$\boldsymbol{p}_k = \frac{\sqrt{P_k}\left(\boldsymbol{I}_N + \sum_{i=1}^{K}\frac{\lambda_i\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{\sigma_{\eta_i}^2}\right)^{-1}\boldsymbol{h}_k^*}{\left\|\left(\boldsymbol{I}_N + \sum_{i=1}^{K}\frac{\lambda_i\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{\sigma_{\eta_i}^2}\right)^{-1}\boldsymbol{h}_k^*\right\|_2}, \tag{49}$$

ISBN 978-3-8007-4177-9

Similarly, to [13], we have now identified the structure of the optimal solution.

The computational complexity of the problem is not reduced, as the computation of the optimal Lagrangian multipliers $\{\lambda_k\}_{k=1}^K$ and the power allocation $\{P_k\}_{k=1}^K$ is still NP-hard, but knowing the structure of the solution, we can derive a suboptimal heuristic solution. In the previous section we found that $\sum_{k=1}^K \lambda_k = \alpha_q \sum_{k=1}^K \|\boldsymbol{p}_k\|_2^2$ and since $\alpha_q \sum_{k=1}^K \|\boldsymbol{p}_k\|_2^2 = P_{\text{tr}}$ at the optimal point,

$$\sum_{k=1}^{K} \lambda_k = P_{\text{tr}}. \tag{50}$$

Instead of finding the optimal Lagrangian multipliers we set $\lambda_k = \lambda = P_{\text{tr}}/K, \forall k$. The resulting beamforming vectors are

$$\begin{aligned}\boldsymbol{p}_{\text{WWFQ},k} &= \frac{\sqrt{P_k}\left[\boldsymbol{I}_N + \sum_{i=1}^{K} \frac{P_{\text{tr}}\left(\tilde{\boldsymbol{H}}_i - \rho_q \,\text{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{K\sigma_{\eta_i}^2}\right]^{-1} \boldsymbol{h}_k^*}{\left\|\left[\boldsymbol{I}_N + \sum_{i=1}^{K} \frac{P_{\text{tr}}\left(\tilde{\boldsymbol{H}}_i - \rho_q \,\text{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{K\sigma_{\eta_i}^2}\right]^{-1} \boldsymbol{h}_k^*\right\|_2} \\ &= \sqrt{P_k}\tilde{\boldsymbol{p}}_k. \end{aligned} \tag{51}$$

We name the solution TxWWFQ, as the beamforming directions are identical to those of TxWFQ [7], but their weights are different. The power assigned to each user P_k still has to be computed. The optimization problem in (33) becomes a power allocation problem

$$\begin{aligned}\max_{P_1,\ldots,P_K} &\sum_{k=1}^{K} w_k \log_2\left(1 + \text{SIDNR}_k\right) \\ \text{s.t. } &\alpha_q \sum_{k=1}^{K} P_k \le P_{\text{tr}} \\ &P_k \ge 0, k = 1, \ldots, K,\end{aligned} \tag{52}$$

where SIDNR_k is given as

$$\begin{aligned}\text{SIDNR}_k &= P_k G_k \\ &= \frac{P_k \left|\alpha_q \boldsymbol{h}_k^{\text{T}} \tilde{\boldsymbol{p}}_k\right|^2}{\sum_{i=1, i\neq k}^{K} P_i \left|\alpha_q \boldsymbol{h}_k^{\text{T}} \tilde{\boldsymbol{p}}_i\right|^2 + \boldsymbol{h}_k^{\text{T}} \boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}} \boldsymbol{h}^* + \sigma_{\eta_k}^2}. \end{aligned} \tag{53}$$

Unfortunately, for the WSR maximization, the power allocation problem is still NP-hard [12]. Therefore, we have to use a heuristic scheme for the power allocation, too. Similarly to [12], by neglecting interference and quantization noise, the problem becomes convex and is easily solved using the waterfilling algorithm.

VIII. Simulation Results

We now wish to compare the algorithms for the maximization of the WSR. We consider the newly proposed heuristic solution taking quantization into account TxWWFQ, the heuristic solution not taking quantization into account, presented in [13], which we will refer to as TxRZFBF, the gradient-based solution and the optimal solution obtained by the BRB algorithm. In our simulations all users have equal weight and equal noise variance.

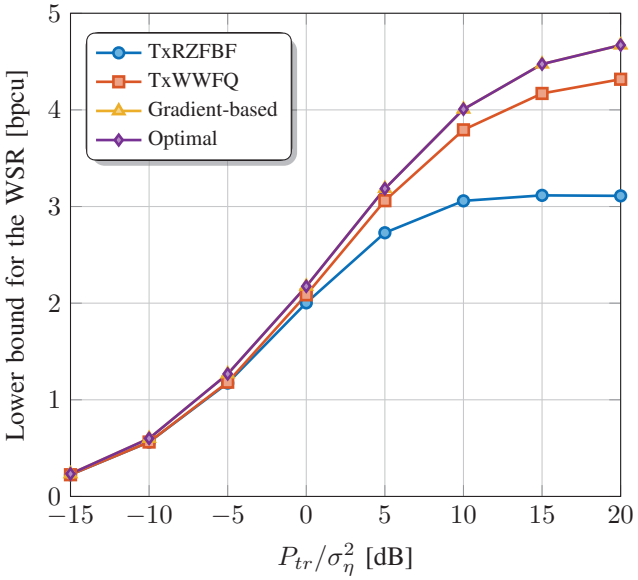

Fig. 5. Lower bound on the sum rate with Gaussian input vs SNR for a MU-MISO system with $N = 4$ transmit antennas, 1-bit DAC and $K = 4$ single-antenna users: TxRZFBF, TxWWFQ, gradient-based solution and optimal solution.

First, we consider a MIMO setup where the transmitter has $N = 4$ antennas and 1-bit DAC and serves $K = 4$ users. In Fig. 5 the lower bound on the sum rate with Gaussian input is plotted as a function of the SNR. The results presented in this figure are for only 1 channel realization, due to the fact that the computation of the optimal solution with the BRB algorithm is extremely time-consuming. Even though this is not necessarily true in general, for the channel realization used in this figure, but also for all the others that we observed, the optimal solution did not outperform the gradient based one. Also, the gradient based solution clearly outperforms TxWWFQ, but this comes at a higher computational cost. In addition, the heuristic solution TxWWFQ, compared to TxZFBF, although having the same computational complexity, offers a significant performance improvement. The observations made here regarding the suboptimal solutions are still valid when averaging over 1000 channel realizations.

Again, for the same setup and 1000 channel realizations, we compare the performance of the filters when a more realistic input distribution is used. Using the toolbox provided in [18], we numerically compute and plot in Fig. 6 the actual sum rate, when the input symbols are drawn from a QPSK constellation. TxWWFQ still performs much better than TxRZFBF, while the gradient based solution offers the highest sum rate.

Keeping all other parameters fixed, we increase the number of transmit antennas to $N = 32$ and we plot again in Fig. 7 the lower bound on the sum rate with Gaussian input, for 1 channel realization. We can see that, except for the TxRZFBF solution, all methods have roughly equal performance. Again,

ISBN 978-3-8007-4177-9

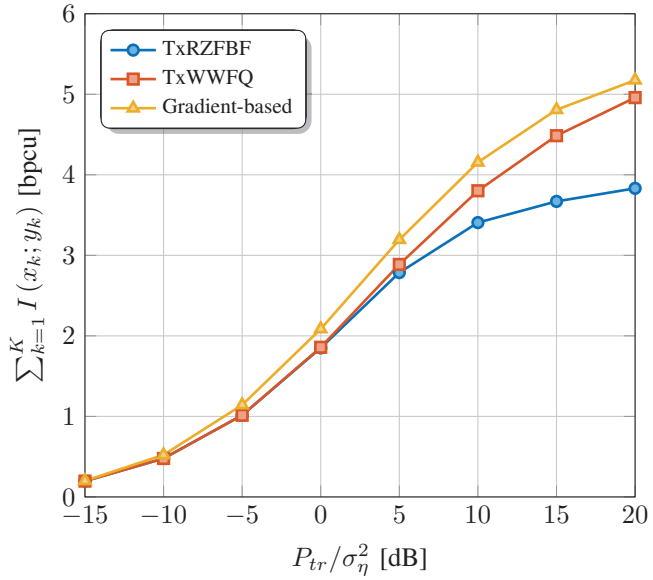

Fig. 6. Sum rate with QPSK input vs SNR for a MU-MISO system with $N = 4$ transmit antennas, 1-bit DAC and $K = 4$ single-antenna users: TxRZFBF, TxWWFQ, gradient-based solution and optimal solution.

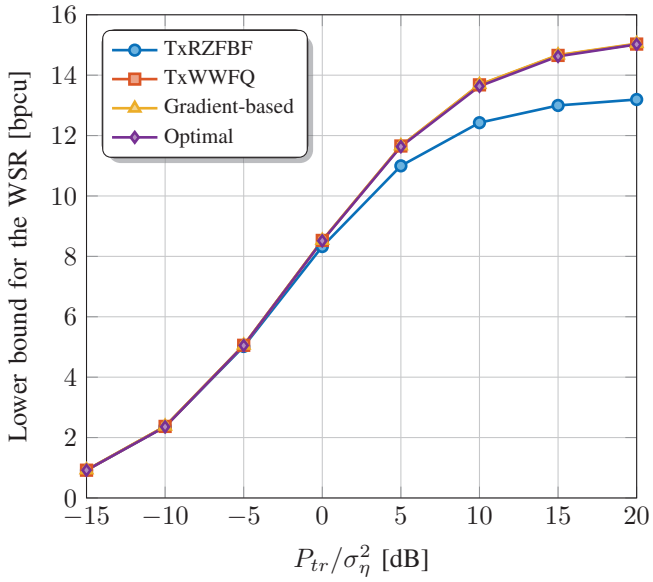

Fig. 7. Lower bound on the sum rate with Gaussian input vs SNR for a MU-MISO system with $N = 32$ transmit antennas, 1-bit DAC and $K = 4$ single-antenna users: TxRZFBF, TxWWFQ, gradient-based solution and optimal solution.

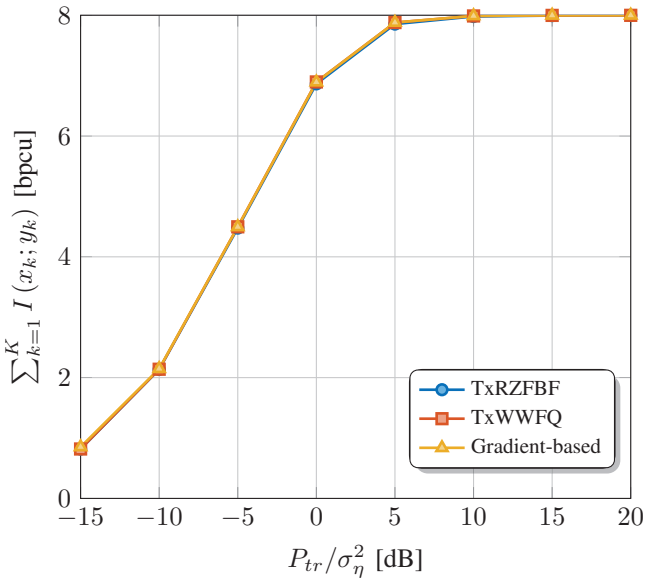

Fig. 8. Sum rate with QPSK input vs SNR for a MU-MISO system with $N = 32$ transmit antennas, 1-bit DAC and $K = 4$ single-antenna users: TxRZFBF, TxWWFQ, gradient-based solution and optimal solution.

this statement is still valid for the suboptimal solutions when the results are averaged over 1000 channel realizations.

Finally, for the same number of channel realizations, we compute the sum rate with QPSK input and plot the results in Fig. 8. Now TxRZFBF achieves the same sum rate as the other two. Therefore, it is suggested from these results that, for QPSK input, when the number of transmit antennas grows larger than the number of users, ignoring quantization in the filter design does not have a negative impact on the performance in terms of the sum rate.

IX. Conclusion

In this paper we have derived two suboptimal beamforming solutions for the maximization of the WSR of a MU-MISO system with low-resolution DACs. We have shown that, when the number of users is close to the number of transmit antennas, the heuristic solution TxWWFQ, which takes quantization into account, offers a significant performance improvement compared to the one that doesn't. When the number of transmit antennas becomes much larger than the number of users, for QPSK input, which is the most natural choice under 1-bit quantization, taking quantization into account doesn't offer any significant gains.

We remind here that the derived solutions aim in maximizing the lower bound on the WSR with Gaussian input, not the sum rate with QPSK input. Therefore, although our results indicate that taking quantization into account becomes less important as the number of antennas increases, further research should be conducted to clarify whether this statement is still valid when an objective function which is more closely related to the QPSK sum rate is used in the filter optimization.

ISBN 978-3-8007-4177-9

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. 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, May 2014.

[2] E. Larsson, O. Edfors, F. Tufvesson, and T. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, February 2014.

[3] T. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. Wong, J. Schulz, M. Samimi, and F. Gutierrez, "Millimeter Wave Mobile Communications for 5G Cellular: It Will Work!" *IEEE Access*, vol. 1, pp. 335–349, May 2013.

[4] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Transactions on Wireless Communications*, vol. 4, no. 5, pp. 2349–2360, Sept 2005.

[5] O. El Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. Heath, "Spatially Sparse Precoding in Millimeter Wave MIMO Systems," *IEEE Transactions on Wireless Communications*, vol. 13, no. 3, pp. 1499–1513, March 2014.

[6] M. Kurras, L. Thiele, and G. Caire, "Interference Mitigation and Multiuser Multiplexing with Beam-Steering Antennas," in *Proceedings of the 19th International ITG Workshop on Smart Antennas; WSA 2015*, March 2015, pp. 1–5.

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

[8] ——, "Tomlinson Harashima Precoding for MIMO Systems with Low Resolution D/A-Converters," in *ITG/IEEE Workshop on Smart Antennas*, February 2008, berlin, Germany.

[9] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Coordinated Beamforming for MISO Interference Channel: Complexity Analysis and Efficient Algorithms," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1142–1157, March 2011.

[10] J. Brehmer and W. Utschick, "Utility Maximization in the Multi-User MISO Downlink with Linear Precoding," in *IEEE International Conference on Communications, 2009. ICC '09.*, June 2009, pp. 1–1.

[11] E. Björnson, G. Zheng, M. Bengtsson, and B. Ottersten, "Robust Monotonic Optimization Framework for Multicell MISO Systems," *IEEE Transactions on Signal Processing*, vol. 60, no. 5, pp. 2508–2523, May 2012.

[12] E. Björnson and E. Jorswieck, "Optimal Resource Allocation in Coordinated Multi-Cell Systems," *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 23, pp. 113–381, 2012. [Online]. Available: http://dx.doi.org/10.1561/0100000069

[13] E. Björnson, M. Bengtsson, and B. Ottersten, "Optimal Multiuser Transmit Beamforming: A Difficult Problem with a Simple Solution Structure [Lecture Notes]," *IEEE Signal Processing Magazine*, vol. 31, no. 4, pp. 142–148, July 2014.

[14] J. Max, "Quantizing for Minimum Distortion," *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7–12, March 1960.

[15] J. G. Proakis, *Digital Communications*. McGraw-Hill, 1995.

[16] J. Bussgang, "Crosscorrelation Functions of Amplitude-Distorted Gaussian Signals," *RLE Technical Reports*, vol. 216, 1952.

[17] S. Diggavi and T. Cover, "The Worst Additive Noise Under a Covariance Constraint," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 3072–3081, Nov 2001.

[18] G. Brown, A. Pocock, M.-J. Zhao, and M. Luján, "Conditional Likelihood Maximisation: A Unifying Framework for Information Theoretic Feature Selection," *Journal of Machine Learning Research*, vol. 13, pp. 27–66, January 2012.

ISBN 978-3-8007-4177-9