Weighted Sum Rate Maximization for Coarsely Quantized Massive MIMO Systems

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). 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 and that the heuristic solution achieves near-optimal performance when a massive Multiple-Input-Multiple-Output setup is considered, namely, when the number of transmit antennas becomes much larger than the number of users.**

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). It becomes an issue that has to be tackled to make these two technologies viable. The power consumed by a DAC has an exponential dependence on the bit resolution b of the converter [4]: $P_{\mathrm{DAC}} \propto 2^b$.

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 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. Our main contribution is the deriva-

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

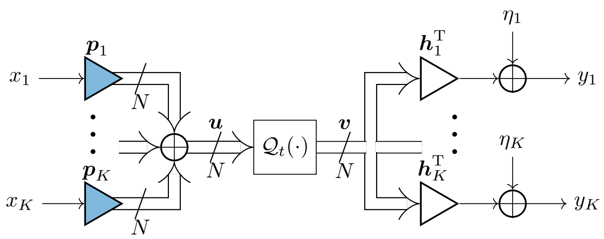

Fig. 1. Downlink beamforming for a MU-MISO system under low resolution DAC.

tion 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 **??** and V. In Sections VI and VII the optimization problem is formulated and solved. Finally, simulation results are presented in Section VIII.

II. Quantized system model

Figure 1 shows the channel model of the downlink of a single-cell scenario, where base station (BS) has N antennas serving K single antenna users. The signal of 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

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

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 x_k is taken as $\mathbb{E}\left[x_k^2\right] = 1$.In our system, the real parts $u_{i,R}$ and the imaginary parts $u_{i,I}$ of the unquantized precoded output $u_i, 1 \leq i \leq N$ are each quantized by b-bit resolution quantizer. Thus the resulting quantized signal is read as:

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

where $\mathcal{Q}\left(.\right)$ denotes the quantization operation. The signal received by the k^{th} user takes the form

$$y_k = \boldsymbol{h}_k^T \boldsymbol{v} + \eta_k. \tag{3}$$

Here $\boldsymbol{v} = \left[v_1, v_2, \ldots, v_N\right] \in \mathbb{C}^N$, $\boldsymbol{h}_k \in \mathbb{C}^N$ 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 $\sigma_{\eta_k}^2 = \mathbb{E}\left[\left|\eta_k\right|^2\right]$.

III. Optimal Quantizer

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

$$\rho_q^{(i,c)} = \frac{\mathbb{E}\left[q_{i,c}^2\right]}{r_{u_{i,c}u_{i,c}}}. \tag{4}$$

Here $r_{u_{i,c}u_{i,c}} = \mathbb{E}\left[u_{i,c}^2\right]$ is the variance of $u_{i,c}$ and the distortion factor $\rho_q^{(i,c)}$ which depends on the number of quantization bits b, the quantizer type (uniform or non-uniform) and the probability density function of $u_{i,c}$. The quantizer is designed based on the minimization of the mean square distortion input $u_{i,c}$ and the output $v_{i,c}$ of each quantizer, i.e.,

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

Under this optimal design of the scalar finite resolution quantizer, whether uniform or not, the following equations hold for all $0 \leq i \leq N, c \in \{R, I\}$ [Ref]:

$$\begin{aligned} \mathbb{E}[q_{i,c}] &= 0 && (6)\\ \mathbb{E}[v_{i,c}q_{i,c}] &= 0 && (7)\\ \mathbb{E}[u_{i,c}q_{i,c}] &= -\rho_q^{(i,c)}\sigma_{u_{i,c}}^2. && (8) \end{aligned}$$

where (8) results from (4) and (7). For the uniform quantizer case, (6) holds only if the probability density function of $u_{i,c}$ is even.

For large number of antennas, the quantizer input signals $u_{i,c}$ are approximately Gaussian distributed and thus, they undergo nearly the same distortion factor ρ_q, i.e., $\rho_q^{(i,c)} = \rho_q, \forall i, \forall c$. Now let, $q_i = q_{i,R} + \mathrm{j}q_{i,I}$ the complex quantization error, complex input $u_i = u_{i,R} + \mathrm{j}u_{i,I}$ and complex output $v_i = v_{i,R} + \mathrm{j}v_{i,I}$. Under the assumption of uncorrelated real and imaginary part of u_i, we easily obtain:

$$\begin{aligned} r_{q_iq_i} &= \mathbb{E}[q_iq_i^*] = \rho_q r_{u_iu_i} && (9)\\ r_{u_iq_i} &= \mathbb{E}[u_iq_i^*] = -\rho_q r_{u_iu_i}. && (10) \end{aligned}$$

The distortion due to quantization operation on all the antennas can be written in vector form as

$$\boldsymbol{q} = \boldsymbol{v} - \boldsymbol{u}. \tag{11}$$

IV. Computation of Covarince Matrices

In order to derive the linearized system model for an uncorrelated quantization noise in (3), we need the convariance matrices involving unquantized input $\boldsymbol{u}$, quantized outptut signal $\boldsymbol{v}$ and distortion $\boldsymbol{q}$. Using (11), the covariance matrices can be written as

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

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

$$\begin{aligned} \left[\boldsymbol{R_{uq}}\right]_{ij} &= r_{u_iq_j} = \mathbb{E}\left[u_iq_j^*\right]\\ &= \mathbb{E}_{u_j}\left[\mathbb{E}\left[u_iq_j^*|u_j\right]\right]\\ &\overset{(a)}{=} \mathbb{E}_{u_j}\left[\mathbb{E}\left[u_i|u_j\right]\mathbb{E}\left[q_j^*|u_j\right]\right]\\ &\overset{(b)}{\approx} \mathbb{E}_{u_j}\left[r_{u_iu_j}r_{u_ju_j}^{-1}u_j\mathbb{E}\left[q_j^*|u_j\right]\right]\\ &= r_{u_iu_j}r_{u_ju_j}^{-1}\mathbb{E}\left[u_jq_j^*\right]\\ &\overset{(c)}{=} -\rho_q r_{u_iu_j}, && (14) \end{aligned}$$

where (a) results from the fact that 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 (8). Hence, from (8) and (14) we get

$$\boldsymbol{R}_{\boldsymbol{uq}} = -\rho_q \boldsymbol{R}_{\boldsymbol{uu}}. \tag{15}$$

Therefore

$$\boldsymbol{R}_{\boldsymbol{uv}} = (1-\rho_q)\,\boldsymbol{R}_{\boldsymbol{uu}} = \alpha_q \boldsymbol{R}_{\boldsymbol{uu}}, \tag{16}$$

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

$$\begin{aligned}
[\boldsymbol{R}_{\boldsymbol{qq}}]_{ij} &= r_{q_i q_j} = \mathbb{E}\left[q_i q_j^*\right] \\
&= \mathbb{E}_{u_j}\left[\mathbb{E}\left[q_i q_j^* | u_j\right]\right] \\
&= \mathbb{E}_{u_j}\left[\mathbb{E}[q_i|u_j]\,\mathbb{E}\left[q_j^*|u_j\right]\right] \\
&\approx \mathbb{E}_{u_j}\left[r_{q_i u_j} r_{u_j u_j}^{-1} u_j \mathbb{E}\left[q_j^*|u_j\right]\right] \\
&= r_{q_i u_j} r_{u_j u_j}^{-1} \mathbb{E}\left[u_j q_j^*\right] \\
&\overset{(4)}{=} -\rho_q r_{q_i u_j} = -\rho_q \mathbb{E}\left[q_i u_j^*\right] = -\rho_q\left(\mathbb{E}[u_j q_i^*]\right)^* \\
&\overset{(14)}{\approx} -\rho_q\left(-\rho_q r_{u_j u_i}\right)^* = \rho_q^2 r_{u_j u_i}^* = \rho_q^2 r_{u_i u_j}.
\end{aligned} \tag{17}$$

So, from (4) and (17)

$$\begin{aligned}
\boldsymbol{R}_{\boldsymbol{qq}} &\approx \rho_q \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{uu}}) + \rho_q^2 \operatorname{nondiag}(\boldsymbol{R}_{\boldsymbol{uu}}) \\
&= \rho_q \boldsymbol{R}_{\boldsymbol{uu}} - (1-\rho_q)\,\rho_q \operatorname{nondiag}(\boldsymbol{R}_{\boldsymbol{uu}}) \\
&= \rho_q \boldsymbol{R}_{\boldsymbol{uu}} - \alpha_q \rho_q \operatorname{nondiag}(\boldsymbol{R}_{\boldsymbol{uu}}).
\end{aligned} \tag{18}$$

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

$$\begin{aligned}
\boldsymbol{R}_{\boldsymbol{vv}} &= \alpha_q\left(\boldsymbol{R}_{\boldsymbol{uu}} - \rho_q \operatorname{nondiag}(\boldsymbol{R}_{\boldsymbol{uu}})\right) \\
&= \alpha_q^2 \boldsymbol{R}_{\boldsymbol{uu}} + \alpha_q \rho_q \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{uu}}).
\end{aligned} \tag{19}$$

V. Linearized System Model Using Bussgang Decomposition

According to the Bussgang theorem [14], a nonlinear function with Gaussian input can be approximated by 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 $\mathcal{Q}(\cdot)$ with input $\boldsymbol{u} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{R}_{\boldsymbol{uu}}) \in \mathbb{C}^N$ we can write

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

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:

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

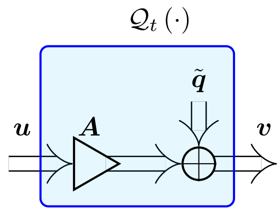

Fig. 2. Bussgang decomposition of the quantizer.

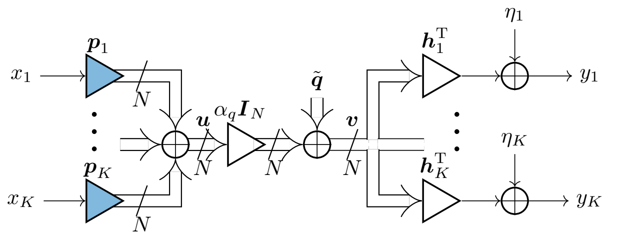

Fig. 3. Linearied model of a coarsely quantized MU-MISO system.

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

$$\begin{aligned}
\boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}} &= \mathbb{E}\left[(\boldsymbol{v}-\boldsymbol{A}\boldsymbol{u})\left(\boldsymbol{v}^{\mathrm{H}} - \boldsymbol{u}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}}\right)\right] \\
&= \boldsymbol{R}_{\boldsymbol{vv}} - {\boldsymbol{R}_{\boldsymbol{uv}}}^{\mathrm{H}}\boldsymbol{A}^{\mathrm{H}} - \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{uv}} + \boldsymbol{A}\boldsymbol{R}_{\boldsymbol{uu}}\boldsymbol{A}^{\mathrm{H}} \\
&\overset{(21)}{=} \boldsymbol{R}_{\boldsymbol{vv}} - {\boldsymbol{R}_{\boldsymbol{uv}}}^{\mathrm{H}}{\boldsymbol{R}_{\boldsymbol{uu}}}^{-1}\boldsymbol{R}_{\boldsymbol{uv}} \\
&= \alpha_q^2 \boldsymbol{R}_{\boldsymbol{uu}} + \alpha_q\rho_q \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{uu}}) - \alpha_q^2 \boldsymbol{R}_{\boldsymbol{uu}} \\
&= \alpha_q \rho_q \operatorname{diag}(\boldsymbol{R}_{\boldsymbol{uu}}).
\end{aligned} \tag{22}$$

Note here that, although the covariance matrix of the distortion $\tilde{\boldsymbol{q}}$ is known, its distribution isn't. The linearized system model (3) with an uncorrelated quantization noise by using Bussgang decomposition yields

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

with

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

and it is depicted in Fig. 3.

VI. Optimization Problem

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

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

We can't write an analytically tractable expression for the rate $I(x_k; y_k) = h(y_k) - h(y_k|x_k)$ 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 [15], 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

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

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

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

$$\max_{\{\boldsymbol{p}_k\}_{k=1}^K} \sum_{k=1}^{K} w_k \log_2\left(1 + \mathrm{SIDNR}_k\right) \quad \text{s.t. } \mathbb{E}\left[\|\boldsymbol{v}\|_2^2\right] \leq P_{\mathrm{tr}}, \tag{30}$$

where

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

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, as the constraint is convex.

VII. Solution to Optimization Problem

In this section, first the globally optimal solution of the problem (30) 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

$$\left[\boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}}\right]_{nn} = \alpha_q \rho_q \left\| \boldsymbol{T}_n \boldsymbol{P} \right\|_{\mathrm{F}}^2, \quad n = 1, \ldots, N, \tag{32}$$

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

$$\begin{aligned} \mathbb{E}\left[\|\boldsymbol{v}\|_2^2\right] &= \mathrm{tr}\left(\alpha_q^2 \boldsymbol{R}_{\boldsymbol{u}\boldsymbol{u}} + \boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}}\right) \\ &= \alpha_q^2 \sum_{k=1}^{K} \|\boldsymbol{p}_k\|_2^2 + \alpha_q \rho_q \sum_{n=1}^{N} \|\boldsymbol{T}_n \boldsymbol{P}\|_{\mathrm{F}}^2. \end{aligned} \tag{33}$$

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 (30) can be reformulated as

$$\max_{\{\boldsymbol{p}_k, \gamma_k\}_{k=1}^K, \{t_n\}_{n=1}^N} \sum_{k=1}^{K} w_k \log_2\left(1 + \gamma_k\right) \tag{34a}$$

$$\text{s.t.} \sqrt{\alpha_q^2 \sum_{i=1}^{K} \left| \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_i \right|^2 + \sum_{n=1}^{N} t_n^2 \left| \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{e}_n \right|^2 + \sigma_{\eta_k}^2} \leq \sqrt{\frac{\gamma_k + 1}{\gamma_k}} \alpha_q \Re\left\{ \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_k \right\}, \forall k, \tag{34b}$$

$$\Im\left\{ \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{p}_k \right\} = 0, \forall k, \tag{34c}$$

$$\alpha_q^2 \sum_{k=1}^{K} \|\boldsymbol{p}_k\|_2^2 + \alpha_q \rho_q \sum_{n=1}^{N} \|\boldsymbol{T}_n \boldsymbol{P}\|_{\mathrm{F}}^2 \leq P_{\mathrm{tr}}, \tag{34d}$$

$$\sqrt{\alpha_q \rho_q} \left\| \boldsymbol{T}_n \boldsymbol{P} \right\|_{\mathrm{F}} \leq t_n, \forall n, \tag{34e}$$

where (34b), (34d) and (34e) are met with equality at the optimal point. As already mentioned, (34) is a non-convex monotonic optimization problem, that can be optimally solved 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 (34). 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

We first rewrite the term corresponding to quantization distortion in the denominator of SIDNR_k as

$$\begin{aligned}\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{R}_{\tilde{\boldsymbol{q}}\tilde{\boldsymbol{q}}}\boldsymbol{h}^* &= \alpha_q\rho_q\boldsymbol{h}_k^{\mathrm{T}}\operatorname{diag}\left(\sum_{i=1}^{K}\boldsymbol{p}_i\boldsymbol{p}_i^{\mathrm{H}}\right)\boldsymbol{h}_k^* \\ &= \alpha_q\rho_q\operatorname{tr}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}\operatorname{diag}\left(\sum_{i=1}^{K}\boldsymbol{p}_i\boldsymbol{p}_i^{\mathrm{H}}\right)\right) \\ &= \alpha_q\rho_q\operatorname{tr}\left(\operatorname{diag}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}\right)\sum_{i=1}^{K}\boldsymbol{p}_i\boldsymbol{p}_i^{\mathrm{H}}\right) \\ &= \alpha_q\rho_q\sum_{i=1}^{K}\boldsymbol{p}_i^{\mathrm{H}}\operatorname{diag}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}\right)\boldsymbol{p}_i, \qquad (35)\end{aligned}$$

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

$$\sigma_{y_k}^2 = \alpha_q\sum_{i=1}^{K}\boldsymbol{p}_i^{\mathrm{H}}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}} - \rho_q\operatorname{nondiag}\left(\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}\right)\right)\boldsymbol{p}_i + \sigma_{\eta_k}^2 \qquad (36)$$

$$\sigma_{\eta_k'}^2 = \sigma_{y_k}^2 - \alpha_q^2\left|\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{p}_k\right|^2. \qquad (37)$$

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

$$S = \sum_{k=1}^{K} w_k\left(\log_2\left(\pi\,\mathrm{e}\,\sigma_{y_k}^2\right) - \log_2\left(\pi\,\mathrm{e}\,\sigma_{\eta_k'}^2\right)\right) \qquad (38)$$

and its gradient is found to be

$$\begin{aligned}\frac{\partial S}{\partial \boldsymbol{p}_k^*} = \frac{1}{\ln 2}\Bigg[&\frac{w_k\alpha_q^2}{\sigma_{\eta_k'}^2}\boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}} + \sum_{i=1}^{K}\frac{w_i\alpha_q\left(\sigma_{\eta_i'}^2 - \sigma_{y_i}^2\right)}{\sigma_{y_i}^2\sigma_{\eta_i'}^2} \\ &\cdot\left(\boldsymbol{h}_i^*\boldsymbol{h}_i^{\mathrm{T}} - \rho_q\operatorname{nondiag}\left(\boldsymbol{h}_i^*\boldsymbol{h}_i^{\mathrm{T}}\right)\right)\Bigg]\boldsymbol{p}_k. \qquad (39)\end{aligned}$$

We express the power of the transmit signal as

$$\begin{aligned}\mathbb{E}\left[\|\boldsymbol{v}\|_2^2\right] &= \operatorname{tr}\left(\alpha_q^2\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}} + \alpha_q\rho_q\operatorname{diag}\left(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\right)\right) \\ &= \operatorname{tr}\left(\alpha_q\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}} - \alpha_q\rho_q\operatorname{nondiag}\left(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\right)\right) \\ &= \alpha_q\operatorname{tr}\left(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\right). \qquad (40)\end{aligned}$$

Using (39) and (40), a locally optimal solution can be obtained through the gradient projection algorithm described in Fig. 4.

Input: $\mu > 0, \epsilon > 0, \boldsymbol{P}, \boldsymbol{P}_{\text{old}} : \|\boldsymbol{P} - \boldsymbol{P}_{\text{old}}\|_{\mathrm{F}} > \epsilon$
while $\|\boldsymbol{P} - \boldsymbol{P}_{\text{old}}\|_{\mathrm{F}} > \epsilon$ **do**
    $\boldsymbol{P}_{\text{old}} = \boldsymbol{P}$
    for $k = 1$ to K **do**
        $\boldsymbol{p}_k \leftarrow \boldsymbol{p}_k + \mu\frac{\partial S}{\partial \boldsymbol{p}_k^*}$
    end for
    $\zeta_n \leftarrow \sqrt{\frac{P_{tr}}{\alpha_q\operatorname{tr}\left(\boldsymbol{P}\boldsymbol{P}^{\mathrm{H}}\right)}}$
    $\boldsymbol{P} \leftarrow \zeta_n\boldsymbol{P}$
end while
Output: $\boldsymbol{P}$

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

C. Heuristic solution

Considering again the optimization problem as posed in (30), with the power of the transmit signal expressed as in (40), the dual feasiblilty KKT condition of the problem is expressed as

$$\frac{\partial S}{\partial \boldsymbol{p}_k^*} - \mu\alpha_q\boldsymbol{p}_k \overset{!}{=} \boldsymbol{0}, \quad \mu \geq 0, \quad k = 1, \ldots, K. \qquad (41)$$

Using (39) and after some algebraic manipulations, we get

$$\boldsymbol{p}_{,k} = \frac{\sqrt{P_k}\left(\boldsymbol{I}_N + \sum_{i=1}^{K}\frac{\lambda_i\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{\sigma_{\eta_i}^2}\right)^{-1}\boldsymbol{h}_k^*}{\left\|\left(\boldsymbol{I}_N + \sum_{i=1}^{K}\frac{\lambda_i\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{\sigma_{\eta_i}^2}\right)^{-1}\boldsymbol{h}_k^*\right\|_2}, \qquad (42)$$

where $\tilde{\boldsymbol{H}}_k = \boldsymbol{h}_k^*\boldsymbol{h}_k^{\mathrm{T}}$ and

$$\lambda_k = \frac{w_k\left(\sigma_{y_k}^2 - \sigma_{\eta_k'}^2\right)\sigma_{\eta_k}^2}{\mu\sigma_{y_k}^2\sigma_{\eta_k'}^2\ln 2} \geq 0 \qquad (43)$$

$$\begin{aligned}P_k = &\left(\frac{w_k\alpha_q\boldsymbol{h}_k^{\mathrm{T}}\boldsymbol{p}_k}{\mu\sigma_{\eta_k'}^2}\right)^2 \\ &\cdot\left\|\left(\boldsymbol{I}_N + \sum_{i=1}^{K}\frac{\lambda_i\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{\sigma_{\eta_i}^2}\right)^{-1}\boldsymbol{h}_k^*\right\|_2^2 \qquad (44)\end{aligned}$$

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 we now know the structure of the solution and we can use it to 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,

$$\sum_{k=1}^{K}\lambda_k = P_{\text{tr}}. \qquad (45)$$

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

$$\begin{aligned}\boldsymbol{p}_{\text{WWFQ},k} &= \frac{\sqrt{P_k}\left[\boldsymbol{I}_N + \sum_{i=1}^{K}\frac{P_{\text{tr}}\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{K\sigma_{\eta_i}^2}\right]^{-1}\boldsymbol{h}_k^*}{\left\|\left[\boldsymbol{I}_N + \sum_{i=1}^{K}\frac{P_{\text{tr}}\left(\tilde{\boldsymbol{H}}_i - \rho_q\operatorname{nondiag}\left(\tilde{\boldsymbol{H}}_i\right)\right)}{K\sigma_{\eta_i}^2}\right]^{-1}\boldsymbol{h}_k^*\right\|_2} \\ &= \sqrt{P_k}\tilde{\boldsymbol{p}}_k. \qquad (46)\end{aligned}$$

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 (30) becomes a power allocation problem

$$\begin{aligned} \max_{P_1,\ldots,P_K} \quad & \sum_{k=1}^{K} w_k \log_2\left(1+\mathrm{SIDNR}_k\right) \\ \text{s.t.} \quad & \alpha_q \sum_{k=1}^{K} P_k \le P_{\mathrm{tr}} \\ & P_k \ge 0, k=1,\ldots,K, \end{aligned} \tag{47}$$

where SIDNR_k is given as

$$\begin{aligned} \mathrm{SIDNR}_k &= P_k G_k \\ &= \frac{P_k \left|\alpha_q \boldsymbol{h}_k^{\mathrm{T}} \tilde{\boldsymbol{p}}_k\right|^2}{\sum_{i=1, i\neq k}^{K} \left|\alpha_q \boldsymbol{h}_k^{\mathrm{T}} \tilde{\boldsymbol{p}}_i\right|^2 + \boldsymbol{h}_k^{\mathrm{T}} \boldsymbol{R}_{\tilde{q}\tilde{q}} \boldsymbol{h}^* + \sigma_{\eta_k}^2} \end{aligned} \tag{48}$$

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], 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 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.

First, a strict MIMO setup is considered, where the transmitter has $N = 4$ antennas, 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. The gradient based solution performs approximately as well as the optimal, but this is not necessarily true in general. Also, it clearly outperforms TxWWFQ, but this comes at a higher computational cost. In addition, the heuristic solution TxWWFQ, compared to TxZFBF, with which the have 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.

Keeping all other parameters fixed, we increase the number of transmit antennas to $N = 32$ and we plot again in Fig. 6 the lower bound on the sum rate with Gaussian input. We can see that, expect for the TxRZFBF solution, all the rest result in a roughly equal performace. Again, this statement is

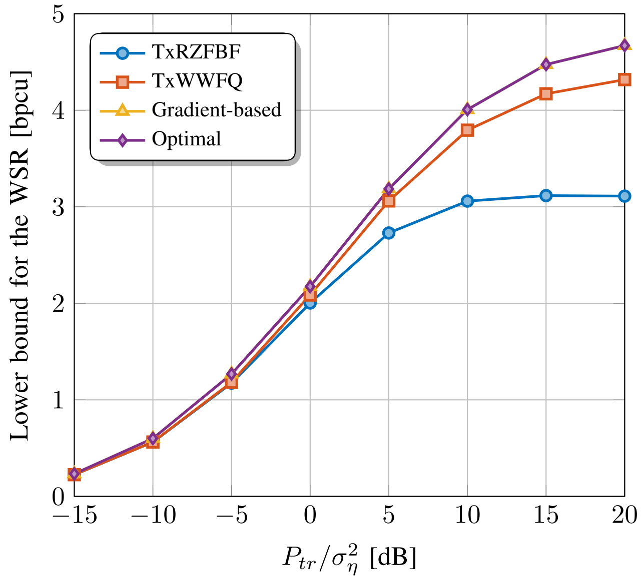

Fig. 5. Lower bound on the sum rate with Gaussian input vs SNR for a MU-MISO system with $N = 4$ transmit antennas, 1-bit DAC and $K = 4$ single-antenna users: TxRZFBF, TxWWFQ, gradient-based solution and optimal solution.

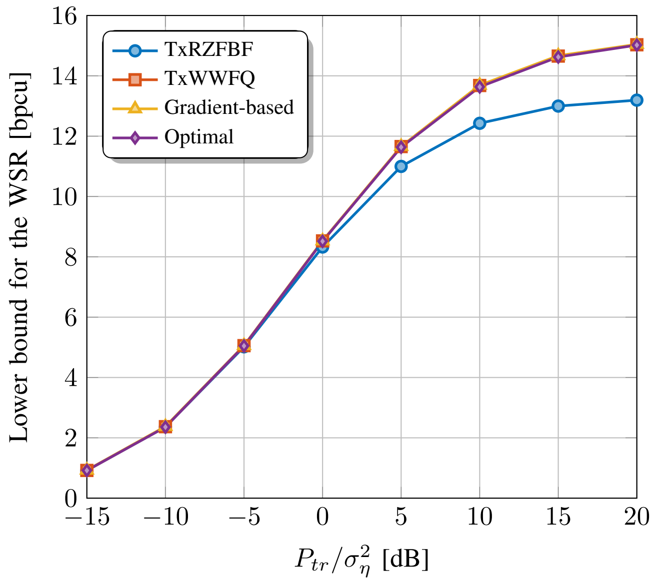

Fig. 6. Lower bound on the sum rate with Gaussian input vs SNR for a MU-MISO system with $N = 32$ transmit antennas, 1-bit DAC and $K = 4$ single-antenna users: TxRZFBF, TxWWFQ, gradient-based solution and optimal solution.

still valid for the suboptimal solutions when the results are averaged over 1000 channel realizations. Hence, for massive MIMO setups, the TxWWFQ solution is the most preferable one, as it offers near-optimal performance with a significantly lower computational cost.

IX. Conclusions and further work

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 the heuristic solution, that takes quantization into account and whose beamforming directions coincide with those of the modified Wiener filter presented in [7], offers a significant performance improvement compared to the one that doesn't take quantization into account. For systems with large transmit antenna arrays this solution achieves a near-optimal performance.

In the final version of this work we plan to include results for the actual WSR of the system, when the input symbols are drawn from a finite constellation. Also, focusing on massive MIMO setups, we aim to compare the linear schemes presented so far with non-linear schemes, such as Dirty Paper Coding. The important question that will be addressed is whether the non-linear precoding strategies can offer higher rates when coarsely quantized systems with very large antenna arrays are considered

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," *Communications Magazine, IEEE*, 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," *Communications Magazine, IEEE*, 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!" *Access, IEEE*, vol. 1, pp. 335–349, 2013.

[4] S. Cui, A. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *Wireless Communications, IEEE Transactions on*, 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," *Wireless Communications, IEEE Transactions on*, 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 *WSA 2015; 19th International ITG Workshop on Smart Antennas; Proceedings of*, 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," *Signal Processing, IEEE Transactions on*, 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 *Communications, 2009. ICC '09. IEEE International Conference on*, June 2009, pp. 1–1.

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

[12] E. Bjrnson 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. Bjornson, 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. Bussgang, "Crosscorrelation Functions of Amplitude-Distorted Gaussian Signals," *RLE Technical Reports*, vol. 216, 1952.

[15] 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.