Branch-and-Bound Precoding for Multiuser MIMO Systems with 1-Bit Quantization

Lukas T. N. Landau, and Rodrigo C. de Lamare
Centro de Estudos em Telecomunicações
Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro CEP 22453-900, Brazil,
Email: {lukas.landau, delamare}@cetuc.puc-rio.br

Abstract—Multiple-antenna systems have been identified as the key technique to serve multiple users in future wireless systems. However, multiple radio front-ends are expensive in terms of hardware complexity and for a large number of antennas the energy consumption of individual components such as digital-to-analog converters can be very high. Moreover, a low peak-to-average ratio is favorable for the utilization of energy efficient power amplifiers with a low dynamic range. For these reasons the consideration of digital-to-analog converters with 1-bit resolution at the transmitter is a promising approach. In this work we propose a precoding design which maximizes the minimum distance to the decision threshold at the receiver. The resulting problem is a scaled version of an integer linear program. We solve the problem exactly with a branch-and-bound strategy and alternatively we approximate the optimum precoder by a conventional relaxation, which corresponds to a linear program. Our results show that the proposed branch-and-bound approach has polynomial complexity. The proposed branch-and-bound and its approximation outperform existing precoding methods in terms of uncoded bit error rate.

Index Terms—Precoding, 1-bit quantization, ADC, MIMO, branch-and-bound.

I. INTRODUCTION

1-bit quantization at the transmitter in a multiuser multiple-input multiple-output (MIMO) communication system [1] is beneficial in terms of energy efficiency as compared to the standard high resolution approach. First of all, this is because the 1-bit digital-to-analog converter itself is more energy efficient in comparison to a standard high resolution converter, which matters especially in systems with a large number of radio front-ends. Moreover, the utilization of a 1-bit quantization corresponds to a low peak-to-average ratio which is favorable for energy efficient power amplifiers. For some applications, e.g., internet of things (IoT), where the receivers have a low-complexity, and limited energy budget it is reasonable to also consider 1-bit quantization at the receiver. In this context, MIMO communication system with 1-bit quantization received increased attention recently. A particular problem is the design of the precoder in the presence of the coarse quantization. In this regard, in [2] and [3] the utilization of linear precoding strategies, such as maximal-ratio transmission and zero-forcing method, with a consecutive quantization has been studied. At the same time a nonlinear approach has been reported in [4], where the precoding vector is obtained based on an optimization method which maximizes the product of the distances to the decision thresholds at the receivers. Very recently another nonlinear precoding strategy is presented in [5] whose computational complexity is only linear with the number of antenna elements.

Unlike prior work, the proposed work focuses on a more traditional method for solving non-convex optimization problems, namely the branch-and-bound method. This strategy has been lately considered for a related problem in [6], namely for discrete receive beamforming with phase shifters having only discrete phases.

In this work, we maximize the minimum distance to the decision thresholds at the receivers which is promising in terms of the uncoded bit error rate (BER). The resulting problem is a scaled version of an integer linear program which we solve exactly by a branch-and-bound strategy. Moreover we approximate the optimal precoder by the relaxed problem formulation which is a linear program. The main part of the paper is organized such that we first introduce the branch-and-bound strategy in general. Subsequently we introduce our sub-optimal design criterion, the bounding method and the actual branch-and-bound algorithm. Finally we test our precoding approach with numerical simulations, which evaluate that even the approximation of our precoding design outperforms the existing approaches known from literature in terms of uncoded BER.

Regarding the notation we want to highlight that the real and imaginary part operator are also applied to vectors and matrices, e.g., $\Re \{x\} = \left[\Re \{x_1\}, \ldots, \Re \{x_M\}\right]^T$ and equivalently for $\Im \{\cdot\}$.

II. SYSTEM MODEL

A multiuser MIMO downlink is considered where the base-station is equipped with $M$ transmit antennas which communicate with $K$ users where each has $N$ antenna elements. The transmit vector is denoted by $x = [x_1, \ldots, x_M]^T$. According to the digital-to-analog converter with 1-bit resolution in the real and imaginary component each element of the transmit vector is taken from a discrete set $x_m \in \mathcal{P}$ which is described by

$$
\mathcal{P} := \left\{ e^{j\pi M^2 \frac{m}{\sqrt{M}}, e^{j\pi M^2 \frac{m}{\sqrt{M}}} / \sqrt{M}, e^{j\pi M^2 / \sqrt{M}}, e^{j\pi M^2 / \sqrt{M}} / \sqrt{M} \right\}
$$

for $m = 1, \ldots, M$. The vector of transmit symbols of length $KN$ is denoted by $s = [s_1, \ldots, s_{KN}]^T$ and $s_l \in$...
\{1 + j, 1 - j, -1 + j, -1 - j\} for \( l = 1, \ldots, KN \). It is assumed a frequency flat fading channel described with independent and identically distributed (i.d.d.) complex Gaussian random variables \( h_{l,m} \) with zero mean, where \( m \) denotes the index of the transmit antenna and the index of the receive antenna is denoted by \( l \). The received signal at the \( nth \) antenna of the \( k \)th user, which corresponds to index \( l = (k - 1)N + n \), is applied to a matched filter, such that the received sample is described by
\[
z_l = r_l + n_l = \sum_{m=1}^{M} h_{m,l} x_m + n_l,
\]
where \( n_l \) represents thermal noise at the receiver which is an i.i.d. complex Gaussian random variable with zero mean and variance \( \sigma_n^2 \). By using vector notation the received \( NK \) samples can be expressed as
\[
z = r + n = Hx + n,
\]
where \( H \) is the channel matrix with dimension \( M \times NK \).

We consider that the received signal is applied to a hard decision detector at the receiver which is in this case equivalent to a 1-bit quantization. The quantization operation is denoted by \( y_l = Q(z_l) \), where \( Q(z_l) = \text{sgn}(\text{Re}\{z_l\}) + j \text{sgn}(\text{Im}\{z_l\}) \), such that \( y_l \in \{1 + j, 1 - j, -1 + j, -1 - j\} \). The quantization operator applies element-wise with \( Q(z) = [Q(z_1), \ldots, Q(z_{KN})]^T \). In the following it is described, how the transmit vector shall be chosen such that transmit symbols \( s \) can be appropriately detected at the receivers.

### III. State-of-the-Art 1-Bit Precoding Methods

The state-of-the-art precoding schemes rely on the computation of a continuous precoding vector \( \hat{x} \) and a consecutive quantization step which ensures that the precoding vector lies in the discrete set \( \mathcal{P}^M \). With this the precoding vector is given by
\[
x = \frac{1}{\sqrt{2M}} Q(\hat{x}).
\]
In this regard, different designs for the computation of the continuous precoding vector \( \hat{x} \) have been presented in literature, which are briefly introduced in the following.

#### A. Zero-Forcing

The zero-forcing design approach for precoding has been studied in [2]. The continuous precoding vector is computed by matrix multiplication of the transmit symbol vector with the precoding matrix \( P \) given by
\[
\hat{x} = Ps,
\]
where the zero-forcing precoding matrix is given by
\[
P = H^H (HH^H)^{-1},
\]
where the pseudo inverse is considered when the inverse of \( (HH^H) \) does not exist.

#### B. Bit Error Oriented

An alternative design objective is presented in [4], which describes the maximization of the product of the distances to the decision thresholds. This approach is promising in terms of bit error rate, because the bit error rate is a function of the distances to the decision thresholds. The optimization problem can be expressed as:
\[
\max_{\hat{x}} \prod_{i=1}^{KN} \text{Re}\{[s_i]\} \text{Re}\{[H\hat{x}]_i\} \text{Im}\{[s_i]\} \text{Im}\{[H\hat{x}]_i\},
\]
s.t. \( |\text{Re}\{\hat{x}\}| \leq c1_M \),
\[
|\text{Im}\{\hat{x}\}| \leq c1_M,
\]
where \( c = 1/\sqrt{2M} \) and \( 1_M \) is a column vector of \( M \) ones. In [4] the problem is solved by an iterative gradient projection method.

### IV. Proposed Branch-and-Bound Precoding

#### A. Concept

We consider a precoding optimization problem in terms of a minimization of an objective function \( f(x, s) \). With this, the optimization problem is given by
\[
x_{\text{opt}} = \min_x f(x, s)
\]
s.t. \( x \in \mathcal{P}^M \),
where \( \mathcal{P} \) denotes a discrete and finite set. In general, by taking into account the constraint the optimization problem is non-convex. A lower bound on the optimal value for the objective function in (8) can be obtained by solving the relaxed problem described by
\[
x_{\text{lb}} = \min_x f(x, s)
\]
s.t. \( |\text{Re}\{x\}| \leq c1_M \),
\[
|\text{Im}\{x\}| \leq c1_M,
\]
where \( c \) is chosen such that all members of the discrete set \( \mathcal{P} \) lie inside or on the boundaries. Moreover, it is considered that the lower-bounding problem can be solved with an acceptable computational burden.

An upper bound on the original problem (8) is given by any valid solution which fulfills the restriction that the entries of the vector \( x \) are taken from the discrete set. For many applications it is promising to consider the vector of the discrete set which has a minimum Euclidean distance to the optimal vector of the continuous problem for computing an upper bound. In this regard, we denote the smallest known upper bound by
\[
\tilde{f} \geq f(x_{\text{opt}}),
\]
where \( x_{\text{opt}} \) denotes the optimal solution of the problem in (8). Considering that \( d \) entries of the precoding vector \( x \) are fixed and taken out of the discrete set. The precoding vector
is then given by $x = [x_1^T, x_2^T]^T$, where $x_1 \in \mathcal{P}^d$. With this, a subproblem can be formulated by
\[ x_{2,\text{lb}} = \min_{x_2} f(x_2, x_1, s) \quad (11) \]
\[ \text{s.t.} \quad |\Re \{x_2\}| \leq c_{1M-d}, \]
\[ |\Im \{x_2\}| \leq c_{1M-d}. \]
If the solution of (11) is larger than a known upper bound $f$ on the solution of the original problem the corresponding vector $x_1$ cannot be part of the optimal solution of the problem. With this, all its evolutions can be excluded from the possible candidates. The branch-and-bound method is effective once there are many exclusions which reduce the number of candidates and once the bounding methods have a low computational burden.

B. Suboptimal Max Min Design Criterion

Instead of considering the BER or the achievable rate as the objective function, we consider a suboptimal design criterion which we have proposed before in [7] and which has been used later also in [8], [9].

The suboptimal design criterion implies that the precoding vector is chosen such that the minimum distance to the decision threshold, denoted by $\epsilon$, is maximized, and in this work, we consider that the decision threshold at the receivers is zero. With this, the optimization problem is given by
\[ x_{\text{opt}} = \min_{x, \epsilon} -\epsilon \quad (12) \]
\[ \text{s.t.} \quad |\Re \{\text{diag} \{s\}\} H x|^{/2} \geq \epsilon_{1K}, \]
\[ |\Im \{\text{diag} \{s\}\} \Im \{H x\}|^{/2} \geq \epsilon_{1K}, \]
\[ x \in \mathcal{P}^M, \]
where we have applied a negation in order to obtain a minimization problem.

C. Lower-Bounding the Objective Function

Relaxing the problem in (12) with respect to the discrete set, yields
\[ x_{\text{lb}} = \min_{x, \epsilon} -\epsilon \quad (13) \]
\[ \text{s.t.} \quad |\Re \{\text{diag} \{s\}\} H x|^{/2} \geq \epsilon_{1K}, \]
\[ |\Im \{\text{diag} \{s\}\} \Im \{H x\}|^{/2} \geq \epsilon_{1K}, \]
\[ |\Re \{x\}| \leq c_{1M}, \]
\[ |\Im \{x\}| \leq c_{1M}, \]
where $c = 1/\sqrt{2M}$. According to the relaxation of the input constraint the optimal value of the optimization problem in (13) is always smaller or equal to the optimal value of (12). In the following we consider the equivalent real-valued problem. The real-valued representations of the precoding vector and transmit symbol vector are given by
\[ x_t = \begin{bmatrix} \Re \{x\} \\ \Im \{x\} \end{bmatrix} \quad (14) \]
\[ s_t = \begin{bmatrix} \Re \{s\} \\ \Im \{s\} \end{bmatrix}, \quad (15) \]
and the real-valued representation of the channel matrix is given by
\[ H_r = \begin{bmatrix} \Re \{H\} & -\Im \{H\} \\ \Im \{H\} & \Re \{H\} \end{bmatrix}. \quad (16) \]

With this, the real-valued representation of the noiseless received vector is given by
\[ r_t = \begin{bmatrix} \Re \{r\} \\ \Im \{r\} \end{bmatrix} = H_x x_t. \quad (17) \]

Based on the real valued quantities the optimization problem is described by
\[ x_{r,\text{lb}} = \min_{x_r, \epsilon} -\epsilon \quad (18) \]
\[ \text{s.t.} \quad |\text{diag} \{s\} H x_r|^{/2} \geq \epsilon_{12KN}, \]
\[ |\{x_r\}| \leq c_{12M}. \]

By introducing the vector $v = [x_t^T, \epsilon]^T$ the problem can be written in a standard form of a linear program
\[ v_{\text{lb}} = \min_v a^T v \quad (19) \]
\[ \text{s.t.} \quad A v \geq 0_{2KN}, \]
\[ |[v_m]| \leq c, \]
\[ \text{for } m = 1, \ldots, 2M, \]
where $a = [0_{2M}, -1]^T$, $A = [\text{diag} \{s\} H_t, -1_{1KN}]$.

D. The Branch and Bound Algorithm

In this section a branch-and-bound algorithm is proposed which converges to the global solution of the problem presented in (12). Considering that the variable vector of the problem in (12) is expressed by $v = [v_1^T, v_2^T]^T$, where $v_1$ has the length $d$, is fixed and taken from the discrete set $\mathcal{P}^d$. The discrete set is described by $\mathcal{P}_d := \{1/\sqrt{2M}, -1/\sqrt{2M}\}$. Accordingly the matrix of the first inequality in (12) or (13) is expressed as $A = [A_1, A_2]$, where $A_1$ contains the first $d$ columns of $A$. The lower bound on the subproblem associated with the residual vector $v_2$ is given by
\[ v_{2,\text{lb}} = \min_{v_2} a_2^T v_2 \quad (20) \]
\[ \text{s.t.} \quad A_2 v_2 \geq b \]
\[ |[v_m]| \leq c, \]
\[ \text{for } m = 1, \ldots, 2M-d, \]
where $a_2 = [0_{2M-d}, -1]^T$ and $b = -A_1 v_1$. The problem described in (20) is a linear program which can be solved by active set methods, which can take advantage of initialization vectors close to the optimum. Because the branch-and-bound strategy implies that the individual subproblems are only slightly different to each other this circumstance can be practically exploited. For the considered problem a breadth-first strategy is proposed, where the individual steps are described in Algorithm. The computation of the optimal precoding vector in each time instance might exceed the computational capacities in
Algorithm 1 Branch-and-Bound Precoding Algorithm

**Initialization:**
Given the channel matrix $H$ and transmit symbols $s$ compute a valid upper bound $\hat{f}$ on the problem, e.g., by solving (13) and subsequently rounding to the closest precoding vector $x_t \in P_t^{2M}$

Define the first level ($d = 1$) of the tree by $G_d := P_t$

for $d = 1 : 2M - 1$

Partition $G_d$ in $x_{1,1}, \ldots, x_{1,|G_d|}$

for $i = 1 : |G_d|$

Given $x_{1,i}$ and $s_t$ determine $v_{2,\text{lb}}$ of (20)

Determine $\epsilon = |v_{2,\text{lb}}|_{2M-d+1}$

Compute the lower bound $\text{lb}(x_{1,i}) := -\epsilon$

Update the best upper bound with the closest valid solution: $\hat{x}_2(x_{2,\text{lb}}) \in P_t^{2M-d}$

$\text{ub}(x_{1,i}) := \min_k \left[ -\text{diag} \{ s_t \} H \begin{bmatrix} x_{1,i} \\ x_2 \end{bmatrix} \right]_k$

$\hat{f} = \min (\hat{f}, \text{ub}(x_{1,i}))$

end for

Eliminate the irrelevant branches $G'_d := \{ x_{1,i} | \text{lb}(x_{1,i}) \leq \hat{f}, i = 1, \ldots, |G_d| \}$

Define the set for the next level in the tree $G_{d+1} := G'_d \times P_t$

end for

Partition $G_{d+1}$ in $x_{1,1}, \ldots, x_{1,|G_{d+1}|}$

$\epsilon(x_{1,i}) := \min_k [\text{diag} \{ s_t \} H, x_{1,i}]_k$

The global optimal solution is

$x_{\text{opt}} = \arg \max_{x_{1,i} \in G_{d+1}} \epsilon(x_{1,i})$

most applications. However, for examples with smaller array sizes and longer coherence time of the channel, a look-up-table can store the entire set of optimal precoding vectors as suggested in [4]. Taking into account the symmetries of the constellation only $4^{K\cdot N-1}$ precoding vectors need to be computed and stored.

### V. Numerical Results

For the numerical evaluation of our proposed methods we consider the uncoded BER. For the BER computations we consider that the signal-to-noise ratio is defined by

$$\text{SNR} = \frac{\text{E} \{ E_{\text{Tx}} \}}{N_0} = \frac{\| x \|_2^2}{\sigma_n^2},$$

where $N_0$ is the noise power density and $\sigma_n^2$ denotes the variance of the noise samples at the output of the matched filter. We assume that the transmitter has perfect channel state information. We have compared our results with the state-of-the-art precoding approaches [2] and [4] summarized in Section III-A. Instead of the gradient projection method, as suggested in [4], we maximize the logarithm of the product (log det) by using CVX [10] which then uses internally a successive approximation technique for solving (7). Fig. 1 shows the BER performance of a scenario with two users with a single antenna and different sizes of the transmit antenna array. As known from the literature the zero-forcing approach [2] shows an error floor whose level decreases with an increasing number of transmit antennas. The proposed max min design criterion shows a superior performance in terms of BER in comparison to the existing methods for $M = 10$ (and below, not shown here) and $KN = 2$. While the branch-and-bound strategy solves the criterion exactly the problem can be approximated by solving only the continuous problem (IV-C) and a consecutive mapping to the closest valid discrete precoding vector $x \in P^M$. For $M = 16$ (and above, not shown here) the approximation and the branch-and-bound strategy show almost identical BER performance.

In what follows we describe by a pessimistic complexity estimation the benefit of the proposed approach in comparison to an exhaustive search method. As explained in Section IV-D the subproblems are linear programs. By considering that a subproblem is solved by interior point methods the complexity of a subproblem is according to a in the community widely accepted estimation in the order of $O(n^{3.5})$, where $n <= 2M + 1$. Moreover the average number of visited branches for the proposed branch-and-bound approach and the setting in Fig. 1 is shown in Table 1 which can be interpreted as a measure for efficiency. Based on the values in Table 1 the number of visited branches is in the order of $\frac{1}{2}(2M)^{2.5}$ for this particular setting. With this, the proposed branch-and-bound strategy has polynomial complexity whereas the exhaustive search has exponential complexity.

From our own experience the utilization of the active set method is more efficient in comparison to the interior point methods although it is known from literature that active set method often does not have tight bounds on the worst-case number of iterations. Moreover we want to remind on the fact that the active set method benefits from initial vectors close to the optimum which can be exploited in the branch-and-bound approach where a series of similar problems is solved. Furthermore the dimension of the subproblems reduces with each iteration of the algorithm.

Finally, we consider larger arrays and more receivers with $KN = 5$ in Fig. 2.
VI. Conclusions

We have proposed a branch-and-bound strategy for the computation of precoding vectors with 1-bit resolution. The design criterion is such that the received signal has a maximized minimum distance to the decision threshold at the receivers. Both, exactly solving the problem and approximating the problem by the continuous relaxation and consecutive mapping to the discrete set yields outstanding performance in terms uncoded bit error rate. The approximation of the problem corresponds to a linear program. Note, that the underlying problem is known as a scaled version of an integer linear program which is discussed in literature very well.

REFERENCES

[1] Q. H. Spencer, C. B. Peel, A. L. Swindlehurst, and M. Haardt, “An introduction to the multi-user MIMO downlink,” IEEE Communications Magazine, vol. 42, no. 10, pp. 60–67, Oct 2004.
[2] A. K. Saxena, I. Fijalkow, and A. L. Swindlehurst, “On one-bit quantized ZF precoding for the multiuser massive MIMO downlink,” in 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), July 2016, pp. 1–5.
[3] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, “Quantized precoding for massive MU-MIMO,” CoRR, vol. abs/1610.07564, 2016. [Online]. Available: http://arxiv.org/abs/1610.07564
[4] H. Jedda, J. A. Nossek, and A. Mezghani, “Minimum BER precoding in 1-bit massive MIMO systems,” in 2016 IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM), July 2016, pp. 1–5.
[5] O. Tirkkonen and C. Studer, “Subset-codebook precoding for 1-bit massive multiuser MIMO,” in Conference on Information Sciences and Systems, 2017.
[6] J. Israel, A. Fischer, J. Martinovic, E. A. Jorswieck, and M. Mesyagutov, “Discrete receive beamforming,” IEEE Signal Processing Letters, vol. 22, no. 7, pp. 958–962, July 2015.
[7] L. Landau, S. Krone, and G. P. Fettweis, “Intersymbol-interference design for maximum information rates with 1-bit quantization and oversampling at the receiver,” in Proc. of the International ITG Conference on Systems, Communications and Coding, Munich, Germany, Jan. 2013.
[8] J. Mo and R. W. Heath Jr, “Capacity analysis of one-bit quantized MIMO systems with transmitter channel state information,” IEEE Transactions on Signal Processing, vol. 63, no. 20, pp. 5498–5512, Oct 2015.
[9] A. Gokceoglu, E. Björnson, E. G. Larsson, and M. Valkama, “Spatio-temporal waveform design for multi-user massive MIMO downlink with 1-bit receivers,” IEEE Journal of Selected Topics in Signal Processing, Nov. 2016.
[10] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” http://cvxr.com/cvx, Mar. 2014.