POKEMON: A NON-LINEAR BEAMFORMING ALGORITHM FOR 1-BIT MASSIVE MIMO

Oscar Castañeda\textsuperscript{1}, Tom Goldstein\textsuperscript{2}, and Christoph Studer\textsuperscript{1}

\textsuperscript{1}School of ECE, Cornell University, Ithaca, NY; e-mail: oc66@cornell.edu, studer@cornell.edu
\textsuperscript{2}Department of CS, University of Maryland, College Park, MD; e-mail: tomg@cs.umd.edu

ABSTRACT

One-bit quantization at the base-station (BS) of a massive multiple-input multiple-output (MIMO) wireless system enables significant power and cost savings. While the 1-bit uplink (users communicate to BS) has gained significant attention, the downlink (BS transmits to users) is far less studied. In this paper, we propose a novel, computationally-efficient 1-bit beamforming algorithm referred to as POKEMON (short for PrOjected downlinK bEaMfOrmiNg), which—after its convolution with the MIMO channel matrix—minimizes multi-user interference (a.k.a. spatial leakage). Our algorithm builds upon the biconvex relaxation (BCR) framework, which efficiently approximates the optimal 1-bit beamforming problem that is of combinatorial nature. Our simulation results show that POKEMON significantly outperforms linear beamformers followed by 1-bit quantization in terms of error-rate performance and recent non-linear beamformers in terms of complexity.

Index Terms— Massive MIMO, beamforming, precoding, quantization, biconvex relaxation, alternating optimization.

1. INTRODUCTION

The base-station (BS) of traditional cellular multiple-input multiple-output (MIMO) wireless systems is typically equipped with high-precision (e.g., 10 bits or more) digital-to-analog converters (DACs). Scaling such a conventional downlink architecture to massive MIMO systems\textsuperscript{[1,2]} with hundreds to thousands of active antenna elements would result in prohibitively high power consumption and system costs\textsuperscript{[3,4,5]}. Instead, the use of low-precision DACs would allow for a significant power reduction and would also enable one to relax the quality requirements on the remaining radio-frequency (RF) circuitry (such as power amplifiers), which further lowers the power consumption and system costs—this is because the impairment strength in the remaining RF circuitry needs to be “just below” the quantization noise floor\textsuperscript{[6]}. Coarsely quantized massive MIMO systems also reduce the raw baseband data-rates that must be transferred from the baseband processing unit to the DACs which may be separated\textsuperscript{[3]}

In order to reduce the power consumption and data rates at the BS of massive MIMO systems, coarse quantization of baseband signals has recently gained significant attention in the massive MIMO uplink (users communicate to the BS); see, e.g.,\textsuperscript{[7,8,9]} and the references therein. In contrast, the quantized massive MIMO downlink (BS communicates with users) is far less explored. Notable exceptions are the papers\textsuperscript{[7,8,9]} which propose to quantize the output of linear beamformers (or precoders) to 1-bit. The proposed methods, however, suffer from a significant error-rate performance loss compared to BS architectures that use high-precision DACs in realistic systems with only hundreds of BS antenna elements. More recently, the papers\textsuperscript{[10,11]} proposed non-linear beamforming algorithms for 1-bit DACs that significantly outperform linear-quantized beamformers in error-rate performance. These methods, however, require significantly higher computational complexity.

1.1. Contributions

In this paper, we propose a novel non-linear 1-bit beamforming algorithm that significantly outperforms linear-quantized algorithms in terms of error-rate performance and requires lower computational complexity than the non-linear methods in\textsuperscript{[10,11]}. Our algorithm, dubbed POKEMON (short for PrOjected downlinK bEaMfOrmiNg), relies on biconvex relaxation (BCR), a framework recently proposed in\textsuperscript{[12]} for approximately solving large semidefinite programs in computer-vision applications. We transform the optimal 1-bit beamforming problem (that is of combinatorial nature) into BCR form, which we then approximately solve using an alternating optimization procedure. We use system simulations to show that BCR approaches the error-rate performance of optimal beamforming methods that use infinite precision DACs. We furthermore compare our method to linear-quantized zero-forcing (ZF) beamforming\textsuperscript{[7,8,9]} and the squared-infinity norm Douglas-Rachford splitting (SQUID) algorithm in\textsuperscript{[10]}.
1.2. Notation

Lowercase boldface letters stand for column vectors; uppercase boldface letters denote matrices. For a matrix $A$, we denote its transpose by $A^T$. The $l^2$-norm of a vector $a$ is $|a|_2 = \sqrt{\sum_k |a_k|^2}$. The identity matrix is $I$. The real and imaginary part of a complex vector $a$ is denoted by $\Re(a)$ and $\Im(a)$, respectively.

1.3. Paper Outline

The rest of the paper is organized as follows. Section 2 introduces the system model and formulates the optimal 1-bit quantized beamforming problem. Section 3 proposes our beamforming algorithm that uses BCR. Section 4 provides numerical simulation results. We conclude in Section 5.

2. DOWNLINK SYSTEM MODEL AND QUANTIZED BEAMFORMING

2.1. Downlink System Model

We consider a massive MIMO downlink system as illustrated in Fig. 1. The BS consists of $B$ antennas and serves $U \ll B$ single-antenna users in the same time-frequency resource. The narrowband, flat-fading input-output relation of the downlink channel is given by

$$y = Hx + n.$$  
(1)

Here, the vector $y \in \mathbb{C}^U$ contains the received signals at all $U$ users. The matrix $H \in \mathbb{C}^{U \times B}$ models the downlink channel, which is assumed to be known. The beamformed vector is denoted by $x \in \mathcal{X}^B$, where $\mathcal{X}$ is the transmit alphabet; this set is $\mathbb{C}$ in the case of infinite-precision transmission, and quaternary in the case of 1-bit DACs. Specifically, we denote the set of possible binary and real-valued quantizer outcomes as $\mathcal{L} = \{-\ell, +\ell\}$ for some $\ell \in \mathbb{R}$ with $\ell > 0$. For each BS antenna $b = 1, \ldots, B$, we transmit $x_b = \ell_R + j\ell_I$ with $\ell_R, \ell_I \in \mathcal{L}$ for $b = 1, \ldots, B$. The vector $n \in \mathbb{C}^U$ in (1) models additive receive-side noise and is assumed to be i.i.d. circularly symmetric complex Gaussian with variance $N_0$.

2.2. Quantized Beamforming

The goal of beamforming in the downlink is to transmit constellation points $s_u \in \mathcal{O}$ for $u = 1, \ldots, U$ to each user $u$, where $\mathcal{O}$ is the transmit constellation (e.g., QPSK or 16-QAM), while minimizing multi-user interference (MUI) \[1, 2\]. Hence, the BS uses channel state information (CSI) to beamform the symbol vector $s \in \mathcal{O}^U$ into a $B$-dimensional beamformed vector $x = P(s)$. Here, $P$ represents the (possibly nonlinear) beamformer, $x$ satisfies an average power constraint $E[\|x\|^2_2] \leq P$, and $P = P/N_0$ is the normalized transmit power.

After transmission of the beamformed vector $x$ over the downlink channel in (1), we consider the following effective input-output relation \[10\]:

$$y = \beta^{-1}s + e + n,$$  
(2)

where the beamforming vector $x$ satisfies $Hx = \beta^{-1}s + e$. We call the quantities $\beta \in \mathbb{R}^+$ the beamforming factor and $e \in \mathbb{C}^U$ the residual error vector, which models MUI and other residual distortion caused by the (quantized) beamformer.

As in \[11\], we assume that the users are able to estimate the beamforming factor $\beta$ and hence, are able to form an estimate

$$\hat{s}_u = \beta y_u = s_u + \beta(c_u + n_u).$$

We assume the users perform minimum distance decoding from their estimates $\hat{s}_u$. Therefore, our beamformer minimizes the mean-squared error (MSE) at the user side

$$MSE = E_n[\|s - \hat{s}\|^2_2] = \|s - \beta Hx\|^2_2 + \beta^2 U N_0.$$  

By replacing the average power constraint with an instantaneous constraint, we obtain the following quantized beamforming (QB) problem \[10\]:

$$\begin{align*}
\text{(QB) } & \quad \text{minimize } x \in \mathcal{X}^B, \beta \in \mathbb{R}^+ \\
& \quad \|s - \beta Hx\|^2_2 + \beta^2 U N_0 \\
& \quad \|x\|^2_2 \leq P,
\end{align*}$$

which is a combinatorial problem in the quaternary-valued vector $x$. For example, for a system with 128 BS antennas...
and 1-bit quantization, an exhaustive search requires one to evaluate the objective function more than $10^{27}$ times, assuming that the optimal beamforming factor $\hat{\beta}$ is known. Hence, such naïve methods are infeasible in massive MIMO systems.

### 3. 1-BIT BEAMFORMING ALGORITHM

#### 3.1. Approximating the QB Problem

To solve the QB problem efficiently for 1-bit DACs, we use the BCR framework put forward in [12], which was initially proposed for solving large semidefinite programs appearing in computer-vision problems. To use this framework, we first simplify the objective function in (QB) by assuming that $N_0 \to 0$, i.e., we assume our system to operate in the high signal-to-noise-ratio (SNR) regime. We furthermore take a leap of faith with the following approximation:

$$\min_{\beta \in \mathbb{R}^+} \|s - \beta Hx\|_2^2 \approx \min_{\alpha \in \mathbb{R}^+} \|\alpha s - Hx\|_2^2$$

(3)

for a given beamforming vector $x$; i.e., we switch sides of the beamforming factor so that $\beta^{-1} \approx \alpha$. We also apply the real-valued decomposition by defining the following quantities:

$$s = \begin{bmatrix} \Re\{s\} \\ \Im\{s\} \end{bmatrix}, \quad \overline{H} = \begin{bmatrix} \Re\{\overline{H}\} - \Im\{\overline{H}\} \\ \Im\{\overline{H}\} + \Re\{\overline{H}\} \end{bmatrix}, \quad \overline{x} = \begin{bmatrix} \Re\{\overline{x}\} \\ \Im\{\overline{x}\} \end{bmatrix}.$$

With these approximations and definitions, we can simplify the QB problem into the following optimization problem:

$$(QB^+) \left\{ \begin{array}{l} \text{minimize} \quad \|\alpha s - \overline{H}x\|_2^2 \\
\text{subject to} \quad \|\pi_b\| \leq \sqrt{P}, \quad b = 1, \ldots, 2B, \end{array} \right.$$ 

whose binary-valued solution vector $\hat{x}$ satisfies the instantaneous power constraint $\|\hat{x}\|_2^2 = P$. The solution to this problem can then be converted to a complex-valued vector and transmitted over the channel as modeled in [1].

#### 3.2. Biconvex Relaxation (BCR)

We next reformulate the above optimization problem (QB+) using BCR [12]. Before we can do so, we start by optimizing the problem in the variable $\alpha \in \mathbb{R}$ while holding the vector $\overline{x}$ fixed. The resulting optimal scaling parameter $\hat{\alpha}$ is given by

$$\hat{\alpha} = \arg \min_{\alpha \in \mathbb{R}} \|\alpha s - \overline{H}x\|_2^2 = s^T \overline{H} \overline{x},$$

where we also allow a negative scaling factor—the sign can be absorbed into $\overline{x}$ thanks to the symmetry of the binary quantization alphabet $\mathcal{L}$. Inserting $\hat{\alpha}$ into (QB+) yields

$$\hat{x} = \arg \min_{\|\pi_b\| \leq \sqrt{P}, \quad b = 1, \ldots, 2B} \|A\overline{x}\|_2^2,$$

with the following auxiliary matrix:

$$A = \left(1 - \frac{s s^T}{\|s\|_2^2}\right) \overline{H},$$

which is a projected version of the real-valued channel matrix $\overline{H}$ on the orthogonal complement of the real-valued vector $s$.

We are now ready to deploy the BCR framework. First, we introduce a copy of the vector $\overline{x} = \overline{q}$, which allows us to rewrite the above optimization problem as

$$\hat{x} = \arg \min_{\|\pi_b\| \leq \sqrt{P}, \quad b = 1, \ldots, 2B} \|A\overline{q}\|_2^2 + \gamma \|\overline{q} - \overline{x}\|_2^2,$$

where $\gamma > 0$ is a (fixed) regularization parameter. We next relax the non-convex alphabet constraint $\|\pi_b\| \leq \sqrt{P}$ to a convex constraint $\|\pi_b\| \leq \sqrt{P}$, $b = 1, \ldots, 2B$, which yields the following approximate solution:

$$\hat{x} = \arg \min_{\|\pi_b\| \leq \sqrt{P}, \quad b = 1, \ldots, 2B} \|A\overline{q}\|_2^2 + \gamma \|\overline{q} - \overline{x}\|_2^2.$$

This problem can be reformulated in a (slightly) more compact form as follows:

$$\hat{x} = \arg \min_{\|\overline{x}\|_\infty \leq \sqrt{P}} \|A\overline{q}\|_2^2 + \gamma \|\overline{q} - \overline{x}\|_2^2.$$

The final step (and the key idea) of BCR is to force the relaxed constraints to be satisfied with equality. This can be accomplished by adding a non-convex term (a negative squared Euclidean norm) in the objective that promotes large values in $\overline{x}$; this final step leads to the BCR problem

$$\overline{x}^{\text{BCR}} = \arg \min_{\|\overline{x}\|_\infty \leq \sqrt{P}} \|A\overline{q}\|_2^2 + \gamma \|\overline{q} - \overline{x}\|_2^2 - \delta \|\overline{x}\|_2^2,$$

(4)

where the two algorithm parameters must satisfy $0 < \delta < \gamma$. Note that these two parameters can be tuned to improve the performance of our algorithm; see [12] for more details.

#### 3.3. Alternating Optimization

As in [12], we solve the BCR problem in (4) using alternating minimization. First, we solve for the vector $\overline{q}$ while holding the vector $\overline{x}$ fixed. Then, we solve for the vector $\overline{x}$ while holding $\overline{q}$ fixed. This results in the following iterative procedure:

$$\overline{q}^{(t+1)} = \arg \min_{\overline{q} \in \mathbb{R}^{2B}} \|A\overline{q}\|_2^2 + \gamma \|\overline{q} - \overline{x}^{(t)}\|_2^2,$$

$$\overline{x}^{(t+1)} = \arg \min_{\|\overline{x}\|_\infty \leq \sqrt{P}} \gamma \|\overline{q}^{(t+1)} - \overline{x}^{(t+1)}\|_2^2,$$

which we initialize with the matched filter $\overline{x}^{(1)} = \overline{H}^T \overline{s}$ at iteration $t = 1$. Note that both steps are convex optimization problems (the optimization problem is biconvex) that can be
solved in closed form. In particular, the algorithm reduces to the following simple two-step procedure:
\[
\begin{align*}
\mathbf{q}^{(t+1)} &= (\mathbf{I} + \gamma^{-1} \mathbf{A}^T \mathbf{A})^{-1} \mathbf{x}^{(t)} \\
\mathbf{x}^{(t+1)} &= \text{proj}(\mathbf{q}^{(t+1)}),
\end{align*}
\]
where the (non-orthogonal) projection operator, designated by \(\text{proj}(\cdot)\), is given by
\[
\mathbf{x}_b = \text{proj}(\mathbf{q}_b^{(t+1)}) = \text{sgn}(\mathbf{q}_b^{(t+1)}) \min \left\{ \frac{\gamma}{\gamma - \delta} |\mathbf{q}_b^{(t+1)}|, \frac{\sqrt{P}}{\sqrt{2B}} \right\}
\]
and operates element-wise on the vectors. After the final iteration \(t_{\text{max}}\), the real-valued vector \(\mathbf{x}^{(t_{\text{max}})}\) is quantized to \(\pm \frac{\sqrt{P}}{\sqrt{2B}}\) and converted back into a complex vector that is transmitted over the downlink channel as modeled in (1). For the lack of a better name, we call this 1-bit beamforming algorithm POKEMON (short for PrOjected downlinK bEaMfOrmiNg).

4. SIMULATION RESULTS

Fig. 2 shows bit error-rate (BER) simulation results with i.i.d. Rayleigh fading channel matrices for two different antenna configurations and modulation schemes. For each normalized transmit power value \(\varrho\), we perform 10,000 Monte-Carlo trials. The POKEMON algorithm parameters are \(\gamma = 1\) and \(\delta = 0.2\), and we perform a maximum of \(t_{\text{max}} = 20\) iterations. We compare our 1-bit beamforming algorithm to that of infinite precision ZF, which is given by \(\mathbf{x}^{\text{ZF}} = \rho \mathbf{H}^H (\mathbf{H} \mathbf{H}^H)^{-1} \mathbf{y}\), where the parameter \(\rho\) is chosen to satisfy the average power constraint \(E[\|\mathbf{x}^{\text{ZF}}\|^2] = P\). We also consider ZF beamforming followed by quantization
\[
\mathbf{x}^{\text{QZF}} = \frac{\sqrt{P}}{\sqrt{2B}} \left( \text{sgn}(\Re\{\mathbf{x}^{\text{ZF}}\}) + j \text{sgn}(\Im\{\mathbf{x}^{\text{ZF}}\}) \right),
\]
as well as the recently-proposed SQUID algorithm, a 1-bit and non-linear beamformer that uses convex relaxation (see [10] for the details); we use 200 SQUID iterations.

We clearly see that the proposed method significantly outperforms linear-quantized methods in the 1-bit DAC case. In addition, POKEMON is able to approach ZF beamforming for infinite precision DACs by a few dB SNR (3 dB for \(B = 128\) using QPSK and 5 dB for \(B = 256\) using 16-QAM at \(10^{-3}\) BER), which demonstrates that massive MIMO enables the use of 1-bit DACs also in the downlink. When compared with a highly-optimized SQUID algorithm implementation from [10], our algorithm achieves similar error-rate performance but at substantially reduced simulation time; an unoptimized MATLAB implementation is \(4.5 \times \) and \(2.7 \times\) faster for the system configuration in Fig. 2 (a) and Fig. 2 (b), respectively. Furthermore, the algorithmic regularity and simplicity of POKEMON would enable more efficient hardware implementations.

\(\footnote{In cases where the matrix \(\mathbf{A}\) is extremely large, the matrix inversion in (5) can be avoided using forward-backward splitting methods [13].}

Fig. 2. Bit error-rate simulations of uncoded massive MIMO using two antenna configurations and modulation schemes. POKEMON significantly outperforms ZF followed by 1-bit quantization, approaches the performance of infinite-precision ZF beamforming, and performs close to SQUID at substantially reduced simulation time.

5. CONCLUSION

We have shown that non-linear beamforming algorithms can significantly outperform linear methods followed by quantization in 1-bit massive MIMO systems. Our algorithm relies upon biconvex relaxation (BCR) [12], which enables us to approach the error-rate performance of infinite-precision ZF beamforming and the recently-proposed non-linear 1-bit beamformer, SQUID [10], at lower computational complexity. Our results demonstrate that 1-bit massive MIMO systems enable reliable communication even with higher order modulation schemes, such as 16-QAM, when combined with sophisticated, non-linear beamforming algorithms. The regularity of POKEMON also paves the way for a very-large-scale integration (VLSI) design that would enable 1-bit beamforming in practical massive MIMO systems.
6. REFERENCES


