

# An Efficient Manifold Algorithm for Constructive Interference Based Constant Envelope Precoding

Fan Liu, *Student Member, IEEE*, Christos Masouros, *Senior Member, IEEE*,  
Pierluigi Vito Amadori, *Student Member, IEEE*, and Huafei Sun

**Abstract**—In this letter, we propose a novel manifold-based algorithm to solve the constant envelope (CE) precoding problem with interference exploitation. For a given power budget, we design the precoded symbols subject to the CE constraints, such that the constructive effect of the multiuser interference is maximized. While the objective function for the original problem is not complex differentiable, we consider the smooth approximation of its real representation, and map it onto a Riemannian manifold. By using the Riemannian conjugate gradient algorithm, a local minimizer can be efficiently found. The complexity of the algorithm is analytically derived in terms of floating-points operations (flops) per iteration. Simulations show that the proposed algorithm outperforms the conventional methods on both symbol error rate and computational complexity.

**Index Terms**—Constant envelope, manifold optimization, massive multi-input multi-output (mMIMO), multiuser multi-input single-output (MU-MISO) downlink.

## I. INTRODUCTION

AS ONE of the most promising approaches in 5G technology, massive multi-input multi-output (mMIMO) communication systems are expected to provide significant benefits over conventional MIMO systems by employing much larger antenna arrays [1], [2]. Nevertheless, such systems face numerous challenges brought by the increasing number of antennas, e.g., higher hardware costs and power consumption, which may delay its deployment in future 5G systems. Hence, cheap and efficient RF power amplifiers (PA) are required for making the technology realizable in practical scenarios.

It is important to note that most of power-efficient PAs are made by nonlinear components; therefore, waveforms with low

peak-to-average-power ratio (PAPR) are needed to avoid signal distortions when the PA is operated at the saturation region [3]. Pioneered by Mohammed and Larsson [4], [5], the constant envelope precoding (CEP) has been proposed as an enabling solution, where the MUI is minimized subject to the CE constraints. The optimization in [5] is a nonconvex nonlinear least square (NLS) problem, and is solved by sequential gradient descent (GD) method, which converges to a local minimum. To further improve the performance, a cross-entropy optimization (CEO) solver is introduced in [6]. Recent contributions exploit the geometric properties of the single-user CE problem and develop approaches for exact phase recovery [7], [8]. More relevant to this letter, by viewing the feasible region of the CE problem as a complex circle manifold, a Riemannian conjugate gradient (RCG) algorithm is proposed in [9], where the NLS problem is solved with much lower complexity than both GD and CEO. While the interference reduction (IR) methods in above-mentioned works are straightforward, their performance is strongly dependent on the constellation energy [5], which is difficult to optimally set in advance. In view of this, a constellation scaling and rotation optimization has been formulated in [10], and solved by semidefinite relaxation (SDR). By noting the fact that MUI is known to the base station (BS), and thus can be utilized as a source of useful power, the previous work [11] considers a novel CEP approach with the concept of constructive interference (CI) [12]–[16], which obtains a significant performance improvement compared to IR approaches. However, both of the above-mentioned methods demand large amount of computations inevitably.

Based on the previous works on manifold optimizations [17], [18], we consider a manifold-based algorithm to solve the CI-CEP problem in this letter. Since the objective function is not complex differentiable, we first equivalently transform the problem into its real representation, and use a smooth upper-bound to obtain a differentiable approximation. By viewing the feasible region as an oblique manifold, an RCG algorithm is employed to find a local minimizer of the problem. Unlike the relaxed convex problem in [11], the proposed algorithm is guaranteed to yield precoded symbols with exactly constant envelopes, and has better performance than the methods of [11] in terms of both symbol error rate (SER) and complexity.

## II. SYSTEM MODEL

We consider a multiuser multi-input single-output (MU-MISO) downlink scenario where an  $N$ -antenna BS transmits signals to  $M$  single-antenna users. The received signal vector is

Manuscript received June 5, 2017; revised July 31, 2017; accepted August 29, 2017. Date of publication September 1, 2017; date of current version September 18, 2017. This work was supported in part by the Engineering and Physical Sciences Research Council under Project EP/M014150/1, in part by the China Scholarship Council, and in part by the National Natural Science Foundation of China under Project No. 61771047. The associate editor coordinating the review of this paper and approving it for publication was Dr. Jiaheng Wang. (Corresponding author: Fan Liu.)

F. Liu is with the School of Information and Electronics, Beijing Institute of Technology, Beijing 100081, China, and is also with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, U.K. (e-mail: liufan92@bit.edu.cn).

C. Masouros and P. V. Amadori are with the Department of Electronic and Electrical Engineering, University College London, London WC1E 7JE, U.K. (e-mail: chris.masouros@ieee.org; uceepva@ucl.ac.uk).

H. Sun is with the School of Mathematics and Statistics, Beijing Institute of Technology, Beijing 100081, China (e-mail: huafeisun@bit.edu.cn).

Color versions of one or more of the figures in this letter are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LSP.2017.2748230

given as

$$\mathbf{y} = \mathbf{H}^T \mathbf{x} + \mathbf{w} \quad (1)$$

where  $\mathbf{y} = [y_1, y_2, \dots, y_M]^T \in \mathbb{C}^{M \times 1}$  with  $y_m$  being the received symbol for the  $m$ th user,  $\mathbf{x} = [x_1, x_2, \dots, x_N]^T \in \mathbb{C}^{N \times 1}$  represents the transmitted symbols,  $\mathbf{w} = [w_1, w_2, \dots, w_M]^T \in \mathbb{C}^{M \times 1} \sim \mathcal{CN}(0, N_0 \mathbf{I})$  is the Gaussian noise, and  $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_M] \in \mathbb{C}^{N \times M}$  is the channel matrix, with  $\mathbf{h}_m$  being the channel vector for the  $m$ th user. Without loss of generality, the channel is assumed to be Rayleigh fading, i.e., each entry of  $\mathbf{H}$  subjects to independent identically distributed complex Gaussian distribution with zero-mean, and is perfectly known to the BS. The transmitted signal is expected to have constant envelope, which is

$$x_n = \sqrt{P_T/N} e^{j\theta_n}, \quad \forall n \quad (2)$$

where  $P_T$  is the total transmit power and  $\theta_n$  is the phase of the  $n$ th transmitted symbol.

Assume that the desired symbol for the  $m$ th user is  $s_m = \sqrt{E_m} e^{j\phi_m}$ , where  $E_m$  and  $\phi_m$  denote the power and the phase of the symbol, respectively. The received symbol for the  $m$ th user can be written as

$$y_m = s_m + \underbrace{(\mathbf{h}_m^T \mathbf{x} - s_m)}_{\text{MUI}} + w_m \quad (3)$$

where the second term represents the interfering signal for the user. The total MUI power is then given by

$$P_{\text{MUI}} = \sum_{m=1}^M (\mathbf{h}_m^T \mathbf{x} - s_m)^2 = \|\mathbf{H}^T \mathbf{x} - \mathbf{s}\|^2 \quad (4)$$

where  $\mathbf{s} = [s_1, s_2, \dots, s_M]^T$  is the desired symbol vector.

### III. PROBLEMS FORMULATION

Aiming at minimizing the MUI power, the conventional CEP approaches are designed to solve the following optimization problem [5]

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{H}^T \mathbf{x} - \mathbf{s}\|^2 \\ \text{s.t.} \quad & |x_n| = \sqrt{P_T/N}, \quad \forall n. \end{aligned} \quad (5)$$

Problem (5) is an NLS problem, which is obviously nonconvex, and has multiple local minima. Fortunately, it has been proven that most of the local minima yield small values [5], and can be obtained by a variety of approaches [5], [6], [9]. However, by treating all the interference as harmful, these techniques ignore the fact that MUI can be employed as a green signal power source to benefit the symbol demodulation. This has been first proposed in [19], where the MUI is classified as constructive and destructive parts. CI-based beamformers aim to minimize destructive and exploiting constructive interference, which enable a relaxed feasible region for the optimization [13]. Based on this, previous work [11] focuses on maximizing the *constructive effect* of the MUI to achieve CE precoding. While the CI approaches have already been applied to quadrature amplitude modulation (QAM) [20], [21] here we recapture the CI-CEP problem with PSK modulations for notational

simplicity as follows [11]:

$$\begin{aligned} \min_{\mathbf{x}} \max_m \quad & |\text{Im}(t_m)| - \text{Re}(t_m) \tan \psi \\ \text{s.t.} \quad & |x_n| = \sqrt{P_T/N} \quad \forall n \\ & t_m = (\mathbf{h}_m^T \mathbf{x} - s_m) e^{-j\phi_m} \quad \forall m \end{aligned} \quad (6)$$

where  $s_m = u e^{j\phi_m}$ ,  $\psi = \pi/L$ ,  $u$  is the amplitude for the PSK symbols, and  $L$  is the PSK modulation order. The above-mentioned problem can be solved by CEO suboptimally, and has been further relaxed as a convex problem by replacing the equality constraints on  $x_n$  as inequalities, i.e.,  $|x_n| \leq \sqrt{P_T/N}$ ,  $\forall n$ . Such a convex approximation problem can be efficiently solved by numerical solvers, e.g., CVX toolbox. The results are then normalized to obtain transmitted symbols with constant envelopes [11]. Nevertheless, using CEO or CVX to solve (6) requires significant computation resources. In the next section, we propose a manifold based optimization technique to solve (6), which has much lower complexity.

### IV. PROPOSED ALGORITHM BASED ON OBLIQUE MANIFOLD

Since  $\text{Re}(\cdot)$  and  $\text{Im}(\cdot)$  are not complex differentiable, we formulate the real representation of (6). First we rewrite  $t_m$  as

$$t_m = (\mathbf{h}_m^T \mathbf{x} - s_m) e^{-j\phi_m} = \tilde{\mathbf{h}}_m^T \mathbf{x} - u \quad (7)$$

where  $\tilde{\mathbf{h}}_m = \mathbf{h}_m e^{-j\phi_m}$ . We then separate the real and imaginary parts of complex notations as follows:

$$\tilde{\mathbf{H}} = \tilde{\mathbf{H}}_R + j\tilde{\mathbf{H}}_I, \quad \tilde{\mathbf{h}}_m = \tilde{\mathbf{h}}_{Rm} + j\tilde{\mathbf{h}}_{Im}, \quad \mathbf{x} = \mathbf{x}_R + j\mathbf{x}_I \quad (8)$$

where  $\tilde{\mathbf{H}} = [\tilde{\mathbf{h}}_1, \tilde{\mathbf{h}}_2, \dots, \tilde{\mathbf{h}}_M]$ . It follows that

$$\begin{aligned} \text{Re}(t_m) &= \tilde{\mathbf{h}}_{Rm}^T \mathbf{x}_R - \tilde{\mathbf{h}}_{Im}^T \mathbf{x}_I - u \\ \text{Im}(t_m) &= \tilde{\mathbf{h}}_{Im}^T \mathbf{x}_R + \tilde{\mathbf{h}}_{Rm}^T \mathbf{x}_I. \end{aligned} \quad (9)$$

By using the fact that  $|a| = \max(a, -a)$ , and denoting  $\beta = \tan \psi$  we have

$$|\text{Im}(t_m)| - \text{Re}(t_m) \tan \psi = \max(g_{2m-1}, g_{2m}) - u\beta \quad (10)$$

where

$$\begin{aligned} g_{2m-1} &= (\tilde{\mathbf{h}}_{Im} - \beta \tilde{\mathbf{h}}_{Rm})^T \mathbf{x}_R + (\tilde{\mathbf{h}}_{Rm} + \beta \tilde{\mathbf{h}}_{Im})^T \mathbf{x}_I \\ g_{2m} &= (\beta \tilde{\mathbf{h}}_{Im} - \tilde{\mathbf{h}}_{Rm})^T \mathbf{x}_I - (\tilde{\mathbf{h}}_{Im} + \beta \tilde{\mathbf{h}}_{Rm})^T \mathbf{x}_R. \end{aligned} \quad (11)$$

Denoting  $\tilde{\mathbf{X}} = \sqrt{\frac{N}{P_T}} [\mathbf{x}_R, \mathbf{x}_I]^T$ , the real representation of the problem (6) can be written compactly as follows:

$$\begin{aligned} \min_{\tilde{\mathbf{X}}} \max_i \quad & g_i \\ \text{s.t.} \quad & (\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})_{nn} = 1, \quad n = 1, 2, \dots, N \end{aligned} \quad (12)$$

where  $i = 1, 2, \dots, 2M$ . It is clear that the feasible region of (12) can be given as

$$\mathcal{M} = \left\{ \tilde{\mathbf{X}} \in \mathbb{R}^{2 \times N} : (\tilde{\mathbf{X}}^T \tilde{\mathbf{X}})_{nn} = 1, \quad \forall n \right\}. \quad (13)$$

We say that  $\mathcal{M}$  forms a *manifold*, and  $\tilde{\mathbf{X}}$  is a point on  $\mathcal{M}$ . To be more specific,  $\mathcal{M}$  is a  $2N$ -dimensional *oblique manifold* [22]. In Riemannian geometry, a manifold is defined as a set of points that endowed with a locally Euclidean structure near each point.

Given a point  $p$  on  $\mathcal{M}$ , a *tangent vector* at  $p$  is defined as the vector that is tangent to any smooth curves on  $\mathcal{M}$  through  $p$ . The set of all such vectors at  $p$  forms the *tangent space*, denoted by  $T_p\mathcal{M}$ , which is an Euclidean space. Specially, the tangent space at  $\tilde{\mathbf{X}}$  is given as [23]

$$T_{\tilde{\mathbf{X}}}\mathcal{M} = \left\{ \mathbf{U} \in \mathbb{R}^{2 \times N} : (\tilde{\mathbf{X}}^T \mathbf{U})_{nn} = 0, \quad \forall n \right\}. \quad (14)$$

If the tangent spaces of a manifold are equipped with a smoothly varying inner product, the manifold is called *Riemannian manifold* [24]. Accordingly, the family of inner products is called *Riemannian metric*, which allows the existence of rich geometric structure on the manifold. Here we use the usual Euclidean inner product as the metric, which is  $\langle \mathbf{U}, \mathbf{V} \rangle_{\tilde{\mathbf{X}}} = \text{tr}(\mathbf{U}^T \mathbf{V})$ , where  $\mathbf{U}, \mathbf{V} \in T_{\tilde{\mathbf{X}}}\mathcal{M}$ .

The algorithm that we employ is the so-called Riemannian conjugate gradient (RCG) algorithm [25], which performs a gradient-dependent line search on the Riemannian manifold rather than the Euclidean space. Since the objective function in (12) is still not differentiable, we consider the well-known smooth log-sum-exp upper-bound  $f(\tilde{\mathbf{X}})$  for the max function [26], which is

$$\begin{aligned} g_{\max} &\leq f(\tilde{\mathbf{X}}) = \varepsilon \log \left( \sum_i \exp(g_i/\varepsilon) \right) \\ &\leq g_{\max} + \varepsilon \log(2M) \end{aligned} \quad (15)$$

where  $\varepsilon > 0$  is some small positive number. The gradient of  $f(\tilde{\mathbf{X}})$  is thus given as

$$\nabla_{\tilde{\mathbf{X}}} f = \left[ \frac{\partial f}{\partial \tilde{\mathbf{x}}_1}, \frac{\partial f}{\partial \tilde{\mathbf{x}}_2}, \dots, \frac{\partial f}{\partial \tilde{\mathbf{x}}_N} \right] \quad (16)$$

where  $\tilde{\mathbf{x}}_n$  is the  $n$ th column of  $\tilde{\mathbf{X}}$ . Noting that  $\mathbf{x}_R = \sqrt{\frac{N}{P_T}} \tilde{\mathbf{X}}^T(:, 1)$ ,  $\mathbf{x}_I = \sqrt{\frac{N}{P_T}} \tilde{\mathbf{X}}^T(:, 2)$ , which are the first and second column of  $\tilde{\mathbf{X}}^T$ , respectively, we have

$$\frac{\partial \mathbf{x}_R}{\partial \tilde{\mathbf{x}}_n} = \sqrt{\frac{N}{P_T}} [\mathbf{e}_n, \mathbf{0}], \quad \frac{\partial \mathbf{x}_I}{\partial \tilde{\mathbf{x}}_n} = \sqrt{\frac{N}{P_T}} [\mathbf{0}, \mathbf{e}_n] \quad (17)$$

where  $\mathbf{e}_n \in \mathbb{R}^{N \times 1}$  have all-zero entries except that its  $n$ th entry equals 1. Based on (17), the  $n$ th column of the gradient is given by

$$\frac{\partial f}{\partial \tilde{\mathbf{x}}_n} = \frac{\sqrt{\frac{N}{P_T}} \sum_{m=1}^M \left( \begin{bmatrix} a_{n,m} & b_{n,m} \\ c_{n,m} & d_{n,m} \end{bmatrix} \begin{bmatrix} \exp\left(\frac{g_{2m-1}}{\varepsilon}\right) \\ \exp\left(\frac{g_{2m}}{\varepsilon}\right) \end{bmatrix} \right)}{\sum_{i=1}^{2M} \exp\left(\frac{g_i}{\varepsilon}\right)} \quad (18)$$

where  $a_{n,m}, b_{n,m}, c_{n,m}$ , and  $d_{n,m}$  denote the  $(n, m)$ th entry of the following matrices:

$$\begin{aligned} \mathbf{A} &= \tilde{\mathbf{H}}_I - \beta \tilde{\mathbf{H}}_R, \mathbf{B} = -\tilde{\mathbf{H}}_I - \beta \tilde{\mathbf{H}}_R \\ \mathbf{C} &= \tilde{\mathbf{H}}_R + \beta \tilde{\mathbf{H}}_I, \mathbf{D} = \tilde{\mathbf{H}}_R - \beta \tilde{\mathbf{H}}_I. \end{aligned} \quad (19)$$

In the RCG algorithm, (16) is called the Euclidean gradient, and can be used to compute the Riemannian gradient, which is defined as the tangent vector belongs to  $T_{\tilde{\mathbf{X}}}\mathcal{M}$  that indicates the steepest ascent direction of  $f(\tilde{\mathbf{X}})$ . It can be viewed as the orthogonal projection of the Euclidean gradient onto the tangent

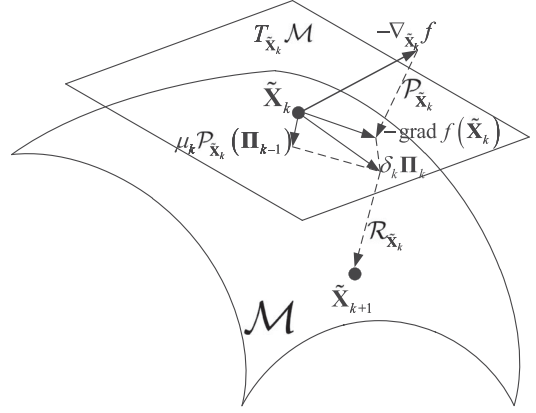


Fig. 1. Riemannian conjugate gradient algorithm.

space [23], [27], which is given as

$$\text{grad } f(\tilde{\mathbf{X}}) = \mathcal{P}_{\tilde{\mathbf{X}}}(\nabla_{\tilde{\mathbf{X}}} f) = \nabla_{\tilde{\mathbf{X}}} f - \tilde{\mathbf{X}} \text{diag}(\tilde{\mathbf{X}}^T \nabla_{\tilde{\mathbf{X}}} f) \quad (20)$$

where  $\mathcal{P}_{\tilde{\mathbf{X}}}(\cdot)$  denotes the projector,  $\text{diag}(\cdot)$  sets all off-diagonal entries of a matrix to zero. At the  $k$ th iteration, the descent direction  $\Pi_k$  is obtained as

$$\Pi_k = -\text{grad } f(\tilde{\mathbf{X}}_k) + \mu_k \mathcal{P}_{\tilde{\mathbf{X}}_k}(\Pi_{k-1}). \quad (21)$$

Here the projector is used as *vector transport*, which maps the vector from one tangent space to another.  $\mu_k$  is given by the Riemannian Polak-Ribière formula, which is [27]

$$\mu_k = \frac{\langle \text{grad } f(\tilde{\mathbf{X}}_k), \text{grad } f(\tilde{\mathbf{X}}_k) - \mathcal{P}_{\tilde{\mathbf{X}}_k}(\text{grad } f(\tilde{\mathbf{X}}_{k-1})) \rangle_{\tilde{\mathbf{X}}_k}}{\langle \text{grad } f(\tilde{\mathbf{X}}_{k-1}), \text{grad } f(\tilde{\mathbf{X}}_{k-1}) \rangle_{\tilde{\mathbf{X}}_{k-1}}}. \quad (22)$$

The  $k+1$ th update is thus given by

$$\tilde{\mathbf{X}}_{k+1} = \mathcal{R}_{\tilde{\mathbf{X}}_k}(\delta_k \Pi_k) \quad (23)$$

where  $\mathcal{R}_{\tilde{\mathbf{X}}_k}(\cdot)$  is called *retraction*, which maps a point on  $T_{\tilde{\mathbf{X}}_k}\mathcal{M}$  to  $\mathcal{M}$  with a local rigidity condition that preserves gradients at  $\tilde{\mathbf{X}}_k$  [23], [27], and is given as

$$\mathcal{R}_{\tilde{\mathbf{X}}_k}(\delta_k \Pi_k) = \left[ \frac{(\tilde{\mathbf{X}}_k + \delta_k \Pi_k)_1}{\|(\tilde{\mathbf{X}}_k + \delta_k \Pi_k)_1\|}, \dots, \frac{(\tilde{\mathbf{X}}_k + \delta_k \Pi_k)_N}{\|(\tilde{\mathbf{X}}_k + \delta_k \Pi_k)_N\|} \right] \quad (24)$$

where  $(\tilde{\mathbf{X}}_k + \delta_k \Pi_k)_n$  is the  $n$ th column of the matrix  $(\tilde{\mathbf{X}}_k + \delta_k \Pi_k)$ , and the stepsize  $\delta_k$  is obtained by backtracking line search algorithms, e.g., Armijo rule [27], [28]. Fig. 1 shows a single iteration of the RCG algorithm on  $\mathcal{M}$ , which has also been summarized in Algorithm 1. According to [29], the solution obtained by RCG for the problem (12) satisfies the Karush–Kuhn–Tucker (KKT) conditions on the manifold.

*Remark:* The complexity of Algorithm 1 mainly comes from the computation of the Euclidean gradient in (16), where  $16MN + 6M$  flops are needed, leading to a total complexity of  $\mathcal{O}(MN)$  for each iteration. By contrast, the complexity of GD and RCG-IR are  $\mathcal{O}(MN^2)$  and  $\mathcal{O}(MN)$  per iteration [5], [9], respectively. For CEO, the complexity is  $\mathcal{O}(KMN)$  in each iteration [11], where  $K$  stands for the number of random samples, which may be quite larger than  $M$  and  $N$ . For clarity, we will

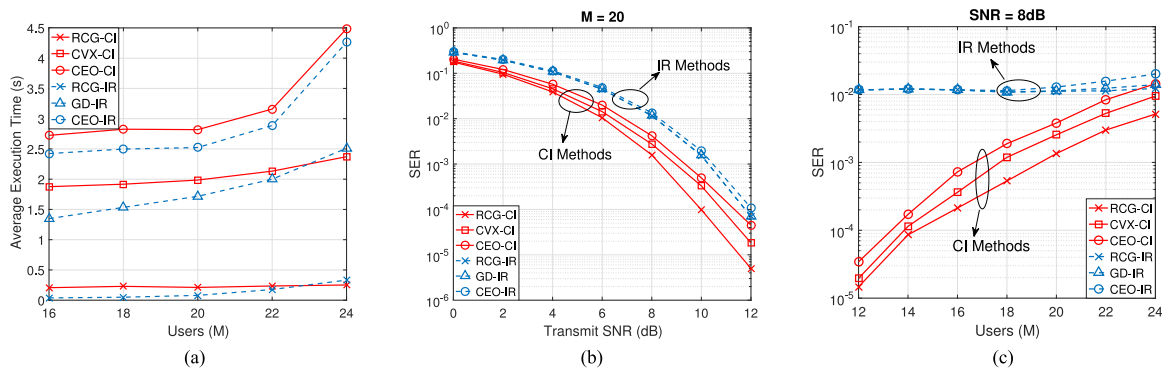


Fig. 2. Numerical results. (a) Average execution time versus number of users for different algorithms. (b) SER versus SNR for different algorithms. (c) SER versus user for different algorithms.

---

### Algorithm 1: RCG for CI-based CEP.

---

**Input:**  $\mathbf{s}$ ,  $\mathbf{H}$ , tolerance threshold  $\Delta = 10^{-6}$  and maximum iteration number  $k_{\max} = 1000$ .

**Output:** Local minimizer  $\tilde{\mathbf{X}}^*$  for (12).

1. Initialize randomly  $\tilde{\mathbf{X}}_0 \in \mathcal{M}$ ,

set  $\Pi_0 = -\text{grad } f(\tilde{\mathbf{X}}_0)$ ,  $k = 0$ ,

**while**  $k \leq k_{\max}$  &  $\|\text{grad } f(\tilde{\mathbf{X}}_k)\|_F \geq \Delta$  **do**

2.  $k = k + 1$ ,

3. Compute stepsize  $\delta_{k-1}$  by Armijo rule, and set

$\tilde{\mathbf{X}}_k$  using the retraction defined in (23),

4. Compute  $\mu_k$  by (22),

5. Compute  $\Pi_k$  by (21).

**end while**

---

show the overall complexity for all the methods numerically in the next section.

## V. NUMERICAL RESULTS

In this section, numerical results based on Monte Carlo simulations have been provided to compare the performance of different algorithms. We consider the following six algorithms:

- 1) the proposed RCG algorithm for CI (RCG-CI);
- 2) convex relaxation for CI (CVX-CI) [11];
- 3) cross-entropy optimization for CI (CEO-CI) [11];
- 4) RCG algorithm for IR (RCG-IR) [9];
- 5) gradient descent algorithm for IR (GD-IR) [5]; and
- 6) cross-entropy optimization for IR (CEO-IR) [6].

Without loss of generality, we use QPSK modulation for all the approaches. We set  $u = 1, \forall m$ , which is a common assumption in related literature for the reason that the optimal  $u$  is difficult to determine for IR methods [5], [6] whereas arbitrary  $u$  can be accepted by CI methods [11]. We also assume that  $P_T = 1, N = 64$  for all the algorithms, and each entry of the channel  $\mathbf{H}$  subjects to standard complex Gaussian distribution, i.e.,  $h_{n,m} \sim \mathcal{CN}(0, 1), \forall n, \forall m$ . For CEO methods, we use the same parameter configuration with [11], which is  $T = 1000$  (the number of iterations),  $K = 500$  (the number of initialized random samples),  $\rho = 0.05$  (quantile), and  $\alpha = 0.08$  (the smooth parameter). For GD-IR, the number of iterations is set as 50.

While the analytic complexity per iteration of the most algorithms has already been given, we compare the overall

complexity in terms of average execution time in Fig. 2(a) since it is difficult to specify the complexity of the CVX-CI approach. The simulation is performed on an Intel Core i7-4790 CPU 32 GB RAM computer with 3.6 GHz. As expected, the RCG methods require least execution time to solve the problem while other methods need much more. Although the proposed RCG-CI algorithm has the same complexity with RCG-IR by each iteration, it needs more iterations for convergence in general. This is because the objective function in the RCG-IR is simpler than that in the RCG-CI. Nevertheless, the total time needed for both methods is still comparable.

In Fig. 2(b), we show the error performance of all six approaches in terms of SER with increased transmit signal-to-noise ratio (SNR), where  $M = 20$  and  $\text{SNR} = P_T/N_0$ . Note that all the IR methods show negligible difference under the given parameter configuration, and all the CI methods outperform the IR methods thanks to the utilization of the MUI power. It is worth noting that the proposed RCG-CI has the best performance among all the six approaches with 2 dB gain over IR methods, and 1 dB gain against the CVX-CI algorithm.

We further consider the error performance with increased number of users in Fig. 2(c), where the SNR is fixed at 8 dB with the number of users ranging from 12 to 24. Once again, we see that the proposed RCG-CI achieves the lowest SER among all the approaches. Note that for the case of small number of users, the IR methods are effective in zeroing interference [5], which results in an almost constant SINR, and leads to the flat SER curves at  $10^{-2}$ , as shown in Fig. 2(c). Nevertheless, by exploiting the MUI as useful power source, the resultant SINR for CI methods increases when the number of users decreases. This gives us a larger gap between the two methods for lower number of users in Fig. 2(c).

## VI. CONCLUSION

A low-complexity manifold optimization algorithm has been introduced to solve the CEP problem with the exploitation of the MUI power. By viewing the feasible region of the optimization as an oblique manifold, the proposed method can efficiently find a near-optimal solution using the Riemannian conjugate gradient algorithm. Numerical results show that the proposed RCG-CI algorithm outperforms the existing five other approaches in terms of error performance, with a comparable complexity to the fastest RCG-IR algorithm.

## REFERENCES

- [1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [3] V. Mancuso and S. Alouf, "Reducing costs and pollution in cellular networks," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 63–71, Aug. 2011.
- [4] S. K. Mohammed and E. G. Larsson, "Single-user beamforming in large-scale MISO systems with per-antenna constant-envelope constraints: The doughnut channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 11, pp. 3992–4005, Nov. 2012.
- [5] S. K. Mohammed and E. G. Larsson, "Per-antenna constant envelope precoding for large multi-user MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 3, pp. 1059–1071, Mar. 2013.
- [6] J. C. Chen, C. K. Wen, and K. K. Wong, "Improved constant envelope multiuser precoding for massive MIMO systems," *IEEE Commun. Lett.*, vol. 18, no. 8, pp. 1311–1314, Aug. 2014.
- [7] J. Pan and W. K. Ma, "Constant envelope precoding for single-user large-scale MISO channels: Efficient precoding and optimal designs," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 982–995, Oct. 2014.
- [8] J. Zhang, Y. Huang, J. Wang, B. Ottersten, and L. Yang, "Per-antenna constant envelope precoding and antenna subset selection: A geometric approach," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6089–6104, Dec. 2016.
- [9] J. C. Chen, "Low-PAPR precoding design for massive multiuser MIMO systems via Riemannian manifold optimization," *IEEE Commun. Lett.*, vol. 21, no. 4, pp. 945–948, Apr. 2017.
- [10] S. Zhang, R. Zhang, and T. J. Lim, "MISO multicasting with constant envelope precoding," *IEEE Wireless Commun. Lett.*, vol. 5, no. 6, pp. 588–591, Dec. 2016.
- [11] P. V. Amadori and C. Masouros, "Constant envelope precoding by interference exploitation in phase shift keying-modulated multiuser transmission," *IEEE Trans. Wireless Commun.*, vol. 16, no. 1, pp. 538–550, Jan. 2017.
- [12] E. Alsusa and C. Masouros, "Adaptive code allocation for interference management on the downlink of DS-CDMA systems," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2420–2424, Jul. 2008.
- [13] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Trans. Signal Process.*, vol. 63, no. 14, pp. 3628–3640, Jul. 2015.
- [14] C. Masouros and E. Alsusa, "Two-stage transmitter precoding based on data-driven code-hopping and partial zero forcing beamforming for MC-CDMA communications," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3634–3645, Jul. 2009.
- [15] F. A. Khan, C. Masouros, and T. Ratnarajah, "Interference-driven linear precoding in multiuser MISO downlink cognitive radio network," *IEEE Trans. Veh. Technol.*, vol. 61, no. 6, pp. 2531–2543, Jul. 2012.
- [16] C. Masouros, M. Sellathurai, and T. Ratnarajah, "Vector perturbation based on symbol scaling for limited feedback MISO downlinks," *IEEE Trans. Signal Process.*, vol. 62, no. 3, pp. 562–571, Feb. 2014.
- [17] X. Duan, H. Sun, L. Peng, and X. Zhao, "A natural gradient descent algorithm for the solution of discrete algebraic Lyapunov equations based on the geodesic distance," *Appl. Math. Comput.*, vol. 219, no. 19, pp. 9899–9905, 2013.
- [18] C. Li, E. Zhang, L. Jiu, and H. Sun, "Optimal control on special euclidean group via natural gradient algorithm," *Sci. China Inf. Sci.*, vol. 59, no. 11, 2016, Art. no. 112203.
- [19] C. Masouros and E. Alsusa, "Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 3, pp. 1396–1404, Mar. 2009.
- [20] A. Li and C. Masouros, "Exploiting constructive mutual coupling in P2P MIMO by analog-digital phase alignment," *IEEE Trans. Wireless Commun.*, vol. 16, no. 3, pp. 1948–1962, Mar. 2017.
- [21] M. T. Kabir, M. R. Khandaker, and C. Masouros, "Interference exploitation in full duplex communications: Trading interference power for both uplink and downlink power savings," arXiv:1703.10666, 2017.
- [22] N. Boumal, B. Mishra, P.-A. Absil, and R. Sepulchre, "Manopt, a MATLAB toolbox for optimization on manifolds," *The J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1455–1459, 2014.
- [23] S. E. Selvan *et al.*, "Descent algorithms on oblique manifold for source-adaptive ICA contrast," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 23, no. 12, pp. 1930–1947, Dec. 2012.
- [24] P. Petersen, *Riemannian Geometry*. New York, NY, USA: Springer-Verlag, 1998, vol. 171.
- [25] N. Boumal, "Optimization and estimation on manifolds," Ph.D. dissertation, Dept. Math. Eng., Université catholique de Louvain, Louvain-la-Neuve, Belgium, Feb. 2014.
- [26] F. Nielsen and K. Sun, "Guaranteed bounds on the Kullback-Leibler divergence of univariate mixtures," *IEEE Signal Process. Lett.*, vol. 23, no. 11, pp. 1543–1546, Nov. 2016.
- [27] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton, NJ, USA: Princeton Univ. Press, 2009.
- [28] J. Nocedal and S. J. Wright, *Numerical Optimization*. New York, NY, USA: Springer-Verlag, 2006.
- [29] W. H. Yang, L.-H. Zhang, and R. Song, "Optimality conditions for the non-linear programming problems on Riemannian manifolds," *Pac. J. Optim.*, vol. 10, no. 2, pp. 415–434, 2014.