FAST FIRST-ORDER METHODS FOR THE MASSIVE ROBUST MULTICAST BEAMFORMING PROBLEM WITH INTERFERENCE TEMPERATURE CONSTRAINTS

Huikang Liu *Peng Wang* *Anthony Man-Cho So*

Department of Systems Engineering and Engineering Management, CUHK, Hong Kong

ABSTRACT

In this paper, we consider the large-scale case of the robust beamforming problem with interference temperature constraints. Previous semidefinite relaxation (SDR) method becomes impracticable because of its expensive computational cost. Even successive convex approximation (SCA) method, the state-of-the-art method, cannot tackle this problem efficiently. Thus, we are motivated to design two efficient first-order methods, multi-block alternating direction method of multipliers (ADMM) and linear programming-assisted subgradient descent (LPA-SD), to solve it. Numerical results demonstrate the potential of our proposed methods in terms of both computational efficiency and solution quality.

Index Terms— Massive robust beamforming, linear programming, subgradient descent, multi-block ADMM, SCA

1. INTRODUCTION

In this paper, we consider the robust transmit beamforming design for secondary multicast transmission in a massive multiple-input multiple-output (MIMO) cognitive radio (CR) network [15]. We assume that the channel state information (CSI) is erroneous or limited at the secondary base station (SBS). Under this scenario, the primary users (PUs) and secondary users (SUs) may interfere with each other. Therefore, we need to limit the interference temperature (IT) associated with the PUs. Because of the imperfect CSI, IT should be satisfied for all possibilities of the channel error, which leads to the robust IT constraints. In our formulation, the worst SUs' signal-to-noise ratio (SNR) will be maximized subject to the transmit power constraint and the robust IT constraints. Similar formulations have been considered in [15, 21, 8, 20, 17, 18].

Generally speaking, this robust transmit beamforming problem and its variants are NP-hard [9, 16]. SDR [14] is well-developed to solve them approximately, in which the robust IT constraints can be eliminated by $\mathcal{S}$-lemma [5, 21]. However, in the massive MIMO system, the computational cost of solving the semidefinite programming (SDP) is unacceptable when either the number of users or antennas is very large. Hence, we are motivated to design some cheap first-order methods to tackle it.

Fortunately, the robust beamforming problem can be reformulated into an equivalent deterministic quadratically constrained quadratic programming (QCQP) with a bunch of difficult constraints[1][5, 15]. Note that many efficient first-order methods have been proposed previously to solve the non-convex QCQP [4, 10, 12, 7, 11, 19]. The most popular one should be successive convex approximation (SCA) and its variants [6, 11, 12]. For SCA, one of the most important issue is how to solve the convex sub-problems quickly. Recently in [11], Konar and Sidiropoulos proposed to use Nesterov smoothing based gradient descent, mirror proximal method and ADMM to solve the subproblems, which are efficient when there is only one simple convex constraint[2] in the QCQP. However, when these techniques are applied to tackle the robust beamforming problem, these complex robust IT constraints still lead to a heavy computational cost.

In this paper, we propose two efficient first-order methods to solve the robust beamforming problem. The first one is the linear programming-assisted subgradient descent (LPA-SD) method, which is motivated by [2, 7]. It is a novel subgradient descent (SD) method. At each iteration, we firstly define two active sets consisting of these *active* objective functions and constraints, respectively. Then we construct a linear programming (LP) to search the descent direction with respect to the objective functions and constraints in the active sets. It is similar to SCA, but we only need to solve a LP in each step. Besides, the size of each sub-problem (LP) is much smaller than the original problem because we only take the active functions and constraints into consideration. Thus, the computational cost in each iteration is much cheaper than SCA.

Another method is the standard multi-block ADMM, which is motivated by [1, 4, 10, 19]. After introducing the auxiliary and slackness variables, we can split these difficult constraints into some easy constraints. Then we can minimize the augmented Lagrangian by solving a sequence of subproblems. Note that the closed-form solution to each subproblem can be found, and only one matrix inverse needs to be computed in the iteration. Thus, the computational cost in each step is much cheaper than SCA and even LPA-SD. With some well-chosen parameters for proximal terms, ADMM

[1]The Euclidean projection has no closed-form solution.
[2]The Euclidean projection has closed-form solution

converges very fast to some critical point. Our numerical results also show both the computational efficiency and the high-level solution quality of the LPA-SD and ADMM.

2. PROBLEM STATEMENT AND REFORMULATION

We consider a massive MIMO system, where the SBS is equipped with N antennas to transmit a common signal to M SUs while limiting the IT to PUs. The transmit beamforming vector $\boldsymbol{w} \in \mathbb{C}^N$ with a power constraint $\|\boldsymbol{w}\| \leq P$, where P^2 is the average transmit power, is employed to transmit the commom information-bearing signal $x \in \mathbb{C}$ to SUs. Since the existence of the noise in signal transmission, the corresponding received signal at ith SU is modeled as

$$y_i = \boldsymbol{h}_i^H \boldsymbol{w} x + n_i, \; i = 1, \ldots, M, \tag{2.1}$$

where $\boldsymbol{h}_i$ is the channel between the SBS and ith SU, and n_i is a complex Gaussian noise with mean zero and variance σ_i^2. We assume that the noise n_i is independent of transmit signal x and channel $\boldsymbol{h}_i$. Based on the model, the received SNR at ith SU is given by $|\boldsymbol{h}_i^H \boldsymbol{w}|^2/\sigma_i^2$. Besides, we model the channel between the SBS and jth PU as

$$\boldsymbol{z}_j = \boldsymbol{a}_j + \boldsymbol{\delta}_j, \; j = 1, \ldots, J, \tag{2.2}$$

where $\boldsymbol{a}_j \in \mathbb{C}^N$ is the estimated channel vector and $\boldsymbol{\delta}_j \in \mathbb{C}^N$ is the error vector. Here, $\boldsymbol{\delta}_j$ is always bounded, i.e., $\|\boldsymbol{\delta}_j\| \leq d_j$ for some given $d_j \geq 0$. Considering the erroneous channel information, we have to limit the IT associated with the PUs:

$$\max_{\|\boldsymbol{\delta}_j\| \leq d_j} |\boldsymbol{w}^H(\boldsymbol{a}_j + \boldsymbol{\delta}_j)|^2 \leq \eta_j^2, \; j = 1, \ldots, J, \tag{2.3}$$

which are the robust IT constraints. Note that we can actually find out the worst $\boldsymbol{\delta}_j$'s which maximize the left hand side of (2.3), so they are equivalent to the following deterministic constraints [5]

$$d_j \|\boldsymbol{w}\| + \|\boldsymbol{a}_j^H \boldsymbol{w}\| \leq \eta_j, \; \text{for } j = 1, \ldots, J. \tag{2.4}$$

Based on the above setting, we can model the robust beamforming problem with interference temperature constraints as follows,

$$\max_{\boldsymbol{w} \in \mathbb{C}^N} \min_{i=1,\ldots,M} |\boldsymbol{h}_i^H \boldsymbol{w}|^2/\sigma_i^2 \tag{2.5a}$$

$$\text{s.t. } d_j \|\boldsymbol{w}\| + \|\boldsymbol{a}_j^H \boldsymbol{w}\| \leq \eta_j, \; \text{for } j = 1, \ldots, J, \tag{2.5b}$$

$$\|\boldsymbol{w}\| \leq P. \tag{2.5c}$$

To simplify the algorithmic development and theoretic analysis, we convert this problem from complex domain to real domain. Assume that $\tilde{\boldsymbol{w}} = \boldsymbol{w}^1 + \boldsymbol{w}^2 i \in \mathbb{C}^N$, then let $\boldsymbol{w} = (\boldsymbol{w}^1, \boldsymbol{w}^2) \in \mathbb{R}^{2N}$ be the corresponding real vector. Similarly, assume that $\boldsymbol{D}_i = \boldsymbol{h}_i/\sigma_i (\boldsymbol{h}_i/\sigma_i)^H \in \mathbb{C}^{N \times N}$, where $\boldsymbol{h}_i = \boldsymbol{h}_i^1 + i\boldsymbol{h}_i^2 \in \mathbb{C}^N$, then the corresponding real matrix is

$$\boldsymbol{H}_i \boldsymbol{H}_i^T = \begin{bmatrix} \mathrm{Re}(\boldsymbol{D}_i) & -\mathrm{Im}(\boldsymbol{D}_i) \\ \mathrm{Im}(\boldsymbol{D}_i) & \mathrm{Re}(\boldsymbol{D}_i) \end{bmatrix}, \tag{2.6}$$

where $\boldsymbol{H}_i = \frac{1}{\sigma_i} \begin{bmatrix} \boldsymbol{h}_i^1 & -\boldsymbol{h}_i^2 \\ \boldsymbol{h}_i^2 & \boldsymbol{h}_i^1 \end{bmatrix}$. Consequently, we can convert (2.5) into the following problem in the real domain,

$$\min_{\boldsymbol{w} \in \mathbb{R}^{2N}} \max_{i=1,\ldots,M} \boldsymbol{w}^T \boldsymbol{G}_i \boldsymbol{w} \tag{2.7a}$$

$$\text{s.t. } d_j \|\boldsymbol{w}\| + \|\boldsymbol{A}_j^T \boldsymbol{w}\| \leq \eta_j, \; \text{for } j = 1, \ldots, J, \tag{2.7b}$$

$$\|\boldsymbol{w}\| \leq P, \tag{2.7c}$$

where $\boldsymbol{G}_i = -\boldsymbol{H}_i \boldsymbol{H}_i^T$ and $\boldsymbol{A}_j = \begin{bmatrix} \mathrm{Re}(\boldsymbol{a}_j) & -\mathrm{Im}(\boldsymbol{a}_j) \\ \mathrm{Im}(\boldsymbol{a}_j) & \mathrm{Re}(\boldsymbol{a}_j) \end{bmatrix}$.

3. ALGORITHM DESIGN AND ANALYSIS

Now, we are ready to present our first-order methods to solve the problem (2.7). The first one is LPA-SD, whose computational cost is insensitive to the number of users M or antennas N. Another one is multi-block ADMM, which can split the difficult constraints in (2.7) and tackle them efficiently.

Before we design the algorithms, some notations need to be clarified. We define $f_i(\boldsymbol{w}) = \boldsymbol{w}^T \boldsymbol{G}_i \boldsymbol{w}$ and $F(\boldsymbol{w}) = \max f_i(\boldsymbol{w})$ where $i = 1, \ldots, M$. Besides, denote that $g_j(\boldsymbol{w}) = d_j \|\boldsymbol{w}\| + \|\boldsymbol{A}_j^T \boldsymbol{w}\|, \; j = 1, \ldots, J$ and $g_{J+1}(\boldsymbol{w}) = \|\boldsymbol{w}\|$ with $\eta_{J+1} = P$. Furthermore, define the k-dimension simplex as $\Delta_k = \{x \in \mathbb{R}^k : e_k^T x = 1, x \geq 0\}$, where $e_k \in \mathbb{R}^k$ denotes a all-ones vector. Furthermore $\nabla f(\boldsymbol{w})$ denotes the gradient or subgradient of f at $\boldsymbol{w}$, and $\boldsymbol{I}$ denotes the identity matrix.

3.1. Linear Programming-Assisted Subgradient Descent

Given any $\boldsymbol{w} \in \mathbb{R}^{2N}$ and $\delta > 0$, let $I_\delta(\boldsymbol{w}) := \{i \in \{1, \ldots, M\} : f_i(\boldsymbol{w}) \geq F(\boldsymbol{w}) - \delta\}$ be the index set consisting of all the δ-active objective functions. Similarly, we can define the index set consisting of all the ϵ-active constraints $I_\epsilon(\boldsymbol{w}) := \{j \in \{1, \ldots, J+1\} : g_j(\boldsymbol{w}) \geq (1-\epsilon)\eta_j\}$ for any $\epsilon \in [0, 1)$. Based on the active constraint set $I_\epsilon(\boldsymbol{w})$, we can define the ϵ-tangent space at $\boldsymbol{w}$,

$$\mathcal{T}_\epsilon(x) = \{x \in \mathbb{R}^{2N} : x^T \nabla g_j(\boldsymbol{w}) = 0, j \in I_\epsilon(\boldsymbol{w})\}. \tag{3.1}$$

Furthermore, the projected gradient of $\boldsymbol{w}$ onto the tangent space $\mathcal{T}_\epsilon(\boldsymbol{w})$ can be computed as follows,

$$\mathrm{grad} f_i(\boldsymbol{w}) = (\boldsymbol{I} - \boldsymbol{U}\boldsymbol{U}^T) \nabla f_i(\boldsymbol{w}), \tag{3.2}$$

where $\boldsymbol{U}$ is the basis matrix of subspace spanned by $\nabla g_j(\boldsymbol{w}), \; j \in I_\epsilon(\boldsymbol{w})$. For simplicity, denote that $p_i(\boldsymbol{w}) = \mathrm{grad} f_i(\boldsymbol{w})$. Then we can apply the following LP to find a good descent direction at iteration point $\boldsymbol{w}$,

$$\max_{t, \lambda} \quad t \tag{3.3a}$$

$$\text{s.t. } p_i(\boldsymbol{w})^T \left(\sum_{i \in I_\delta(\boldsymbol{w})} \lambda_i p_i(\boldsymbol{w}) \right) \geq t, \tag{3.3b}$$

$$i \in I_\delta(\boldsymbol{w}), \; \lambda \in \Delta_{|I_\delta(\boldsymbol{w})|}. \tag{3.3c}$$

Compactly, we can rewrite the above LP as follows:

$$\max_{t,\lambda} t \quad \text{s.t.} \qquad \boldsymbol{B}^T\boldsymbol{B}\lambda \geq te,\ \lambda \in \Delta_{|I_\delta(\boldsymbol{w})|}, \tag{3.4}$$

where $\boldsymbol{B} \in \mathbb{R}^{2N\times|I_\delta(\boldsymbol{w})|}$, and the columns of $\boldsymbol{B}$ consists of $\{\boldsymbol{p}_i(\boldsymbol{w}),\ i \in I_\delta(\boldsymbol{w})\}$. Let (t^*, λ^*) be the optimal solution of the above LP. Hence, the descent direction we find is $\boldsymbol{d} = \boldsymbol{B}\lambda^*$. Furthermore, to guarantee sufficient decrease of LPA-SD at each iteration, an Armijo-type line-search rule is employed to find the stepsize:

$$\bar{\theta} = \max_{l\geq 0}\{\theta^l : F\left(c(\boldsymbol{w} - \theta^l\boldsymbol{d})\right) \leq F(\boldsymbol{w}) - \gamma\theta^l t^*\}, \tag{3.5}$$

where $c > 0$, $0 < \theta < 1$ and $0 < \gamma \leq 0.5$. Here, we introduce c as the scaling parameter to ensure the feasibility of $\tilde{\boldsymbol{w}} := \boldsymbol{w} - \theta^l\boldsymbol{d}$. A simple choice of c is given by

$$c = 1/\max\left\{\frac{g_1(\tilde{\boldsymbol{w}})}{\eta_1}, \ldots, \frac{g_{J+1}(\tilde{\boldsymbol{w}})}{\eta_{J+1}}, 1\right\}. \tag{3.6}$$

It's easy to check that any limit point of Algorithm 1 is a KKT point of Problem (2.7). A similar proof can be found in [7].

Algorithm 1 LPA-SG for robust beamforming

Input: $\boldsymbol{w}^0 \in \mathbb{R}^{2N}$, $\delta_0, \epsilon_0, \theta \in (0,1)$ and $\gamma \in (0, 0.5]$.
for $k = 0, 1, 2, \ldots$ **do**
  Compute $F(\boldsymbol{w}^k)$ and the active index set $I_{\delta_k}(\boldsymbol{w}^k)$.
  Compute the active constraint set $I_{\epsilon_k}(\boldsymbol{w}^k)$ and the basis matrix $\boldsymbol{U}$ via *QR decomposition*.
  Compute the projected gradient $\boldsymbol{p}_i(\boldsymbol{w}), i \in I_{\delta_k}(\boldsymbol{w}^k)$.
  Solve the LP (3.4) to get (t^*, λ^*), then set $\boldsymbol{d}^k = \boldsymbol{B}\lambda^*$.
  if $\|\boldsymbol{d}^k\| \leq \delta^k$ **then**
    $\delta^{k+1} = \delta^k/2$, $\epsilon^{k+1} = \epsilon^k/2$
  end if
  Choose $\bar{\theta}$, c via (3.5), (3.6) and $\boldsymbol{w}^{k+1} = c(\boldsymbol{w}^k - \bar{\theta}\boldsymbol{d}^k)$.
end for

3.2. Multi-block ADMM

Here, we consider to use multi-block ADMM to solve (2.7) directly. Firstly, we reformulate (2.7) as follows,

$$\min_{\boldsymbol{w},t} \quad t \tag{3.7a}$$
$$\text{s.t. } \|\boldsymbol{H}_i^T\boldsymbol{w}\| \geq \sqrt{-t},\ i = 1, \ldots, M, \tag{3.7b}$$
$$d_j\|\boldsymbol{w}\| + \|\boldsymbol{A}_j^T\boldsymbol{w}\| \leq \eta_j,\ j = 1, \ldots, J, \tag{3.7c}$$
$$\|\boldsymbol{w}\| \leq P. \tag{3.7d}$$

Introducing the auxiliary and slackness variables to split the constraints,

$$\min_{\boldsymbol{w},t,x,y,z,u,v} \quad t \tag{3.8a}$$
$$\text{s.t. } \|x_i\| - u_i = \sqrt{-t},\ i = 1, \ldots, M, \tag{3.8b}$$
$$d_j\|y\| + \|z_j\| + v_j = \eta_j,\ j = 1, \ldots, J, \tag{3.8c}$$
$$x_i = \boldsymbol{H}_i^T\boldsymbol{w},\ i = 1, \ldots, M, \tag{3.8d}$$
$$y = \boldsymbol{w},\ z_j = \boldsymbol{A}_j^T\boldsymbol{w},\ j = 1, \ldots, J, \tag{3.8e}$$
$$\|y\| \leq P,\ u \geq 0,\ v \geq 0. \tag{3.8f}$$

Here, the augmented Lagrangian for (3.8) is defined as following,

$$\begin{aligned}&\mathcal{L}(\boldsymbol{w}, t, x, y, z, u, v; \lambda, \gamma, \zeta, \alpha, \beta) = t + \\ &\sum_{i=1}^{M}\lambda_i(\|x_i\| - u_i - \sqrt{-t}) + \sum_{j=1}^{J}\gamma_j(d_j\|y\| + \|z_j\| + v_j - \eta_j) + \\ &\sum_{i=1}^{M}\zeta_i^T(x_i - \boldsymbol{H}_i^T\boldsymbol{w}) + \alpha^T(y - \boldsymbol{w}) + \sum_{j=1}^{J}\beta_j^T(z_j - \boldsymbol{A}_j^T\boldsymbol{w}) + \frac{\rho_1}{2} \\ &\sum_{i=1}^{M}(\|x_i\| - u_i - \sqrt{-t})^2 + \frac{\rho_2}{2}\sum_{j=1}^{J}(d_j\|y\| + \|z_j\| + v_j - \eta_j)^2 + \\ &\frac{\rho_3}{2}\sum_{i=1}^{M}\|x_i - \boldsymbol{H}_i^T\boldsymbol{w}\|^2 + \frac{\rho_4}{2}\|y - \boldsymbol{w}\|^2 + \frac{\rho_5}{2}\sum_{j=1}^{J}\|z_j - \boldsymbol{A}_j^T\boldsymbol{w}\|^2.\end{aligned}$$

Then we can minimize the augmented Lagrangian with fixed positive $\rho_1, \rho_2, \rho_3, \rho_4$ and ρ_5. The update for the multi-block ADMM follows the traditional update of ADMM. For each subproblem, the closed-form solution can be computed. Because of the limitation of the space, we only provide closed-form solutions for some tough subproblems,

$$t^{k+1} = -\left(\frac{1}{\rho_1 M - 2}\max\left\{0, \rho_1\sum_{i=1}^{M}(\|x_i^k\| - u_i^k) + \sum_{i=1}^{M}\lambda_i^k\right\}\right)^2.$$

$$x_i^{k+1} = \min\left\{\frac{b_1 - \|c_1\|}{2a_1}, 0\right\}\frac{c_1}{\|c_1\|}, \text{ where } a_1 = \frac{1}{2}(\rho_1 + \rho_3),$$

$$b_1 = \lambda_i^k - \rho_1(u_i^{k+1} + \sqrt{-t^{k+1}}), c_1 = \xi_i^k - \rho_3\boldsymbol{H}_i\boldsymbol{w}^k.$$

$$y^{k+1} = \begin{cases}\boldsymbol{0}, & \text{if } \|c_2\| < b_2,\\ \frac{b_2 - \|c_2\|}{2a_2}\frac{c_2}{\|c_2\|}, & \text{if } b_2 \leq \|c_2\| \leq 2a_2 + b_2,\\ -\frac{c_2}{\|c_2\|}, & \text{if } \|c_2\| > 2a_2 + b_2,\end{cases}$$

where $a_2 = \frac{\rho_2}{2}\sum_{j=1}^{J}d_j^2 + \frac{\rho_4}{2},\ c_2 = \alpha^k - \rho_4\boldsymbol{w}^k,$

$$b_2 = (\gamma^k)^T d + \rho_2\sum_{j=1}^{J}(\|z_j^k\| + v_j^{k+1} - \eta_j)d_j.$$

$$z_j^{k+1} = \min\left\{\frac{b_3 - \|c_3\|}{2a_3}, 0\right\}\frac{c_3}{\|c_3\|}, \text{ where } a_3 = \frac{1}{2}(\rho_2 + \rho_5),$$

$$b_3 = \gamma_j^{k+1} + \rho_2(d_j\|y^{k+1}\| + v_j^{k+1} - \eta_j),\ c_3 = \beta_j^k - \rho_5\boldsymbol{A}_j^T\boldsymbol{w}^k.$$

In the update of t, we choose $\rho_1 > \frac{2}{M}$. Note that in the update of $\boldsymbol{w}^k$, the matrix inverse can be computed in advance for fixed ρ_3, ρ_4, ρ_5, thus all the updates are very cheap.

4. NUMERICAL RESULTS

In this section, we provide our numerical results to demonstrate the solution quality and efficiency of the LPA-SD and multi-block ADMM. Firstly, we compare both the value of SNR and computation time of LPA-SD, ADMM with SCA method under the massive MIMO setting. Secondly, we compare LPA-SD, ADMM, SCA with SDR to check the solution quality under the small-scale case. Here, our codes are implemented in MATLAB R2018a and the tests are conducted on

64-bit Windows desktop with Intel(R) Core(TM) i5 (3.4GHz) and 8GB of RAM.

We generate the data as follows: the SU's actual channel $\boldsymbol{h}_i$ follows the standard complex normal distribution $\boldsymbol{h}_i \sim \mathcal{CN}(0, \boldsymbol{I})$, and PU's estimated channel $\boldsymbol{a}_j \sim \mathcal{CN}(0, \boldsymbol{I}/\sqrt{N})$. Besides, noise variance $\sigma_i = 1$ for all users, and the transmit power $P = 1$. Furthermore, the bounds of PU channel errors $\boldsymbol{d}$ are generated from the uniform distribution $d \sim U(0, 1)$, and the upper bounds of IT are set as $\eta = d + n$, where n also follows the uniform distribution $n \sim U(0, 1)$.

We randomly choose the same initial point for LPA-SD, ADMM and SCA, and use the same stopping criterion as

$$|F(\boldsymbol{w}^k) - F(\boldsymbol{w}^{k-5})|/|F(\boldsymbol{w}^k)| \leq 10^{-4}, \; k \geq 5, \qquad (4.1)$$

i.e., we check the change of function value every 5 iterations, and once the change is small, we stop the algorithm. We repeat the tests 100 times and take the average to avoid the influence of randomness.

4.1. Computational Efficiency

Here, we will compare our LPA-SD, multi-block ADMM with standard SCA algorithm (SCA-MOSEK). In detail, we linearize the piece $f_i(w)$ at the point w^k in the iteration $k \geq 0$, just as the technique in [11], and employ MOSEK to solve the sub-problem. In LPA-SD algorithm, we set the parameters as $\delta_0 = \epsilon_0 = \theta = 0.5$, $\bar{\delta} = 10^{-4}$ and $\gamma = 1/2$. In ADMM, we choose $\rho_1 = \rho_2 = \rho_3 = \rho_5 = 0.5$ and $\rho_4 = 1$ in the first case, while $\rho_1 = \rho_2 = \rho_3 = \rho_5 = 0.25$ and $\rho_4 = 0.5$ in the second case.

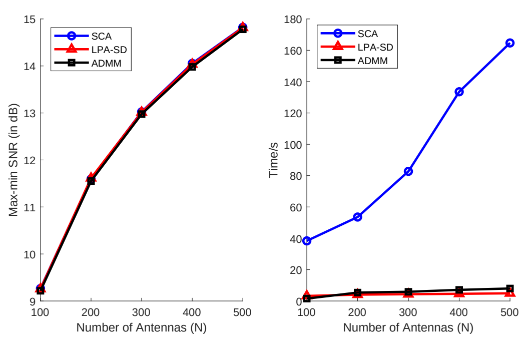

Fig. 1. The worst SUs' SNR and computation time scale with N while $M = 50$ and $J = 3$.

In Fig. 1, we fix the number of users $M = 50$ and the number of constraints $J = 3$, and let the number of antennas N change from 100 to 500; While in Fig. 2, we fix $N = 30$ and $J = 3$, and let M change from 100 to 500. According the results in both figure, we can see that LPA-SD and ADMM can achieve the similar SNR as the SCA through less computational time. Besides, we can observe that LPA-SD is insensitive to the number of users and antennas. With the increasing number of users or antennas, the advantages of LPA-SD and ADMM is becoming more and more obvious.

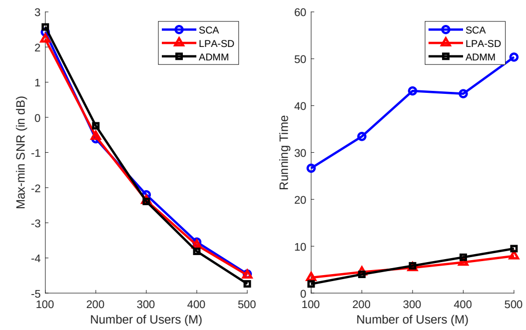

Fig. 2. The worst SUs' SNR and computation time scale with M while $N = 30$ and $J = 3$.

4.2. Solution Quality

In this part, we use the SDR and SCA-MOSEK as the benchmark to check the solution quality of LPA-SD and multi-block ADMM in terms of SNR. In the first test, we fix the number of users $M = 5$ and the number of PUs $J = 3$, then increase the number of antennas N from 5 to 25; While in the second test, we fix $N = 5$ and $J = 3$, then increase M from 5 to 25. In SDR, the number of Gaussian randomization is set as 10000. The parameters setting in LPA-SD is same as before. In ADMM, we choose $\rho_1 = \rho_2 = \rho_3 = \rho_5 = 5$ and $\rho_4 = 10$ in the first cases, while $\rho_4 = 20$, and others are same in the second case. It is easy to see that both LPA-SD and ADMM return quite good solutions under the above scenarios.

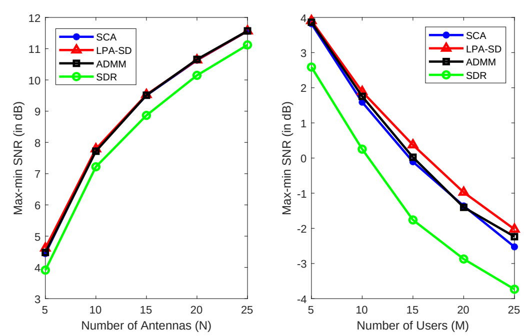

Fig. 3. Small-size robust MIMO

5. CONCLUSION

In this paper, we proposed two efficient and effective first-order methods, LPA-SD and multi-block ADMM, to solve the robust beamforming problem with interference temperature constraints. The experimental results corroborate that our algorithms have a great potential in comparison with SDR and SCA in terms of both computational efficiency and accuracy. An interesting future direction is to applying these first-order methods to solve other robust problems in the MIMO system, like robust multi-group multicast beamforming [3, 13].

6. REFERENCES

[1] J. Bolte, S. Sabach, and M. Teboulle. Nonconvex Lagrangian-Based Optimization: Monitoring Schemes and Global Convergence. *ArXiv e-prints*, Jan. 2018.

[2] J. V. Burke, F. E. Curtis, A. S. Lewis, M. L. Overton, and L. E. Simões. Gradient sampling methods for nonsmooth optimization. *arXiv preprint arXiv:1804.11003*, 2018.

[3] Z. Chen, W. Zhang, and G. Wei. Robust transmit beamforming for multigroup multicasting. In *Vehicular Technology Conference (VTC Fall), 2012 IEEE*, pages 1–5. IEEE, 2012.

[4] K. Huang and N. D. Sidiropoulos. Consensus-admm for general quadratically constrained quadratic programming. *IEEE Transactions on Signal Processing*, 64(20):5297–5310, 2016.

[5] Y. Huang, Q. Li, W.-K. Ma, and S. Zhang. Robust multicast beamforming for spectrum sharing-based cognitive radios. *IEEE Transactions on Signal Processing*, 60(1):527, 2012.

[6] M. S. Ibrahim, A. Konar, M. Hong, and N. D. Sidiropoulos. Mirror-prox sca algorithm for multicast beamforming and antenna selection. *arXiv preprint arXiv:1803.00678*, 2018.

[7] R. Jiang, H. Liu, and A. M. So. Lpa-sd: An efficient first-order method for single-group multicast beamforming. In *2018 IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, pages 1–5, June 2018.

[8] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo. Convex transmit beamforming for downlink multicasting to multiple co-channel groups. In *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, volume 5, pages V–V. IEEE, 2006.

[9] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo. Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups. *IEEE Transactions on Signal Processing*, 56(3):1268–1279, 2008.

[10] A. Konar and N. D. Sidiropoulos. A fast approximation algorithm for single-group multicast beamforming with large antenna arrays. In *Signal Processing Advances in Wireless Communications (SPAWC), 2016 IEEE 17th International Workshop on*, pages 1–5. IEEE, 2016.

[11] A. Konar and N. D. Sidiropoulos. Fast approximation algorithms for a class of non-convex qcqp problems using first-order methods. *IEEE Transactions on Signal Processing*, 65(13):3494–3509, July 2017.

[12] A. Konar and N. D. Sidiropoulos. First-order methods for fast feasibility pursuit of non-convex qcqps. *IEEE Transactions on Signal Processing*, 65(22):5927–5941, Nov 2017.

[13] K. L. Law, I. Wajid, and M. Pesavento. Robust downlink beamforming in multi-group multicasting using trace bounds on the covariance mismatches. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 3229–3232. IEEE, 2012.

[14] Z.-Q. Luo, W.-K. Ma, A. M.-C. So, Y. Ye, and S. Zhang. Semidefinite relaxation of quadratic optimization problems. *IEEE Signal Processing Magazine*, 27(3):20–34, 2010.

[15] K. T. Phan, S. A. Vorobyov, N. D. Sidiropoulos, and C. Tellambura. Spectrum sharing in wireless networks via qos-aware secondary multicast beamforming. *IEEE Transactions on signal processing*, 57(6):2323–2335, 2009.

[16] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo. Transmit beamforming for physical-layer multicasting. *IEEE Trans. Signal Processing*, 54(6-1):2239–2251, 2006.

[17] S. X. Wu, Q. Li, W.-K. Ma, and A. M.-C. So. Robust transmit designs for an energy harvesting multicast system. In *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*, pages 4748–4752. IEEE, 2014.

[18] S. X. Wu, M.-C. Yue, A. M.-C. So, and W.-K. Ma. Sdr approximation bounds for the robust multicast beamforming problem with interference temperature constraints. In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, pages 4054–4058. IEEE, 2017.

[19] L. Yang, T. K. Pong, and X. Chen. Alternating direction method of multipliers for a class of nonconvex and nonsmooth problems with applications to background/foreground extraction. *SIAM Journal on Imaging Sciences*, 10(1):74–110, 2017.

[20] L. Zhang, Y.-C. Liang, Y. Xin, and H. V. Poor. Robust cognitive beamforming with partial channel state information. *IEEE Transactions on Wireless Communications*, 8(8), 2009.

[21] G. Zheng, K.-K. Wong, and B. E. Ottersten. Robust cognitive beamforming with bounded channel uncertainties. *IEEE Trans. Signal Processing*, 57(12):4871–4881, 2009.