

# Convex-Concave Procedure for Weighted Sum-Rate Maximization in a MIMO Interference Network

Seungil You

Lijun Chen

Youjian (Eugene) Liu

**Abstract**—The weighted sum-rate maximization in a general multiple-input multiple-output (MIMO) interference network has known to be a hard nonconvex problem, mainly due to the interference between different links. In this paper, by exploring the special structure of the sum-rate function being a difference of concave functions, we apply the convex-concave procedure to the weighted sum rate maximization to handle non-convexity. With the introduction of a certain damping term, we establish the monotonic convergence of the proposed algorithm. Numerical examples show that the introduced damping term, which may be needed for convergence, slows down the convergence but helps with finding a better solution in the network with high interference. Even though our algorithm has a slower convergence than some existing ones, it has the guaranteed convergence and can handle more general constraints and thus provides a general solver that can find broader applications.

**Index Terms**—Convex-concave procedure, weighted sum-rate maximization, interference networks, multiple-input multiple-output (MIMO).

## I. INTRODUCTION

Channel capacity is a fundamental quantity in information theory that specifies the hard limit in communications [1]. It is well-known that the optimal joint transmit signals that maximize the sum capacity for parallel Gaussian channels have the water-filling structure. However, in general MIMO additive Gaussian channels, finding the optimal joint transmit signals is a hard and open problem, as the interference between channels may make the problem highly nonconvex.

In this paper, we consider the joint transmit signals design for the weighted sum-rate maximization in a MIMO B-MAC network, a combination of multiple interfering broadcast channels (BC) and multiaccess channels (MAC), with the transmit signals and noises being Gaussian and the channel states known at the transmitters [2]. Many algorithms have been proposed for this problem; see, *e.g.*, the iterative weighted MMSE (Minimization of Mean Squared Error) algorithms that exploit the relation between the mutual information and the MMSE [3], [4], the polite water-filling algorithm that exploits a variant of water-filling structure at the optimum to solve the KKT conditions [2], and the references therein. These algorithms have various limitations, *e.g.*, some of them are not able to handle general constraints or do not have guaranteed convergence.

We have recently exploited the minimax Lagrangian duality to design an efficient algorithm for the weighted sum-rate maximization [5], which is inspired by our new algorithm [6]. But at this point we only handle the affine constraints. The aim of this paper is to propose a general solver to handle convex constraints such as nonlinear power constraints and the condition number of power covariance matrix, etc.

S. You is with the Division of Engineering and Applied Science, California Institute of Technology, Pasadena, CA 91125, USA (email: syou@caltech.edu). L. Chen and Y. Liu are with the College of Engineering and Applied Science, University of Colorado, Boulder, CO 80309, USA (emails: {lijun.chen, eugeneliu}@colorado.edu).

By exploring the special structure of the sum-rate function being a *difference of concave functions*, we apply the convex-concave procedure to the weighted sum-rate maximization to handle non-convexity [7]. Specifically, we add and subtract an appropriate quadratic term, termed a damping term, to the sum-rate function and then apply the convex-concave procedure to obtain an algorithm with guaranteed convergence. The resulting algorithm generates a sequence of convex optimization problems to obtain a (local) optimal solution to the weighted sum-rate maximization, and the convex problem at each iteration can be solved efficiently using fast (polynomial time) and numerically stable (robust to the numerical errors) methods such as the interior-point method. Numerical examples show that the introduced damping term, which may be needed for convergence, slows down the convergence but helps with finding a better solution in the network with high interference.

Even though our algorithm has a slower convergence than some well-known existing algorithms such as the iterative weighted MMSE algorithm [4] and the polite water-filling algorithm [2], it has the guaranteed convergence and can handle more general constraints and thus provides a general solver that can find broader applications.

The similar idea has been proposed in [8], and [9] to solve the weighted sum-rate maximization with the per-link power constraint, and [10] in terms of soft interference nulling scheme. To some extent, this paper can be seen as a generalization of the results in [9], [10]. The main contribution of this paper is to provide a general algorithm that can handle general convex constraints (not necessarily affine) and ensure the monotonic convergence of the algorithm by augmenting a damping term which leads to better convergence and in many cases better solutions. From this perspective, the aforementioned works can be seen as special cases of this paper.

## II. SYSTEM MODEL AND PROBLEM FORMULATION

Consider a MIMO B-MAC network, introduced in [2], that consists of a set  $\mathcal{N}$  of interfering data links, with link  $n \in \mathcal{N}$  equipped with  $T_n$  antennas at the transmitter and  $R_n$  antennas at the receiver. The received signal at the receiver of the link  $n$  is given by

$$\mathbf{y}_n = \sum_{k=1}^N \mathbf{H}_{nk} \mathbf{x}_k + \mathbf{w}_n,$$

where  $\mathbf{x}_k \in \mathbb{C}^{T_k}$  is a circularly symmetric complex Gaussian transmit signal of the  $k$ th link,  $\mathbf{H}_{nk} \in \mathbb{C}^{R_n \times T_k}$  is the channel matrix between the transmitter of the  $k$ th link and the receiver of the  $n$ th link, and  $\mathbf{w}_n \in \mathbb{C}^{R_n}$  is circularly symmetric complex Gaussian noise vector with identity covariance matrix.

If the channel matrices are known at the transmitter, then an achievable rate of link  $n \in \mathcal{N}$  is given by

$$R_n(\boldsymbol{\Sigma}, \boldsymbol{\Omega}) = \log |\boldsymbol{\Omega}_n + \mathbf{H}_{nn} \boldsymbol{\Sigma}_n \mathbf{H}_{nn}^+| - \log |\boldsymbol{\Omega}_n|, \quad (1)$$

where  $\Sigma = (\Sigma_n; n \in \mathcal{N})$  with  $\Sigma_n = \mathbb{E}[\mathbf{x}_n \mathbf{x}_n^+]$  the power covariance matrix at the link  $n$ ;  $\Omega = (\Omega_n; n \in \mathcal{N})$  with

$$\Omega_n = \mathbf{I} + \sum_{k \in \mathcal{N} \setminus \{n\}} \mathbf{H}_{nk} \Sigma_k \mathbf{H}_{nk}^+$$

the interference-plus-noise covariance matrix at the link  $n$ ; and ‘+’ denotes the Hermitian of a matrix. We aim to maximize the weighted sum-rate under general convex constraints

$$\begin{aligned} & \underset{\Sigma, \Omega}{\text{maximize}} && f_0(\Sigma, \Omega) = \sum_{n \in \mathcal{N}} w_n R_n(\Sigma, \Omega) \\ & \text{subject to} && \Sigma_n \succeq \mathbf{0}, n \in \mathcal{N}, \\ & && f_i(\Sigma) \leq 0, i = 1, \dots, n_f, \\ & && h_j(\Sigma) = 0, j = 1, \dots, n_h, \\ & && \Omega_n = \mathbf{I} + \sum_{k \in \mathcal{N} \setminus \{n\}} \mathbf{H}_{nk} \Sigma_k \mathbf{H}_{nk}^+, n \in \mathcal{N}. \end{aligned} \quad (2)$$

Here  $f_i$  are convex functions of  $\Sigma$ ,  $h_j$  are affine functions of  $\Sigma$ , and  $n_f, n_h$  are the number of inequality constraints and equality constraints, respectively. The Linear Matrix Inequality  $\mathbf{X} \succ \mathbf{0}$  ( $\succeq \mathbf{0}$ ) means that  $\mathbf{X}$  is a positive (semi)definite matrix.

As the objective function is not a concave function in general, the weighted sum-rate maximization (2) is not a convex program; see, e.g., [11], [12] for the theory of convex optimization and its applications to communications, especially the role of the semidefinite program (SDP) and the linear matrix inequality (LMI).

For notational convenience, we denote the feasible set of the weighted sum-rate maximization problem (2) as  $\mathcal{S}$ , which is a convex set, and its relative interior  $\text{ri}(\mathcal{S}) = \{\Sigma | \Sigma_n \succ \mathbf{0}, f_i(\Sigma) < 0, h_j(\Sigma) = 0, \Omega_n = \mathbf{I} + \sum_{k \in \mathcal{N} \setminus \{n\}} \mathbf{H}_{nk} \Sigma_k \mathbf{H}_{nk}^+\}$ , i.e., the set of feasible points that satisfy the strict inequalities.

**Example 1: The Total Power Constraint.** Suppose that the total power of all transmitters is bounded by  $P_T$ . The corresponding constraint function is given by

$$f(\Sigma) = \sum_{n \in \mathcal{N}} \text{Tr}(\Sigma_n) - P_T,$$

which is an affine function of  $\Sigma$ . The constraints of the weighted sum-rate maximization (2) reduce to

$$\begin{aligned} \Sigma_n \succeq \mathbf{0}, n \in \mathcal{N}, & \quad \sum_{n \in \mathcal{N}} \text{Tr}(\Sigma_n) - P_T \leq 0 \\ \Omega_n = \mathbf{I} + \sum_{k \in \mathcal{N} \setminus \{n\}} \mathbf{H}_{nk} \Sigma_k \mathbf{H}_{nk}^+, & \quad n \in \mathcal{N}. \end{aligned}$$

The feasible set  $\mathcal{S}$  is compact, and its relative interior  $\text{ri}(\mathcal{S})$  is not empty, since there exist positive definite matrices  $\Sigma_n = \epsilon \mathbf{I}$ ,  $0 < \epsilon < P_T / \sum_{n=1}^N \text{Tr}(\Sigma_n)$  satisfying the strict inequality  $\sum_{n=1}^N \text{Tr}(\Sigma_n) - P_T < 0$ .

**Example 2: The Per-Link Power Constraints.** Suppose that the transmitter power of the  $n$ th link is bounded by  $P_n$ . The corresponding constraint functions are given by

$$f_n(\Sigma) = \text{Tr}(\Sigma_n) - P_n, \quad n \in \mathcal{N},$$

which are all affine. The constraints of the weighted sum-rate maximization (2) reduce to

$$\begin{aligned} \Sigma_n \succeq \mathbf{0}, \text{Tr}(\Sigma_n) - P_n \leq 0, & \quad n \in \mathcal{N} \\ \Omega_n = \mathbf{I} + \sum_{k \in \mathcal{N} \setminus \{n\}} \mathbf{H}_{nk} \Sigma_k \mathbf{H}_{nk}^+, & \quad n \in \mathcal{N}. \end{aligned}$$

Similarly, the feasible set  $\mathcal{S}$  of the above problem is compact and its relative interior  $\text{ri}(\mathcal{S})$  is not empty.

**Example 3: Maximum Beam Power Constraints.** Suppose that the total power of all transmitters is bounded by  $P_T$ , and moreover, the power of each beam at each link is bounded by  $P_n$ . This constraint is equivalent to all eigenvalues of power covariance matrix  $\Sigma_n$  being bounded by  $P_n$ , which is nonlinear but convex. The constraints of the weighted sum-rate maximization (2) reduce to

$$\begin{aligned} \sum_{n \in \mathcal{N}} \text{Tr}(\Sigma_n) - P_T \leq 0, \quad \mathbf{0} \preceq \Sigma_n \preceq P_n \mathbf{I}_n, & \quad n \in \mathcal{N} \\ \Omega_n = \mathbf{I} + \sum_{k \in \mathcal{N} \setminus \{n\}} \mathbf{H}_{nk} \Sigma_k \mathbf{H}_{nk}^+, & \quad n \in \mathcal{N}, \end{aligned}$$

where  $\mathbf{I}_n$  is the identity matrix with proper dimension. We can easily verify that the feasible set  $\mathcal{S}$  is compact and its relative interior  $\text{ri}(\mathcal{S})$  is not empty.

**Example 4: Condition Number Constraints.** Suppose that the total power of all transmitters is bounded by  $P_T$ , and moreover, we want to fully utilize the signal space. This can be done by imposing a conditional number constraint on power covariance matrix  $\Sigma_n$ , say,  $\frac{\lambda_{\max}(\Sigma_n)}{\lambda_{\min}(\Sigma_n)} \leq \kappa_n$ , where  $\lambda_{\max}, \lambda_{\min}$  are the maximal and the minimal eigenvalues respectively. We can easily see that this constraint forces to the ratio between the maximum beam power and the minimum beam power at link  $n$  is bounded by  $\kappa_n$ . Although this constraint is non-convex in its ‘naive’ form, it can be represented as LMI by introducing a slack variable. The constraints of the weighted sum-rate maximization (2) can be written as

$$\begin{aligned} \sum_{n \in \mathcal{N}} \text{Tr}(\Sigma_n) - P_T \leq 0, \quad \gamma_n \mathbf{I}_n \preceq \Sigma_n \preceq \kappa_n \gamma_n \mathbf{I}_n, & \quad n \in \mathcal{N} \\ \Omega_n = \mathbf{I} + \sum_{k \in \mathcal{N} \setminus \{n\}} \mathbf{H}_{nk} \Sigma_k \mathbf{H}_{nk}^+, & \quad n \in \mathcal{N}, \end{aligned}$$

where  $\gamma_n > 0$  is a scalar slack variable (an additional decision variable, but does not change the structure of the problem). Again, we can easily verify the feasible set  $\mathcal{S}$  is compact and its relative interior  $\text{ri}(\mathcal{S})$  is not empty.

Note that the link data rate (1) is the difference of two concave functions in  $(\Sigma, \Omega)$ . So, the weighted sum-rate maximization (2) is a *difference of convex optimization* problem. For non-convex problems with such a structure, a general solution approach – the convex-concave procedure – has been proposed in [7]. We will explore the application of the convex-concave procedure to the weighted sum-rate maximization.

### III. CONVEX-CONCAVE PROCEDURE

For a maximization problem, the basic idea of the convex-concave procedure is to linearize the convex terms of the objective function to obtain a concave objective, so as to generate a sequence of convex problems that approximately solves the original non-convex problem [7]. This linearization generates the best concave approximation of the target objective function at a given point, which is a lower bound in the entire region. Therefore the convex-concave procedure maximizes the best lower bound iteratively.

Under proper conditions, the convergence of the convex-concave procedure to a stationary point is guaranteed. Moreover, there are a few efficient polynomial time and numerically stable solvers such as the interior-point method [13] that can

be used to find the optimal solution of the convex problem of each iteration. Therefore, the convex-concave procedure gives us a theoretically as well as practically good algorithm.

Note that for any given optimization problem, we can formulate various equivalent problems [11], and these different equivalent problems may have different algorithmic advantages or limitations. That said, the way we represent the objective function (as well as the constraints) changes the resulting algorithm from the convex-concave procedure. For example, suppose  $f$  and  $g$  are concave, and we want to maximize  $f - g$ . If  $g$  is concave, then  $(f + h) - (g + h)$  is the same objective function with the same structure, *i.e.*,  $f + h$  and  $g + h$  are concave. Since there are infinite number of convex functions, it is obvious that a difference of concave (convex) functions representation of any given function is not unique; see, *e.g.*, [14], [15] and the references therein for more discussion.

Consider the following equivalent problem with additional quadratic terms

$$\begin{aligned} & \underset{\Sigma, \Omega}{\text{maximize}} && \sum_{n \in \mathcal{N}} \left( w_n \left( \log |\Omega_n + \mathbf{H}_{nn} \Sigma_n \mathbf{H}_{nn}^*| \right. \right. \\ & && \left. \left. - \log |\Omega_n| \right) - \rho \left( \|\Sigma_n\|_F^2 - \|\Sigma_n^{(i)}\|_F^2 \right) \right) \quad (3) \\ & \text{subject to} && (\Sigma, \Omega) \in \mathcal{S}, \end{aligned}$$

where  $\|\Sigma_n\|_F$  is the Frobenius norm, *i.e.*,  $\|\Sigma_n\|_F = \sqrt{\text{Tr}(\Sigma_n^* \Sigma_n)}$ . By including  $-\rho \|\Sigma_n\|_F^2$  in the concave part of the objective function and  $+\rho \|\Sigma_n^{(i)}\|_F^2$  in the convex part of the objective function, we can derive different algorithms from the convex-concave procedure by varying  $\rho$ .

Applying the convex-concave procedure, we obtain the following algorithm: Let  $\Sigma^{(i)}, \Omega^{(i)}$  be the optimal solution at the  $i$ th iteration, and solve the the convex problem

$$\begin{aligned} & \underset{\Sigma, \Omega}{\text{maximize}} && \sum_{n \in \mathcal{N}} \left( w_n \left( \log |\Omega_n + \mathbf{H}_{nn} \Sigma_n \mathbf{H}_{nn}^*| \right. \right. \\ & && \left. \left. - \text{Tr} \left( \Omega_n^{(i)-1} \Omega_n \right) \right) - \rho \|\Sigma_n - \Sigma_n^{(i)}\|_F^2 \right) \quad (4) \\ & \text{subject to} && (\Sigma, \Omega) \in \mathcal{S} \end{aligned}$$

iteratively until the convergence is detected.<sup>1</sup> The detailed derivation can be found in the Appendix VII-A

On one hand, a larger  $\rho$  makes the point generated at each iteration closer to the point from the previous iteration. Since the first order approximation is valid around the initial point (*i.e.*, the point generated at the previous iteration), this leads to a more accurate approximation at each iteration but possibly a larger number of iterations. On the other hand, the damping term  $-\rho \|\Sigma_n - \Sigma_n^{(i)}\|_F^2$  with larger  $\rho$  makes the objective in (4) a steeper function and thus an easier convex problem at each iteration.

The proposed algorithm is summarized as follows.

#### Convex-Concave Procedure:

- 1) Initialize  $(\Sigma^{(0)}, \Omega^{(0)}) \in \mathcal{S}$ , set  $i = 0$ .
- 2) Generate  $(i + 1)$ th  $(\Omega, \Sigma)$  by solving (4), and set  $i = i + 1$ .
- 3) Repeat 2 until the convergence is detected.

<sup>1</sup>When  $\rho = 0$  and  $\mathcal{S}$  only contains the per-link power constraints, the above algorithm recovers [9], so our algorithm includes an existing scheme as a special case.

## IV. CONVERGENCE ANALYSIS

The convex-concave procedure generates a sequence of feasible points  $\{\Sigma^{(i)}, \Omega^{(i)}\}$ . In this section, we investigate the convergence of this sequence. All the proofs can be found in the Appendix. The first result shows the convergence of our algorithm.

**Proposition 1.** *Suppose  $\mathcal{S}$  is compact and  $\text{ri}(\mathcal{S}) \neq \emptyset$ , then the sequence  $\{\Sigma^{(i)}\}$  under the proposed algorithm converges to a stationary point  $\Sigma^{(\infty)}$  of the weighted sum-rate maximization problem (2) for  $\rho > 0$ .*

Here a stationary point refers to a (local) optimal point that satisfies the KKT conditions. Note that the convergence does not depend on the initial point, but the stationary point may depend on the initial point.

The requirements for the feasible set  $\mathcal{S}$  (compact, nonempty relative interior) hold in many practical cases, *e.g.*, all examples in the Section II which include the network with the total power constraint, and the network with the per-link power constraints

The technical condition  $\rho > 0$  ensures strict concavity (convexity) in the concave (convex) part of the objective function. If  $\rho = 0$ , this strict concavity (convexity) may not be guaranteed, depending on the channel matrices  $\mathbf{H}_{ij}$ . Without this property, the sequence of points generated under the convex-concave procedure can exhibit the limit cycle behavior, *i.e.*, oscillate among a set of stationary points, which is a weaker result than Proposition 1.

**Proposition 2.** *Suppose  $\mathcal{S}$  is compact and  $\text{ri}(\mathcal{S}) \neq \emptyset$ , then the sequence  $\{\Sigma^{(i)}\}$  under the proposed algorithm converges to a set  $\mathcal{A}$  of stationary points of the weighted sum-rate maximization problem (2), all of them achieve the same weighted sum-rate.*

Suppose  $\mathcal{A}$  contains two points,  $\Sigma^{(\infty)}$  and  $\tilde{\Sigma}^{(\infty)}$ . The above result implies that the weighted sum-rate achieved at both points is the same, but the sequence of points under convex-concave procedure may oscillate between these two points. This oscillatory behavior does not occur if  $\rho > 0$ . Therefore, the introduction of the damping term  $-\rho \|\Sigma_n - \Sigma_n^{(i)}\|_F^2$  can avoid oscillation and guarantee convergence (even though it may slow down the convergence as discussed in Section V).

Moreover, the proposed algorithm based on the convex-concave procedure generates a monotone increasing sequence in the weighted sum-rate.

**Proposition 3.** *The weighted sum-rate is always non-decreasing under the proposed algorithm, *i.e.*,  $f_0(\Sigma^{(i)}, \Omega^{(i)}) \leq f_0(\Sigma^{(i+1)}, \Omega^{(i+1)})$  for all  $i = 0, 1, \dots$ .*

## V. NUMERICAL EXAMPLES

In this section, we provide numerical examples to complement the analysis in the previous sections. Consider a network with  $|\mathcal{N}| = 10$  links, corresponding to 10 transmitter-receiver pairs that interfere with each other. The number of antennas at the transmitter and receiver of each link is uniformly drawn from  $\{2, 3, 4\}$ . The channel matrices have zero-mean, unit-variance, *i.i.d.* complex Gaussian entries. We will consider and compare the networks with low, moderate, and high interference, which are characterized by scaling the interference channel matrices  $\mathbf{H}_{ij}$ ,  $i \neq j$  with a factor of 0.1, 1, and 10 respectively. The weights  $w_n$ 's are uniformly

drawn from  $[0.5, 1]$ . Due to space limit, we only report part of numerical examples in this paper.

Each iteration of the proposed algorithm involves solving a max-det problem [16], for which we use SDPT3 [17] combined with the problem parser YALMIP [18].

#### A. The network with the per-link power constraints

We consider Example 2 in Section II. Figure 1 shows the convergence of the proposed algorithm, where the per-link power budgets  $P_n$ 's are uniformly drawn from  $\{1, 2, \dots, 10\}$ . We see that the larger  $\rho$  is, the slower the convergence, as the  $\rho$  term introduce an additional damping term in the optimization at each iteration. In the network with high interference, however, this damping term may help with finding a better solution which has larger weighted sum-rate. Roughly speaking, when the interference is high the weighted sum-rate maximization becomes highly non-convex, which makes the problem harder; but the additional term helps with finding a better solution by jumping out of "traps" of local minima or maxima. This non-convexity also affects the convergence speed, and the stronger the inference, the slower the convergence is.

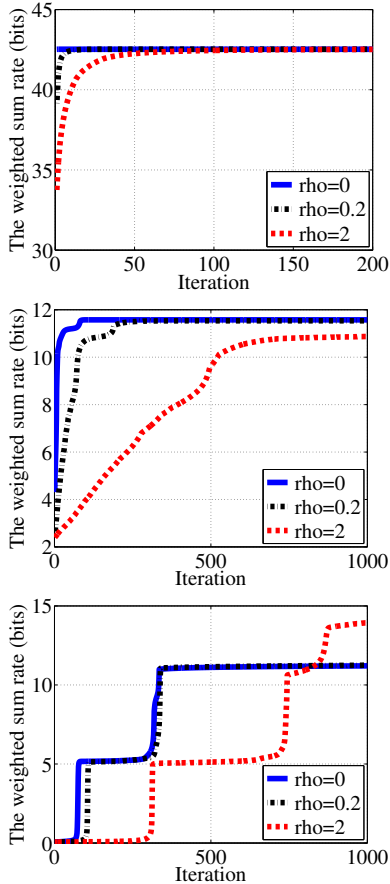


Fig. 1. The network with low (top panel), moderate (middle panel), and high (bottom panel) interference and per-link power constraint.

Note that our algorithm shows a slower convergence than some existing algorithms [2], [4], [6]. This is mainly because these existing algorithms exploit the particular structures of a restrictive class of power covariance constraints (e.g., the affine power covariance constraints). However, the convex-concave procedure can handle more general constraints as in

the last two examples described in Section II, so our algorithm provides a more general solver. That said, our algorithm trades off efficiency for generality.

#### B. The network with nonlinear constraints

We consider Example 3 and Example 4 in Section II. Figure 2 shows the monotonic convergence of the proposed algorithms in a network with low interference, where the total power budget  $P_T = 10$ , per-beam power bounds  $P_n = 2$ , and per-link conditional numbers  $\kappa_n = 2$ .

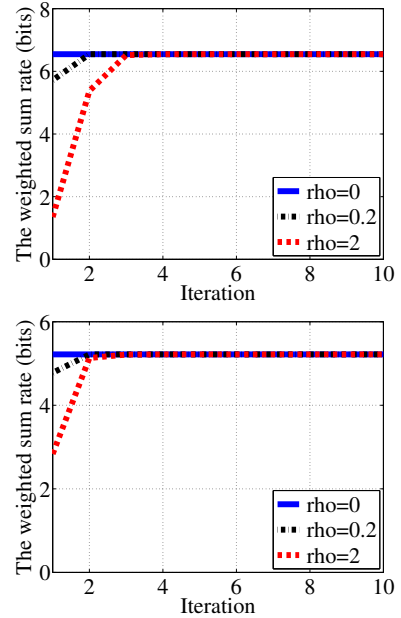


Fig. 2. The network with maximum beam power constraints (top panel) and condition number constraints (bottom panel).

## VI. CONCLUSION

We have applied the convex-concave procedure to the weighted sum-rate maximization in the MIMO B-MAC network. By adding and subtracting a quadratic term, we obtain an algorithm that guarantees monotonic convergence. Numerical examples confirm the monotonic convergence of the proposed algorithm and show the effect of the introduced damping term. The damping term slows down the convergence but guarantees the convergence, and may help with finding a better solution in the network with high interference.

Even though our algorithm has a slower convergence than the some of existing ones, it can handle more general constraints and thus provides a general solver that can find broader applications; and we are currently exploring its practical applications. Moreover, as the objective function at each iteration is separable, the algorithm naturally admits a distributed implementation as in [9], which we will investigate in the future.

## VII. APPENDIX

#### A. Linearization

Without loss of generality, assume that we linearize the function around the point  $(\mathbf{\Omega}^{(0)}, \mathbf{\Sigma}^{(0)})$ . First, consider the

logdet term

$$-\log |\mathbf{\Omega}| = -\log |\mathbf{\Omega}^{(0)}| - \text{Tr} \left( \left( \mathbf{\Omega}^{(0)} \right)^{-1} (\mathbf{\Omega} - \mathbf{\Omega}^{(0)}) \right) + O(\|\mathbf{\Omega} - \mathbf{\Omega}^{(0)}\|_F^2).$$

Now consider the quadratic term (the Frobenius norm)

$$\begin{aligned} & -\|\mathbf{\Sigma}\|_F^2 + \|\mathbf{\Sigma}\|_F^2 \\ &= -\|\mathbf{\Sigma}\|_F^2 + \|\mathbf{\Sigma}^{(0)}\|_F^2 + \text{Tr} \left( \mathbf{\Sigma}^{(0)*} (\mathbf{\Sigma} - \mathbf{\Sigma}^{(0)}) \right) \\ & \quad + \text{Tr} \left( (\mathbf{\Sigma} - \mathbf{\Sigma}^{(0)})^* \mathbf{\Sigma}^{(0)} \right) + O(\|\mathbf{\Sigma} - \mathbf{\Sigma}^{(0)}\|_F^2) \\ &= -\|\mathbf{\Sigma} - \mathbf{\Sigma}^{(0)}\|_F^2 + 2\|\mathbf{\Sigma}^{(0)}\|_F^2 + O(\|\mathbf{\Sigma} - \mathbf{\Sigma}^{(0)}\|_F^2). \end{aligned}$$

Combining the linearization of both terms and dropping the constant terms, we obtain the proposed algorithm from the convex-concave procedure.

### B. Proof of Proposition 1

*Proof:* We apply Theorem 8 in [19] for the proof. From the compactness of  $\mathcal{S}$  and non emptiness of  $\text{ri}(\mathcal{S})$ , we can conclude that all the requirements in Theorem 8 in [19] are satisfied except for the strict concavity of the decomposition of the weighted sum-rate function with the damping term. To show strict concavity, consider the concave part of the objective function  $u(\mathbf{\Omega}, \mathbf{\Sigma}) = \log |\mathbf{\Omega} + \mathbf{H}\mathbf{\Sigma}\mathbf{H}^*| - \rho\|\mathbf{\Sigma}\|_F^2$ . From concavity, for any two points in the feasible set,  $(\mathbf{\Omega}_1, \mathbf{\Sigma}_1)$  and  $(\mathbf{\Omega}_2, \mathbf{\Sigma}_2)$ , and  $0 < t < 1$ , we have

$$\begin{aligned} & u(t\mathbf{\Omega}_1 + (1-t)\mathbf{\Omega}_2, t\mathbf{\Sigma}_1 + (1-t)\mathbf{\Sigma}_2) \\ & \stackrel{\text{(ineq:1)}}{\geq} t \log |\mathbf{\Omega}_1 + \mathbf{H}\mathbf{\Sigma}_1\mathbf{H}^*| + (1-t) \log |\mathbf{\Omega}_2 + \mathbf{H}\mathbf{\Sigma}_2\mathbf{H}^*| \\ & \quad - \rho\|t\mathbf{\Sigma}_1 + (1-t)\mathbf{\Sigma}_2\|_F^2 \\ & \stackrel{\text{(ineq:2)}}{\geq} t \log |\mathbf{\Omega}_1 + \mathbf{H}\mathbf{\Sigma}_1\mathbf{H}^*| + (1-t) \log |\mathbf{\Omega}_2 + \mathbf{H}\mathbf{\Sigma}_2\mathbf{H}^*| \\ & \quad - \rho t\|\mathbf{\Sigma}_1\|_F^2 - \rho(1-t)\|\mathbf{\Sigma}_2\|_F^2. \end{aligned}$$

Note that in the feasible set,  $\mathbf{\Omega} \succeq \mathbf{I}$ . Now suppose  $\mathbf{\Sigma}_1 \neq \mathbf{\Sigma}_2$ , then from strict convexity of  $\|\mathbf{\Sigma}\|_F^2$ , (ineq:2) becomes strict, provided  $\rho > 0$ . If  $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_2$  and  $\mathbf{\Omega}_1 \neq \mathbf{\Omega}_2$ , then from the strict convexity of the logdet function, (ineq:1) becomes strict. Therefore, for any pair  $(\mathbf{\Omega}_1, \mathbf{\Sigma}_1) \neq (\mathbf{\Omega}_2, \mathbf{\Sigma}_2)$ , and  $0 < t < 1$ , we have  $u(t\mathbf{\Omega}_1 + (1-t)\mathbf{\Omega}_2, t\mathbf{\Sigma}_1 + (1-t)\mathbf{\Sigma}_2) > tu(\mathbf{\Omega}_1, \mathbf{\Sigma}_1) + (1-t)u(\mathbf{\Omega}_2, \mathbf{\Sigma}_2)$ , which implies strict concavity of  $u$  over the feasible set. Similar argument applies to  $v(\mathbf{\Omega}, \mathbf{\Sigma}) = \log |\mathbf{\Omega}| - \rho\|\mathbf{\Sigma}\|_F^2$ , which implies strict concavity of  $v$  thus strict convexity of  $-v$ . ■

### C. Proof of Proposition 2

*Proof:* By the same argument as in the proof of Theorem 1, we can apply the Theorem 4 in [19] to prove the result. ■

### D. Proof of Proposition 3

*Proof:* Here we show that the weighted sum-rate always increases in each iteration.

$$\begin{aligned} & f_0(\mathbf{\Omega}^{(i+1)}, \mathbf{\Sigma}^{(i+1)}) \\ & \stackrel{(1)}{\geq} \sum_{n \in \mathcal{N}} w_n \left( \log |\mathbf{\Omega}_n^{(i+1)} + \mathbf{H}_{nn} \mathbf{\Sigma}_n^{(i+1)} \mathbf{H}_{nn}^*| - \log |\mathbf{\Omega}^{(i)}| \right. \\ & \quad \left. - \text{Tr} \left( \mathbf{\Omega}_n^{(i)-1} (\mathbf{\Omega}_n^{(i+1)} - \mathbf{\Omega}_n^{(i)}) \right) - \rho\|\mathbf{\Sigma}_n^{(i+1)}\|_F^2 \right. \\ & \quad \left. + \rho \left( \|\mathbf{\Sigma}_n^{(i)}\|_F^2 + 2\text{Tr} \left( \mathbf{\Sigma}_n^{(i)*} (\mathbf{\Sigma}^{(i+1)} - \mathbf{\Sigma}^{(i)}) \right) \right) \right) \end{aligned}$$

$$\begin{aligned} & \stackrel{(2)}{\geq} \sum_{n \in \mathcal{N}} w_n \left( \log |\mathbf{\Omega}_n^{(i)} + \mathbf{H}_{nn} \mathbf{\Sigma}_n^{(i)} \mathbf{H}_{nn}^*| - \log |\mathbf{\Omega}^{(i)}| \right. \\ & \quad \left. - \text{Tr} \left( \mathbf{\Omega}_n^{(i)-1} (\mathbf{\Omega}_n^{(i)} - \mathbf{\Omega}_n^{(i)}) \right) - \rho\|\mathbf{\Sigma}_n^{(i)}\|_F^2 \right. \\ & \quad \left. + \rho \left( \|\mathbf{\Sigma}_n^{(i)}\|_F^2 + 2\text{Tr} \left( \mathbf{\Sigma}_n^{(i)*} (\mathbf{\Sigma}^{(i)} - \mathbf{\Sigma}^{(i)}) \right) \right) \right) \\ & = f_0(\mathbf{\Omega}^{(i)}, \mathbf{\Sigma}^{(i)}), \end{aligned}$$

where the inequality (1) is from the first order condition of convex (concave) functions, and the inequality (2) is from the fact that  $(\mathbf{\Omega}^{(i+1)}, \mathbf{\Sigma}^{(i+1)})$  is a global maximum of the lower bound from the inequality (1). ■

## REFERENCES

- [1] T. M. Cover and J. A. Thomas, *Elements of information theory*. John Wiley & Sons, 2012.
- [2] A. Liu, Y. Liu, H. Xiang, and W. Luo, "Duality, polite water-filling, and optimization for MIMO B-MAC interference networks and iTrees networks," *arXiv preprint arXiv:1004.2484*, 2010.
- [3] S. S. Christensen, R. Agarwal, E. D. Carvalho, and J. M. Cioffi, "Weighted Sum-Rate Maximization using Weighted MMSE for MIMO-BC Beamforming Design," *IEEE Transactions on Wireless Communications*, vol. 7, no. 12, pp. 4792 – 4799, December 2008.
- [4] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *Signal Processing, IEEE Transactions on*, vol. 59, no. 9, pp. 4331–4340, 2011.
- [5] L. Chen and S. You, "The weighted sum rate maximization in mimo interference networks: The minimax lagrangian duality and algorithm," *CoRR*, vol. abs/1309.4034, 2013.
- [6] X. Li, S. You, L. Chen, and Y. Liu, "A New Algorithm for the Weighted Sum Rate Maximization in MIMO Interference Networks," *submitted to ISIT 2014*, January 2014.
- [7] A. L. Yuille and A. Rangarajan, "The concave-convex procedure," *Neural Computation*, vol. 15, no. 4, pp. 915–936, 2003.
- [8] C. Shi, R. A. Berry, and M. L. Honig, "Monotonic convergence of distributed interference pricing in wireless networks," in *Information Theory, 2009. ISIT 2009. IEEE International Symposium on*. IEEE, 2009, pp. 1619–1623.
- [9] S.-J. Kim and G. B. Giannakis, "Optimal resource allocation for mimo ad hoc cognitive radio networks," *Information Theory, IEEE Transactions on*, vol. 57, no. 5, pp. 3117–3131, 2011.
- [10] C. T. Ng and H. Huang, "Linear precoding in cooperative MIMO cellular networks with limited coordination clusters," *Selected Areas in Communications, IEEE Journal on*, vol. 28, no. 9, pp. 1446–1454, 2010.
- [11] S. P. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [12] Z.-Q. Luo and W. Yu, "An introduction to convex optimization for communications and signal processing," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 8, pp. 1426–1438, 2006.
- [13] Y. Nesterov, A. S. Nemirovskii, and Y. Ye, *Interior-point polynomial algorithms in convex programming*. SIAM, 1994, vol. 13.
- [14] P. D. Tao and L. T. H. An, "Convex analysis approach to DC programming: theory, algorithms and applications," *Acta Mathematica Vietnamica*, vol. 22, no. 1, pp. 289–355, 1997.
- [15] P. D. Tao *et al.*, "The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems," *Annals of Operations Research*, vol. 133, no. 1-4, pp. 23–46, 2005.
- [16] L. Vandenberghe, S. Boyd, and S.-P. Wu, "Determinant maximization with linear matrix inequality constraints," *SIAM journal on matrix analysis and applications*, vol. 19, no. 2, pp. 499–533, 1998.
- [17] K.-C. Toh, M. J. Todd, and R. H. Tütüncü, "SDPT3—a MATLAB software package for semidefinite programming, version 1.3," *Optimization Methods and Software*, vol. 11, no. 1-4, pp. 545–581, 1999.
- [18] J. Lofberg, "YALMIP: A toolbox for modeling and optimization in MATLAB," in *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*. IEEE, 2004, pp. 284–289.
- [19] G. R. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Advances in neural information processing systems*, 2009, pp. 1759–1767.