Abstract—The weighted sum-rate maximization in a general multiple-input multiple-output (MIMO) interference network has known to be a hard highly 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], and the polite water-filling algorithm that exploits a variant of water-filling structure at the optimum to solve iteratively for the KKT conditions [2]. Many of these algorithms have various limitations; e.g., some of them are not able to handle general constraints or do not even have guaranteed convergence.

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

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]. The convex-concave procedure is one of state-of-the-art techniques handling non-convex problems with general constraints. 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 high interference network.

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.

For the single-input single-output (SISO) case, this convex-concave procedure approach has been exploited in [8] and [9] which also includes distributed implementation. Despite the importance of these previous works, the convex approximation in the SISO case cannot be generalized to the MIMO case directly. For example, exponential transformation is used in [9], but the matrix exponential is not a convex function, therefore the same technique does not yield a convex program anymore. This suggests that the MIMO case requires more careful treatment. We resolve this situation by adding auxiliary variables corresponding to the interference-plus-noise covariance matrices. This transformation results in a straightforward derivation of the algorithm. This is important, because apply-

1In optimization theory, it is very often that a seemingly trivial equivalent problem formulation has significant implications in characterizing the optimality condition and designing the optimization algorithm. It is an important technique in optimization theory to explore and exploit different transformations and equivalent problem formulations.
ing the convex-concave procedure to an arbitrary function is complicated in general.

For the MIMO case we consider, the iterative convex approximation has been proposed in [10], [11] with the per-link power constraint, and [12] in terms of soft interference nulling scheme. To some extent this paper can be seen as a generalization of the results in [11], [12]. The main contribution of this paper is to provide a general algorithm that can handle nonlinear but convex constraints, and establish 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 a special case 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

\[
y_n = \sum_{k=1}^{N} H_{nk} x_k + w_n,
\]

where \( x_k \in \mathbb{C}^{T_k} \) is a circularly symmetric complex Gaussian transmit signal of the \( k \)th link, \( H_{nk} \in \mathbb{C}^{R_n \times T_k} \) is the channel matrix between the transmitter of the \( k \)th link and the receiver of the \( n \)th link, and \( w_n \in \mathbb{C}^{R_n} \) 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

\[
R_n(\Sigma; \Omega) = \log |\Omega_n + H_{nn} \Sigma_n H_{nn}^+| - \log |\Omega_n|,
\]

where \( \Sigma = (\Sigma_n; n \in \mathcal{N}) \) with \( \Sigma_n = \mathbb{E}[x_n x_n^+] \) the power covariance matrix at the link \( n \);\(^{2}\) \( \Omega_n = (\Omega_n; n \in \mathcal{N}) \) with

\[
\Omega_n = I + \sum_{k \in \mathcal{N}\setminus\{n\}} H_{nk} \Sigma_k H_{nk}^+
\]

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

\[
\max_{\Sigma, \Omega} \quad f_0(\Sigma, \Omega) = \sum_{n \in \mathcal{N}} w_n R_n(\Sigma, \Omega)
\]

subject to

\[
\begin{align*}
\Sigma_n &\succeq 0, \quad n \in \mathcal{N}, \\
f_i(\Sigma) &\leq 0, \quad i = 1, \ldots, n_f, \\
h_j(\Sigma) &= 0, \quad j = 1, \ldots, n_h, \\
\Omega_n &= I + \sum_{k \in \mathcal{N}\setminus\{n\}} H_{nk} \Sigma_k H_{nk}^+, \quad n \in \mathcal{N}.
\end{align*}
\]

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 generalized inequality \( X \succeq 0 \) means that \( X \) is a positive (semi)definite matrix.

Since the objective function is not a concave function in general, the weighted sum-rate maximization (2) is not a convex program; see, e.g., [14], [15], for the theory of convex optimization and its applications to communications.

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 as \( \text{ri}(\mathcal{S}) = \{ \Sigma | \Sigma_n \succeq 0, f_i(\Sigma) < 0, h_j(\Sigma) = 0, \Omega_n = I + \sum_{k \in \mathcal{N}\setminus\{n\}} H_{nk} \Sigma_k H_{nk}^+ \} \), i.e., the set of 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

\[
\begin{align*}
\Sigma_n &\succeq 0, \quad n \in \mathcal{N}, \\
\sum_{n \in \mathcal{N}} \text{Tr}(\Sigma_n) - P_T &\leq 0, \\
\Omega_n &= I + \sum_{k \in \mathcal{N}\setminus\{n\}} H_{nk} \Sigma_k H_{nk}^+, \quad n \in \mathcal{N}.
\end{align*}
\]

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

\[
f_n(\Sigma) = \text{Tr}(\Sigma_n) - P_n, \quad n = 1, \ldots, N,
\]

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

\[
\begin{align*}
\Sigma_n &\succeq 0, \quad \text{Tr}(\Sigma_n) - P_n \leq 0, \quad n \in \mathcal{N}, \\
\Omega_n &= I + \sum_{k \in \mathcal{N}\setminus\{n\}} H_{nk} \Sigma_k H_{nk}^+, \quad n \in \mathcal{N}.
\end{align*}
\]

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 \( P_T \). Moreover, the power of each beam at each link is bounded by \( P_n \). This constraint is equivalent to all eigenvalues of each covariance matrix \( \Sigma_n \) is bounded by \( P_n \). Henceforth, the constraints reduce to

\[
\sum_{n \in \mathcal{N}} \text{Tr}(\Sigma_n) - P_T \leq 0, \quad 0 \preceq \Sigma_n \preceq P_n I_n, \quad n \in \mathcal{N}
\]

\[
\Omega_n = I + \sum_{k \in \mathcal{N}\setminus\{n\}} H_{nk} \Sigma_k H_{nk}^+, \quad n \in \mathcal{N},
\]

where \( I_n \) is the identity matrix with proper dimension. We can easily show the feasible set \( \mathcal{S} \) is compact and its relative interior \( \text{ri}(\mathcal{S}) \) is not empty. Clearly, this example contains
Example 4: Condition Number Constraints. Suppose that the total power of all transmitters is bounded by \( P_T \). Moreover, suppose we want to fully utilize the signal space. This can be done by imposing a conditional number constraint on the transmit covariance matrix \( \Sigma_n \), say \( \lambda_{\min}(\Sigma_n) \leq \kappa_n \), where \( \lambda_{\max}, \lambda_{\min} \) are the maximum and the minimum eigenvalue respectively. From the definition of the condition number we can directly conclude that the ratio between the power consumption among any pair of beams is bounded by \( \kappa_n \). In this case, the constraints reduce to:

\[
\sum_{n \in \mathcal{N}} \text{Tr}(\Sigma_n) - P_T \leq 0, \quad \gamma_n I_n \preceq \Sigma_n \preceq \kappa_n \gamma_n I_n, \quad n \in \mathcal{N}
\]

where \( I_n \) is the identity matrix with proper dimension, and \( \gamma_n \) is a scalar slack variable. Again, we can easily show the feasible set \( \mathcal{S} \) is compact and its relative interior \( \text{ri}(\mathcal{S}) \) is not empty. This example also contains nonlinear, but convex inequalities: \( \gamma_n I_n \preceq \Sigma_n \preceq \kappa_n \gamma_n I_n \).

### III. Convex-Concave Procedure

The basic idea of the convex-concave procedure is to linearize the convex terms of the objective function to obtain a concave objective for a maximization problem, 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 we maximize 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 [16] 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 [14], 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 convex, and we want to minimize \( f - g \). If \( h \) is convex, then \( (f + h) - (g + h) \) is the same objective function with the same structure, i.e., \( f + h \) and \( g + h \) are convex. Since there are infinite number of convex functions, it is obvious that a difference of convex (concave) functions representation of any given function is not unique; see, e.g., [17], [18] and the references therein for more discussion.

Consider the following equivalent problem with additional quadratic terms

\[
\begin{align*}
\max_{\Sigma, \Omega} & \quad \sum_{n \in \mathcal{N}} \left( w_n \left( \log |\Omega_n + H_{nn} \Sigma_n H_{nn}^*| \right) - \log |\Omega_n| \right) - \rho \left( ||\Sigma_n||_F^2 - ||\Sigma_n||_F^2\right) \\
\text{subject to} & \quad (\Sigma, \Omega) \in \mathcal{S},
\end{align*}
\]

where \( ||\Sigma_n||_F \) is the Frobenius norm, i.e., \( ||\Sigma_n||_F = \sqrt{\text{Tr}(\Sigma_n^T \Sigma_n)} \). By including \(-\rho ||\Sigma_n||_F^2 \) in the convex part of the objective function and \(+\rho ||\Sigma_n||_F^2 \) 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

\[
\begin{align*}
\max_{\Sigma, \Omega} & \quad \sum_{n \in \mathcal{N}} \left( w_n \left( \log |\Omega_n + H_{nn} \Sigma_n H_{nn}^*| \right) - \text{Tr}\left( \Omega_n^{(i-1)} \Omega_n \right) \right) - \rho \left( ||\Sigma_n - \Sigma^{(i)}||_F^2 \right) \\
\text{subject to} & \quad (\Sigma, \Omega) \in \mathcal{S}
\end{align*}
\]

iteratively until the convergence is detected.\(^3\) On the one hand, a larger \( \rho \) makes the point generated at each iteration closer to that at 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 \(-\rho ||\Sigma_n - \Sigma^{(i)}||_F^2 \) 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; see the Appendix for the detailed derivation.

**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.

### 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 \( \text{ri}(\mathcal{S}) \) is not empty, then the sequence \( \{\Sigma^{(i)}\} \) under the proposed algorithm converges to a stationary point \( \Sigma^{*} \) 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

\(^3\)When \( \rho = 0 \), and \( \mathcal{S} \) only contains the per-link power constraint, the above algorithm recovers [11], so our algorithm includes an existing scheme as a special case.
does not depend on the initial point. The requirements for the feasible set $S$ (compact, nonempty relative interior) hold in many practical cases, e.g., 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) is not guaranteed, and depends on the channel matrices $H_{ij}$. 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 Theorem 1.

**Proposition 2.** Suppose $S$ is compact and $\text{ri}(S)$ is not empty, then the sequence $\{\Sigma^{(i)}\}$ under the proposed algorithm converges to a set $A$ of stationary points of the weighted sum-rate maximization problem (2), all of them achieve the same weighted sum-rate.

Suppose $A$ contains two points, $\Sigma^{(\infty)}$ and $\Sigma^{(\infty)}$. 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 $-\rho ||\Sigma_n - \Sigma^{(i)}||_F^2$ 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 $i = 0, 1, \cdots$.

### 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 $H_{ij}$, $i \neq j$ with a factor of 0.1, 1, and 10 respectively. The weights $w_n$’s are uniformly drawn from $[0.5, 1]$, for the case with total power constraint $P_T = 10$, and for the case with the per-link power constraints $P_n$’s are uniformly drawn from $\{1, 2, \cdots, 10\}$.

Each iteration of the proposed algorithm involves solving a max-det problem [19], for which we use SDPT3 [20] combined with the problem parser YALMIP [21].

**A. The network with the total power constraint**

In each iteration, we solve the following max-det problem:

$$\begin{align*}
\text{maximize} & \quad \Sigma, \Omega \\
\text{subject to} & \quad \sum_{n \in \mathcal{N}} \left( w_n \left( \log |\Omega_n + H_{nn}, \Sigma_n, H_{nn}^*| \right) \\
& \quad - \text{Tr} \left( \Omega_n^{-1} \Sigma_n \right) - \rho \|\Sigma_n - \Sigma^{(i)}\|_F^2 \right) \\
& \quad \Omega_n = I + \sum_{k \in \mathcal{N} \setminus \{n\}} H_{nk} \Sigma_k H_{nk}^*, n \in \mathcal{N}.
\end{align*}$$

(Figure 1 shows the convergence of the proposed algorithm for a network with the total power constraint. We see that a large $\rho$ makes the convergence slow in the network with low and moderate interference, 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 (see, especially, the next subsection). 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.

Note that our algorithm shows a slower convergence than existing algorithms [2], [4], [6]. This is mainly because aforementioned algorithms exploit the special structures in the total power constraint. However, the convex-concave procedure can handle more general constraints as in the last two examples, so our algorithm provides a more general solver. That said, our algorithm trades off efficiency for generality.

B. The network with the per-link power constraints

In each iteration, we solve the following max-det problem:

\[
\begin{align*}
\text{maximize} \quad & \sum_{n \in N} \left( w_n \left( \log \| \Omega_n + H_{nn} \Sigma_n H_{nn}^* \| \right) \\
- & \text{Tr} \left( \Omega_n^{(i)} \right) - \rho \| \Sigma_n - \Sigma_n^{(i)} \|_F \right) \\
\text{subject to} \quad & \Sigma_n \succeq 0, \quad \text{Tr} (\Sigma_n) \leq P_n, \quad n \in N \\
\Omega_n = I + \sum_{k \in N \setminus \{n\}} H_{nk} \Sigma_k H_{nk}^+, \quad n \in N.
\end{align*}
\]

(6)

Figure 2 shows the convergence of the proposed algorithm for a network with the per-link power constraints. We observe similar patterns of convergence as in the network with the total power constraint. The larger the \( \rho \) value, the slower the convergence is; however, when the interference is high, the damping term helps with finding a better solution.

C. The network with nonlinear constraints

Here we consider the example 3 and the example 4 in the low interference regime to show the convergence of the algorithm. For simplicity we set \( P_n = 2 \), and \( \kappa_n = 2 \). In figure 3, we can see the monotonic convergence.

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 for typical cases such as the networks with the total power constraint and with the per-link power constraints 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 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 working towards practical engineering applications of the algorithm. Moreover, as the objective function at each iteration is separable, it naturally admits a distributed implementation as in [11], which we will investigate in the future. Lastly, the reference [22] establishes the convergence of the convex-concave procedure under weak assumptions, and provides a blockwise convex-concave procedure for distributed implementation. We will seek to weaken the assumptions in our setting and provide a blockwise/distributed version of our algorithm as well.
AKNOWLEDGMENT
We thank Mingyi Hong for helpful discussion.

APPENDIX

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

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

Since \(- \log |\Omega|\) is a convex function, \(O(||\Omega - \Omega^{(0)}||^2_F) \geq 0\). Therefore, for any pair \((\Omega, \Sigma)\) of the weighted sum-rate function with the damping term.

B. Proof of Theorem 1
Proof: We apply Theorem 8 in [23] for the proof. From the compactness of \(S\) and non emptiness of \(\text{ri}(S)\), we can conclude that all the requirements in Theorem 8 in [23] are satisfied except for the strict concavity of the decomposed weighted sum-rate function with the damping term. To show strict concavity, consider the concave part of the objective function \(u(\Omega, \Sigma) = \log |\Omega + H \Sigma H^*| - \rho ||\Sigma||^2_F\). From concavity, for any two points in the feasible set, \((\Omega_1, \Sigma_1)\) and \((\Omega_2, \Sigma_2)\), and \(0 < t < 1\), we have

\[
\begin{align*}
&u(t \Omega_1 + (1-t) \Omega_2, t \Sigma_1 + (1-t) \Sigma_2) \\
\geq & t \log |\Omega_1 + H \Sigma_1 H^*| + (1-t) \log |\Omega_2 + H \Sigma_2 H^*| \\
&- \rho t ||\Sigma_1||^2_F - (1-t) ||\Sigma_2||^2_F.
\end{align*}
\]

Note that in the feasible set, \(\Omega \succeq I\). Now suppose \(\Sigma_1 \neq \Sigma_2\), then from strict convexity of \(||\Sigma||^2_F\), (ineq:1) becomes strict, provided \(\rho > 0\). If \(\Sigma_1 = \Sigma_2\) and \(\Omega_1 \neq \Omega_2\), then from the strict convexity of the logdet function, (ineq:1) becomes strict. Therefore, for any pair \((\Omega_1, \Sigma_1) \neq (\Omega_2, \Sigma_2)\), and \(0 < t < 1\), we have \(u(t \Omega_1 + (1-t) \Omega_2, t \Sigma_1 + (1-t) \Sigma_2) > tu(\Omega_1, \Sigma_1) + (1-t)u(\Omega_2, \Sigma_2)\), which implies strict concavity of \(u\) over the feasible set. Similar argument applies to \(u(\Omega, \Sigma) = \log |\Omega - \rho \Sigma||^2_F\), which implies strictly concavity of \(-v\).

C. Proof of Theorem 2
Proof: By the same argument as in the proof of Theorem 1, we can apply the Theorem 4 in [23] to prove the result.

D. Proof of Theorem 3
Proof: Here we show that the weighted sum-rate always increases in each iteration. From convexity, we have

\[
\begin{align*}
&\sum_{n \in \mathcal{N}} w_n \left( \log |\Omega_n + H_n \Sigma_n H_n^*| - \log |\Omega_n| \right) \\
&- \rho \left( ||\Sigma_n||^2_F - ||\Sigma_n||^2_F \right) \\
\geq & \sum_{n \in \mathcal{N}} w_n \left( \log |\Omega_n + H_n \Sigma_n H_n^*| - \log |\Omega_n| \right) - \rho \left( ||\Sigma_n||^2_F - 2||\Sigma_n||^2_F \right) + 2 \text{Tr} \left( \Sigma_n^{(0)^*} (\Sigma - \Sigma^{(0)}) \right).
\end{align*}
\]

Since \((\Omega^{(i+1)}, \Sigma^{(i+1)})\) is a global maximum of this lower bound, we have

\[
\begin{align*}
&f_0(\Omega^{(i+1)}, \Sigma^{(i+1)}) \\
\geq & \sum_{n \in \mathcal{N}} w_n \left( \log |\Omega_n^{(i+1)} + H_n \Sigma_n^{(i+1)} H_n^*| - \log |\Omega_n| \right) - \rho \left( ||\Sigma_n^{(i+1)}||^2_F - 2||\Sigma_n^{(i+1)}||^2_F \right) + 2 \text{Tr} \left( \Sigma_n^{(0)^*} (\Sigma - \Sigma^{(0)}) \right).
\end{align*}
\]

This concludes the proof.

REFERENCES
distributed interference pricing in wireless networks,” in *Information

ad hoc cognitive radio networks,” *Information Theory, IEEE Transac-

cellular networks with limited coordination clusters,” *Selected Areas in

for interference alignment: Complexity and computation,” *Information


munications and signal processing,” *Selected Areas in Communications,


programming: theory, algorithms and applications,” *Acta Mathematica

and DCA revisited with DC models of real world nonconvex optimiza-

with linear matrix inequality constraints,” *SIAM Journal on matrix

ware package for semidefinite programming, version 1.3.” *Optimization


analysis of block successive minimization methods for nonsmooth
optimization,” *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–
1153, 2013.

the concave-convex procedure,” in *Advances in neural information