Penalty Dual Decomposition Method For Nonsmooth Nonconvex Optimization—Part II: Applications

Qingjiang Shi, Mingyi Hong, Xiao Fu, Tsung-Hui Chang

***Abstract*—In Part I of this paper, we proposed and analyzed a novel algorithmic framework, termed penalty dual decomposition (PDD), for the minimization of a nonconvex nonsmooth objective function, subject to difficult coupling constraints. Part II of this paper is devoted to evaluation of the proposed methods in the following three applications, ranging from communication networks to data analytics: i) the max-min rate fair multicast beamforming problem; ii) the sum-rate maximization problem in multi-antenna relay broadcast networks; and iii) the volume-min based structured matrix factorization problem, which is often used in document topic modeling. By exploiting the structure of the aforementioned problems, we develop a new class of algorithms based on the PDD framework. Differently from the state-of-the-art algorithms, they are proven to achieve convergence to stationary solutions of the aforementioned nonconvex problems. Numerical results validate the efficacy of the proposed schemes.**

***Index Terms*—Penalty dual decomposition, multicast beamforming, sum-rate maximization, matrix factorization.**

I. Introduction

In Part I of this paper, we have proposed a generic algorithm for optimizing the following nonconvex problem with coupling constraints:

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathcal{X},\boldsymbol{y}} \quad & F(\boldsymbol{x},\boldsymbol{y}) \triangleq f(\boldsymbol{x},\boldsymbol{y}) + \sum_{j=1}^{n_y} \tilde{\phi}(\boldsymbol{y}_j) \\ \text{s.t.} \quad & \boldsymbol{h}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{0}, \\ & \boldsymbol{g}_i(\boldsymbol{x}_i) \le 0, \forall i \end{aligned} \tag{P}$$

where $\boldsymbol{x} \triangleq (\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n)$ and $\boldsymbol{y} \triangleq (\boldsymbol{y}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_{n_y})$; $\tilde{\phi}(\boldsymbol{y}_j) = \phi_j(s_j(\boldsymbol{y}_j))$ is a composite function, with $s_j(\boldsymbol{y}_j)$ being a *convex* but possibly nondifferentiable function while $\phi_j(x)$ a *nondecreasing* and *continuously differentiable* function; the feasible set $\mathcal{X}$ is the Cartesian product of n simple *closed convex* sets $\mathcal{X}_i$'s with $\boldsymbol{x}_i \in \mathcal{X}_i$, $\forall i$; $f(\boldsymbol{x},\boldsymbol{y})$ and each component of the vector functions $\boldsymbol{h}(\boldsymbol{x},\boldsymbol{y})$ and $\boldsymbol{g}_i(\boldsymbol{x}_i)$'s are all continuously differentiable functions.

Our proposed algorithm, termed penalty dual decomposition (PDD), is a combination of primal dual based augmented Lagrangian method, block-coordinate-descent-type algorithm, and the penalty method. Under certain constraint qualification (CQ) named Robinson's condition, we show that every limit point generated by the PDD method is a KKT solution of problem (P). As will be shown in Part II of the paper, the main advantage of the PDD method is that, it is capable of exploiting the problem structure in a way which results in computationally lightweight algorithms. Particularly, by introducing appropriate auxiliary variables, the subproblem involved in the PDD method can be efficiently solved. As a result, the PDD method could be very efficient in dealing with nonconvex problems with difficult coupling constraints.

Many engineering problems can be formulated as problem (P). Some examples include dictionary learning and compressive sensing [2]–[4], volume-minization based matrix factorization [5]–[7], joint transceiver optimization of wireless systems [8]–[14], waveform design for radar systems [15]–[17], cross-layer design of wireless networks [18]–[21], and sensor network localization [22]–[24], etc.. Among the potential applications stated above, Part II of this paper focuses on performing case-studies on the following three important signal processing applications:

1) **Maxi-Min Fair Multicast beamforming**. Multicast beamforming is an important component of the evolved multimedia broadcast multicast service (eMBMS) in the long-term evolution (LTE) standard [25]. In multicast beamforming, a base station (BS) with multi-antennas transmits common information to multiple groups of users. For efficient multicasting, the BS chooses different weights for different streams based on the channel state information (CSI) to steer the transmit power in the directions of each group of users while limiting inter-group interference. To guarantee the rate fairness among the users, we often design bemforming weights to maximize the minimum user rate subject to a BS power constraint. This problem is known to be NP-hard [10] and has received a lot of attention from the research community [10], [12], [13]. A well-known approach to dealing with the multicast beamforming problem is using the celebrated semidefinite relaxation (SDR) method [10], [12]. However, to recover a high-quality suboptimal solution, *Gaussian randomization procedure* is needed after solving the SDR problem, resulting in high computational complexity. In this paper, we propose a PDD-based iterative multicast beamforming algorithm which achieves a higher max-min user rate but lower complexity than the SDR method.

2) **Sum-rate maximization for relay broadcast channel**. Relay-based cooperative communication has been adopted in LTE-Advanced standard as a key technology for future generation wireless communication systems [26]. In a relay-assisted cellular downlink system, the link quality between the BS and cell-edge users would benefit from deploying a relay

Part of this paper has been presented in IEEE ICASSP 2017.

Q. Shi is with the College of EIE, Nanjing Aeronautics and Astronautics, Nanjing 210016, China, Email: qing.j.shi@gmail.com

M. Hong is with the Dept. of Electrical and Computer Engineering, University of Minnesota, MN 55454, USA, Email: mhong@umn.edu

X. Fu is with the School of EECS, Oregon State University, Corvallis, OR, 97330, Email: xiao.fu@oregonstate.edu

T.-H. Chang is with The Chinese University of Hong Kong, Shenzhen 518172, China, Email: changtsunghui@cuhk.edu.cn

as well as joint source (i.e., BS)-relay design. However, the relay transmission introduces a coupling between the source precoder and the relay precoder in relay power constraints, which poses a fundamental challenge in joint source-relay design. In fact, such a challenge exists in various relay-assisted communication systems, e.g., multi-hop relay networks [27], [28], two-way relay networks [29], [30], relay interference networks [31], [32], etc.. Despite having extensive research on relay systems, there is still a lack of efficient optimization method to address the difficulty arising from the intrinsic coupling between the source precoder and the relay precoder. A promising way to address the coupling of two precoders [31], [33]–[35] is using alternating optimization (AO), i.e., alternatingly optimizes one precoder while fixing the other. However, the AO scheme can easily get trapped in some non-stationary solutions, which can have very low system throughput. In this paper, by applying PDD, we propose an efficient optimization framework to deal with the coupling between the source precoder and the relay precoder in joint design of relay systems.

3) **Volume-min based matrix factorization**. Structured factorization for given data matrices has many applications in signal processing and machine learning [36], [37]. As one important criterion for structured matrix factorization, volume minimization (VolMin) finds the minimum-volume *simplex* that embraces all the given data points [36]. This criterion can guarantee the identifiability of the factor matrices under mild conditions that are realistic in a wide variety of applications [7]. Hence, it recently attracted considerable interest in document clustering [5], blind separation of power spectral for dynamic spectrum access [38], and remote sensing [6], etc.. Due to the nature of matrix factorization, there exists a coupling of two matrix factors in the VolMin problem, making the VolMin problem quite challenging. In the literature, the VolMin problem is first transformed into the dimension-reduced space and then solved using alternating optimization or penalty method [6]. However, the existing algorithms cannot guarantee stationary solution to the VolMin problem. Moreover, in applications with additional constraints (e.g., nonnegativity of matrix factors), the problem has to be solved in the *original* space, rather than in the dimension-reduced space where the existing algorithms do no work. In this paper, we propose a PDD-based VolMin algorithm which works well in the original data space with guaranteed convergence.

The key to applying the PDD framework to nonconvex problems is to properly reformulate the problems at hand as problem (P), so that the corresponding augmented Lagrange problems can be easily solved via block-coordinate-desent (BCD)-type algorithms. In this paper, we present some reformulations of the aforementioned three problems in the form of problem (P) and develop a new class of algorithms for the reformulations by applying the PDD framework. By applying BCD-type algorithms to these reformulations, we can fully exploit the problem structures of the aforementioned problems, and significantly alleviate the challenging nonconvexity arising from either objective functions (e.g., the max-min structure in the max-min fair multicasting beamforming problem, and the volume function in the VolMin problem) or the constraints (e.g., the relay power constraint in the joint source-relay design problem that couples the source precoder and the relay preocder, and the matrix factorization equality constraint in the VolMin problem).

The developed algorithms enjoy several desirable features. First, differently from the state-of-the-art algorithms, they are proven to achieve convergence to stationary solutions of the aforementioned problems, and in practical they outperform the state-of-the-art algorithms in a number of performance metrics. Second, the iterations of the algorithms have closed-form and simple updates, therefore they are relatively easy to implement. Third, the algorithms are quite flexible and they are applicable to some generalizations of the aforementioned problems as well. For instance, the PDD algorithm can be easily generalized to dealing with the VolMin problem in the original space with a nonnegativity constraint on the basis factors.

The remainder of this paper is organized as follows. In Section II-IV, we apply the PDD method to the the aforementioned problems. Specifically, in each of three sections, we first reformulate the three problems, then show how the corresponding augmented Lagrangian problem is solved by BCD-type algorithms, followed by some simulations to compare the performance of the PDD-based algorithms with the state-of-art algorithms. Section V concludes the paper.

Notations: Besides the notations specified in Part I of this paper, we use the following notations. $\mathbb{C}^n$ (or $\mathbb{C}^{m\times n}$) denotes the n (or $m\times n$)-dimensional space of complex number. For a matrix $\mathbf{X}$, $\mathbf{X}^H$ and $\sigma_i(\mathbf{X})$ denote its conjugate transpose and its i-th largest singular value, respectively. For a vector $\boldsymbol{x}$, $\mathrm{diag}\{\boldsymbol{x}\}$ denotes a diagonal matrix with the elements of $\boldsymbol{x}$ being its diagonal entries. $\Re e\{x\}$ and $\Im m\{x\}$ denote the real part and the imaginary part of a complex number x, respectively, and x^* denotes the conjugate of x. The notation $\mathbf{A}\otimes\mathbf{B}$ means the Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$. $\mathbf{A}\succeq 0$ (or $\succ 0$) means that $\mathbf{A}$ is a positive semidefinite (or definite) matrix. $\mathbb{E}\{\cdot\}$ denotes expectation operation.

II. Maxmin-Rate Fairness Multi-cast Beamforming

Signal-to-interference-plus-noise ratio is an important performance metric used in signal design. It is generally in a *quadratic ratio* form with respect to the designed variables. On the other hand, max-min fairness is a popular resource allocation criterion that is widely adopted in wireless communication and signal processing [10], [12], [13], [39], [40]. As a result, we are often faced with the following problem

$$\max_{\boldsymbol{x}\in\mathcal{X}}\min_{k\in\mathcal{K}}\quad \frac{\boldsymbol{x}^H\mathbf{A}_k\boldsymbol{x}}{\boldsymbol{x}^H\mathbf{B}_k\boldsymbol{x}} \tag{1}$$

where $\boldsymbol{x}$ is a design variable which is constrained to a set $\mathcal{X}$; $\mathbf{A}_k$'s and $\mathbf{B}_k$'s are known matrices with $\mathbf{B}_k\succ 0$. Several examples of (1) can be found in max-min fariness precoding for wireless networks [14], [41], waveform design for radar systems [39], [40], and robust classification in machine learning [42], etc.. Problem (1) is challenging due to the nonlinear and nondifferentiable max-min ratio structure. In this section,

as an important example, we illustrate the application of PDD to multi-cast beamforming for achieving max-min rate fairness [10].

A. Problem Formulation

Consider a single-cell multi-user multiple-input-single-output (MISO) downlink system, where a base station (BS) equipped with N_t antennas transmits $n_g > 1$ independent data streams to n_g group of users over a common frequency band. Suppose that the i-th group, denoted by $\mathcal{G}_i$, has m_i single-antenna users, each of which is interested in receiving a common data stream. Let s_i denote the data stream for group $\mathcal{G}_i$, $i = 1, 2, \ldots, n_g$ and $\boldsymbol{w}_i \in \mathbb{C}^{N_t}$ be the beamforming weight for the i-th group. The transmitted signal at the BS is given by $\sum_{i=1}^{n_g} \boldsymbol{w}_i s_i$. Let $\boldsymbol{h}_k \in \mathbb{C}^{N_t}$ denote the conjugated channel between the BS and the receiver $k \in \mathcal{G}_i$. Then the received signal at receiver $k \in \mathcal{G}_i$ is given by

$$r_k = \boldsymbol{h}_k^H \boldsymbol{w}_i s_i + \sum_{j \neq i} \boldsymbol{h}_k^H \boldsymbol{w}_j s_j + z_k, \quad k \in \mathcal{G}_i \tag{2}$$

where z_k denotes additional Gaussian white noise (AWGN) with variance σ_k^2.

Assume that s_i's are i.i.d complex Gaussian random variable with zero mean and unit variance, and moreover s_i's and z_k's are independent of each other. Then the signal-to-interference-plus-noise-ratio (SINR) can be expressed as

$$\text{SINR}_k = \frac{\boldsymbol{w}_i^H \mathbf{R}_k \boldsymbol{w}_i}{\sum_{j \neq i} \boldsymbol{w}_j^H \mathbf{R}_k \boldsymbol{w}_j + \sigma_k^2}, k \in \mathcal{G}_i, i = 1, 2, \ldots, n_g \tag{3}$$

where $\mathbf{R}_k \triangleq \boldsymbol{h}_k \boldsymbol{h}_k^H$.

To achieve rate fairness among users, a popular criterion for beamforming design is to maximize the minimum user rate subject to the BS power constraint $\sum_{i=1}^{n} \|\boldsymbol{w}_i\|^2 \leq P_{BS}$, where P_{BS} denotes the total available power at the BS. Since the power constraint must be active at the optimality, we can write the max-min rate fairness multi-cast beamforming problem equivalently as

$$\begin{aligned} \max_{\{\boldsymbol{w}_i\}} \min_{i} \min_{k \in \mathcal{G}_i} \quad & \log_2 \left(1 + \frac{\boldsymbol{w}^H \mathbf{A}_{i_k} \boldsymbol{w}}{\boldsymbol{w}^H \mathbf{B}_{i_k} \boldsymbol{w}}\right), \\ \text{s.t.} \quad & \|\boldsymbol{w}\|^2 = 1 \end{aligned} \tag{4}$$

where $\boldsymbol{w} = (\boldsymbol{w}_i)_i$, $\mathbf{A}_{i_k} = \text{diag}\{\boldsymbol{e}_i\} \otimes \mathbf{R}_k$, and

$$\mathbf{B}_{i_k} = (\mathbf{I} - \text{diag}\{\boldsymbol{e}_i\}) \otimes \mathbf{R}_k + \frac{\sigma_k^2}{P_{BS}} \mathbf{I}.$$

This problem is known to be NP-hard [10]. After solving (4), we need to scale $\boldsymbol{w}$ such the power constraint. A popular method to address this problem is using semidefinite relaxation method coupled with bisection method [10], referred to as BisecSDR method, where, in each bisection, it is required to solve a semidefinite programming, requiring complexity at most $O\left(I_{bsc} \log(\frac{1}{\epsilon_{ip}}) \sqrt{n_g N_t} (n_g^3 N_t^6 + n_g N_t^2 K\right)$. Here $K \triangleq \sum_{i=1}^{n_g} m_i$, the parameter ϵ_{ip} represents the solution accuracy at the interior-point algorithm's termination, and I_{bsc} denotes the number of bisections.

B. PDD-based Algorithm

For convenience, let us consider a more general but equivalent formulation of problem (4), which is given by

$$\boxed{\max_{\boldsymbol{w}} \min_{k \in \mathcal{K}} \quad \frac{\boldsymbol{w}^H \mathbf{A}_k \boldsymbol{w}}{\boldsymbol{w}^H \mathbf{B}_k \boldsymbol{w}}, \quad \text{s.t.} \quad \|\boldsymbol{w}\|^2 = 1} \tag{P1}$$

where $\mathcal{K} \triangleq \{1, 2, \ldots, K\}$, and the matrices $\mathbf{A}_k$'s are all positive semidefinite and $\mathbf{B}_k$'s are all positive definite. In what follows, we present the PDD-based algorithm for problem $(P1)$.

First, we recast problem $(P1)$ as follows

$$\begin{aligned} \max_{\boldsymbol{t} \geq 0, \boldsymbol{w}} \quad & \min_k t_k \\ \text{s.t.} \quad & \|\mathbf{A}_k^{\frac{1}{2}} \boldsymbol{w}\| = t_k \|\mathbf{B}_k^{\frac{1}{2}} \boldsymbol{w}\|, \quad \forall k, \\ & \|\boldsymbol{w}\|^2 = 1, \end{aligned} \tag{5}$$

which is a special case of problem (P). In problem (5), the first K equality constraints are difficult coupling constraints. By moving these constraints into the objective, we obtain the corresponding augmented Lagrangian problem as follows

$$\begin{aligned} \max_{\boldsymbol{t} \geq 0, \boldsymbol{w}} \quad & \min_k t_k - \frac{1}{2\rho} \sum_{k=1}^{K} \left(\|\mathbf{A}_k^{\frac{1}{2}} \boldsymbol{w}\| - t_k \|\mathbf{B}_k^{\frac{1}{2}} \boldsymbol{w}\| + \rho \lambda_k\right)^2 \\ \text{s.t.} \quad & \|\boldsymbol{w}\|^2 = 1. \end{aligned} \tag{6}$$

where ρ is a penalty parameter and λ_k is a Lagrange multiplier associated with the k-th constraint.

The key to using the PDD method is to find appropriate locally tight lower bounds for the objective function, so that BSUM [43] can be applied to optimize the AL. For problem (6), we can simply decouple the variables into two blocks $\boldsymbol{w}$ and $\boldsymbol{t}$, leading to two subproblems: i.e., 1) solve (6) for $\boldsymbol{t}$ while fixing $\boldsymbol{w}$, and 2) solve (6) for $\boldsymbol{w}$ while fixing $\boldsymbol{t}$, which are respectively referred to as *$\boldsymbol{t}$-subproblem and $\boldsymbol{w}$-subproblem*. The $\boldsymbol{t}$-subproblem is strictly convex and thus has a unique solution, which can be easily solved by exploiting the problem structure; see Appendix A for a detailed derivation. The main difficulty lies in solving the $\boldsymbol{w}$-subproblem given by

$$\begin{aligned} \min_{\boldsymbol{w}} \quad & \vartheta(\boldsymbol{w}) \triangleq \sum_{k=1}^{K} \left(\|\mathbf{A}_k^{\frac{1}{2}} \boldsymbol{w}\| - t_k \|\mathbf{B}_k^{\frac{1}{2}} \boldsymbol{w}\| + \rho \lambda_k\right)^2 \\ \text{s.t.} \quad & \|\boldsymbol{w}\|^2 = 1. \end{aligned}$$

Apparently, the $\boldsymbol{w}$-subproblem is difficult to solve. Instead of exactly minimizing $\vartheta(\boldsymbol{w})$, we try to find a locally tight upper bound $u(\boldsymbol{w}; \tilde{\boldsymbol{w}})$ for $\vartheta(\boldsymbol{w})$ and minimize this upper bound to update $\boldsymbol{w}$ given $\boldsymbol{t}$. Observing the constraint $\|\boldsymbol{w}\| = 1$, we expect the upper bound to be a *homogeneous quadratic function* in the form of $\boldsymbol{w}^H \mathbf{C} \boldsymbol{w}$ or $\boldsymbol{w}_{eq}^T \mathbf{C} \boldsymbol{w}_{eq}$ where $\boldsymbol{w}_{eq} \triangleq (\Re e\{\boldsymbol{w}\}, \Im m\{\boldsymbol{w}\})$, so that the resulting problem is an easily solvable eigenvalue problem.

By expanding $\vartheta(\boldsymbol{w})$, we can find that $\vartheta(\boldsymbol{w})$ includes the following four kinds of terms: **1)** $\boldsymbol{w}^H \mathbf{A}_k \boldsymbol{w} + t_k^2 \boldsymbol{w}^H \mathbf{B}_k \boldsymbol{w}$; **2)** $-2t_k \|\mathbf{A}_k^{\frac{1}{2}} \boldsymbol{w}\| \|\mathbf{B}_k^{\frac{1}{2}} \boldsymbol{w}\|$; **3)** $2\rho\lambda_k \|\mathbf{A}_k^{\frac{1}{2}} \boldsymbol{w}\|$; **4)**$-2\rho\lambda_k t_k \|\mathbf{B}_k^{\frac{1}{2}} \boldsymbol{w}\|$. Clearly, we need to make efforts to bound the last three terms with homogenous quadratic functions. Unfortunately, since the multiplier λ_k's could be either negative or positive,

TABLE I
ALGORITHM 1: BSUM FOR PROBLEM (6)

0. initialize $\boldsymbol{w}$ and $\boldsymbol{t}$
1. **repeat**
2. $\quad a_k \longleftarrow \frac{\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\|^2}{2\rho}$
3. $\quad b_k \longleftarrow \frac{\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\|+\rho\lambda_k}{\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\|}$.
4. $\quad$ update $\boldsymbol{t}$ by solving problem (48)
5. $\quad$ compute $\mathbf{C}=\sum_{k=1}^K \mathbf{C}_k$ via (57) with $\tilde{\boldsymbol{w}}=\boldsymbol{w}$
6. $\quad$ update $\boldsymbol{w} \longleftarrow \boldsymbol{v}_{\min}(\mathbf{C})$
7. **until** some termination criterion is met

it is challenging to bound the last two terms with homogenous quadratic functions. Thanks to the fact that $\|\boldsymbol{w}\|=1$, we can modify the third term as $2\rho\lambda_k\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\|\,\|\boldsymbol{w}\|$ when $\lambda_k<0$ and the fourth term as $-2\rho\lambda_k t_k\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\|\,\|\boldsymbol{w}\|$ when $\lambda_k>0$. Hence, essentially, $\vartheta(\boldsymbol{w})$ includes two kinds of terms in the forms of $\|\mathbf{Q}_1\boldsymbol{w}\|$ and $-\|\mathbf{Q}_1\boldsymbol{w}\|\,\|\mathbf{Q}_2\boldsymbol{w}\|$ with some appropriate $\mathbf{Q}_1$ and $\mathbf{Q}_2$. To bound these two terms, we resort to the following lemma.

Lemma 2.1: For real vectors $\boldsymbol{x}$, $\boldsymbol{y}$, $\tilde{\boldsymbol{x}}$, $\tilde{\boldsymbol{y}}$, the following inequalities

1) $\|\boldsymbol{x}\|\,\|\boldsymbol{y}\| \geq \frac{1}{\|\tilde{\boldsymbol{x}}\|\|\tilde{\boldsymbol{y}}\|}\boldsymbol{x}^T\tilde{\boldsymbol{x}}\tilde{\boldsymbol{y}}^T\boldsymbol{y},\ \forall\ \tilde{\boldsymbol{x}}\neq\mathbf{0},\tilde{\boldsymbol{y}}\neq\mathbf{0},\boldsymbol{x},\boldsymbol{y}$;
2) $\|\boldsymbol{x}\| \leq \frac{1}{2\|\tilde{\boldsymbol{x}}\|}\|\boldsymbol{x}\|^2+\frac{1}{2}\|\tilde{\boldsymbol{x}}\|,\ \forall\ \tilde{\boldsymbol{x}}\neq\mathbf{0},\boldsymbol{x}$

hold true with equality satisfied at $\boldsymbol{x}=\tilde{\boldsymbol{x}}$ and $\boldsymbol{y}=\tilde{\boldsymbol{y}}$.

Proof: Part 1) follows directly from the Cauchy-Schwartz inequality, while Part 2) follows from the property of concave function by noting that $\|\boldsymbol{x}\|=\sqrt{\|\boldsymbol{x}\|^2}$ is a concave function of $\|\boldsymbol{x}\|^2$. ■

In terms of the above analysis and using Lemma 2.1, we can obtain $\vartheta(\boldsymbol{w}) \leq u(\boldsymbol{w},\tilde{\boldsymbol{w}}) \triangleq \boldsymbol{w}_{eq}^T\mathbf{C}\boldsymbol{w}_{eq}+const$ in the real domain, where $\mathbf{C}$ is a $2n_gN_t$ by $2n_gN_t$ matrix function of $\tilde{\boldsymbol{w}}$ whose detailed derivation is shown in Appendix B. Moreover, it can be verified that $u(\boldsymbol{w},\tilde{\boldsymbol{w}})$ is a locally tight upper bound [43] of $\vartheta(\boldsymbol{w})$ over the set $\{\boldsymbol{w}\mid\|\boldsymbol{w}\|=1\}$. With such an upper bound function, we update $\boldsymbol{w}$ by solving the following eigenvalue problem, i.e., $\min_{\boldsymbol{w}_{eq}}\boldsymbol{w}_{eq}^T\mathbf{C}\boldsymbol{w}_{eq}$, s.t. $\|\boldsymbol{w}_{eq}\|=1$. Denote by $\boldsymbol{v}_{min}(\mathbf{C})$ the eigenvector of $\mathbf{C}$ corresponding to its minimum eigenvalue. Once we get $\boldsymbol{v}_{min}(\mathbf{C})$, we can construct the corresponding $\boldsymbol{w}$.

To summarize, the BSUM algorithm for addressing problem (6) is presented in TABLE I. It can be shown that the most costly step of the BSUM algorithm lies in calculating $\boldsymbol{v}_{min}(\mathbf{C})$, requiring complexity of $O(Kn_g^2N_t^2)+O(n_g^3N_t^3)$, where the first term corresponds to the computation of $\mathbf{C}$ while the second term corresponds to the eigenvalue decomposition. It is easily seen that the PDD method has lower complexity than the BisecSDR method in [10].

C. Numerical Results

In the simulations, the noise power is set to unit for all receivers and $P_{BS}=10$ dB. For convenience, we denote by (N_t,n_g,m_g) a multi-user multi-cast network with N_t BS antennas, n_g multi-cast groups each with m_g single-antenna users, hence $K=n_gm_g$ users in total. Furthermore, unless otherwise specified, we set $\rho_0=0.5K$, $\epsilon_0=1e{-}3$, $\epsilon_O=1e{-}4$, as well as $\rho_k=c\rho_{k-1}$ and $\epsilon_k=\epsilon_{k-1}c$ with $c=0.6$ for the PDD method in all of our simulations. Moreover, to avoid numerical instability, we set the maximum number of inner BSUM iterations of the PDD method as 100 in practical implementation.

We compare the PDD method with the BisecSDR method in [10] and the penalty-BSUM method[1] proposed in [45] (abbreviated as "Penalty" in the plot). The basic idea of the BisecSDR method is as follows. First, by applying semidefinite relaxation (SDR), problem $(P1)$ is relaxed as

$$\begin{aligned}\max_{t,\mathbf{W}}\quad & t\\ \text{s.t.}\quad & \text{Tr}\left((\mathbf{A}_k-t\mathbf{B}_k)\mathbf{W}\right)\geq 0,\forall k,\\ & \mathbf{W}\succeq 0.\end{aligned}\tag{7}$$

Second, by searching over t using Bisection method, we can obtain the optimal $\mathbf{W}$ by solving a sequence of semidefinite programmings (obtained by fixing t in the above problem). Third, given the optimal $\mathbf{W}$, we can find a suboptimal solution $\boldsymbol{w}$ by checking all candidate solutions including the principal eigenvector of $\mathbf{W}$ and those obtained by performing Gaussian randomization procedure (GRP). Note that, the optimal value of problem (7) can serve as an upper bound for the achievable maxmin user rate. Particularly, when the SDR is tight, the maxmin user rate coincides with the upper bound. In simulations, the semidefinite programmings are solved by interior-point method, e.g., using the off-the-shelf package SeDuMi [46] for efficiency. The Bisection procedure is terminated when the relative size of the bisection interval is smaller than $1e{-}3$. In addition, the penalty-BSUM shares the same parameter setting with the PDD.

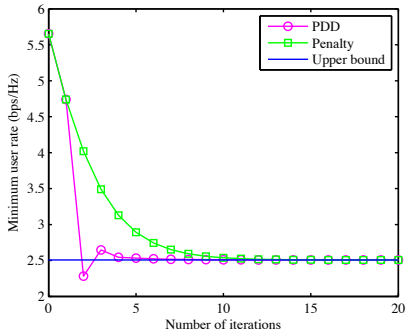
The average convergence behavior of the algorithm over ten randomly generated examples is illustrated in Fig. 1, where the minimum user rate is normalized by the upper bound value. It is seen that the PDD method exhibits better convergence behavior than the penalty-BSUM method in terms of both the objective value and the optimality gap, while both achieving similar constraint violation. Here the optimality gap measures how well the solution $\boldsymbol{w}$ satisfies the KKT condition of problem $(P1)$, which is defined by the optimal value of the following convex optimization problem[2]

$$\begin{aligned}\min_{\{\lambda_k\}}\quad & \left\|\sum_{k=1}^K\lambda_k\boldsymbol{f}_k+\lambda_0\boldsymbol{w}\right\|\\ \text{s.t.}\quad & \sum_{k=1}^K\lambda_k=1,\lambda_k\geq 0,k=1,2,\ldots,K\end{aligned}\tag{8}$$

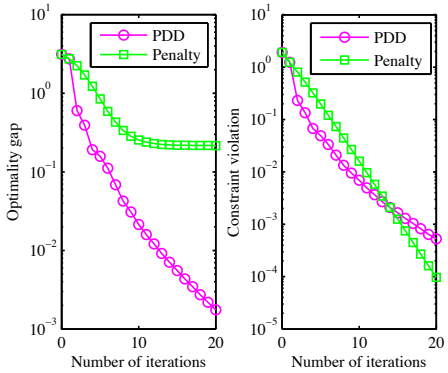
[1]The penalty-BSUM algorithm is similar to the PDD method but does not include the dual update as in the PDD method. The algorithm in [44] is in essence the penalty-BSUM algorithm, with the only difference in that some fixed penalty parameter was used in [44] while the penalty-BSUM algorithm uses increasing penalty. However, fixed penalty parameter cannot guarantee a KKT solution. Moreover, it is generally difficult to choose a penalty parameter which works well for all cases. Hence, we modify the algorithm in [44] to the exact penalty-BSUM algorithm by using increasing penalty.

[2]By KKT analysis, it can be shown that problem (8) having a zero optimal value is a necessary optimality condition for problem $(P1)$.

where $\boldsymbol{f}_k$ is the gradient of the function $\frac{\boldsymbol{w}^H \mathbf{A}_k \boldsymbol{w}}{\boldsymbol{w}^H \mathbf{B}_k \boldsymbol{w}}$ with respect to $\boldsymbol{w}$. Moreover, the PDD can achieve the upper bound value in this example, implying the excellent performance of the PDD method. In addition, one can see that both the feasibility gap and the optimality gap (i.e., constraint violation defined in Part I) decrease at the same time. Although the zero feasibility gap does not necessarily imply the zero optimality gap, it is much easier to evaluate the former than the latter. Hence, in our later simulations, we only examine the feasibility gap for simplicity.



(a) The minimum user rate.



(b) The optimality and feasibility gap.

Fig. 1. The convergence behavior of the PDD method for network $(8, 4, 2)$.

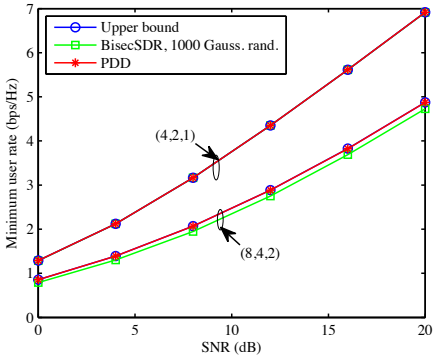


Fig. 2. The average minimum user rate achieved by various methods versus SNR for different networks.

Fig. 2 illustrates the max-min rate performance of the PDD method versus the BS power budget P_{BS} as compared to the upper bound provided by the BisecSDR method (i.e., the optimal value of problem (7)) and the performance of the BisecSDR method with 1000 Gaussian randomizations. The results in the plot are averaged over 100 random channel realizations. For the network $(4, 2, 1)$, it is known that the SDR is tight in this case and thus the upper bound is exactly the same as the optimal max-min user rate. From the figure, it is observed that the performance of the PDD method coincides with the upper bound for both networks, which is better than the performance of the BisecSDR method for the network $(8, 4, 2)$.

For a clear illustration, Table 2 compares the performance of three methods in terms of the cpu time and the achieved minimum rate averaged over 100 random channel realizations. In the table, R_{UB}, R_{PDD}, R_{SDR}, and $R_{Penalty}$ denote the upper bound value, and the minimum rate achieved by the PDD method, the BisecSDR method with 1000 Gaussian randomizations, and the penalty-BSUM method, respectively, while T_{PDD}, T_{SDR}, and $T_{Penalty}$ denote the corresponding cpu time required by three methods. It can be observed that the PDD method requires less cpu time than the BisecSDR method while achieving almost global optimality. Moreover, it performs more efficiently than the penalty-BSUM method in terms of the consumed cpu time.

TABLE II
THE AVERAGE CPU TIME AND MIN. RATE COMPARISON

Network	$\frac{R_{PDD}}{R_{UB}}$	$\frac{R_{PDD}}{R_{SDR}}$	$\frac{T_{SDR}}{T_{PDD}}$	$\frac{R_{PDD}}{R_{Penalty}}$	$\frac{T_{Penalty}}{T_{PDD}}$
$(2, 2, 2)$	99.97%	100.36%	3.38	100.26%	2.05
$(4, 2, 2)$	99.98%	100.28%	3.83	100.14%	1.98
$(8, 4, 2)$	99.98%	102.06%	3.44	100.13%	1.96
$(8, 2, 4)$	99.93%	101.78%	3.02	100.15%	1.74
$(16, 4, 4)$	99.92%	102.93%	2.76	100.18%	1.70

III. JOINT SOURCE-RELAY DESIGN FOR MULTI-ANTENNA RELAY BROADCAST SYSTEMS

Wireless relaying in cellular networks has attracted considerable attention due to its advantage of coverage extension and throughput improvement. It is well-known that, joint source-relay design can further enhance the system throughput performance for multi-antenna relay systems. However, the relay power constraint results in the coupling between the source precoder and the relay precoder, therefore the resulting joint source-relay design problem is very challenging to solve. In this section, by applying PDD, we present a joint source-relay design method which can reach at least stationary solutions. Note that our method is developed for a multi-antenna relay broadcasting channel but its basic idea can be extended to joint source-relay design of other relay systems.

A. Problem formulation

Consider a sum-rate maximization problem for a multi-antenna relay broadcasting channel, where a multi-antenna source (e.g, base station), equipped with N_s antennas, sends signal to K single-antenna users with the aid of a multi-antenna relay equipped with N_r antennas. The received signal

at each user can be expressed as

$$\boldsymbol{y}_k=\boldsymbol{g}_k^H\mathbf{F}\left(\mathbf{H}\sum_{j=1}^{K}\boldsymbol{v}_j s_j+\boldsymbol{n}_R\right)+n_k,\quad k=1,2,\ldots,K. \quad (9)$$

where $\boldsymbol{v}_j$ and s_j denote transmit beamformer employed by the source and the transmitted symbol intended for user j, respectively; the term $\sum_{j=1}^{K}\boldsymbol{v}_j s_j$ is the transmit signal of the source; $\mathbf{H}\in\mathbb{C}^{N_r\times N_s}$ represents the channel between the source and the relay; $\boldsymbol{n}_R$ and n_k denote the AWGN at the relay and user k, respectively; $\mathbf{F}\in\mathbb{C}^{N_r\times N_r}$ is the precoder employed by the relay to process the received signal (i.e., the bracketed term) from the source; $\boldsymbol{g}_k\in\mathbb{C}^{N_r}$ denotes the conjugated channel between the relay and user k.

Suppose that the transmitted symbols and noises are independent of each other. Moreover, let σ_R^2 and σ_k^2 denote the noise power at the relay and user k, and define the source precoder $\mathbf{V}\triangleq[\boldsymbol{v}_1\ \boldsymbol{v}_2\ \ldots\ \boldsymbol{v}_K]\in\mathbb{C}^{N_s\times K}$. Then the SINR γ_k at user k is given by

$$\gamma_k(\mathbf{V},\mathbf{F})\triangleq\frac{|\boldsymbol{g}_k^H\mathbf{F}\mathbf{H}\boldsymbol{v}_k|^2}{\sum_{j\neq k}|\boldsymbol{g}_k^H\mathbf{F}\mathbf{H}\boldsymbol{v}_j|^2+\sigma_R^2\left\|\boldsymbol{g}_k^H\mathbf{F}\right\|^2+\sigma_k^2}. \quad (10)$$

Furthermore, the source power consumption is given by $\mathrm{Tr}\left(\mathbf{V}\mathbf{V}^H\right)$ and the relay power consumption is $\|\mathbf{F}\mathbf{H}\mathbf{V}\|_F^2+\sigma_R^2\|\mathbf{F}\|_F^2$.

We are interested in maximizing the weighted sum-rate subject to the source and relay power constraints, which can be mathematically formulated as follows

$$\boxed{\begin{aligned}&\max_{\mathbf{V},\mathbf{F}}\sum_{k=1}^{K}\alpha_k\log\left(1+\gamma_k(\mathbf{V},\mathbf{F})\right)\\&\text{s.t. }\mathrm{Tr}\left(\mathbf{V}\mathbf{V}^H\right)\le P_S,\\&\qquad\|\mathbf{F}\mathbf{H}\mathbf{V}\|_F^2+\sigma_R^2\|\mathbf{F}\|_F^2\le P_R.\end{aligned}} \quad \text{(P2)}$$

where α_k denotes the weight measuring the priority of user k, P_S and P_R denote the source and relay power budget. The problem is hard to solve due mainly to the coupling of the source precoder $\mathbf{V}$ and relay precoder $\mathbf{F}$ at the relay power constraint. Note that such coupling is common to a number of joint source-relay designs well beyond problem (P2). We here aim to provide a way to deal with such coupling constraints.

B. PDD-based algorithm

We start by reformulating problem $(P2)$ so that the PDD algorithm can be easily applied. Introducing a set of auxiliary variables $\{\mathbf{X},\bar{\mathbf{V}},\bar{F},\bar{\mathbf{X}}\}$, and defining $\mathcal{X}\triangleq\{\mathbf{V},\mathbf{F},\mathbf{X},\bar{\mathbf{V}},\bar{F},\bar{\mathbf{X}}\}$ for notational simplicity, we can recast problem $(P2)$ as

$$\begin{aligned}&\max_{\mathcal{X}}\sum_{k=1}^{K}\alpha_k\log\left(1+\frac{|\boldsymbol{g}_k^H\boldsymbol{x}_k|^2}{\sum_{j\neq k}|\boldsymbol{g}_k^H\boldsymbol{x}_j|^2+\sigma_R^2\left\|\boldsymbol{g}_k^H\mathbf{F}\right\|^2+\sigma_k^2}\right)\\&\text{s.t. }\mathrm{Tr}\left(\bar{\mathbf{V}}\bar{\mathbf{V}}^H\right)\le P_S,\\&\quad\left\|\bar{\mathbf{X}}\right\|_F^2+\sigma_R^2\left\|\bar{\mathbf{F}}\right\|_F^2\le P_R,\\&\quad\mathbf{X}=\mathbf{F}\mathbf{H}\mathbf{V},\ \sigma_R\mathbf{F}=\sigma_R\bar{\mathbf{F}},\\&\quad\mathbf{X}=\bar{\mathbf{X}},\ \mathbf{V}=\bar{\mathbf{V}}\end{aligned} \quad (11)$$

where having σ_R in the constraint $\sigma_R\mathbf{F}=\sigma_R\bar{\mathbf{F}}$ facilitates the solution of the subproblem involving $\{\bar{\mathbf{X}},\bar{\mathbf{F}}\}$. This point will become clear shortly. Now we can see that the reformulation has separable inequality constraints. By building all the equality constraints into the objective, we can obtain the augmented Lagrangian problem as follows

$$\begin{aligned}&\max_{\mathcal{X}}\sum_{k=1}^{K}\alpha_k\log\left(1+\frac{|\boldsymbol{g}_k^H\boldsymbol{x}_k|^2}{\sum_{j\neq k}|\boldsymbol{g}_k^H\boldsymbol{x}_j|^2+\sigma_R^2\left\|\boldsymbol{g}_k^H\mathbf{F}\right\|^2+\sigma_k^2}\right)\\&\qquad-P_\rho(\mathcal{X})\\&\text{s.t. }\mathrm{Tr}\left(\bar{\mathbf{V}}\bar{\mathbf{V}}^H\right)\le P_S,\\&\quad\left\|\bar{\mathbf{X}}\right\|_F^2+\sigma_R^2\left\|\bar{\mathbf{F}}\right\|_F^2\le P_R\end{aligned} \quad (12)$$

where

$$P_\rho(\mathcal{X})\triangleq\frac{1}{2\rho}\Big(\|\mathbf{X}-\mathbf{F}\mathbf{H}\mathbf{V}+\rho\mathbf{Z}\|^2+\left\|\sigma_R\mathbf{F}-\sigma_R\bar{\mathbf{F}}+\rho\mathbf{Z}_f\right\|^2+\left\|\mathbf{X}-\bar{\mathbf{X}}+\rho\mathbf{Z}_x\right\|^2+\left\|\mathbf{V}-\bar{\mathbf{V}}+\rho\mathbf{Z}_v\right\|^2\Big),$$

and $\mathbf{Z}$, $\mathbf{Z}_f$, $\mathbf{Z}_x$ and $\mathbf{Z}_v$ are the dual variables associated with the equality constraints of problem (11).

Next, we show how to solve problem (12) using BSUM. The key to apply BSUM to (12) is to find a tractable locally tight lower bound for the objective of (12). To do so, we resort to the well-known WMMSE method [47]. First, by the theory of the WMMSE method, we have the following lemma.

Lemma 3.1: For each k, we have

$$\begin{aligned}&\log\left(1+\frac{|\boldsymbol{g}_k^H\boldsymbol{x}_k|^2}{\sum_{j\neq k}|\boldsymbol{g}_k^H\boldsymbol{x}_j|^2+\sigma_R^2\left\|\boldsymbol{g}_k^H\mathbf{F}\right\|^2+\sigma_k^2}\right)\\&=\max_{u_k,w_k}\log(w_k)-w_k e_k(u_k,\mathbf{X},\mathbf{F})+1\end{aligned} \quad (13)$$

where $e_k(u_k,\mathbf{X},\mathbf{F})\triangleq|1-u_k^*g_k^H\boldsymbol{x}_k|^2+\sum_{j\neq k}\alpha_k|u_k^*g_k^H\boldsymbol{x}_j|^2+\sigma_R^2\|u_k^*\boldsymbol{g}_k^H\mathbf{F}\|^2+\sigma_k^2|u_k|^2$.
This lemma can be easily proven by checking the first-order optimality condition of the problem on the right-hand-side (rhs) of (13), leading to the optimal u_k and w_k (given $\mathbf{X}$ and $\mathbf{F}$) as follows

$$u_k(\mathbf{X},\mathbf{F})=\frac{\boldsymbol{g}_k^H\boldsymbol{x}_k}{\sum_{k=1}^{K}|\boldsymbol{g}_k^H\boldsymbol{x}_j|^2+\sigma_R^2\|\boldsymbol{g}_k^H\mathbf{F}\|^2+\sigma_k^2}, \quad (14)$$

$$\begin{aligned}w_k(\mathbf{X},\mathbf{F})&=\frac{1}{e_k(u_k(\mathbf{X},\mathbf{F}),\mathbf{X},\mathbf{F})}\\&=\frac{1}{1-u_k^*(\mathbf{X},\mathbf{F})\boldsymbol{g}_k^H\boldsymbol{x}_k}\\&=1+\frac{|\boldsymbol{g}_k^H\boldsymbol{x}_k|^2}{\sum_{j\neq k}|\boldsymbol{g}_k^H\boldsymbol{x}_j|^2+\sigma_R^2\|\boldsymbol{g}_k^H\mathbf{F}\|^2+\sigma_k^2},\end{aligned} \quad (15)$$

where we have denoted the optimal u_k and w_k as $u_k(\mathbf{X},\mathbf{F})$ and $\boldsymbol{w}_k(\mathbf{X},\mathbf{F})$ for a clear illustration of their dependence on $\mathbf{X}$ and $\mathbf{F}$. As a direct result of Lemma 3.1, we have

$$\begin{aligned}&\log\left(1+\frac{|\boldsymbol{g}_k^H\boldsymbol{x}_k|^2}{\sum_{j\neq k}|\boldsymbol{g}_k^H\boldsymbol{x}_j|^2+\sigma_R^2\left\|\boldsymbol{g}_k^H\mathbf{F}\right\|^2+\sigma_k^2}\right)\\&\ge\log(w_k(\tilde{\mathbf{X}},\tilde{\mathbf{F}}))-w_k(\tilde{\mathbf{X}},\tilde{\mathbf{F}})e_k(u_k(\tilde{\mathbf{X}},\tilde{\mathbf{F}}),\mathbf{X},\mathbf{F})+1,\\&\quad\forall\tilde{\mathbf{X}},\tilde{\mathbf{F}},\mathbf{X},\mathbf{F}.\end{aligned} \quad (16)$$

Moreover, it can be easily verified that the rhs of (16) is a locally tight lower bound of the rate function shown on the lhs of (16). With such a tractable locally tight lower bound, we can easily apply BSUM to (12) with the block variables separated as 1) $\mathbf{F}$, 2) $\mathbf{X}$, 3) $\{\bar{\mathbf{V}}, \bar{\mathbf{X}}, \bar{\mathbf{F}}\}$, and 4) $\mathbf{V}$. Specifically, by applying the lower bound shown in (16), we propose to solve the following problem

$$\begin{aligned} \min_{\mathcal{X}} \ & \sum_{k=1}^{K} w_k \alpha_k e_k(u_k, \mathbf{X}, \mathbf{F}) + P_\rho(\mathcal{X}) \\ \text{s.t. } & \mathrm{Tr}\left(\bar{\mathbf{V}}\bar{\mathbf{V}}^H\right) \le P_S, \\ & \left\|\bar{\mathbf{X}}\right\|_F^2 + \sigma^2 \left\|\bar{\mathbf{F}}\right\|_F^2 \le P_R \end{aligned} \tag{17}$$

where w_k and u_k are given. Further, by simple manipulations, we can rewrite the above problem compactly as

$$\begin{aligned} \min_{\mathcal{X}} \ & \mathrm{Tr}\left(\mathbf{X}^H \mathbf{G}_w \mathbf{X}\right) - 2\Re e\left\{\mathrm{Tr}\left(\mathbf{X}^H \mathbf{G} \mathbf{D}_w\right)\right\} \\ & + \sigma_R^2 \mathrm{Tr}\left(\mathbf{F}^H \mathbf{G}_w \mathbf{F}\right) + P_\rho(\mathcal{X}) \\ \text{s.t. } & \mathrm{Tr}\left(\bar{\mathbf{V}}\bar{\mathbf{V}}^H\right) \le P_S, \\ & \left\|\bar{\mathbf{X}}\right\|_F^2 + \sigma_R^2 \left\|\bar{\mathbf{F}}\right\|_F^2 \le P_R \end{aligned} \tag{18}$$

where

$$\mathbf{G}_w \triangleq \sum_{k=1}^{K} w_k \alpha_k |u_k|^2 \boldsymbol{g}_k \boldsymbol{g}_k^H \text{ and } \mathbf{D}_w \triangleq \mathrm{diag}\left\{(w_k \alpha_k u_k)_k\right\}. \tag{19}$$

Now we are ready to show the BSUM iteration for problem (18), which consists of the following four steps.

1) Step 1: solving (18) *for* $\mathbf{F}$ *given* $\{\mathbf{V}, \bar{\mathbf{F}}\}$*:* The $\mathbf{F}$-subproblem is an unconstrained quadratic optimization problem. By the first order optimality condition, we obtain

$$\begin{aligned} & \sigma_R^2(2\rho\mathbf{G}_w + \mathbf{I})\mathbf{F} + \mathbf{F}\mathbf{H}\mathbf{V}\mathbf{V}^H\mathbf{H}^H \\ =& \sigma_R(\sigma_R\bar{\mathbf{F}} - \rho\mathbf{Z}_f) + (\mathbf{X} + \rho\mathbf{Z})\mathbf{V}^H\mathbf{H}^H \end{aligned} \tag{20}$$

which is the so-called Sylvester equation and admits efficient unique solution [48].

2) Step 2: solving (18) *for* $\{\bar{\mathbf{X}}, \bar{\mathbf{F}}, \bar{\mathbf{V}}\}$ *given* $\{\mathbf{X}, \mathbf{F}\}$*:* The subproblem with respect to $\{\bar{\mathbf{X}}, \bar{\mathbf{F}}, \bar{\mathbf{V}}\}$ can be further divided into two *independent* problems: one is with respect to $\bar{\mathbf{V}}$ while the other is with respect to $\{\bar{\mathbf{F}}, \bar{\mathbf{V}}\}$. Both problems are equivalent to projection of a point onto a ball centered at the origin, which can be solved in closed-form. Specifically, we have

$$\bar{\mathbf{V}} = \mathcal{P}_{P_S}\{\mathbf{V} + \rho\mathbf{Z}_v\}, \tag{21}$$

$$[\bar{\mathbf{X}} \ \ \sigma_R\bar{\mathbf{F}}] = \mathcal{P}_{P_R}\left\{[\mathbf{X} + \rho\mathbf{Z}_x \ \ \sigma_R\mathbf{F} + \rho\mathbf{Z}_f]\right\}. \tag{22}$$

where $\mathcal{P}_{\mathcal{X}}(\boldsymbol{x})$ denotes the projection of $\boldsymbol{x}$ onto the convex set $\mathcal{X}$. From (22), we can obtain the optimal $\bar{\mathbf{F}}$. It is worth mentioning that, the σ_R in the constraint $\sigma_R\mathbf{F} = \sigma_R\bar{\mathbf{F}}$ is introduced to make the subproblem with respect to $\{\bar{\mathbf{F}}, \bar{\mathbf{X}}\}$ have a closed-form solution; otherwise, we need to solve a quadratic equation to get the optimal Lagrange multiplier associated with the relay power constraint.

TABLE III
ALGORITHM 2: BSUM ALGORITHM FOR PROBLEM (12)

0. initialize $\{\mathbf{F}, \mathbf{V}\}$ such that the power constraints
1. set $\mathbf{X} = \mathbf{F}\mathbf{H}\mathbf{V}$, $\bar{\mathbf{X}} = \mathbf{X}$, $\bar{\mathbf{F}} = \mathbf{F}$, $\bar{\mathbf{V}} = \mathbf{V}$
2. **repeat**
3. compute $\boldsymbol{u}$ and $\boldsymbol{w}$ via (14) and (15)
4. compute $\mathbf{G}_w$ and $\mathbf{D}_w$ via (19)
5. update $\mathbf{F}$ by solving Eq. (20)
6. update $\bar{\mathbf{V}}$ via Eq. (21)
7. update $\bar{\mathbf{X}}$ and $\bar{\mathbf{F}}$ via Eq. (22)
8. update $\mathbf{X}$ via (23)
9. update $\mathbf{V}$ via (24)
10. **until** some termination criterion is met

3) Step 3: solving (18) *for* $\mathbf{X}$ *given* $\{\bar{\mathbf{X}}, \mathbf{V}, \bar{\mathbf{F}}\}$*:* The $\mathbf{X}$-subproblem is also an unconstrained quadratic optimization problem. Again, by the first order optimality condition, we obtain a unique closed-form solution as follows

$$\mathbf{X} = \frac{1}{2}(\rho\mathbf{G}_w + \mathbf{I})^{-1}(2\rho\mathbf{G}\mathbf{D}_w + (\mathbf{F}\mathbf{H}\mathbf{V} - \rho\mathbf{Z}) + (\bar{\mathbf{X}} - \rho\mathbf{Z}_x)). \tag{23}$$

4) Step 4: solving (18) *for* $\mathbf{V}$ *given* $\{\bar{\mathbf{V}}, \mathbf{X}, \bar{\mathbf{F}}\}$*:* The $\mathbf{V}$-subproblem is also an unconstrained quadratic optimization problem. Similarly, we obtain a unique closed-form solution as follows by applying the first-order optimality condition

$$\mathbf{V} = (\mathbf{I} + \mathbf{H}^H\mathbf{F}^H\mathbf{F}\mathbf{H})^{-1}(\bar{\mathbf{V}} - \rho\mathbf{Z}_v + \mathbf{H}^H\mathbf{F}^H(\mathbf{X} + \rho\mathbf{Z})). \tag{24}$$

In sum, every step of the BSUM iteration has a unique closed-form solution. Combining the steps for computing the lower bound, i.e., (14) and (15), we summarize the BSUM algorithm for (12) in Table III.

C. Numerical results

This section presents some numerical results to evaluate the performance of the proposed PDD method by comparing the alternating optimization (AO) method. In the AO method applied to problem $(P2)$, we alternatingly update the source precoder $\mathbf{V}$ and $\mathbf{F}$ while fixing the other. Specifically, in each iteration of the AO method, after updating $\boldsymbol{u}$ and $\boldsymbol{w}$ via (14) and (15), we alternatingly optimize $\mathbf{F}$ and $\mathbf{V}$ by solving the following problem for the optimized variable

$$\begin{aligned} \min \ & \mathrm{Tr}\left(\mathbf{V}^H\mathbf{H}^H\mathbf{F}^H\mathbf{G}_w\mathbf{F}\mathbf{H}\mathbf{V}\right) \\ & - 2\Re e\left\{\mathrm{Tr}\left(\mathbf{V}^H\mathbf{H}^H\mathbf{F}^H\mathbf{G}\mathbf{D}_w\right)\right\} + \sigma_R^2\mathrm{Tr}\left(\mathbf{F}^H\mathbf{G}_w\mathbf{F}\right) \\ \text{s.t. } & \mathrm{Tr}\left(\mathbf{V}\mathbf{V}^H\right) \le P_S, \\ & \left\|\mathbf{F}\mathbf{H}\mathbf{V}\right\|_F^2 + \sigma_R^2\left\|\mathbf{F}\right\|_F^2 \le P_R \end{aligned} \tag{25}$$

leading to $\mathbf{F}$-subproblem and $\mathbf{V}$-subproblem. The $\mathbf{F}$-subproblem can be solved using Bisection method while the $\mathbf{V}$-subproblem can be solved by interior-point method, e.g., using the off-the-shelf package SeDuMi. Due to the coupling between the source precoder and the relay precoder in the relay power constraint, the AO method is not necessarily converge to KKT solutions of problem $(P2)$. Specifically, for this problem, the AO method can easily get trapped in some inefficient feasible point, as shown below.

In the simulations, we set $\alpha_k = 1, \forall k$, and the noise power is set to unit for all receivers (i.e., $\sigma_k^2 = \sigma_R^2 = 1, \forall k$). It is further assumed that the relay and the source have the same power budget P for simplicity, i.e., $P_S = P_R = P$, and define $SNR \triangleq 10\log_{10}(P)$. Furthermore, each channel coefficient in both $\mathbf{H}$ and $\mathbf{G}$ is generated from the zero mean complex Gaussian distribution with unit variance. Moreover, for convenience, we denote by (N_s, N_r, K) a relay BC network with N_s source antenna, N_r relay antennas, and K users. For the PDD method, we set the initial penalty parameter $\rho_0 = \frac{500K}{2KN_r+N_r^2+KN_s}$.

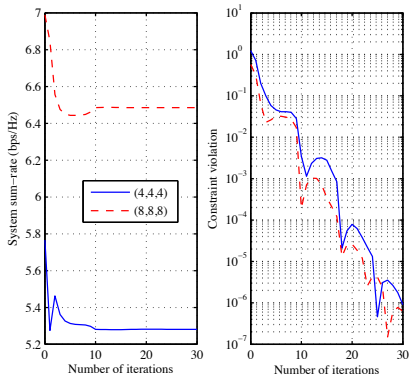


Fig. 3. The convergence performance of the PDD method for different relay BC networks.

The average convergence behavior of the PDD method over ten randomly generated examples is illustrated in Fig. 3, compared with the AO method in Fig. 4. It is observed from Fig. 3 that the PDD method exhibits excellent convergence performance in terms of both the objective value and the feasibility gap. In general, the PDD can converge in 20 iterations for both the network $(4, 4, 4)$ and the network $(8, 8, 8)$. Moreover, it is seen from Fig. 4 that the AO method is not only slow but also gets trapped in inefficient solutions whose objective values are much smaller than that achieved by the PDD method (i.e., right Y-axis versus left Y-axis).

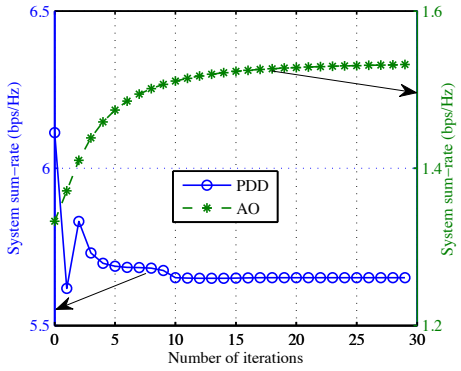


Fig. 4. The convergence performance of the PDD method and the AO method for network $(4, 4, 4)$.

Fig. 5 shows the average sum-rate performance of the PDD method as compared to the AO method for three different networks. Each result of the plot is averaged over 100 channels. One can see that, the PDD method always significantly outperforms the AO mehtod. his is mainly due to the fac that the AO method gets often trapped in some inefficient solutions due to the coupling between the source precoder and the relay precoder. In particular, as the variable dimension grows with the network size, the nonlinear coupling between the variables becomes more heavy and the AO method exhibits worse performance. In addition, we find that the PDD method is always more efficient than the AO method in terms of the cpu time required for convergence, as shown in the caption of Fig. 5.

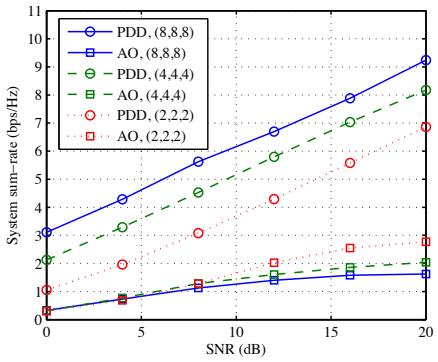


Fig. 5. The sum-rate performance of the PDD method and the AO method. The average ratio between the cpu time required by the AO method and by the PDD method is respectively 26.83, 39.06, and 12.82 for three networks.

IV. VolMin-based matrix factorization

As a popular tool in signal processing and machine learning, matrix factorization (MF) has attracted considerable interest in recent years. In addition to the most popular nonnegative matrix factorization (NMF) [37], various matrix factorization models have been proposed in the literature. Among them, volume-minimization (VolMin)-based matrix factorization is an important class of matrix factorizations where the columns of one factor matrix are constrained to lie in the unit simplex [36]. Compared to NMF, VolMin-based matrix factorization is computationally more challenging. This section considers application of PDD to VolMin-based matrix factorization and provides an alterative VolMin algorithm which can work in the original data space.

A. Problem formulation

Consider the following data measurement model:

$$\boldsymbol{a}[\ell] = \mathbf{X}\boldsymbol{s}[\ell], \ell = 1, 2, \ldots, L, \tag{26}$$

where $\boldsymbol{a}[\ell] \in \mathbb{R}^N$ is a measured data vector indexed by ℓ, $\mathbf{X} \in \mathbb{R}^{N\times K}$ denotes a basis which is assumed to have full column-rank, $\boldsymbol{s}[\ell] \in \mathbb{R}^K$ is the weight vector lying in a probability simplex, i.e.,

$$\boldsymbol{s}[\ell] \geq 0, \mathbf{1}^T\boldsymbol{s}[\ell] = 1, \forall \ell. \tag{27}$$

Define $\mathbf{S} \triangleq [\boldsymbol{s}[1]\ \boldsymbol{s}[2]\ \ldots\ \boldsymbol{s}[L]]$ and $\mathbf{A} \triangleq [\boldsymbol{a}[1]\ \boldsymbol{a}[2]\ \ldots\ \boldsymbol{a}[L]]$. Then the signal model (26) can be compactly written as

$$\mathbf{A} = \mathbf{X}\mathbf{S}. \tag{28}$$

An important motivating example of this model is hyperspectral remote sensing [6], where $\boldsymbol{a}[\ell]$ represents a remotely sensed pixel using sensors of high spectral resolution, the columns of $\mathbf{X}$ denote K different spectral signatures of materials that comprise the pixels, $\boldsymbol{s}_k[\ell]$ denotes the portion of material k contained in pixel $\boldsymbol{x}[\ell]$. Recovering $\mathbf{X}$ is helpful in recognition of the underlying materials in a hyperspectral image. Other applications of this model can be found in document clustering, multi-sensor array processing and blind separation of power spectra for dynamic spectrum access [7], [38], [49].

Given the data measurements $\mathbf{A}$, there possibly exist many combinations of factors $\mathbf{X}$ and $\mathbf{S}$ such $\mathbf{A} = \mathbf{XS}$. The notable works [7] and [50] showed that, under some realistic conditions, unique loading factors (up to column permutations) can be obtained by finding a minimum-volume enclosing simplex of the data vectors. Formally, the VolMin problem can be formulated as[3] [7]

$$\begin{aligned} \min_{\mathbf{X},\mathbf{S}} \ & \log\det(\mathbf{X}^T\mathbf{X}) \\ \text{s.t. } & \mathbf{A} = \mathbf{XS}, \\ & \mathbf{S}^T\mathbf{1} = \mathbf{1}, \mathbf{S} \geq 0. \end{aligned} \tag{29}$$

Problem (29) is challenging due to the nonconvex objective function and the presence of the coupling constraint $\mathbf{A}$=$\mathbf{XS}$. Moreover, it is readily seen that the objective function is not well-defined for rank-deficient $\mathbf{X}$, which may produce numerical instability for iterative algorithms that cannot guarantee full-rank of $\mathbf{X}$ during iterations. To make it well-defined, we modify the objective function as $f_\epsilon(\mathbf{X}^T\mathbf{X})$, defined by

$$f_\epsilon(\mathbf{Y}) \triangleq \sum_{i=1}^{n} \log(g_\epsilon(\sigma_i(\mathbf{Y})))$$

where $g_\epsilon(\cdot)$ is given by

$$g_\epsilon(x) \triangleq \begin{cases} x & \text{if } \ |x| \geq \epsilon \\ \dfrac{1}{2\epsilon}x^2 + \dfrac{\epsilon}{2} & \text{otherwise} \end{cases} \tag{30}$$

where ϵ is a small scalar, e.g., $1e{-}2$, to control the approximation accuracy. With the modified objective function, we obtain an approximation of problem (29)

$$\boxed{\begin{aligned} \min_{\mathbf{X},\mathbf{S}} \ & f_\epsilon(\mathbf{X}^T\mathbf{X}) \\ \text{s.t. } & \mathbf{A} = \mathbf{XS}, \\ & \mathbf{S}^T\mathbf{1} = \mathbf{1}, \mathbf{S} \geq 0. \end{aligned}} \tag{P3}$$

It is noted from (30) that $f_\epsilon(\mathbf{X}^T\mathbf{X}) = \log\det(\mathbf{X}^T\mathbf{X})$ when the smallest singular value of $\mathbf{X}$ is equal or larger than $\sqrt{\epsilon}$. As a result, problem (P3) is equivalent to problem (29) when all the singular values of the optimal solution $\mathbf{X}$ to problem (29) are no smaller than $\sqrt{\epsilon}$. Note that, since the singular value of the optimal $\mathbf{X}$ can be easily ensured to be larger than $\sqrt{\epsilon}$ by appropriately scaling up $\mathbf{A}$, the approximation (P3) could incur no loss of optimality.

[3]It is worth mentioning that, we can also apply the PDD to an equivalent problem of (29) in the reduced-dimension domain (see (45) below). However, we here focus on the original data domain because it allows incorporating constraints on $\mathbf{X}$ (e.g., nonnegativity constraints or other box constraints) into (29) easily, and aims to shed lights on algorithm design for VolMin-related problems in the original data domain.

B. PDD-based algorithm

In this subsection, we develop PDD-based algorithm to address problem $(P3)$. First we reformulate $(P3)$ as follows

$$\begin{aligned} \min_{\mathbf{X},\mathbf{S},\mathbf{Y}} \ & f_\epsilon(\mathbf{X}^T\mathbf{X}) \\ \text{s.t. } & \mathbf{A} = \mathbf{YS} \\ & \mathbf{X} - \mathbf{Y} = 0, \\ & \mathbf{S}^T\mathbf{1} = \mathbf{1}, \mathbf{S} \geq 0. \end{aligned} \tag{31}$$

By building the first two equality constraints into the objective, we obtain the augmented Lagrangian problem of the above reformulation as follows

$$\begin{aligned} \min_{\mathbf{X},\mathbf{S},\mathbf{Y}} \ & f_\epsilon(\mathbf{X}^T\mathbf{X}) + \frac{1}{2\rho}\|\mathbf{A} + \rho\mathbf{P} - \mathbf{YS}\|^2 \\ & \qquad + \frac{1}{2\rho}\|\mathbf{X} + \rho\mathbf{Q} - \mathbf{Y}\|^2 \\ \text{s.t. } & \mathbf{S}^T\mathbf{1} = \mathbf{1}, \mathbf{S} \geq 0. \end{aligned} \tag{32}$$

where $\mathbf{P}$ and $\mathbf{Q}$ are the Lagrange multipliers associated with the first two equality constraints of problem (31), respectively.

Next, we present the BSUM algorithm for problem (32), which consists of the following four steps.

1) Step 1: Update $\mathbf{Y}$ given $\mathbf{X}$ and $\mathbf{S}$: : Fixing $\mathbf{X}$ and $\mathbf{S}$ in (32), we obtain the subproblem with respect to $\mathbf{Y}$ as follows

$$\min_{\mathbf{Y}} \|\mathbf{A} + \rho\mathbf{P} - \mathbf{YS}\|^2 + \|\mathbf{X} + \rho\mathbf{Q} - \mathbf{Y}\|^2 \tag{33}$$

It is a quadratic optimization problem which admits a closed-form solution as follows

$$\mathbf{Y} = \big((\mathbf{A} + \rho\mathbf{P})\mathbf{S}^T + (\mathbf{X} + \rho\mathbf{Q})\big)\,(\mathbf{I} + \mathbf{SS}^T)^{-1}. \tag{34}$$

2) Step 2: Update $\mathbf{S}$ given $\mathbf{X}$ and $\mathbf{Y}$: : By fixing $\mathbf{X}$ and $\mathbf{Y}$ in (32), we obtain the subproblem with respect to $\mathbf{S}$ as follows

$$\begin{aligned} \min_{\mathbf{S}} \ & \|\mathbf{YS} - (\mathbf{A} + \rho\mathbf{P})\|^2 \\ \text{s.t. } & \mathbf{1}^T\mathbf{S} = 1, \mathbf{S} \geq 0. \end{aligned} \tag{35}$$

The above problem is a convex problem which can be globally solved by using some iterative algorithms. To obtain an efficient update for $\mathbf{S}$, we consider updating $\mathbf{S}$ by minimizing a locally tight upper bound of the objective function of (35), i.e., solving

$$\begin{aligned} \min_{\mathbf{S}} \ & \|\mathbf{YS} - (\mathbf{A} + \rho\mathbf{P})\|^2 + \|\mathbf{S} - \tilde{\mathbf{S}}\|^2_{\mathbf{W}} \\ \text{s.t. } & \mathbf{1}^T\mathbf{S} = 1, \mathbf{S} \geq 0. \end{aligned} \tag{36}$$

where $\tilde{\mathbf{S}}$ is the value of $\mathbf{S}$ obtained in the last iteration and $\mathbf{W}$ is a positive definite matrix such that $\mathbf{Y} + \mathbf{W} = \beta\mathbf{I}$ with $\beta > (\sigma_1(\mathbf{Y}))^2$. With simple manipulations, problem (36) can be equivalently written as

$$\begin{aligned} \min_{\mathbf{S}} \ & \|\mathbf{S} - \bar{\mathbf{S}}\|^2, \\ \text{s.t. } & \mathbf{1}^T\mathbf{S} = 1, \mathbf{S} \geq 0. \end{aligned} \tag{37}$$

where $\bar{\mathbf{S}} = \frac{1}{\beta}\left(\mathbf{Y}^T(\mathbf{A}+\rho\mathbf{P}) + (\beta\mathbf{I} - \mathbf{Y}^T\mathbf{Y})\tilde{\mathbf{S}}\right)$. Problem (37) can be decomposed into L independent subproblems which are known as the problem of *projection onto the probability simplex* and admit very efficient semi-closed-form solutions (see [51]).

3) Update $\mathbf{X}$ *given* $\mathbf{Y}$ *and* $\mathbf{S}$*:* By fixing $\mathbf{Y}$ and $\mathbf{S}$ in (32), we obtain the subproblem with respect to $\mathbf{X}$ as follows

$$\min_{\mathbf{X}} \ f_\epsilon(\mathbf{X}^T\mathbf{X}) + \frac{1}{2\rho}\|\mathbf{X} - \bar{\mathbf{X}}\|^2 \tag{38}$$

where $\bar{\mathbf{X}} \triangleq \mathbf{Y} - \rho\mathbf{Q}$. By the definition of $f_\epsilon(\cdot)$, it is known that $f_\epsilon(\mathbf{X}^T\mathbf{X})$ is only related to the singular values of $\mathbf{X}$. Furthermore, by *Von Neumann's trace inequality* $|\mathrm{Tr}\left(\mathbf{X}^T\bar{\mathbf{X}}\right)| \leq \sum_{i=1}^n \sigma_i(\mathbf{X})\sigma_i(\bar{\mathbf{X}})$, it is easily known that the singular vectors of the optimal $\mathbf{X}$ should be aligned with those of $\bar{\mathbf{X}}$. Hence, letting $\bar{\mathbf{U}}\bar{\mathbf{\Sigma}}\bar{\mathbf{V}}^H$ be the thin SVD of $\bar{\mathbf{X}}$, the optimal $\mathbf{X}$ is structured as $\mathbf{X} = \bar{\mathbf{U}}\mathbf{\Sigma}\bar{\mathbf{V}}^H$. As a result, problem (38) reduces to

$$\min_{\mathbf{\Sigma}\geq 0} \ f_\epsilon(\mathbf{\Sigma}^2) + \frac{1}{2\rho}\|\mathbf{\Sigma} - \bar{\mathbf{\Sigma}}\|^2 \tag{39}$$

which can be decomposed into a set of independent subproblems with the i-th subproblem in the form of

$$\min_{\sigma_i\geq 0} \ \log(g_\epsilon(\sigma_i^2)) + \frac{1}{2\rho}(\sigma_i - \bar{\sigma}_i)^2 \tag{40}$$

where σ_i is the i-th singular value of $\mathbf{X}$. The above problem is difficult to solve and thus we devote our efforts to minimizing a locally tight upper bound of the objective function. By the concavity of the $\log(\cdot)$ function, we have $\log(x) \leq \log(\tilde{x}) + \frac{1}{\tilde{x}}(x - \tilde{x})$, $\forall x, \tilde{x} > 0$. Using such an upper bound for $\log(g_\epsilon(\sigma_i^2))$, we update σ_i by solving

$$\min_{\sigma_i\geq 0} \ \frac{1}{\tilde{g}_\epsilon} g_\epsilon(\sigma_i^2) + \frac{1}{2\rho}(\sigma_i - \bar{\sigma}_i)^2 \tag{41}$$

where $\tilde{g}_\epsilon = g_\epsilon\left((\tilde{\sigma}_i)^2\right)$ and $\tilde{\sigma}_i = \sigma_i(\tilde{\mathbf{X}})$ with $\tilde{\mathbf{X}}$ being the value of $\mathbf{X}$ obtained in the last iteration.

It can be shown that the objective function of problem (41) is convex with respect to $\sigma_i \geq 0$. Hence, problem (41) is a convex problem and can be globally solved. Specifically, we solve it by considering two cases. The first case is when $\sigma_i^2 \geq \epsilon$. In this case, problem (41) reduces to

$$\min_{\sigma_i\geq\sqrt{\epsilon}} \ \frac{1}{\tilde{g}_\epsilon}\sigma_i^2 + \frac{1}{2\rho}(\sigma_i - \bar{\sigma}_i)^2 \tag{42}$$

which admits a closed-form solution as follows

$$\sigma_i = \max\left(\frac{\tilde{g}_\epsilon}{2\rho + \tilde{g}_\epsilon}\bar{\sigma}_i, \sqrt{\epsilon}\right)$$

In the second case when $\sigma_i^2 \leq \epsilon$, problem (41) reduces to

$$\min_{0\leq\sigma_i\leq\sqrt{\epsilon}} \ \frac{1}{2\epsilon\tilde{g}_\epsilon}\sigma_i^4 + \frac{1}{2\rho_k}(\sigma_i - \bar{\sigma}_i)^2 \tag{43}$$

which admits a closed-form solution as $\sigma_i = [\sigma_i^\star]_0^{\sqrt{\epsilon}}$ where $\sigma_i^\star$ is the unique solution to the following cubic equation

$$\frac{2}{\epsilon\tilde{g}_\epsilon}\sigma_i^3 + \frac{1}{\rho_k}\sigma_i - \frac{1}{\rho_k}\bar{\sigma}_i = 0.$$

By comparing the objective values of the above two cases, we can obtain the optimal σ_i for problem (41). After obtaining $\mathbf{\Sigma}$, we finally obtain the optimal solution $\mathbf{X}$ to problem (38), i.e., $\mathbf{X} = \bar{\mathbf{U}}\mathbf{\Sigma}\bar{\mathbf{V}}^H$. We omit the detailed implementation of the BSUM algorithm for problem (32) due to space limitation.

4) Numerical examples: We here present numerical examples to illustrate the performance of the PDD-based VolMin algorithm by comparing with the state-of-art VolMin algorithm—SISAL [6]. SISAL works for the equivalent problem of problem $(P3)$ in the reduced-dimension domain [6],

$$\begin{aligned}\max_{\mathbf{Q}\in\mathbb{R}^{K\times K}} \ & \log|\det(\mathbf{Q})| \\ \text{s.t. } & \mathbf{Q}^T\mathbf{1} = (\mathbf{A}_r\mathbf{A}_r^T)^{-1}\mathbf{A}_r\mathbf{1}, \mathbf{Q}\mathbf{A}_r \geq 0.\end{aligned} \tag{44}$$

where $\mathbf{A}_r \triangleq \mathbf{U}_r^T\mathbf{A}$ and $\mathbf{U}_r \in \mathbb{R}^{N\times K}$ consists of the left-sigular vectors of $\mathbf{A}$ (which can be obtained by performing thin SVD on $\mathbf{A}$). In SISAL, the inequality constraints are penalized to the objective by using *hinge-loss* function, leading to a penalized problem [6]

$$\begin{aligned}\max_{\mathbf{Q}\in\mathbb{R}^{K\times K}} \ & \log|\det(\mathbf{Q})| - \eta\sum_{i,j}\max(-[\mathbf{Q}\mathbf{A}_r]_{ij}, 0) \\ \text{s.t. } & \mathbf{Q}^T\mathbf{1} = (\mathbf{A}_r\mathbf{A}_r^T)^{-1}\mathbf{1}.\end{aligned} \tag{45}$$

where η is a penalty parameter. The SISAL algorithm aims to solve the penalized problem by using successive second-order approximation and variable splitting technique[4]. The algorithm is lightweight but its convergence is unclear. Moreover, the SISAL algorithm does not apply to the original data space.

In our simulations, we generate the elements of $\mathbf{X}$ from the uniform distribution between zero and one, and generate $\boldsymbol{s}[\ell]$ on the unit simplex and with $\max_i s_i[\ell] \leq \gamma$, where $\gamma = 0.8$ is given, which results in a so-called 'no-pure-pixel case' in the context of remote sensing and is known to be challenging to handle [6], [49]. Moreover, we use the mean-square-error (MSE) of $\mathbf{X}$ as a measure of estimation performance (instead of achieved volume which is less physically meaningful), defined by

$$MSE = \min_{\boldsymbol{\pi}\in\Pi}\frac{1}{K}\sum_{k=1}^K\left\|\frac{\boldsymbol{x}_k}{\|\boldsymbol{x}_k\|} - \frac{\hat{\boldsymbol{x}}_{\boldsymbol{\pi}_k}}{\|\hat{\boldsymbol{x}}_{\boldsymbol{\pi}_k}\|}\right\|^2 \tag{46}$$

where Π is the set of all permutations of $\{1, 2, \ldots, K\}$, and $\hat{\boldsymbol{x}}_k$ is the estimate of $\boldsymbol{x}_k$. For the PDD method, the initial penalty parameter ρ_0 is set to $L/100$.

We first show in Fig. 6 the convergence performance of the PDD method. In the plot, the results are averaged over ten randomly generated examples with random initialization for two cases: $(N, K, L) = (10, 3, 200)$ and $(N, K, L) = (50, 3, 2000)$. It can be observed that the PDD method achieve approximate feasibility in tens of iterations. Particularly, it can quickly reach a good estimation accuracy; the MSE could be less than -35 dB in twenty outer iterations.

We then use two illustrative examples to show the effectiveness of the proposed PDD-based VolMin algorithm. In these two examples, we again set $(N, K, L) = (10, 3, 200)$ and $(N, K, L) = (50, 3, 2000)$. To visualize the results, we

[4]The code of SISAL can be found from http://www.lx.it.pt/~bioucas/code.htm.

project the data points, the ground truth, and the estimates to a two-dimensional plane. In Figs. 7 and 8, we see that, the PDD method can provide an estimate as good as the SISAL method's if the latter is particularly initialized from the estimate of VCA method [52]. Moreover, we observe from simulations that the PDD method is less sensitive to random initialization than the SISAl method, as shown in the Figs. 7-8 and also Fig. 9 below.

To further demonstrate the performance of the PDD method under random initialization, we randomly generate 100 examples and evaluate the estimation performance of various methods. Motivated by the observations from Fig. 6, we simply set the maximum outer iterations of the PDD method as 30 in this set of simulations. Moreover, to test the performance of the PDD method in a noisy environment, zero-mean white Gaussian noise $\boldsymbol{v}[\ell]$ is added to each generated data. We define the signal-to-noise ratio (SNR) as $SNR{=}10\log_{10}\left(\frac{\mathbb{E}\{\|\mathbf{X}\boldsymbol{s}[\ell]\|^2\}}{\mathbb{E}\{\|\boldsymbol{v}[\ell]\|^2\}}\right)$, and set $SNR{=}40$ dB for each example. Fig. 9 illustrates the estimation performance of various methods for 100 examples. We see that, the PDD is much more robust to random initializations than SISAL. In addition to random initialization, the performance of SISAL is also impacted by the choice of η. Moreover, the PDD method[5] with three random initializations can provide very high estimation performance that is comparable with the SISAL method when the latter is set with a finely tuned penalty parameter η and particularly initialized from VCA.

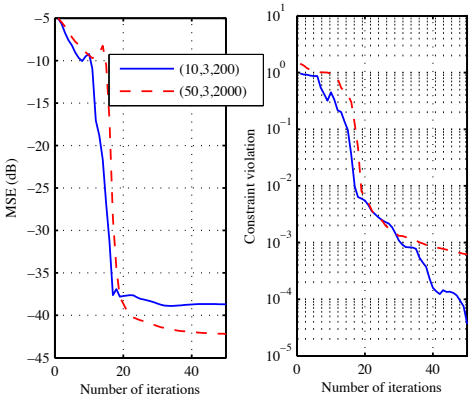


Fig. 6. The convergence performance of the PDD method.

V. Conclusions

In this two-part paper, we proposed and analyzed a new optimization framework for optimizing nonconvex nonsmooth functions subject to nonconvex coupling constraints. Part I developed the general framework and investigated its convergence properties. In this Part II, we customized our PDD framework to three challenging problems in signal processing and machine learning. Our algorithms guarantee convergence to stationary solutions of the three problems[6] and were shown numerically to be able to yield better solutions than the state-of-the-art schemes in the literature. We remark that, our framework finds applications also in other areas, such as optimal power flow in smart grids, user scheduling in wireless communications, cross-layer design of networks, etc..

[5]To combat against the impact of initialization, we run PDD with three random initializations and pick the best one as the output.

[6]The verification of constraint qualification for the three problems is relegated to Appendix C.

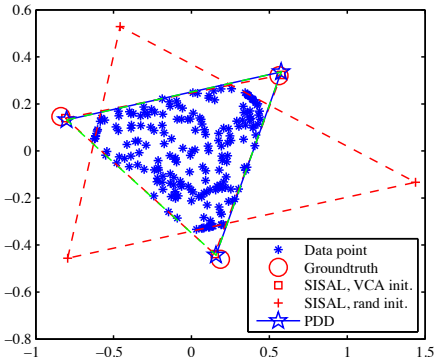


Fig. 7. A geometric illustration of estimation results by different methods with $N = 10$, $K = 3$, and $L = 200$.

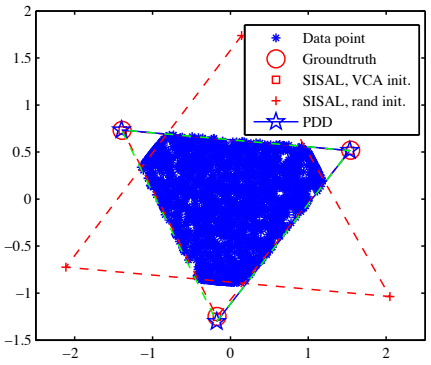


Fig. 8. A geometrical illustration of estimation results by different methods with $N = 50$, $K = 3$, and $L = 2000$.

Appendix A
Solving Problem (6) for $\boldsymbol{t}$

This appendix shows how the $\boldsymbol{t}$-subproblem is globally solved. For a clear illustration, we write the $\boldsymbol{t}$-subproblem

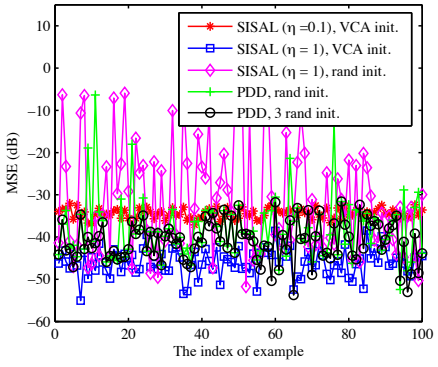


Fig. 9. The MSE performance of different methods with $N = 50$, $K = 3$, and $L = 1000$

explicitly as follows

$$\max_{\boldsymbol{t}\geq 0}\ \min_{1\leq k\leq K} t_k - \frac{1}{2\rho}\sum_{k=1}^{K}\left(\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\| - t_k\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\| + \rho\lambda_k\right)^2 \quad (47)$$

It can be compactly written as

$$\max_{\boldsymbol{t}\geq 0}\ \min_{k} t_k - \sum_{k=1}^{K} a_k(t_k - b_k)^2 \quad (48)$$

where $a_k \triangleq \frac{\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\|^2}{2\rho}$ and $b_k \triangleq \frac{\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\| + \rho\lambda_k}{\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\|}$. Further, by introducing an auxiliary variable s, the above problem can be equivalently written as

$$\begin{aligned}\max_{\boldsymbol{t}\geq 0, s}\ & s - \sum_{k=1}^{K} a_k(t_k - b_k)^2 \quad (49)\\ \text{s.t. }& t_k \geq s \geq 0, \forall k.\end{aligned}$$

We solve the above problem by exploiting its problem structure. To do so, let t_k^*'s and s^* denote the optimal solution to the above problem. It is seen from (49) that, if $t_k^* > s^* \geq 0$, then we must have $t_k^* = b_k$ to maximize the objective. Hence, by assuming without loss of generality $b_1 \geq b_2 \geq \ldots \geq b_K$, we infer that, there exists $\bar{k}$ such that $t_k^* = b_k$, $\forall k \leq \bar{k}$, and $t_k^* = s^*$, $\forall k > \bar{k}$, where s^* is given by

$$\begin{aligned}s^* &= \arg\max_{s\geq 0}\ s - \sum_{k>\bar{k}} a_k(s - b_k)^2 \quad (50)\\ &= \max\left(\frac{1 + 2\sum_{k>\bar{k}} a_k b_k}{2\sum_{k>\bar{k}} a_k}, 0\right) \triangleq \tau(\bar{k}). \quad (51)\end{aligned}$$

As a result, it is equivalent to find $\bar{k} > 0$ such that $b_{\bar{k}}^* > \tau(\bar{k})$ if such a $\bar{k}$ exists; otherwise we have $\bar{k} = 0$ (i.e., $t_k^* = s^* = \tau(0)$, $\forall k$). Given $\bar{k}$, we can derive the optimal solution as stated above.

Appendix B
The expression of matrix $\mathbf{C}$

We here derive a homogeneous quadratic upper bound for $\vartheta(\boldsymbol{w})$ over the constraint $\|\boldsymbol{w}\| = 1$. Since $\vartheta(\boldsymbol{w})$ is in the form of

$$\vartheta(\boldsymbol{w}) = \sum_{k=1}^{K}\left(\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\| - t_k\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\| + \rho\lambda_k\right)^2,$$

we only need to bound each summand $\vartheta_k(\boldsymbol{w}) \triangleq \left(\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\| - t_k\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\| + \rho\lambda_k\right)^2$, $k = 1, 2, \ldots, K$.

First, $\vartheta_k(\boldsymbol{w})$ can be expressed as

$$\begin{aligned}\vartheta_k(\boldsymbol{w}) = \boldsymbol{w}^H\mathbf{A}_k\boldsymbol{w} + t_k^2\boldsymbol{w}^H\mathbf{B}_k\boldsymbol{w} + \rho^2\lambda_k^2\\ - 2t_k\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\|\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\| + 2\rho g_k(\boldsymbol{w})\end{aligned} \quad (52)$$

where

$$g_k(\boldsymbol{w}) \triangleq \begin{cases} -\lambda_k t_k\|\boldsymbol{w}\|\left\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\right\| + \lambda_k\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\|, \text{if } \lambda_k \geq 0\\ -\lambda_k t_k\left\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\right\| + \lambda_k\|\boldsymbol{w}\|\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\|, \text{otherwise}\end{cases} \quad (53)$$

Note that we have used the fact $\|\boldsymbol{w}\| = 1$ in the definition of $g_k(\boldsymbol{w})$, so that $g_k(\boldsymbol{w})$ has similar forms for both cases of λ_k, which can be easily upper bounded. In what follows, without loss of generality, we consider only the case when $\lambda_k \geq 0$.

To bound $\vartheta_k(\boldsymbol{w})$, let us define $\boldsymbol{w}_{eq} \triangleq (\Re e\{\boldsymbol{w}\}, \Im m\{\boldsymbol{w}\})$, $\tilde{\boldsymbol{w}}_{eq} \triangleq (\Re e\{\tilde{\boldsymbol{w}}\}, \Im m\{\tilde{\boldsymbol{w}}\})$, and

$$\begin{aligned}\mathbf{A}_{k,eq} &\triangleq \begin{pmatrix} \Re e\{\mathbf{A}_k\} & -\Im m\{\mathbf{A}_k\}\\ \Im m\{\mathbf{A}_k\} & \Re e\{\mathbf{A}_k\}\end{pmatrix},\\ \mathbf{B}_{k,eq} &\triangleq \begin{pmatrix} \Re e\{\mathbf{B}_k\} & -\Im m\{\mathbf{B}_k\}\\ \Im m\{\mathbf{B}_k\} & \Re e\{\mathbf{B}_k\}\end{pmatrix}.\end{aligned}$$

Then, by applying part 1) of Lemma 2.1, we have

$$\begin{aligned}&2t_k\left\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\right\|\left\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\right\| + 2\rho\lambda_k t_k\|\boldsymbol{w}\|\left\|\mathbf{B}_k^{\frac{1}{2}}\boldsymbol{w}\right\|\\ =&2t_k\left\|\mathbf{A}_{k,eq}^{\frac{1}{2}}\boldsymbol{w}_{eq}\right\|\left\|\mathbf{B}_{k,eq}^{\frac{1}{2}}\boldsymbol{w}_{eq}\right\| + 2\rho\lambda_k t_k\|\boldsymbol{w}_{eq}\|\left\|\mathbf{B}_{k,eq}^{\frac{1}{2}}\boldsymbol{w}_{eq}\right\|\\ \geq&\boldsymbol{w}_{eq}^T\boldsymbol{\Omega}_k(\tilde{\boldsymbol{w}})\boldsymbol{w}_{eq}\end{aligned}$$

where

$$\begin{aligned}\boldsymbol{\Omega}_k(\tilde{\boldsymbol{w}}) \triangleq& \frac{t_k}{\left\|\mathbf{A}_{k,eq}^{\frac{1}{2}}\tilde{\boldsymbol{w}}_{eq}\right\|\left\|\mathbf{B}_{k,eq}^{\frac{1}{2}}\tilde{\boldsymbol{w}}_{eq}\right\|}\left(\mathbf{A}_{k,eq}\tilde{\boldsymbol{w}}_{eq}\tilde{\boldsymbol{w}}_{eq}^T\mathbf{B}_{k,eq}\right.\\ &\left.+\mathbf{B}_{k,eq}\tilde{\boldsymbol{w}}_{eq}\tilde{\boldsymbol{w}}_{eq}^T\mathbf{A}_{k,eq}\right)\\ &+\frac{\rho\lambda_k t_k}{\|\tilde{\boldsymbol{w}}_{eq}\|\left\|\mathbf{B}_{k,eq}^{\frac{1}{2}}\tilde{\boldsymbol{w}}_{eq}\right\|}\left(\tilde{\boldsymbol{w}}_{eq}\tilde{\boldsymbol{w}}_{eq}^T\mathbf{B}_{k,eq} + \mathbf{B}_{k,eq}\tilde{\boldsymbol{w}}_{eq}\tilde{\boldsymbol{w}}_{eq}^T\right).\end{aligned} \quad (54)$$

Furthermore, by applying part 2) of Lemma 2.1, we have

$$\begin{aligned}&2\left\|\mathbf{A}_k^{\frac{1}{2}}\boldsymbol{w}\right\| = 2\left\|\mathbf{A}_{k,eq}^{\frac{1}{2}}\boldsymbol{w}_{eq}\right\|\\ &\leq \frac{1}{\left\|\mathbf{A}_{k,eq}^{\frac{1}{2}}\tilde{\boldsymbol{w}}_{eq}\right\|}\left\|\mathbf{A}_{k,eq}^{\frac{1}{2}}\boldsymbol{w}_{eq}\right\|^2 + \left\|\mathbf{A}_{k,eq}^{\frac{1}{2}}\tilde{\boldsymbol{w}}_{eq}\right\|\end{aligned} \quad (55)$$

As a result, we can obtain a locally tight quadratic upper bound for $\vartheta_k(\boldsymbol{w})$ given by

$$\vartheta_k(\boldsymbol{w}) \leq \boldsymbol{w}_{eq}^T\mathbf{C}_k\boldsymbol{w}_{eq} + const \quad (56)$$

where

$$\mathbf{C}_k \triangleq \left(1 + \frac{\rho\lambda_k}{\|\mathbf{A}_k^{\frac{1}{2}}\tilde{\boldsymbol{w}}\|}\right)\mathbf{A}_{k,eq} + t_k^2\mathbf{B}_{k,eq} - \boldsymbol{\Omega}_k(\tilde{\boldsymbol{w}}). \quad (57)$$

Finally, we have

$$\vartheta(\boldsymbol{w}) \leq u(\boldsymbol{w}, \tilde{\boldsymbol{w}}) \triangleq \boldsymbol{w}_{eq}^T\mathbf{C}\boldsymbol{w}_{eq} + const \quad (58)$$

where $\mathbf{C} \triangleq \sum_{k=1}^{K}\mathbf{C}_k$.

Appendix C
Constraint qualification of problems (5), (11), and (31)

In this appendix, we verify the constraint qualification of problems (5), (11), and (31) by considering Mangasarian-Fromovitz constraint qualification (MFCQ) (which is equivalent to Robinson's condition for these problems).

First, let us introduce MFCQ for the constraints of the following problem

$$\begin{aligned}&\min\ f(\boldsymbol{x})\\ &\text{s.t. } \boldsymbol{h}(\boldsymbol{x}) = 0,\\ &\quad \boldsymbol{g}(\boldsymbol{x}) \leq 0,\end{aligned} \quad (59)$$

where the functions $f : \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{g} : \mathbb{R}^n \to \mathbb{R}^p$ and $\boldsymbol{h} : \mathbb{R}^n \to \mathbb{R}^m$ are continuously differentiable. The feasible set is $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{h}(\boldsymbol{x}) = 0, \boldsymbol{g}(\boldsymbol{x}) \leq 0\}$. Given $\bar{\boldsymbol{x}} \in \Omega$, $\mathcal{A}(\bar{\boldsymbol{x}})$ is the set of the inequality active constraint indices, that is

$$\mathcal{A}(\bar{\boldsymbol{x}}) = \{i \in \{1, 2, \ldots, p\} \mid \boldsymbol{g}_i(\bar{\boldsymbol{x}}) = 0\}. \tag{60}$$

For problem (59), we say that *MFCQ holds at $\bar{\boldsymbol{x}}$ when the equality constraint gradients are linearly independent and there exists a vector $\boldsymbol{d} \in \mathbb{R}^n$ such that*

$$\nabla \boldsymbol{h}(\bar{\boldsymbol{x}})\boldsymbol{d} = \boldsymbol{0}, \tag{61}$$

$$\nabla g_j(\bar{\boldsymbol{x}})^T \boldsymbol{d} < 0, \forall j \in \mathcal{A}(\bar{\boldsymbol{x}}). \tag{62}$$

Here, $\nabla \boldsymbol{h}(\bar{\boldsymbol{x}})$ denotes the Jacobian matrix of $\boldsymbol{h}(\boldsymbol{x})$ and $\nabla g_j(\bar{\boldsymbol{x}})$ is the gradient of $g_j(\boldsymbol{x})$. Thus, the equality constraint gradients are given by the columns of $\nabla \boldsymbol{h}(\bar{\boldsymbol{x}})^T$.

Remark C.1: By the first-order approximation, we have $\boldsymbol{h}(\bar{\boldsymbol{x}} + \boldsymbol{d}) \approx \boldsymbol{h}(\bar{\boldsymbol{x}}) + \nabla \boldsymbol{h}(\bar{\boldsymbol{x}})\boldsymbol{d}$. Hence, we can obtain the term $\nabla \boldsymbol{h}(\bar{\boldsymbol{x}})\boldsymbol{d}$ by using first-order approximation without need of computing the Jacobian matrix (or the gradient). This observation will facilitate the MFCQ verification in the case when $\boldsymbol{h}$ and $\boldsymbol{x}$ are both matrices.

Next, let us check the MFCQ of three problems one by one.

A. CQ verification for problem (5)

For problem (5), we have

Lemma C.1: MFCQ holds for problem (5) at any feasible point $(\boldsymbol{w}, \boldsymbol{t})$.

Proof: It is readily seen that the inequality constraints $\boldsymbol{t} \geq 0$ of problem (5) must be inactive. Hence, we only need to check the equality constraints which can be compactly written as $\boldsymbol{h}(\boldsymbol{w}, t) = \boldsymbol{0}$ with the following definition

$$\boldsymbol{h}(\boldsymbol{w}, \boldsymbol{t}) \triangleq \begin{bmatrix} \|\mathbf{A}_1^{\frac{1}{2}}\boldsymbol{w}\| - t_1\|\mathbf{B}_1^{\frac{1}{2}}\boldsymbol{w}\| \\ \|\mathbf{A}_2^{\frac{1}{2}}\boldsymbol{w}\| - t_2\|\mathbf{B}_2^{\frac{1}{2}}\boldsymbol{w}\| \\ \vdots \\ \|\boldsymbol{w}\|^2 - 1 \end{bmatrix}. \tag{63}$$

The equality constraint gradients are computed in (64) (see the top of the next page). By noting $\boldsymbol{w} \neq \boldsymbol{0}$, it is readily known that the columns of $\nabla \boldsymbol{h}(\boldsymbol{w}, t)^T$ are linearly independent. Furthermore, by simply setting $\boldsymbol{d} = \boldsymbol{0}$, MFCQ holds for problem (5) at any its feasible point. ■

B. CQ verification for problem (11)

For problem (11), we have

Lemma C.2: MFCQ holds for problem (11) at any *nonzero* feasible point $(\mathbf{V}, \mathbf{F}, \mathbf{X})$ with $\mathbf{V} \neq \mathbf{0}$, and $\mathbf{F} \neq \mathbf{0}$ or $\mathbf{X} \neq 0$.

Proof: The constraints of problem (11) are written as follows

$$g_1 \triangleq \mathrm{Tr}\left(\bar{\mathbf{V}}\bar{\mathbf{V}}^H\right) - P_S \leq 0, \tag{65a}$$

$$g_2 \triangleq \left\|\bar{\mathbf{X}}\right\|_F^2 + \sigma_R^2 \left\|\bar{\mathbf{F}}\right\|_F^2 - P_R \leq 0, \tag{65b}$$

$$\boldsymbol{\Theta}_1 \triangleq \mathbf{X} - \mathbf{FHV} = \mathbf{0}, \tag{65c}$$

$$\boldsymbol{\Theta}_2 \triangleq \mathbf{F} - \bar{\mathbf{F}} = \mathbf{0}, \tag{65d}$$

$$\boldsymbol{\Theta}_3 \triangleq \mathbf{X} - \bar{\mathbf{X}} = \mathbf{0}, \tag{65e}$$

$$\boldsymbol{\Theta}_4 \triangleq \mathbf{V} - \bar{\mathbf{V}} = \mathbf{0}. \tag{65f}$$

As we can see, $\boldsymbol{\Theta}_1$, $\boldsymbol{\Theta}_2$ and $\boldsymbol{\Theta}_3$ do not contain variable $\bar{\mathbf{V}}$ but $\boldsymbol{\Theta}_4$ does. Thus, the gradients of the components of $\boldsymbol{\Theta}_1$, $\boldsymbol{\Theta}_2$, $\boldsymbol{\Theta}_3$, and $\boldsymbol{\Theta}_4$ are linearly dependent if and only if those of the components of $\boldsymbol{\Theta}_1$, $\boldsymbol{\Theta}_2$, and $\boldsymbol{\Theta}_3$ are linearly dependent. Similarly, since $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$ do not contain $\bar{\mathbf{X}}$, the gradients of the components of $\boldsymbol{\Theta}_1$, $\boldsymbol{\Theta}_2$, and $\boldsymbol{\Theta}_3$ are linearly dependent if and only if those of the components of $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$ are linearly dependent. However, the gradients of the components of $\boldsymbol{\Theta}_1$ and $\boldsymbol{\Theta}_2$ are linearly independent because $\boldsymbol{\Theta}_1$ does not contain $\bar{\mathbf{F}}$ but $\boldsymbol{\Theta}_2$ does. Therefore, the equality constraint gradients of problem (11) are linearly independent.

Given the above gradient independence result, we are left to show that, there exists $\{\mathbf{D}_{\mathbf{X}}, \mathbf{D}_{\mathbf{F}}, \mathbf{D}_{\mathbf{V}}, \mathbf{D}_{\bar{\mathbf{X}}}, \mathbf{D}_{\bar{\mathbf{F}}}, \mathbf{D}_{\bar{\mathbf{V}}}\}$ such

$$\Re e\left\{\mathrm{Tr}\left(\bar{\mathbf{V}}\mathbf{D}_{\mathbf{V}}^H\right)\right\} < 0, \tag{66a}$$

$$\Re e\left\{\mathrm{Tr}\left(\bar{\mathbf{X}}\mathbf{D}_{\bar{\mathbf{X}}}^H\right) + \sigma_R^2 \Re e\left\{\mathrm{Tr}\left(\bar{\mathbf{F}}\mathbf{D}_{\bar{\mathbf{F}}}^H\right)\right\}\right\} < 0, \tag{66b}$$

$$\mathbf{D}_{\mathbf{X}} - \mathbf{D}_{\mathbf{F}}\mathbf{HV} - \mathbf{FHD}_{\mathbf{V}} = \mathbf{0}, \tag{66c}$$

$$\mathbf{D}_{\mathbf{F}} - \mathbf{D}_{\bar{\mathbf{F}}} = \mathbf{0}, \tag{66d}$$

$$\mathbf{D}_{\mathbf{X}} - \mathbf{D}_{\bar{\mathbf{X}}} = \mathbf{0}, \tag{66e}$$

$$\mathbf{D}_{\mathbf{V}} - \mathbf{D}_{\bar{\mathbf{V}}} = \mathbf{0}, \tag{66f}$$

which are derived using first-order approximation according to Remark C.1. Note that we here consider only the case when (65a) and (65b) are active. Other cases (i.e., both are inactive, and either of (65a) and (65b) is active) can be simply treated.

It can be shown that, Eq. (66) is satisfied by taking $\{\mathbf{D}_{\mathbf{X}}, \mathbf{D}_{\mathbf{F}}, \mathbf{D}_{\mathbf{V}}, \mathbf{D}_{\bar{\mathbf{X}}}, \mathbf{D}_{\bar{\mathbf{F}}}, \mathbf{D}_{\bar{\mathbf{V}}}\} = \{-2\mathbf{X}, -\mathbf{F}, -\mathbf{V}, -2\mathbf{X}, -\mathbf{F}, -\mathbf{V}\}$ with $\mathbf{V} \neq \mathbf{0}$, and $\mathbf{F} \neq \mathbf{0}$ or $\mathbf{X} \neq 0$. This completes the proof.

■

C. CQ verification for problem (31)

For problem (31), we have

Lemma C.3: MFCQ holds for problem (31) at any feasible point $(\mathbf{X}, \mathbf{S}, \mathbf{Y})$.

Proof: Similarly as for problem (11), we can show that the linear independence of the equality constraint gradients of problem (31). So our main efforts are paid to show that, there exists $\{\mathbf{D}_{\mathbf{X}}, \mathbf{D}_{\mathbf{S}}, \mathbf{D}_{\mathbf{Y}}\}$ such

$$\mathbf{YD}_{\mathbf{S}} + \mathbf{D}_{\mathbf{Y}}\mathbf{S} = \mathbf{0}, \tag{67a}$$

$$\mathbf{D}_{\mathbf{X}} - \mathbf{D}_{\mathbf{Y}} = \mathbf{0}, \tag{67b}$$

$$\mathbf{D}_{\mathbf{S}}^T\mathbf{1} = \mathbf{0}, \tag{67c}$$

$$[\mathbf{D}_{\mathbf{S}}]_{i,j} > 0, \forall (i,j) \in \mathcal{S}_0, \tag{67d}$$

where $\mathcal{S}_0$ is the set of zero entry indices of $\mathbf{S}$, $[\mathbf{D}_{\mathbf{S}}]_{i,j}$ is the (i,j)-th entry of $\mathbf{D}_{\mathbf{S}}$.

Let us check Eq. (67) with the point $(\mathbf{D}_{\mathbf{X}}, \mathbf{D}_{\mathbf{Y}}, \mathbf{D}_{\mathbf{S}})$ given by

$$\mathbf{D}_{\mathbf{S}} = \frac{1}{L}\mathbf{1}\mathbf{1}^T - \mathbf{S}, \mathbf{D}_{\mathbf{X}} = \mathbf{D}_{\mathbf{Y}} = \mathbf{Y}(\mathbf{I} - \frac{1}{L}\mathbf{1}\mathbf{1}^T). \tag{68}$$

where $\mathbf{1}$ is an all-one vector of dimension L. Obviously, Eqs. (67b) and (67d) are true. Furthermore, we have

$$\begin{aligned} \mathbf{D}_{\mathbf{S}}^T\mathbf{1} =& (\frac{1}{L}\mathbf{1}\mathbf{1}^T - \mathbf{S})\mathbf{1} \\ =& \mathbf{1} - \mathbf{S}^T\mathbf{1} = 0 \end{aligned} \tag{69}$$

$$\nabla \boldsymbol{h}(\boldsymbol{w},t)=\begin{bmatrix} \frac{\boldsymbol{w}^T\mathbf{A}_1^T}{\|\mathbf{A}_1\boldsymbol{w}\|}-\frac{t_1\boldsymbol{w}^T\mathbf{B}_1^T}{\|\mathbf{B}_1\boldsymbol{w}\|} & -\|\mathbf{B}_1\boldsymbol{w}\| & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\boldsymbol{w}^T\mathbf{A}_K^T}{\|\mathbf{A}_K\boldsymbol{w}\|}-\frac{t_K\boldsymbol{w}^T\mathbf{B}_K^T}{\|\mathbf{B}_K\boldsymbol{w}\|} & 0 & \vdots & -\|\mathbf{B}_K\boldsymbol{w}\| \\ 2\boldsymbol{w}^T & \mathbf{0}^T & \mathbf{0}^T & \mathbf{0}^T \end{bmatrix}. \tag{64}$$

where the last equality follows from the feasibility. Substituting (68) into (67a), we obtain

$$\begin{aligned} &\mathbf{Y}(\frac{1}{L}\mathbf{1}\mathbf{1}^T-\mathbf{S})+\mathbf{D}_\mathbf{Y}\mathbf{S} \\ &=\frac{1}{L}\mathbf{Y}\mathbf{1}\mathbf{1}^T-\mathbf{Y}\mathbf{S}+\mathbf{D}_\mathbf{Y}\mathbf{S} \\ &=\frac{1}{L}\mathbf{Y}\mathbf{1}\mathbf{1}^T\mathbf{S}-\mathbf{Y}\mathbf{S}+\mathbf{D}_\mathbf{Y}\mathbf{S} \\ &=[\mathbf{D}_\mathbf{Y}-\mathbf{Y}(\mathbf{I}-\frac{1}{L}\mathbf{1}\mathbf{1}^T)]\mathbf{S}=\mathbf{0}, \end{aligned} \tag{70}$$

where the second equality is due to $\mathbf{S}^T\mathbf{1}=\mathbf{1}$. Therefore, MFCQ holds for problem (31) at any feasible point. ■

REFERENCES

[1] H. T. Wai, T. H. Chang, and A. Scaglione, "A consensus-based decentralized algorithm for non-convex optimization with application to dictionary learning," in *IEEE ICASSP*, April 2015, pp. 3546–3550.

[2] H. Zayyani, M. Korki, and F. Marvasti, "Dictionary learning for blind one bit compressed sensing," *IEEE Signal Processing Letters*, vol. 23, no. 2, pp. 187–191, Feb 2016.

[3] Z. Yang and L. Xie, "Enhancing sparsity and resolution via reweighted atomic norm minimization," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 995–1006, Feb 2016.

[4] X. Fu, W. K. Ma, K. Huang, and N. D. Sidiropoulos, "Blind separation of quasi-stationary sources: Exploiting convex geometry in covariance domain," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2306–2320, May 2015.

[5] J. M. Bioucas-Dias, "A variable splitting augmented lagrangian approach to linear spectral unmixing," in *2009 First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing*, Aug 2009, pp. 1–4.

[6] X. Fu, W. K. Ma, K. Huang, and N. D. Sidiropoulos, "Blind separation of quasi-stationary sources: Exploiting convex geometry in covariance domain," *IEEE Trans. Signal Process.*, vol. 63, no. 9, pp. 2306–2320, May 2015.

[7] Q. Shi, C. Peng, W. Xu, M. Hong, and Y. Cai, "Energy efficiency optimization for miso swipt systems with zero-forcing beamforming," *IEEE Trans. Signal Process.*, vol. 64, no. 4, pp. 842–854, Feb 2016.

[8] Q. Shi, M. Razaviyayn, M. Hong, and Z. Q. Luo, "SINR constrained beamforming for a MIMO multi-user downlink system: Algorithms and convergence analysis," *IEEE Trans. Signal Process.*, vol. 64, no. 11, pp. 2920–2933, June 2016.

[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 Trans. Signal Process.*, vol. 56, no. 3, pp. 1268–1279, March 2008.

[10] M. Hong, R. Sun, H. Baligh, and Z. Q. Luo, "Joint base station clustering and beamformer design for partial coordinated transmission in heterogeneous networks," *IEEE J. Sel. Areas in Commun.*, vol. 31, no. 2, pp. 226–240, February 2013.

[11] N. D. Sidiropoulos, T. N. Davidson, and Z.-Q. Luo, "Transmit beamforming for physical-layer multicasting," *IEEE Trans. Signal Process.*, vol. 54, no. 6, pp. 2239–2251, June 2006.

[12] O. Mehanna, N. D. Sidiropoulos, and G. B. Giannakis, "Joint multicast beamforming and antenna selection," *IEEE Trans. Signal Process.*, vol. 61, no. 10, pp. 2660–2674, May 2013.

[13] Y. F. Liu, Y. H. Dai, and Z. Q. Luo, "Max-min fairness linear transceiver design for a multi-user mimo interference channel," *IEEE Trans. Signal Process.*, vol. 61, no. 9, pp. 2413–2423, May 2013.

[14] P. Setlur and M. Rangaswamy, "Waveform design for radar stap in signal dependent interference," *IEEE Trans. Signal Process.*, vol. 64, no. 1, pp. 19–34, Jan 2016.

[15] G. Cui, H. Li, and M. Rangaswamy, "MIMO radar waveform design with constant modulus and similarity constraints," *IEEE Trans. Signal Process.*, vol. 62, no. 2, pp. 343–353, Jan 2014.

[16] Y. C. Wang, X. Wang, H. Liu, and Z. Q. Luo, "On the design of constant modulus probing signals for mimo radar," *IEEE Trans. Signal Process.*, vol. 60, no. 8, pp. 4432–4438, Aug 2012.

[17] W. Xu, Q. Shi, X. Wei, Z. Ma, X. Zhu, and Y. Wang, "Distributed optimal rate-reliability-lifetime tradeoff in time-varying wireless sensor networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 9, pp. 4836–4847, Sept 2014.

[18] W. C. Liao, M. Hong, H. Farmanbar, X. Li, Z. Q. Luo, and H. Zhang, "Min flow rate maximization for software defined radio access networks," *IEEE Journal on Selected Areas in Communications*, vol. 32, no. 6, pp. 1282–1294, June 2014.

[19] L. Liu and W. Yu, "Cross-layer design for downlink multi-hop cloud radio access networks with network coding," *IEEE Trans. Signal Process.*, vol. PP, no. 99, pp. 1–1, 2017.

[20] W. Xu, Y. Zhang, Q. Shi, and X. Wang, "Energy management and cross layer optimization for wireless sensor network powered by heterogeneous energy sources," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2814–2826, May 2015.

[21] Q. Shi and C. He, "A sdp approach for range-free localization in wireless sensor networks," in *IEEE ICC*, May 2008, pp. 4214–4218.

[22] Q. Shi, C. He, H. Chen, and L. Jiang, "Distributed wireless sensor network localization via sequential greedy optimization algorithm," *IEEE Trans. Signal Process.*, vol. 58, no. 6, pp. 3328–3340, June 2010.

[23] S. Ji, K.-F. Sze, Z. Zhou, A. M.-C. So, and Y. Ye, "Beyond convex relaxation: A polynomial-time non-convex optimization approach to network localization." in *IEEE INFOCOM*, 2013, pp. 2499–2507.

[24] D. Lecompte and F. Gabin, "Evolved multimedia broadcast/multicast service (embms) in LTE-advanced: overview and Rel-11 enhancements," *IEEE Commun. Mag.*, vol. 50, no. 11, pp. 68–74, Nov. 2012.

[25] Y. Yang, H. Hu, J. Xu, and G. Mao, "Relay technologies for WiMax and LTE-advanced mobile systems," *IEEE Commun. Mag.*, vol. 47, no. 10, pp. 100–105, Oct. 2009.

[26] Z. He, S. Guo, Y. Ou, and Y. Rong, "Multiuser multihop mimo relay system design based on mutual information maximization," *IEEE Trans. Signal Process.*, vol. 62, no. 21, pp. 5725–5733, Nov 2014.

[27] C. Xing, S. Ma, Z. Fei, Y. C. Wu, and H. V. Poor, "A general robust linear transceiver design for multi-hop amplify-and-forward mimo relaying systems," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1196–1209, March 2013.

[28] R. Wang and M. Tao, "Joint source and relay precoding designs for mimo two-way relaying based on mse criterion," *IEEE Trans. Signal Process.*, vol. 60, no. 3, pp. 1352–1365, Mar. 2012.

[29] C. Sun, C. Yang, Y. Li, and B. Vucetic, "Transceiver design for multi-user multi-antenna two-way relay cellular systems," *IEEE Trans. Commun.*, vol. 60, no. 10, pp. 2893–2903, Oct. 2012.

[30] K. T. Truong, P. Sartori, and R. W. Heath, "Cooperative algorithms for mimo amplify-and-forward relay networks," *IEEE Trans. Signal Process.*, vol. 61, no. 5, pp. 1272–1287, March 2013.

[31] K. X. Nguyen, Y. Rong, and S. Nordholm, "Mmse-based transceiver design algorithms for interference mimo relay systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6414–6424, Nov 2015.

[32] R. Zhang, C. C. Chai, and Y. C. Liang, "Joint beamforming and power control for multiantenna relay broadcast channel with QoS constraints," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 726–737, Feb 2009.

[33] H. Wan, W. Chen, and X. Wang, "Joint source and relay design for mimo relaying broadcast channels," *IEEE Commun. Lett.*, vol. 17, no. 2, pp. 345–348, Feb. 2013.

[34] M. R. A. Khandaker and Y. Rong, "Joint transceiver optimization for multiuser mimo relay communication systems," *IEEE Trans. Signal Process.*, vol. 60, no. 11, pp. 5977–5986, Nov 2012.

[35] W. K. Ma, J. M. Bioucas-Dias, T. H. Chan, N. Gillis, P. Gader, A. J. Plaza, A. Ambikapathi, and C. Y. Chi, "A signal processing perspective on hyperspectral unmixing: Insights from remote sensing," *IEEE Signal Processing Magazine*, vol. 31, no. 1, pp. 67–81, Jan 2014.

[36] D. D. Lee and H. S. Seung, "Learning the parts of objects by nonnegative matrix factorization," *Nature*, vol. 401, pp. 788–791, 1999.

[37] X. Fu, N. D. Sidiropoulos, and W. K. Ma, "Power spectra separation via structured matrix factorization," *IEEE Trans. Signal Process.*, vol. 64, no. 17, pp. 4592–4605, Sept 2016.

[38] A. D. Maio, Y. Huang, and M. Piezzo, "A doppler robust max-min approach to radar code design," *IEEE Trans. Signal Process.*, vol. 58, no. 9, pp. 4943–4947, Sept 2010.

[39] M. M. Naghsh, M. Soltanalian, P. Stoica, M. Modarres-Hashemi, A. D. Maio, and A. Aubry, "A doppler robust design of transmit sequence and receive filter in the presence of signal-dependent interference," *IEEE Trans. Signal Process.*, vol. 62, no. 4, pp. 772–785, Feb 2014.

[40] W. Wang, S. Jin, and F. C. Zheng, "Maximin snr beamforming strategies for two-way relay channels," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1006–1009, July 2012.

[41] S. Kim, A. Magnani, and S. P. Boyd, "Robust fisher discriminant analysis," in *NIPS*, Dec 2005, pp. 659–666.

[42] M. Razaviyayn, M. Hong, and Z.-Q. Luo, "A unified convergence analysis of block successive minimization methods for nonsmooth optimization," *SIAM Journal on Optimization*, vol. 23, no. 2, pp. 1126–1153, 2013.

[43] M. Soltanalian, A. Gharanjik, M. R. B. Shankar, and B. Ottersten, "Grab-n-Pull: An optimization framework for fairness-achieving networks," in *IEEE ICASSP*, March 2016, pp. 3301–3305.

[44] Q. Shi, M. Hong, X. Gao, E. Song, Y. Cai, and W. Xu, "Joint source-relay design for full–duplex mimo af relay systems," *IEEE Trans. Signal Process.*, vol. 62, no. 23, Dec. 2016.

[45] J. F. Sturm, "Using SeDuMi 1.02, a matlab toolbox for optimization over symmetric cones," 1998.

[46] 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," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, Sept 2011.

[47] G. Golub, S. Nash, and C. V. Loan, "A hessenberg-schur method for the problem AX + XB= C," *IEEE Trans. Auto. Contr.*, vol. 24, no. 6, pp. 909–913, Dec 1979.

[48] X. Fu, K. Huang, B. Yang, W. K. Ma, and N. D. Sidiropoulos, "Robust volume minimization-based matrix factorization for remote sensing and document clustering," *IEEE Trans. Signal Process.*, vol. 64, no. 23, pp. 6254–6268, Dec 2016.

[49] C. H. Lin, W. K. Ma, W. C. Li, C. Y. Chi, and A. Ambikapathi, "Identifiability of the simplex volume minimization criterion for blind hyperspectral unmixing: The no-pure-pixel case," *IEEE Transactions on Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5530–5546, Oct 2015.

[50] W. Wang and M. Á. Carreira-Perpiñán, "Projection onto the probability simplex: An efficient algorithm with a simple proof, and an application," *CoRR*, vol. abs/1309.1541, 2013. [Online]. Available: http://arxiv.org/abs/1309.1541

[51] J. M. P. Nascimento and J. M. B. Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan 2005.