Cross-Layer Optimization of MIMO-Based Mesh Networks Under Orthogonal Channels

Jia Liu*, Tae Yoon Park*, Y. Thomas Hou*, Yi Shi*, and Hanif D. Sherali†
* Bradley Department of Electrical and Computer Engineering
† Grado Department of Industrial and Systems Engineering
Virginia Polytechnic Institute and State University, Blacksburg, VA 24061

***Abstract*— MIMO-based systems have great potential to improve network capacity for wireless mesh networks (WMNs). Due to unique physical layer characteristics associated with MIMO systems, network performance is tightly coupled with mechanisms at physical layer and link layer. So far, research on MIMO-based WMNs is still in its infancy and little results are available in this important area. In this paper, we consider the problem of jointly optimizing power and bandwidth allocation at each node and multihop/multipath routing in a MIMO-based WMN where links operate in orthogonal channels. To solve this problem, we develop a mathematical solution procedure, which combines Lagrangian dual decomposition, gradient projection, and cutting-plane methods. We provide theoretical insights in deriving gradient projection and cutting plane methods. We also use simulations to verify the efficacy of our algorithm.**

I. Introduction

Since Winters's [1], Telatar's [2] and Foschini's [3] pioneering works predicting the potential of high spectral efficiency provided by multiple antenna systems, the last decade has witnessed the soar of research activity on Multiple-Input Multiple-Output (MIMO) technologies. Without costs of extra spectrum, MIMO technology, which exploits the rich scattering characteristic of wireless channels, is able to increase channel capacities substantially than conventional communication systems.

However, compared to the research on the capacity of single-user MIMO, for which many results are available (see [4] and [5] and references therein), the capacity issue of *multiuser* MIMO systems is much less studied and many fundamental problems remain unsolved [5]. With the emergence of wireless mesh networks (WMNs), which are multiuser and multihop in nature, the need to extend the MIMO communication concept from single-user systems to multiuser systems has become increasingly compelling. In a WMN, however, applying MIMO technique becomes far more complicated. Power control and power allocation at each link as well as multihop/multipath routing across the network all interact with one and another, and cross-layer optimization is not only desirable, but also necessary.

In this paper, we study the problem of cross-layer optimization on multihop/multi-path routing, power control, power allocation, and bandwidth allocation (CRPBA) for MIMO-based mesh network where links operate in orthogonal channels. Specifically, we consider how to support a set of user communication sessions by jointly optimizing power control, power allocation, bandwidth allocation, and flow routing such that some network utility (e.g., proportional fairness) is maximized. This problem, to the best of the authors' knowledge, has not been studied thus far.

A. Main Contribution

The main contribution of this paper are the following:

1) We developed a mathematical solution procedure to solve CRPBA by combining Lagrangian decomposition, gradient projection, and cutting-plane algorithms.
2) For the challenging link layer subproblem, we develop a rigorous gradient projection method as opposed to the heuristic one in [6].
3) Our proposed cutting-plane method can not only solve the Lagrangian dual, but also easily recover the optimal primal feasible solutions, thus circumventing a major difficulty of the popular subgradient-based approaches for solving Lagrangian dual problems.

B. Paper Organization

The remainder of this paper is organized as follows. In Section II, we review related work. In Section III, we discuss the network model and problem formulation. Section IV introduces the key components in our solution procedure, including gradient projection, cutting-plane algorithm, and the recovering of optimal primal feasible solutions. Numerical results are presented in Section VI. Section VII concludes this paper.

II. Related Work

Research on applying MIMO to WMN is still in its infancy and results remain extremely limited. In this section, we provide a synopsis of related work on the MIMO research evolution from single-hop ad hoc networks to mesh networks. For single-hop MIMO ad hoc networks, in [6], Ye and Blum introduced a gradient projection method to find a suboptimal solution for the nonconvex optimization problem for single-hop ad hoc networks. However, the way they handled gradient projection is based on heuristic: in solving the constrained Lagrangian dual problem in projection, they simply set the first derivative to zero to get the solution. Such a method does not work for general constrained optimization problem. For multihop WMNs, Hu and Zhang [7] examined the problem of joint medium access control and routing, and in particular, considered the optimal hop distance to minimize the end-to-end delay. However, power control and power allocation were not considered in this work. In [8], the authors designed different routing protocols for WMN to explore the tradeoff between multiplexing gain and diversity gain [9]. However, this work is largely based on simulation observations.

III. Network Model

We first introduce notation for matrices, vectors, and complex scalars in this paper. We use boldface to denote matrices and vectors. For a matrix $\mathbf{A}$, $\mathbf{A}^{\dagger}$ denotes the conjugate transpose. $\mathrm{Tr}\{\mathbf{A}\}$ denotes the trace of $\mathbf{A}$. $\mathbf{Diag}\begin{bmatrix} \mathbf{A}_1 & \ldots & \mathbf{A}_n \end{bmatrix}$ represents the block diagonal matrix with matrices $\mathbf{A}_1, \ldots, \mathbf{A}_n$ on its main diagonal. We let $\mathbf{I}$ denote the identity matrix with dimension determined from context. $\mathbf{A} \succeq 0$ represents that $\mathbf{A}$ is Hermitian and positive semidefinite (PSD). $\mathbf{1}$ and $\mathbf{0}$ denote vectors whose elements are all ones and zeros, respectively, and their dimensions are determined from context. $(\mathbf{v})_m$ represents the m^{th} entry of vector $\mathbf{v}$. For a real vector $\mathbf{v}$ and a real matrix $\mathbf{A}$, $\mathbf{v} \geq \mathbf{0}$ and $\mathbf{A} \geq \mathbf{0}$ mean that all entries in $\mathbf{v}$ and $\mathbf{A}$ are nonnegative, respectively. We let $\mathbf{e}_i$ be the unit column vector where the i^{th} entry is 1 and all other entries are 0. The dimension of $\mathbf{e}_i$ is determined from context as well. The operator "$\langle , \rangle$" represents vector or matrix inner product operation.

A. Link Capacity Model

In this paper, it is assumed that the system has perfect channel knowledge, that is, the transmitters have perfect channel state information (CSI). Let the matrix $\mathbf{H}_l \in \mathbb{C}^{n_r \times n_t}$ represent the wireless channel gain matrix from the transmitting node to the receiving node of link l, where n_t and n_r are the numbers of transmitting and receiving antenna elements at each node, respectively. Although wireless channels in reality are time-varying, we consider a "constant" channel model in this paper, i.e., $\mathbf{H}_l$'s coherence time is larger than the transmission period we consider. This simplification is of much interest for the insight it provides and its application in finding the ergodic capacity for block-wise fading channels [5]. The received complex base-band signal vector for MIMO link l with n_t transmitting antennas and n_r receiving antennas in a Gaussian channel is given by

$$\mathbf{r}_l = \sqrt{\rho_l}\mathbf{H}_l\mathbf{t}_l + \mathbf{n}_l, \tag{1}$$

where $\mathbf{r}_l$ and $\mathbf{t}_l$ represent the received and transmitted signal vectors, $\mathbf{n}_l$ is the normalized additive white Gaussian noise vector, ρ_j captures path-loss effect. By adopting the path-loss model with path-loss exponent being equal to α, ρ_l can be computed as [10] $\rho_l = (\frac{G_t G_r \lambda^2}{(4\pi)^2})/(N_0 W D_l^{\alpha})$, where D_l denotes the length of link l, N_0 represents the power spectral density of white Gaussian noise, and W denotes the communication bandwidth, G_t and G_r are transmit and receive antenna gains, respectively, which are assumed to be 1 in this paper, λ is the wavelength of the transmitted signal.

Let matrix $\mathbf{Q}_l$ represent the covariance matrix of a zero-mean Gaussian input symbol vector $\mathbf{t}_l$ at link l, i.e., $\mathbf{Q}_l = \mathbb{E}\left\{\mathbf{t}_l \cdot \mathbf{t}_l^{\dagger}\right\}$. The definition of $\mathbf{Q}_l$ implies that it is Hermitian and $\mathbf{Q}_l \succeq 0$. Physically, $\mathbf{Q}_l$ represents the power allocation in different antenna elements in link l's transmitter and correlations between each of these elements. In this paper, we use the complex matrix $\mathbf{Q} \triangleq \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 & \ldots & \mathbf{Q}_L \end{bmatrix} \in \mathbb{C}^{n_t \times (n_t \cdot L)}$ to denote the collection of all input covariance matrices. The link capacity of a MIMO link l in an AWGN channel can be written as

$$\Phi_l(W_l, \mathbf{Q}_l) \triangleq W_l \log_2 \det\left(\mathbf{I} + \rho_l \mathbf{H}_l \mathbf{Q}_l \mathbf{H}_l^{\dagger}\right), \tag{2}$$

where W_l represents the communication bandwidth of link l. It can be readily verified that $\Phi_l(\mathbf{Q}_l)$ is a monotone increasing concave function in W_l and $\mathbf{Q}_l$.

B. Data Routing and Network Flows

One of the most challenging aspects of a WMN is that its connectivity is highly dependent on the transmission power at each node. As a result, the network topology is not fixed. Our approach to handle this elusive network topology is to denote the network by a directed graph consisting of N nodes and L *possible* MIMO links. By saying that a link is possible we mean the distance between its transmitting node and receiving node is less than or equal to the transmission range by using the maximum transmission power. We assume that the graph is always connected. The network topology can be represented by a *node-arc incidence matrix* (NAIM) [11] $\mathbf{A} \in \mathbb{R}^{N \times L}$, whose entry a_{nl} associating with node n and link l is defined as

$$a_{nl} = \begin{cases} 1 & \text{if } n \text{ is the transmitting node of link } l \\ -1 & \text{if } n \text{ is the receiving node of link } l \\ 0 & \text{otherwise.} \end{cases}$$

We define $\mathcal{O}(n)$ and $\mathcal{I}(n)$ as the sets of links that are outgoing from and incoming to node n, respectively. We use a multicommodity flow model for the routing of data packets across the WMN.

Suppose that there are F sessions in total in the network, representing F different commodities. The source and destination nodes of session f, $1 \leq f \leq F$, are denoted as $\mathrm{src}(f)$ and $\mathrm{dst}(f)$, respectively. For each session, we define a *source-destination vector* $\mathbf{s}_f \in \mathbb{R}^N$, whose entries, other than at the positions of $\mathrm{src}(f)$ and $\mathrm{dst}(f)$, are all zeros. In addition, from the flow conservation law, we must have $(\mathbf{s}_f)_{\mathrm{src}(f)} = -(\mathbf{s}_f)_{\mathrm{dst}(f)}$. Without loss of generality, we let$(\mathbf{s}_f)_{\mathrm{src}(f)} \geq 0$ and simply denote it by a scalar s_f. Therefore, we can further write the source-destination vector of session f as

$$\mathbf{s}_f = s_f \begin{bmatrix} \cdots & 1 & \cdots & -1 & \cdots \end{bmatrix}^T, \tag{3}$$

where the dots represent zeros, and 1 and -1 are in the positions of $\mathrm{src}(f)$ and $\mathrm{dst}(f)$, respectively[1]. Using the notation "$=_{x,y}$" to represent the component-wise equality of a vector except at the x^{th} and the y^{th} entries, we have $\mathbf{s}_f =_{\mathrm{src}(f),\mathrm{dst}(f)} \mathbf{0}$. Using matrix $\mathbf{S} \triangleq \begin{bmatrix} \mathbf{s}_1 & \mathbf{s}_2 & \ldots & \mathbf{s}_F \end{bmatrix} \in \mathbb{R}^{N \times F}$ to

[1]Note that for the source-destination vector of a flow f, 1 does not necessarily appear before -1 as in (3), which is only for illustrative purpose.

denote the collection of all source-sink vectors $\mathbf{s}_f$, we further have

$$\mathbf{S}\mathbf{e}_f =_{\mathrm{src}(f),\mathrm{dst}(f)} \mathbf{0}, \quad 1 \leq f \leq F, \tag{4}$$

$$\langle \mathbf{1}, \mathbf{S}\mathbf{e}_f \rangle = 0, \qquad 1 \leq f \leq F, \tag{5}$$

$$(\mathbf{S}\mathbf{e}_f)_{\mathrm{src}(f)} = s_f, \quad 1 \leq f \leq F, \tag{6}$$

where $\mathbf{e}_f$ is the f^{th} unit column vector.

On each link l, we let $x_l^{(f)} \geq 0$ be the amount of flow of session f. We define $\mathbf{x}^{(f)} \in \mathbb{R}^L$ as the *flow vector* for session f. At each node n, components of the flow vector and source-destination vector for the same session satisfy the flow conservation law:

$$\sum_{l \in \mathcal{O}(n)} x_l^{(f)} - \sum_{l \in \mathcal{I}(n)} x_l^{(f)} = (\mathbf{s}_f)_n, \quad 1 \leq n \leq N, \ 1 \leq f \leq F.$$

With NAIM, the flow conservation law across the whole network can be compactly written as $\mathbf{A}\mathbf{x}^{(f)} = \mathbf{s}_f, \quad 1 \leq f \leq F$. We use matrix $\mathbf{X} \triangleq \begin{bmatrix} \mathbf{x}^{(1)} & \mathbf{x}^{(2)} & \ldots & \mathbf{x}^{(F)} \end{bmatrix} \in \mathbb{R}^{L \times F}$ to denote the collection of flow vectors $\mathbf{x}^{(f)}$. With $\mathbf{X}$ and $\mathbf{S}$, the flow conservation law can be further compactly written as

$$\mathbf{A}\mathbf{X} = \mathbf{S}. \tag{7}$$

C. Problem Formulation

The goal of this paper is to design an algorithm that performs the cross-layer optimization on multihop/multipath routing, power control, power allocation, and bandwidth allocation (CRPBA) for a MIMO-based WMN. We consider an FDMA MIMO-based WMN, where each node has been assigned non-overlapping (possibly reused) frequency bands for its incoming and outgoing links so that each node can simultaneously transmit and receive, and cause no interference to other nodes. How to perform channel assignments is a huge research topic by itself, and there are a vast amount of literature that discuss channel assignment problems. Thus in this paper, we focus on how to jointly optimize routing in the network layer and power control/allocation as well as bandwidth allocation in the link layer.

We adopt the well-known proportional fairness utility function, i.e., $\ln(s_f)$ for flow f [12]. We wish to perform the cross-layer optimization such that the sum of all utilities of flows is maximized. Since the network flows in a link cannot exceed the link's capacity limit, we have $\sum_{f=1}^{F} x_l^{(f)} \leq \Phi_l(W_l, \mathbf{Q}_l)$, for $1 \leq l \leq L$. Using matrix-vector notations, it can be further compactly written as

$$\langle \mathbf{1}, \mathbf{X}^T \mathbf{e}_l \rangle \leq \Phi_l(W_l, \mathbf{Q}_l), \quad 1 \leq l \leq L. \tag{8}$$

In the link layer, since the total transmit power of each node is subject to a maximum power constraint, we have $\sum_{l \in \mathcal{O}(n)} \mathrm{Tr}\{\mathbf{Q}_l\} \leq P_{\max}^{(n)}$, $1 \leq n \leq N$, where $P_{\max}^{(n)}$ represents the maximum transmit power of node n. Also, the sum of bandwidths of all the outgoing links for a node n cannot exceed the assigned bandwidth for node n, i.e., $\sum_{l \in \mathcal{O}(n)} W_l \leq B_n$, $1 \leq n \leq N$, where B_n is the assigned transmission band for node n. We use the matrix $\mathbf{W} = \begin{bmatrix} W_1 & W_2 & \ldots & W_L \end{bmatrix}^T \in \mathbb{R}^{L \times 1}$ to denote the collection of the bands for links from 1 to L. Coupling the MIMO link capacity model in Section III-A and the network flow model in Section III-B, we have the problem formulation for CRPBA as in (9).

$$\begin{array}{lll} \textbf{CRPBA}: & \text{Maximize} & \sum_{f=1}^{F} \ln(s_f) \\ & \text{subject to} & \mathbf{A}\mathbf{X} = \mathbf{S} \\ & & \mathbf{X} \geq \mathbf{0} \\ & & \langle \mathbf{1}, \mathbf{X}^T \mathbf{e}_l \rangle \leq \Phi_l(W_l, \mathbf{Q}_l) \quad \forall\, l \\ & & \mathbf{S}\mathbf{e}_f =_{\mathrm{src}(f),\mathrm{dst}(f)} \mathbf{0} \quad \forall\, f \\ & & \langle \mathbf{1}, \mathbf{S}\mathbf{e}_f \rangle = 0 \quad \forall\, f \\ & & (\mathbf{S}\mathbf{e}_f)_{\mathrm{src}(f)} = s_f \quad \forall\, f \\ & & \sum_{l \in \mathcal{O}(n)} \mathrm{Tr}\{\mathbf{Q}_l\} \leq P_{\max}^{(n)} \quad \forall\, n \\ & & \mathbf{Q}_l \succeq 0 \quad \forall\, l \\ & & \sum_{l \in \mathcal{O}(n)} W_l \leq B_n \quad \forall\, n \\ & \text{Variables:} & \mathbf{S}, \mathbf{X}, \mathbf{Q}, \mathbf{W} \end{array} \tag{9}$$

IV. Solution Procedure

It can be observed that the CRPBA possesses a special structure: the network layer variables and the link layer variables are coupled through the link capacity constraints $\langle \mathbf{1}, \mathbf{X}^T \mathbf{e}_l \rangle \leq \Phi_l(W_l, \mathbf{Q}_l)$. Thus, we can exploit this special structure using Lagrangian dual decomposition to solve CRPBA efficiently. In [13], the authors used a similar decomposition technique to solve simultaneous routing and resource allocation problems. However, their routing setting was very different to our work and their link layer was not MIMO-based. Due to the spatial dimension resulted from MIMO, the link layer subproblem in this paper is completely different and substantially more challenging. Generally, given a nonlinear programming problem, several different Lagrangian dual problems can be constructed depending on which constraints are associated with Lagrangian dual variables [14]. For CRPBA, we associate Lagrangian multipliers u_i to the link capacity coupling constraints $\langle \mathbf{1}, \mathbf{X}^T \mathbf{e}_l \rangle \leq \Phi_l(W_l, \mathbf{Q}_l)$. Hence, the Lagrangian can be written as [14]

$$\Theta(\mathbf{u}) = \sup_{\mathbf{S},\mathbf{X},\mathbf{Q},\mathbf{W}} \left\{ L(\mathbf{S}, \mathbf{X}, \mathbf{Q}, \mathbf{W}, \mathbf{u}) | (\mathbf{S}, \mathbf{X}, \mathbf{Q}, \mathbf{W}) \in \Gamma \right\},$$

where

$$L(\mathbf{S}, \mathbf{X}, \mathbf{Q}, \mathbf{W}, \mathbf{u}) = \sum_f \ln(s_f) + \sum_l u_l \big(\Phi_l(W_l, \mathbf{Q}_l) - \langle \mathbf{1}, \mathbf{X}^T \mathbf{e}_l \rangle \big) \tag{10}$$

and Γ is defined as

$$\Gamma \triangleq \left\{ (\mathbf{S}, \mathbf{X}, \mathbf{Q}, \mathbf{W}) \;\middle|\; \begin{array}{ll} \mathbf{A}\mathbf{X} = \mathbf{S} & \\ \mathbf{X} \geq \mathbf{0} & \\ \mathbf{S}\mathbf{e}_f =_{\mathrm{src}(f),\mathrm{dst}(f)} \mathbf{0} & \forall\, f \\ \langle \mathbf{1}, \mathbf{S}\mathbf{e}_f \rangle = 0 & \forall\, f \\ (\mathbf{S}\mathbf{e}_f)_{\mathrm{src}(f)} = s_f & \forall\, f \\ \sum_{l \in \mathcal{O}(n)} \mathrm{Tr}\{\mathbf{Q}_l\} \leq P_{\max}^{(n)} & \forall\, n \\ \mathbf{Q}_l \succeq 0 & \forall\, l \\ \sum_{l \in \mathcal{O}(n)} W_l \leq B_n & \forall\, n \end{array} \right\}$$

The Lagrangian dual problem of CRPBA can thus be written as [14]:

$$\begin{array}{lll} \mathbf{D}^{\mathrm{CRPBA}}: & \text{Minimize} & \Theta(\mathbf{u}) \\ & \text{subject to} & \mathbf{u} \geq \mathbf{0}. \end{array}$$

It is easy to recognize that, for any given Lagrangian multiplier $\mathbf{u}$, the Lagrangian in (10) can be separated into two terms:

$$\Theta(\mathbf{u}) = \Theta_{\rm net}(\mathbf{u}) + \Theta_{\rm link}(\mathbf{u}),$$

where $\Theta_{\rm net}$ and $\Theta_{\rm link}$ are two subproblems respectively corresponding to network layer and link layer:

$$\begin{array}{rll} \mathbf{D}_{\rm net}^{\rm CRPBA} : \Theta_{\rm net}(\mathbf{u}) \triangleq & \text{Maximize} \sum_f \ln(s_f) & \\ & \quad - \sum_l u_l \langle \mathbf{1}, \mathbf{X}^T \mathbf{e}_l \rangle & \\ \text{subject to} & \mathbf{A}\mathbf{X} = \mathbf{S} & \\ & \mathbf{X} \geq \mathbf{0} & \\ & \mathbf{S}\mathbf{e}_f =_{\mathrm{src}(f),\mathrm{dst}(f)} \mathbf{0} & \forall f \\ & \langle \mathbf{1}, \mathbf{S}\mathbf{e}_f \rangle = 0 & \forall f \\ & (\mathbf{S}\mathbf{e}_f)_{\mathrm{src}(f)} = s_f & \forall f \\ \text{Variables:} & \mathbf{S}, \mathbf{X} & \end{array}$$

$$\begin{array}{rll} \mathbf{D}_{\rm link}^{\rm CRPBA} : \Theta_{\rm link}(\mathbf{u}) \triangleq & \text{Maximize} \sum_l u_l \Phi_l(W_l, \mathbf{Q}_l) & \\ \text{subject to} & \sum_{l \in \mathcal{O}(n)} \mathrm{Tr}\{\mathbf{Q}_l\} \leq P_{\max}^{(n)} & \forall n \\ & \sum_{l \in \mathcal{O}(n)} W_l \leq B_n & \forall n \\ & \mathbf{Q}_l \succeq 0 & \forall l \\ \text{Variables:} & \mathbf{Q}, \mathbf{W} & \end{array}$$

The CRPBA Lagrangian dual problem can be thus transformed into the following *master* dual problem:

$$\begin{array}{rll} \mathbf{MD}^{\rm CRPBA}: & \text{Minimize} & \Theta_{\rm net}(\mathbf{u}) + \Theta_{\rm link}(\mathbf{u}) \\ & \text{subject to} & \mathbf{u} \geq \mathbf{0} \end{array}$$

Now, the task of solving the decomposed Lagrangian dual problem boils down to how to evaluate the subproblems $\mathbf{D}_{\rm net}^{\rm CRPBA}$ and $\mathbf{D}_{\rm link}^{\rm CRPBA}$, and how to handle the master problem. Note that in the network layer subproblem $\mathbf{D}_{\rm net}^{\rm CRPBA}$, the objective function is concave and all constraints are affine. Therefore, $\mathbf{D}_{\rm net}^{\rm CRPBA}$ is readily solvable by using many polynomial time convex programming methods. However, solving $\mathbf{D}_{\rm link}^{\rm CRPBA}$ is not trivial because the objective function and constraints involve many complex matrices variables, even though it can be shown that $\mathbf{D}_{\rm link}^{\rm CRPBA}$ is a convex problem. In the following subsections, we will discuss each techniques we use to solve the link layer subproblem and the master problem in detail.

A. Modified Gradient Projection Method (MGP)

In this paper, we propose a modified "gradient projection" (MGP) method to solve the link subproblem. Gradient projection, originally proposed by Rosen [15], is a classical nonlinear programming method aiming at solving constrained optimization problems. But its formal convergence proof has not been established until very recently [14]. The framework of MGP is shown in Algorithm 1. Due to the complexity of the objective function, we cannot afford the luxury of performing an exact line search which requires the expense of excessive objective function evaluations. Therefore, we adopt the "Armijo rule" inexact line search method [14], which still enjoys provable convergence. The basic idea of Armijo rule is that at each step of the line search, we sacrifice accuracy for efficiency as long as we have sufficient improvement. According to Armijo rule, we choose $s_k = 1$ and $\alpha_k = \beta^{m_k}$

Algorithm 1 Modified Gradient Projection Method

Initialization:
Choose the initial conditions $\mathbf{W}^{(0)} = [W_1^{(0)}, W_2^{(0)}, \ldots, W_L^{(0)}]^T$, $\mathbf{Q}^{(0)} = [\mathbf{Q}_1^{(0)}, \mathbf{Q}_2^{(0)}, \ldots, \mathbf{Q}_L^{(0)}]^T$. Let $k = 0$.

Main Loop:

1. Calculate the gradients $G_{W_l}^{(k)} = \nabla_{W_l} \Theta_{\rm link}(\mathbf{u}, \mathbf{W}^{(k)}, \mathbf{Q}^{(k)})$ and $\mathbf{G}_{\mathbf{Q}_l}^{(k)} = \nabla_{\mathbf{Q}_l} \Theta_{\rm link}(\mathbf{u}, \mathbf{W}^{(k)}, \mathbf{Q}^{(k)})$, for $l = 1, 2, \ldots, L$.
2. Choose an appropriate step size s_k. Let $W_l^{(k)'} = W_l^{(k)} + s_k G_{W_l}^{(k)}$, $\mathbf{Q}_l^{(k)'} = \mathbf{Q}_l^{(k)} + s_k \mathbf{G}_{\mathbf{Q}_l}^{(k)}$, for $l = 1, 2, \ldots, L$.
3. Let $[\bar{\mathbf{W}}_n^{(k)}, \bar{\mathbf{Q}}_n^{(k)}]^T$ be the projection of $[\mathbf{W}_n^{(k)'}, \mathbf{Q}_n^{(k)'}]^T$ onto $\Omega_+(n)$, where $\Omega_+(n) \triangleq \{(W_l, \mathbf{Q}_l) | l \in \mathcal{O}(n), W_l \geq 0, \mathbf{Q}_l \succeq 0, \sum_{l \in \mathcal{O}(n)} W_l \leq B_n, \sum_{l \in \mathcal{O}(n)} \mathrm{Tr}\{\mathbf{Q}_l\} \leq P_{\max}^{(n)}\}$.
4. Choose appropriate step size α_k. Let $W_l^{(k+1)} = W_l^{(k)} + \alpha_k (\bar{W}_l^{(k)} - W_l^{(k)})$, $\mathbf{Q}_l^{(k+1)} = \mathbf{Q}_l^{(k)} + \alpha_k (\bar{\mathbf{Q}}_l^{(k)} - \mathbf{Q}_l(k))$, $l = 1, 2, \ldots, L$.
5. $k = k+1$. If the maximum absolute value of the elements in $\mathbf{Q}_l^{(k)} - \mathbf{Q}_l^{(k-1)} < \epsilon$ and $W_l^{(k)} - W_l^{(k-1)} < \epsilon$, for $l = 1, 2, \ldots, L$, then stop; else go to step 1.

(the same as in [6]), where m_k is the first non-negative that satisfies

$$\begin{aligned} &\Theta_{\rm link}(\mathbf{Q}^{(k+1)}) - \Theta_{\rm link}(\mathbf{Q}^{(k)}) \\ &\geq \sigma \beta^{m_k} \sum_{l=1}^{L} \mathrm{Tr}\left[\nabla_{\mathbf{Q}_l} \Theta_{\rm link}(\mathbf{Q}^{(k)})^\dagger \left(\bar{\mathbf{Q}}_l^{(k)} - \mathbf{Q}_l^{(k)} \right) \right], \end{aligned}$$

where $0 < \beta < 1$ and $0 < \sigma < 1$ are fixed scalars. It is evident that the gradient $G_{W_l} \triangleq \nabla_{W_l} \Theta_{\rm link} = u_l \log_2 \det(\mathbf{I} + \rho_l \mathbf{H}_l \mathbf{Q}_l \mathbf{H}_l^\dagger)$. By using the formula $\frac{\partial}{\partial \mathbf{X}} \ln \det(\mathbf{A} + \mathbf{BXC}) = \left[\mathbf{C}(\mathbf{A} + \mathbf{BXC})^{-1} \mathbf{B} \right]^T$ [6], [16], where matrices $\mathbf{A} \in \mathbb{C}^{p \times p}$, $\mathbf{B} \in \mathbb{C}^{p \times m}$, $\mathbf{X} \in \mathbb{C}^{m \times n}$, $\mathbf{C} \in \mathbb{C}^{n \times p}$, and $(\mathbf{A} + \mathbf{BXC})$ is invertible , we are able to derive the gradient $\mathbf{G}_{\mathbf{Q}_l} \triangleq \nabla_{\mathbf{Q}_l} \Theta_{\rm link}$ as follows [17]:

$$\mathbf{G}_{\mathbf{Q}_l} = \frac{2 W_l u_l \rho_l}{\ln 2} \mathbf{H}_l^\dagger \left(\mathbf{I} + \rho_l \mathbf{H}_l \mathbf{Q}_l \mathbf{H}_l^\dagger \right)^{-1} \mathbf{H}_l. \tag{11}$$

Noting that $\mathbf{G}_{\mathbf{Q}_l}$ are Hermitian, we have that $\mathbf{Q}_l^{'}(k)$ is Hermitian as well. Then, for a node n having $|\mathcal{O}(n)|$ outgoing links, the projection problem becomes how to simultaneously project the $|\mathcal{O}(n)|$ W-scalars and $|\mathcal{O}(n)|$ $\mathbf{Q}$-covariance matrices onto $\Omega_+(n) \triangleq \{(W_l, \mathbf{Q}_l) | l \in \mathcal{O}(n), W_l \geq 0, \mathbf{Q}_l \succeq 0, \sum_{l \in \mathcal{O}(n)} W_l \leq B_n, \sum_{l \in \mathcal{O}(n)} \mathrm{Tr}\{\mathbf{Q}_l\} \leq P_{\max}^{(n)}\}$.

We construct a block diagonal matrix $\mathbf{D}_n$ as follows:

$$\mathbf{D}_n = \left[\begin{array}{c|c} \mathbf{W}_n & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{Q}_n \end{array} \right] \in \mathbb{C}^{|\mathcal{O}(n)|(n_t+1) \times |\mathcal{O}(n)|(n_t+1)}$$

where $\mathbf{W}_n \triangleq \mathbf{Diag}\left[W_l : l \in \mathcal{O}(n) \right] \in \mathbb{C}^{|\mathcal{O}(n)| \times |\mathcal{O}(n)|}$, and $\mathbf{Q}_n \triangleq \mathbf{Diag}\left[\mathbf{Q}_l : l \in \mathcal{O}(n) \right] \in \mathbb{C}^{|\mathcal{O}(n)| n_t \times |\mathcal{O}(n)| n_t}$. Moreover, we introduce two more matrices $\mathbf{E}_1^{(n)}$ and $\mathbf{E}_2^{(n)}$ as follows:

$$\mathbf{E}_1^{(n)} = \left[\begin{array}{c|c} \mathbf{I}_{|\mathcal{O}(n)|} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array} \right] \in \mathbb{C}^{|\mathcal{O}(n)|(n_t+1) \times |\mathcal{O}(n)|(n_t+1)},$$

$$\mathbf{E}_2^{(n)} = \left[\begin{array}{c|c} \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{I}_{|\mathcal{O}(n)| n_t} \end{array} \right] \in \mathbb{C}^{|\mathcal{O}(n)|(n_t+1) \times |\mathcal{O}(n)|(n_t+1)}.$$

It is easy to recognize that if $\mathbf{D}_n \in \Omega_+(n)$, we have $\mathrm{Tr}(\mathbf{E}_1^{(n)} \mathbf{D}_n) = \sum_{l \in \mathcal{O}(n)} W_l \leq B_n$, $\mathrm{Tr}(\mathbf{E}_2^{(n)} \mathbf{D}_n) =$

$\sum_{l\in\mathcal{O}(n)} \mathrm{Tr}\left(\mathbf{Q}_l\right) \leq P_{\max}^{(n)}$, and $\mathbf{D}_n \succeq 0$. In our projection, given a block diagonal matrix $\mathbf{D}_n$, we wish to find a matrix $\tilde{\mathbf{D}}_n \in \Omega_+(n)$ such that $\tilde{\mathbf{D}}_n$ minimizes $\|\tilde{\mathbf{D}}_n - \mathbf{D}_n\|_F$, where $\|\cdot\|_F$ denotes Frobenius norm. For more convenient algebraic manipulations, we instead study the following equivalent optimization problem:

$$\begin{array}{ll} \text{Minimize} & \frac{1}{2}\|\tilde{\mathbf{D}}_n - \mathbf{D}_n\|_F^2 \\ \text{subject to} & \mathrm{Tr}(\mathbf{E}_1^{(n)}\tilde{\mathbf{D}}_n) \leq B_n \\ & \mathrm{Tr}(\mathbf{E}_2^{(n)}\tilde{\mathbf{D}}_n) \leq P_{\max}^{(n)} \\ & \tilde{\mathbf{D}}_n \succeq 0 \end{array} \quad (12)$$

Notice that the problem is a convex minimization problem and we can solve this minimization problem by solving its Lagrangian dual. Associating Hermitian matrix $\mathbf{X}$ to the constraint $\tilde{\mathbf{D}}_n \succeq 0$, ν to the constraint $\mathrm{Tr}(\mathbf{E}_1^{(n)}\tilde{\mathbf{D}}_n) \leq B_n$, and μ to the constraint $\mathrm{Tr}(\mathbf{E}_2^{(n)}\tilde{\mathbf{D}}_n) \leq P_{\max}^{(n)}$, we can write the Lagrangian as

$$g(\mathbf{X},\nu,\mu) = \min_{\tilde{\mathbf{D}}_n} \Big\{ (1/2)\|\tilde{\mathbf{D}}_n - \mathbf{D}_n\|_F^2 - \mathrm{Tr}(\mathbf{X}^\dagger \tilde{\mathbf{D}}_n) + \nu\left(\mathrm{Tr}[\mathbf{E}_1^{(n)}\tilde{\mathbf{D}}_n] - B_n\right) + \mu\left(\mathrm{Tr}[\mathbf{E}_2^{(n)}\tilde{\mathbf{D}}_n] - P_{\max}^{(n)}\right)\Big\}. \quad (13)$$

Since $\tilde{\mathbf{D}}_n$ becomes unconstrained after removing its positive semidefinite constraint (correspondingly, adding a penalty term to the Lagrangian), we can compute the minimizer of (13) by simply setting the derivative of (13) to zero. Thus, we have $\tilde{\mathbf{D}}_n = \mathbf{D}_n + \mathbf{X} - \nu\mathbf{E}_1^{(n)} - \mu\mathbf{E}_2^{(n)}$. Substituting $\tilde{\mathbf{D}}_n$ back into (13), and after some algebraic simplifications, we can rewrite the Lagrangian dual problem as

$$\begin{array}{ll} \text{Maximize} & -\frac{1}{2}\|\mathbf{D}_n - \nu\mathbf{E}_1^{(n)} - \mu\mathbf{E}_2^{(n)} + \mathbf{X}\|_F^2 - \\ & \nu\mathbf{B}_n - \mu P_{\max}^{(n)} + \frac{1}{2}\|\mathbf{D}_n\|^2 \\ \text{subject to} & \mathbf{X} \succeq 0, \nu \geq 0, \mu \geq 0. \end{array} \quad (14)$$

Eq. (14) belongs the class of so-called *matrix nearness problems*, which are not easy to solve in general (see [18], [19] and references therein). However, based on the special structure in $\mathbf{E}_1^{(n)}$ and $\mathbf{E}_2^{(n)}$, we are able to design a polynomial time algorithm to solve (14) Due to space limitation, we only give the pseudo-codes in Algorithm 2 and Algorithm 3 and refer readers to [17] for more details.

Algorithm 2 Projection onto $\Omega_+(n)$

1. Construct a block diagonal matrix $\mathbf{D}_n$. Perform eigenvalue decomposition $\mathbf{D}_n = \mathbf{U}_n\mathbf{\Lambda}_n\mathbf{U}_n^\dagger$, separate the eigenvalues in two groups corresponding to $\mathbf{W}_n$ and $\mathbf{Q}_n$, and sort them in non-increasing order within each group, respectively.
2. For each group of eigenvalues, call Algorithm 3 to find the optimal dual variable ν^* and μ^*.
3. Compute $\tilde{\mathbf{D}}_n = \mathbf{U}_n(\mathbf{\Lambda}_n - \nu^*\mathbf{E}_1^{(n)} - \mu^*\mathbf{E}_2^{(n)})_+\mathbf{U}^\dagger$.

V. Cutting-Plane Method for Solving $\mathbf{D}^{\mathrm{CRPBA}}$

Compared to the popular subgradient-based approaches for solving Lagrangian dual problems, the attractive feature of the cutting-plane method is its speed of convergence and its simplicity in recovering optimal primal feasible solutions. As

Algorithm 3 Search the Optimal Dual Variable

Initiation:
Introduce $\lambda_0 = \infty$ and $\lambda_K = -\infty$. Let $\hat{I} = 0$. Let endpoint objective value $\psi_{\hat{I}}(\lambda_0) = 0$, $\phi^* = \psi_{\hat{I}}(\lambda_0)$, and $\mu^* = \lambda_0$.
Main Loop:
1. If $\hat{I} > K$, return μ^*; else let $\mu_{\hat{I}}^* = (\sum_{j=1}^{\hat{I}} \lambda_j - P)/\hat{I}$.
2. If $\mu_{\hat{I}}^* \in [\lambda_{\hat{I}+1}, \lambda_{\hat{I}}] \cap \mathbb{R}_+$, then let $\mu^* = \mu_{\hat{I}}^*$ and return μ^*.
3. Compute $\psi_{\hat{I}}(\lambda_{\hat{I}+1})$. If $\psi_{\hat{I}}(\lambda_{\hat{I}+1}) < \phi^*$, then return μ^*; else let $\mu^* = \lambda_{\hat{I}+1}$, $\phi^* = \psi_{\hat{I}}(\lambda_{\hat{I}+1})$, $\hat{I} = \hat{I} + 1$ and continue.

opposed to the cumbersomeness of subgradient method, in cutting-plane method, primal optimal feasible solutions can be exactly computed by averaging all the primal solutions (may or may not be primal feasible) using the dual variables as weights [17].

We briefly introduce the basic idea of cutting-plane method as follows. Letting $z = \Theta(\mathbf{u})$, the inequality $z \geq \sum_f \ln(s_f) + \sum_l u_l \left(\Phi_l(W_l, \mathbf{Q}_l) - \langle \mathbf{1}, \mathbf{X}^T\mathbf{e}_l\rangle\right)$ must hold for all $(\mathbf{S},\mathbf{X},\mathbf{Q},\mathbf{W}) \in \Gamma$. Thus, the dual problem is equivalent to

$$\begin{array}{ll} \text{Minimize} & z \\ \text{subject to} & z \geq \sum_f \ln(s_f) + \\ & \quad \sum_l u_l \left(\Phi_l(W_l,\mathbf{Q}_l) - \langle \mathbf{1}, \mathbf{X}^T\mathbf{e}_l\rangle\right) \\ & \mathbf{u} \geq 0, \end{array} \quad (15)$$

where $(\mathbf{S},\mathbf{X},\mathbf{Q},\mathbf{W}) \in \Gamma$. Although (15) is a linear program with infinite constraints not known explicitly, we can consider the following approximating problem:

$$\begin{array}{ll} \text{Minimize} & z \\ \text{subject to} & z \geq \sum_f \ln\left(s_f^{(j)}\right) + \\ & \quad \sum_l u_l \left(\Phi_l(W_l^{(j)},\mathbf{Q}_l^{(j)}) - \langle \mathbf{1}, \mathbf{X}^{(j)T}\mathbf{e}_l\rangle\right) \\ & \mathbf{u} \geq 0, \end{array} \quad (16)$$

where the points $(\mathbf{S}^{(j)},\mathbf{X}^{(j)}),\mathbf{Q}^{(j)},\mathbf{W}^{(j)}) \in \Gamma$, $j = 1,\ldots,k-1$. The problem in (16) is a linear program with a finite number of constraints and can be solved efficiently. Let $(z^{(k)},\mathbf{u}^{(k)})$ be an optimal solution to the approximating problem, which we refer to as the *master program*. If the solution is feasible to (15), then it is an optimal solution to the Lagrangian dual problem. To check the feasibility, we consider the following *subproblem*:

$$\begin{array}{ll} \text{Maximize} & \sum_f \ln(s_f) + \sum_l u_l^{(k)} \left(\Phi_l(W_l,\mathbf{Q}_l) - \langle \mathbf{1}, \mathbf{X}^T\mathbf{e}_l\rangle\right) \\ \text{subject to} & (\mathbf{S},\mathbf{X},\mathbf{Q},\mathbf{W}) \in \Gamma \end{array} \quad (17)$$

Suppose that $(\mathbf{S}^{(k)},\mathbf{X}^{(k)},\mathbf{Q}^{(k)},\mathbf{W}^{(k)})$ is an optimal solution to the subproblem (17) and $\Theta^*(\mathbf{u}^{(k)})$ is the corresponding optimal objective value. If $z_k \geq \Theta^*(\mathbf{u}^{(k)})$, then $\mathbf{u}^{(k)}$ is an optimal solution to the Lagrangian dual problem. Otherwise, for $\mathbf{u} = \mathbf{u}^{(k)}$, the inequality constraint in (15) is not satisfied for $(\mathbf{S}^{(j)},\mathbf{X}^{(j)},\mathbf{Q}^{(j)},\mathbf{W}^{(j)})$. Thus, we can add the constraint

$$z \geq \sum_f \ln\left(s_f^{(k)}\right) + \sum_l u_l \left(\Phi_l(W_l^{(k)},\mathbf{Q}_l^{(k)}) - \langle \mathbf{1}, \mathbf{X}^{(k)T}\mathbf{e}_l\rangle\right) \quad (18)$$

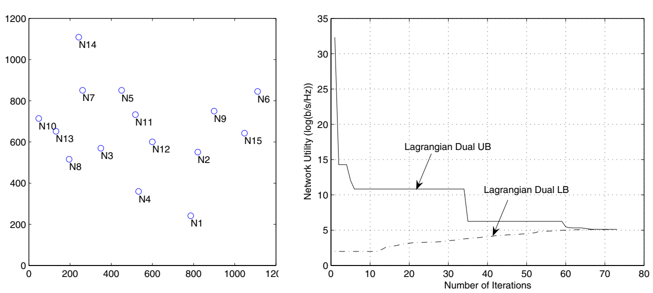
to (16), and re-solve the master linear program. Obviously, $(z^{(k)}, \mathbf{u}^{(k)})$ violates (18) and will be cut off by (18). The cutting plane algorithm is summarized in Algorithm 4.

Algorithm 4 Cutting Plane Algorithm for Solving $\mathbf{D}^{\text{CRPBA}}$

Initialization:
Find a point $(\mathbf{S}^{(0)}, \mathbf{X}^{(0)}, \mathbf{Q}^{(0)}, \mathbf{W}^{(0)}) \in \Gamma$. Let $k = 1$.
Main Loop:
1. Solve the master program in (16). Let $(z^{(k)}, \mathbf{u}^{(k)})$ be an optimal solution.
2. Solve the subproblem in (17). Let $(\mathbf{S}^{(k)}, \mathbf{X}^{(k)}, \mathbf{Q}^{(k)}, \mathbf{W}^{(k)})$ be an optimal point, and let $\Theta^*(\mathbf{u}^{(k)})$ be the corresponding optimal objective value.
3. If $z^{(k)} \geq \Theta(\mathbf{u}^{(k)})$, then stop with $\mathbf{u}^{(k)}$ as the optimal dual solution. Otherwise, add the constraint (18) to the master program, replace k by $k+1$, and go to step 1.

VI. Numerical Results

We present some numerical results through simulations to provide further insights on solving CRPBA. We use a 15-node network example, as shown in Fig. 1(a), to show the convergence process of the cutting-plane algorithm for solving $\mathbf{D}^{\text{CRPBA}}$. Each node in the network is equipped with two antennas and assigned a unit transmit bandwidth. In this example, there are three flows transmitting across the network: N14 to N1, N6 to N10, and N5 to N4, respectively. The convergence process is illustrated in Fig. 1(b). It is seen that the cutting-plane method only takes 72 iterations to converge.



(a) Network topology. (b) Convergence process.

Fig. 1. A simple example illustrating the need of scheduling to improve performance.

VII. Conclusion

In this paper, we investigated the problem of cross-layer optimization of routing, power control, power allocation, and bandwidth allocation for MIMO-based mesh networks. We developed a mathematical solution procedure, which combines Lagrangian decomposition, gradient projection, and cutting-plane methods. We provided the theoretical insights of our proposed algorithms and conducted simulations to verify their efficacy. Our results show that the nice decoupled structure and the high efficiency of our proposed algorithm make it an attractive method for optimizing the performance of MIMO-based mesh networks.

Acknowledgement

The work of Y.T. Hou, J. Liu, and Y. Shi has been supported in part by NSF Grant CNS-0347390. The work of H.D. Sherali has been supported in part by NSF Grant 0094462.

References

[1] J. Winters, "On the capacity of radio communication systems with diversity in a Rayleigh fading environment," *IEEE J. Select. Areas Commun.*, vol. 5, no. 5, pp. 871–878, June 1987.
[2] I. E. Telatar, "Capacity of multi-antenna Gaussian channels," *European Trans. Telecomm.*, vol. 10, no. 6, pp. 585–596, Nov. 1999.
[3] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Commun.*, vol. 6, pp. 311–355, Mar. 1998.
[4] D. Gesbert, M. Shafi, D. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: An overview of MIMO space-time coded wireless systems," *IEEE J. Select. Areas Commun.*, vol. 21, no. 3, pp. 281–302, Apr. 2003.
[5] A. Goldsmith, S. A. Jafar, N. Jindal, and S. Vishwanath, "Capacity limits of MIMO channels," *IEEE J. Select. Areas Commun.*, vol. 21, no. 1, pp. 684–702, June 2003.
[6] S. Ye and R. S. Blum, "Optimized signaling for MIMO interference systems with feedback," *IEEE Trans. Signal Processing*, vol. 51, no. 11, pp. 2839–2848, Nov. 2003.
[7] M. Hu and J. Zhang, "MIMO ad hoc networks: Medium access control, saturation throughput, and optimal hop distance," *Special Issue on Mobile Ad Hoc Networks, Journal of Communications and Networks*, pp. 317–330, Dec. 2004.
[8] K. Sundaresan and R. Sivakumar, "Routing in ad hoc networks with MIMO links," in *Proc. IEEE ICNP*, Boston, MA, U.S.A., Nov. 2005, pp. 85–98.
[9] L. Zheng and D. N. C. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inform. Theory*, vol. 49, no. 5, pp. 1073–1096, May 2003.
[10] J. G. Proakis, *Digital Communications*. New York: McGraw-Hill, 2000.
[11] M. S. Bazaraa, J. J. Jarvis, and H. D. Sherali, *Linear Programming and Network Flows*. New York-Chichester-Brisbane-Toronto-Singapore: John Wiley & Sons Inc., 1990.
[12] F. P. Kelly, A. K. Malullo, and D. K. H. Tan, "Rate control in communications networks: Shadow prices, proporotional fairness and stablity," *Journal of the Operational Research Society*, vol. 49, pp. 237–252, 1998.
[13] L. Xiao, M. Johansson, and S. P. Boyd, "Simultaneous routing and resource allocation via dual decomposition," *IEEE Trans. Commun.*, vol. 52, no. 7, pp. 1136–1144, July 2004.
[14] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed. New York, NY: John Wiley & Sons Inc., 2006.
[15] J. B. Rosen, "The gradient projection method for nonlinear programming, Part I, Linear constraints," *SIAM J. Applied Mathematics*, vol. 8, pp. 181–217, 1960.
[16] J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Economics*. New York: Wiley, 1999.
[17] J. Liu, Y. T. Hou, Y. Shi, and H. D. Sherali, "Cross-layer optimization on routing and power control of MIMO-based mesh networks," *Technical Report, Deptment of ECE, Virginia Tech*, Sept. 2006. [Online]. Available: http://filebox.vt.edu/users/kevinlau/publications/TR_CL_MIMO.pdf
[18] S. Boyd and L. Xiao, "Least-squares covariance matrix adjustment," *SIAM Journal on Matrix Analysis and Applications*, vol. 27, no. 2, pp. 532–546, Nov. 2005.
[19] J. Malick, "A dual approach to semidefinite least-squares problems," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 1, pp. 272–284, Sept. 2005.