Matrix-Lifting SDP for Decoding in Multiple Antenna Systems

Amin Mobasher, and Amir K. Khandani

Electrical and Computer Engineering Department
University of Waterloo, Waterloo, ON, Canada
Email:{amin, khandani}@cst.uwaterloo.ca

Technical Report UW-E&CE#2007-06

Feb 16, 2007

National Library of Canada | Bibliothèque nationale du Canada

2007-06 (549)

PUBLICATION FOR LISTING IN CANADIANA - CANADA'S NATIONAL BIBLIOGRAPHY
PUBLICATION À INSCRIRE DANS CANADIANA - BIBLIOGRAPHIE NATIONALE DU CANADA

The information requested is used for cataloguing Canadian publications. If this title was submitted for Canadian Cataloguing in Publication (CIP) Data, only show changes or information not previously submitted.
Les renseignements servent au catologage des publications canadiennes. Si le titre a déjà été soumis en vue du programme CIP (Catalogage avant publication) canadien, n'indiquer que les changements ou les renseignements nouveaux.

Authors' names in full - Nom et prénom des auteurs	Birthdate - Date de naissance	Canadian Citizen / Citoyen canadien	Landed immigrant / Immigrant reçu	Non-Canadian / Étranger
Amin Mobasher	March 29, 1978	☐	☒	☐
Amir K. Khandani	Dec. 2, 1959	☒	☐	☐
		☐	☐	☐

Title of publication - Titre de la publication
Matrix-Lifting SDP for Decoding in Multiple Antenna Systems

Publisher - Maison d'édition
University of Waterloo

Where can booksellers and libraries obtain copies? Où les librairies et les bibliothèques peuvent-elles se procurer des exemplaires?
University of Waterloo
Dept. of Electrical and Computer Engineering
200 University Avenue West
Waterloo, ON N2L 3G1
CANADA

ISBN number - Numéro d'ISBN

Retail sale price - Prix de détail

Date published - Date d'édition: Dec. 21, 2007

Print run - Tirage: 1st run

Is distribution limited? La distribution est-elle restreinte? ☐ Yes (*to whom?*) Oui (*à qui?*) ☒ No - Non

Signature: J. Reimer per A. Khandani
Date: December 21, 2007

Telephone no. - Nº de téléphone: 519-885-1211
Fax no. - Nº de télécopieur: 519-746-3077
Internet: www.ece.uwaterloo.ca

This form should be completed and attached to the copies sent for legal deposit purposes.
Annexer la formule dûment remplie aux exemplaires envoyés conformément au règlement sur le dépôt légal.

Canada

NL/BN 15 (95-03)

Sent to Nat'l Library
Dec. 21, 2007
Janet Reimer

COPY

Matrix-Lifting SDP for Decoding in Multiple Antenna Systems

Amin Mobasher† and Amir K. Khandani†

Abstract

Recently, a lot of quasi-maximum likelihood decoding methods have been introduced to solve the decoding problem in multiple antenna systems. The general method proposed in [1] has a near optimal performance for M-ary QAM or PSK constellation. The advantage of this algorithm is that it can be implemented for any constellation with an arbitrary binary labeling, say Gray labeling. However, it is more complex compared to some other methods that specialized their algorithm for a limited scenario (also with degraded performance). In this paper, we introduce a new general algorithm based on matrix-lifting Semi-Definite Programming (SDP). The new relaxation introduces a small degradation in the performance; however, the reduction in the complexity is prohibitively large. The number of variables is decreased from $\mathcal{O}(N^2K^2)$ to $\mathcal{O}((N+K)^2)$. Moreover, it can be implemented for any constellation and labeling method.

I. Introduction

Recently, there has been a considerable interest in Multi-Input Multi-Output (MIMO) antenna systems due to achieving a very high capacity as compared to single-antenna systems [2]. One of the important problems in MIMO systems is Decoding. Decoding concerns the operation of recovering the transmitted vector from the received signal, which is known to be an NP-hard problem.

† Coding & Signal Transmission Laboratory (www.cst.uwaterloo.ca), Dept. of Elec. and Comp. Eng., University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1, Tel: 519-884-8552, Fax: 519-888-4338, e-mail: {amin, khandani}@cst.uwaterloo.ca

In the last decade, Sphere Decoder (SD)[1] is introduced as a Maximum Likelihood (ML) decoding method for MIMO systems with near-optimal performance [5]. In the SD method, the lattice points inside a hyper-sphere are generated and the closest lattice point to the received signal is determined. In [6], an exponential lower bound is derived on the average complexity of SD, and it is shown that the worst case complexity is exponential [7], [8]. However, it is experienced that over certain ranges of rate, Signal to Noise Ratio (SNR) and dimension the average complexity is polynomial.

To overcome the complexity issue, a variety of sub-optimum polynomial time algorithms based on Semi-Definite Programming (SDP) are suggested in the literature for decoding [1], [9]–[12].

In [9], a quasi-maximum likelihood method for decoding is introduced. Each signal constellation is expressed by its binary representation and the decoding is transformed into a quadratic minimization problem [9]. Then, the resulting problem is solved using a relaxation for rank-one matrices in Semi-Definite Programming (SDP) context. It is shown that this method has a near optimum performance and a polynomial time worst case complexity. However, the method proposed in [9] is limited to scenarios that the constellation points are expressed as a linear combination of bit labels. A typical example is the case of natural labeling in conjunction with PSK constellation [10]. Another quasi-maximum likelihood decoding method is introduced in [11] for larger PSK constellations with near ML performance and low complexity.

Another quasi-maximum likelihood decoding method in introduced in [12] for the MIMO systems employing 16-QAM. They replace any finite constellation by a polynomial constraint, e.g. if $x \in \{a, b, c\}$, then $(x-a)(x-b)(x-c) = 0$. Then, by introducing some slack variables, the constraints are expressed in terms of quadratic polynomials. Finally, the SDP relaxation is resulted by dropping the rank-one constraint. The work in [12] , in its current form, is restricted to MIMO systems employing 16-QAM. However, it can be generalized for larger constellations at the cost of defining more slack variables,

[1]This technique is introduced in the mathematical literature several years ago [3], [4].

increasing the complexity, and significantly decreasing the performance.

In [1], an efficient approximate ML decoder for MIMO systems is developed based on SDP. The transmitted vector is expanded as a linear combination (with zero-one coefficients) of all the possible constellation points in each dimension. Using this formulation, the distance minimization in Euclidean space is expressed in terms of a binary quadratic minimization problem. The minimization of this problem is over the set of all binary rank-one matrices with column sums equal to one. In order to solve this minimization problem, two relaxation models is presented, providing a trade-off between the computational complexity and the performance (both models can be solved with polynomial-time complexity). Simulation results show that the performance of the last model is near optimal for M-ary QAM or PSK constellation (with an arbitrary binary labeling, say Gray labeling). Therefore, the decoding algorithm built on the proposed model in [1] has a near-ML performance with polynomial computational complexity.

The general method proposed in [1] has a near optimal performance for M-ary QAM or PSK constellation. However, it is more complex compared to some other methods that specialized their algorithm for a limited scenarios [1], [9]–[12]. In this paper, we introduce a new general algorithm based on matrix-lifting Semi-Definite Programming (SDP) [13], [14]. The new relaxation introduces a small degradation in the performance; however, the reduction in the complexity is prohibitively large. The number of variables is decreased from $\mathcal{O}(N^2K^2)$ to $\mathcal{O}((N+K)^2)$. Moreover, it can be implemented for any constellation and labeling method.

Following notations are used in the sequel. The space of $N \times K$ (resp. $N \times N$) real matrices is denoted by $\mathcal{M}_{N\times K}$ (resp. $\mathcal{M}_N$), and the space of $N \times N$ symmetric matrices is denoted by $\mathcal{S}_N$. For a $N \times K$ matrix $\mathbf{X} \in \mathcal{M}_{N\times K}$ the (i, j)th element is represented by x_{ij}, where $1 \leq i \leq N,\ 1 \leq j \leq K$, i.e. $\mathbf{X} = [x_{ij}]$. We use $\mathrm{trace}(\mathbf{A})$ to denote the trace of a square matrix $\mathbf{A}$. The space of symmetric matrices is considered with the trace inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{trace}(\mathbf{AB})$. For $\mathbf{A}, \mathbf{B} \in \mathcal{S}_N$, $\mathbf{A} \succeq 0$ (resp. $\mathbf{A} \succ 0$) denotes positive semi-definiteness (resp. positive definiteness), and $\mathbf{A} \succeq \mathbf{B}$ denotes $\mathbf{A} - \mathbf{B} \succeq 0$. For two matrices $\mathbf{A}, \mathbf{B} \in \mathcal{M}_N$, $\mathbf{A} \geq \mathbf{B}$, ($\mathbf{A} > \mathbf{B}$) means $a_{ij} \geq b_{ij}$, $(a_{ij} > b_{ij})$ for all i, j.

The Kronecker product of two matrices $\mathbf{A}$ and $\mathbf{B}$ is denoted by $\mathbf{A} \otimes \mathbf{B}$ (for definition see [15]).

For $\mathbf{X} \in \mathcal{M}_{N\times K}$, $\mathrm{vec}(\mathbf{X})$ denotes the vector in $\mathbb{R}^{NK}$ (real NK-dimensional space) that is formed from the columns of the matrix $\mathbf{X}$. For $\mathbf{X} \in \mathcal{M}_N$, $\mathrm{diag}(\mathbf{X})$ is a vector of the diagonal elements of $\mathbf{X}$. We use $\mathbf{e}_N \in \mathbb{R}^N$ (resp. $\mathbf{0}_N \in \mathbb{R}^N$) to denote the $N \times 1$ vector of all ones (resp. all zeros), $\mathbf{E}_{N\times K} \in \mathcal{M}_{N\times K}$ to denote the matrix of all ones, and $\mathbf{I}_N$ to denote the $N \times N$ Identity matrix. For $\mathbf{X} \in \mathcal{M}_{N\times K}$, the notation $\mathbf{X}(1:i, 1:j)$, $i < K$ and $j < N$ denotes the sub-matrix of $\mathbf{X}$ containing the first i rows and the first j columns.

The rest of the paper is organized as follows. The problem formulation is introduced in Section II. Section III is the review of the vector-lifting semi-definite programming presented in [1]. In Section IV, we propose our new algorithm based on matrix-lifting semi-definite programming. we use the geometry of the relaxation to find a projected relaxation which has a better performance. Section V is devoted to the methods that can be used to solve the SDP problem. An augmented lagrangian method is proposed for the special structure of the problem. In Section VI, we present an optimization method, based on matrix nearness, on how we can find the integer solution of the original decoding problem from the relaxed optimization problem.

II. Problem Formulation

A MIMO system can be modeled by [1]

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{n}, \tag{1}$$

where $\mathbf{y}$ is the $M \times 1$ received vector, $\mathbf{H}$ is $M \times N$ real channel matrix, $\mathbf{n}$ is $N \times 1$ additive white gaussian noise vector, and $\mathbf{x}$ is $N \times 1$ data vector whose components are selected from the set $\{s, \cdots, s_K\}$.

Noting $x_i \in \{s, \cdots, s_K\}$, for $i = 1, \cdots, N$, we have

$$x_i = u_{i,1}s_1 + u_{i,2}s_2 + \cdots + u_{i,K}s_K, \tag{2}$$

where

$$u_{i,j} \in \{0,1\} \quad \text{and} \quad \sum_{j=1}^{K} u_{i,j} = 1, \quad \forall\, i = 1, \cdots, N. \tag{3}$$

Let

$$\mathbf{U} = \begin{bmatrix} u_{1,1} & \cdots & u_{1,K} \\ u_{2,1} & \cdots & u_{2,K} \\ \vdots & \ddots & \vdots \\ u_{N,1} & \cdots & u_{N,K} \end{bmatrix} \quad and \quad \mathbf{s} = \begin{bmatrix} s_1 \\ \vdots \\ s_K \end{bmatrix}.$$

Therefore, the transmitted vector is $\mathbf{x} = \mathbf{U}\mathbf{s}$ and $\mathbf{U}\mathbf{e}_K = \mathbf{e}_N$.

At the receiver, the Maximum-Likelihood (ML) decoding rule is given by

$$\hat{\mathbf{x}} = \arg \min_{x_i \in \{s, \cdots, s_K\}} \|\hat{\mathbf{y}} - \mathbf{H}\mathbf{x}\|^2, \tag{4}$$

where $\hat{\mathbf{x}}$ is the most likely input vector and $\hat{\mathbf{y}}$ is the received vector. Noting $\mathbf{x} = \mathbf{U}\mathbf{s}$, this problem is equivalent to

$$\begin{aligned} &\min_{\mathbf{U}\mathbf{e}=\mathbf{e}} \|\hat{\mathbf{y}} - \mathbf{H}\mathbf{U}\mathbf{s}\|^2 \equiv \\ &\min_{\mathbf{U}\mathbf{e}=\mathbf{e}} \mathbf{s}^T\mathbf{U}^T\mathbf{H}^T\mathbf{H}\mathbf{U}\mathbf{s} - 2\hat{\mathbf{y}}^T\mathbf{H}\mathbf{U}\mathbf{s}. \end{aligned} \tag{5}$$

Therefore, the decoding problem can be formulated as

$$\begin{aligned} \min \quad & \mathbf{s}^T\mathbf{U}^T\mathbf{H}^T\mathbf{H}\mathbf{U}\mathbf{s} - 2\hat{\mathbf{y}}^T\mathbf{H}\mathbf{U}\mathbf{s} \\ s.t. \quad & \mathbf{U}\mathbf{e}_K = \mathbf{e}_N \\ & u_{i,j} \in \{0,1\}. \end{aligned} \tag{6}$$

Let $\mathbf{Q} = \mathbf{H}^T\mathbf{H}$, $\mathbf{S} = \mathbf{s}\mathbf{s}^T$, $\mathbf{C} = -\mathbf{s}\mathbf{y}^T\mathbf{H}$, and let $\mathcal{E}_{K\times N}$ denote the set of all binary matrices in $\mathcal{M}_{K\times N}$ with row sums equal to one, i.e.

$$\mathcal{E}_{N\times K} = \{\mathbf{U} \in \mathcal{M}_{N\times K} : \mathbf{U}\mathbf{e}_K = \mathbf{e}_N, u_{ij} \in \{0,1\}, \forall i,j\}. \tag{7}$$

Therefore, the minimization problem (6) is

$$\begin{aligned} \min \quad & \operatorname{trace}\left(\mathbf{S}\mathbf{U}^T\mathbf{Q}\mathbf{U} + 2\mathbf{C}\mathbf{U}\right) \\ s.t. \quad & \mathbf{U} \in \mathcal{E}_{N\times K} \end{aligned} \tag{8}$$

III. VECTOR-LIFTING SEMI-DEFINITE PROGRAMMING

In order to solve the optimization problem (8), the authors in [1] proposed a *quadratic vector optimization* solution by defining $\mathbf{u} = \mathrm{vec}(\mathbf{U}^T), \mathbf{U} \in \mathcal{E}_{N\times K}$. By using this notation, the objective function is replaced by $\mathbf{u}^T(\mathbf{Q}\otimes\mathbf{S})\mathbf{u} + 2\mathrm{vec}(\mathbf{C}^T)^T\mathbf{u}$. To solve this vector quadratic problem, the quadratic form is linearized using the vector $\begin{bmatrix} 1 \\ \mathbf{u} \end{bmatrix}$, i.e.

$$\begin{aligned}\mathbf{Z}_\mathbf{u} &= \begin{bmatrix} 1 \\ \mathbf{u} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{u}^T \end{bmatrix} \\ &= \begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{u}\mathbf{u}^T \end{bmatrix} = \begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{Y} \end{bmatrix}, \end{aligned} \tag{9}$$

where $\mathbf{Y} = \mathbf{u}\mathbf{u}^T$ and it is relaxed to $\mathbf{Y} \succeq \mathbf{u}\mathbf{u}^T$, or equivalently, by the Schur complement, to the lifted constraint $\begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{Y} \end{bmatrix} \succeq 0$. Note that this matrix is selected from the set

$$\mathcal{F} := \mathrm{conv}\left\{\mathbf{Z}_\mathbf{u} : \mathbf{u} = \mathrm{vec}(\mathbf{U}^T),\ \mathbf{U} \in \mathcal{E}_{N\times K}\right\}. \tag{10}$$

Therefore, the decoding problem using *vector lifting semi-definite programming* can be represented by

$$\begin{aligned} \mathrm{trace} \quad & \begin{bmatrix} 0 & \mathrm{vec}(\mathbf{C}^T)^T \\ \mathrm{vec}(\mathbf{C}^T) & \mathbf{Q} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{Y} \end{bmatrix} \\ s.t. \quad & \begin{bmatrix} 1 & \mathbf{u}^T \\ \mathbf{u} & \mathbf{Y} \end{bmatrix} \in \mathcal{F}, \end{aligned} \tag{11}$$

which can be solved by usual SDP techniques. For more details, we refer the reader to [1].

Note that the optimization variable is a matrix in $\mathcal{S}_{NK+1}$. This leads to $NK+1$ by $NK+1$ matrix variables, which has $\mathcal{O}(N^2K^2)$ variables and it is prohibitively large for computations. However, the best approach is to keep and exploit the structure of the original optimization problem (8).

IV. Matrix-Lifting Semi-Definite Programming

In order to keep the matrix $\mathbf{U}$ in its original form in (8), the idea is to use the constraint $\mathbf{Y} = \mathbf{U}^T\mathbf{U}$ instead of $\mathbf{Y} = \mathbf{u}\mathbf{u}^T$. Now the relaxation is $\mathbf{Y} \succeq \mathbf{U}^T\mathbf{U}$, or equivalently, by the Schur complement, $\begin{bmatrix} \mathbf{I}_N & \mathbf{U} \\ \mathbf{U}^T & \mathbf{Y} \end{bmatrix} \succeq 0$. This is known as matrix-lifting semi-definite programming.

Define the new variable $\mathbf{V} = \mathbf{U}\mathbf{S}$. Since the matrix $\mathbf{S}$ is symmetric, the objective function in (8) can be represented as the Quadratic Matrix Program [14]

$$\begin{aligned} & \mathrm{trace}\ \left(\begin{bmatrix} \mathbf{U}^T & \mathbf{V}^T \end{bmatrix} \begin{bmatrix} \mathbf{0} & \frac{1}{2}\mathbf{Q} \\ \frac{1}{2}\mathbf{Q} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} + 2\mathbf{C}\mathbf{U} \right) \\ = & \ \mathrm{trace}\ \left(\begin{bmatrix} \mathbf{0} & \frac{1}{2}\mathbf{Q} \\ \frac{1}{2}\mathbf{Q} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{U}^T & \mathbf{V}^T \end{bmatrix} + 2\mathbf{C}\mathbf{U} \right) \\ = & \ \mathrm{trace}\ \left(\mathcal{L}_{\mathbf{Q}} \mathbf{W}_{\mathbf{U}} \right), \end{aligned} \tag{12}$$

where

$$\mathcal{L} = \begin{bmatrix} \mathbf{0} & \mathbf{C}^T & \mathbf{0} \\ \mathbf{C} & \mathbf{0} & \frac{1}{2}\mathbf{Q} \\ \mathbf{0} & \frac{1}{2}\mathbf{Q} & \mathbf{0} \end{bmatrix}$$

and

$$\mathbf{W}_{\mathbf{U}} = \begin{bmatrix} \mathbf{I} & \mathbf{U}^T & \mathbf{V}^T \\ \mathbf{U} & \mathbf{U}\mathbf{U}^T & \mathbf{U}\mathbf{V}^T \\ \mathbf{V} & \mathbf{V}\mathbf{U}^T & \mathbf{V}\mathbf{V}^T \end{bmatrix}.$$

In order to linearize $\mathbf{W}_{\mathbf{U}}$ consider the matrix

$$\begin{bmatrix} \mathbf{U} \\ \mathbf{V} \end{bmatrix} \begin{bmatrix} \mathbf{U}^T & \mathbf{V}^T \end{bmatrix} = \begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y} & \mathbf{Z} \end{bmatrix}, \tag{13}$$

where $\mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathcal{S}^N$. This equality can be relaxed to

$$\begin{bmatrix} \mathbf{U}\mathbf{U}^T & \mathbf{U}\mathbf{V}^T \\ \mathbf{V}\mathbf{U}^T & \mathbf{V}\mathbf{V}^T \end{bmatrix} - \begin{bmatrix} \mathbf{X} & \mathbf{Y} \\ \mathbf{Y} & \mathbf{Z} \end{bmatrix} \preceq 0. \tag{14}$$

It can be shown that this relaxation is convex in the Löwner partial order [13] and it is equivalent to the linear constraint [13]

$$\mathbf{W} \triangleq \begin{bmatrix} \mathbf{I} & \mathbf{U}^T & \mathbf{V}^T \\ \mathbf{U} & \mathbf{X} & \mathbf{Y} \\ \mathbf{V} & \mathbf{Y} & \mathbf{Z} \end{bmatrix} \succeq 0. \tag{15}$$

On the other hand, the feasible set in (8) is the set of binary matrices in $\mathcal{M}_{N\times K}$ with row sum equal to one, the set $\mathcal{E}_{N\times K}$ in (7). By relaxing the rank-one constraint for the matrix variable in (12), we have a tractable SDP problem. The feasible set for the objective function in (12) is approximated by

$$\begin{aligned} \mathcal{F}_{\mathcal{M}} = \operatorname{conv}\{\mathbf{W}_{\mathbf{U}} \mid \mathbf{U} \in \mathcal{M}_{N\times K} : \mathbf{U}\mathbf{e}_K = \mathbf{e}_N, \\ u_{ij} \in \{0,1\}, \forall i,j; \mathbf{V} = \mathbf{U}\mathbf{S}\} \end{aligned} \tag{16}$$

Therefore, the decoding problem can be represented by

$$\begin{aligned} \min \quad & \operatorname{trace}(\mathcal{L}\mathbf{W}) \\ s.t. \quad & \mathbf{W} \in \mathcal{F}_{\mathcal{M}} \end{aligned} \tag{17}$$

Note that the size of matrix $\mathbf{W}$ is $(2N+K)\times(2N+K)$, compared to $(NK+1)\times(NK+1)$ in [1], which is a huge reduction in the size of the problem.

Although the constraint in (13) is relaxed, we still can add/consider some linear constraints that have been removed. These constraints are valid for the non-convex rank-constrained decoding problem. However, we force the SDP problem to satisfy these constraints. Consider the auxiliary matrix $\mathbf{V}$ and the symmetric matrices $\mathbf{X}, \mathbf{Y}$ and $\mathbf{Z}$ in matrix $\mathbf{W}$.

Since $\mathbf{U} \in \mathcal{E}_{N\times K}$ and $\sum_{j=1}^{N} u_{ij}^2 = 1$, it is clear that $\operatorname{diag}(X) = \mathbf{e}_N$. Also, $\mathbf{Y}$ represents $\mathbf{U}\mathbf{S}\mathbf{U}^T$ and $\mathbf{Z}$ represents $\mathbf{U}\mathbf{S}^{\mathbf{2}}\mathbf{U}^T$. In each row of $\mathbf{U}$, there is only one 1 and the rest are zero. Therefore, we have

$$\operatorname{diag}(\mathbf{Y}) = \mathbf{U}\operatorname{diag}(\mathbf{S}) \quad \text{and} \quad \operatorname{diag}(\mathbf{Z}) = \mathbf{U}\operatorname{diag}(\mathbf{S}^{\mathbf{2}}). \tag{18}$$

Moreover, $\mathbf{S} = \mathbf{s}\mathbf{s}^T$ (rank-one matrix) and $\mathbf{S}^2 = (\sum_{1=i}^{K} s_i^2)\mathbf{S}$, we have a stronger results for $\mathbf{Z}$, i.e. $\mathbf{Z} = (\sum_{1=i}^{K} s_i^2)\mathbf{Y}$. Therefore, we have

$$\begin{aligned}
\min \quad & \operatorname{trace}\left(\mathcal{L}\begin{bmatrix} \mathbf{I} & \mathbf{U}^T & \mathbf{V}^T \\ \mathbf{U} & \mathbf{X} & \mathbf{Y} \\ \mathbf{V} & \mathbf{Y} & \mathbf{Z} \end{bmatrix}\right) \\
s.t. \quad & \mathbf{U}\mathbf{e}_K = \mathbf{e}_N \;\; ; \;\; \mathbf{U} \geq 0 \\
& \mathbf{V} = \mathbf{U}\mathbf{S} \\
& \operatorname{diag}(\mathbf{X}) = \mathbf{e}_N \\
& \operatorname{diag}(\mathbf{Y}) = \mathbf{U}\operatorname{diag}(\mathbf{S}) \\
& \mathbf{Z} = (\sum_{1=i}^{K} s_i^2)\mathbf{Y} \\
& \begin{bmatrix} \mathbf{I} & \mathbf{U}^T & \mathbf{V}^T \\ \mathbf{U} & \mathbf{X} & \mathbf{Y} \\ \mathbf{V} & \mathbf{Y} & \mathbf{Z} \end{bmatrix} \succeq 0 \\
& \mathbf{U}, \mathbf{V} \in \mathcal{M}_{N\times K}, \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathcal{S}^N
\end{aligned} \tag{19}$$

Remark 1: In communication applications it is always assumed that the average of the constellation points is zero, in order to reduce the transmitted energy. This results in $\mathbf{s}^T\mathbf{e}_K = 0$. Therefore, we have $\mathbf{V}\mathbf{e}_K = \mathbf{U}\mathbf{S}\mathbf{e}_K = \mathbf{U}\mathbf{s}\mathbf{s}^T\mathbf{e}_K = \mathbf{0}_N$.

Remark 2: The equation in (18) determines the diagonal elements of $\mathbf{Y}$. This property is hidden in the special structure of $\mathbf{U}$, i.e. $\mathbf{U} \in \mathcal{E}$. By using this property, we can even add more constraints. The equation $\mathbf{Y} = \mathbf{U}\mathbf{S}\mathbf{U}^T$ implies that $Y_{ij} = S_{kl}$ for some k and l. Therefore, the value of Y_{ij} is between the minimum and maximum elements of $\mathbf{S}$.

Remark 3: It can be easily shown that, in communication applications, $\mathbf{S}$, $\mathbf{Y}$, and $\mathbf{Z}$ are diagonal dominant matrices (since $\mathbf{s}^T\mathbf{e}_K = 0$). This property can also be used to add more constraints.

A. Geometry of the Relaxation

In this section, we eliminate the constraints defining $\mathbf{Ue} = \mathbf{e}$ by providing a tractable representation of the linear manifold spanned by this constraint. This method is called *gradient projection* or *reduced gradient method* [16]. The following lemma is on the representation of matrices having sum of the elements in each row equal to one. This lemma is used in our reduced gradient method.

Lemma 1: Let

$$\mathbf{G} = \left[\begin{array}{c|c} \mathbf{I}_{K-1} & -\mathbf{e}_{K-1} \end{array} \right] \in \mathcal{M}_{(K-1)\times K} \tag{20}$$

and

$$\mathbf{F} = \frac{1}{K}\left(\mathbf{E}_{N\times K} - \mathbf{E}_{N\times(K-1)}\mathbf{G}\right) \in \mathcal{M}_{N\times K}. \tag{21}$$

A matrix $\mathbf{U} \in \mathcal{M}_{N\times K}$ with the property that the summation of its elements in each row is equal to one, i.e. $\mathbf{U}\mathbf{e}_K = \mathbf{e}_N$, can be written as

$$\mathbf{U} = \mathbf{F} + \hat{\mathbf{U}}\mathbf{G}, \tag{22}$$

where $\hat{\mathbf{U}} = \mathbf{U}(\mathbf{1}:\mathbf{N}, \mathbf{1}:(\mathbf{K}-\mathbf{1}))$.

Proof: see [1]. ■

Corollary 4: $\forall \mathbf{U} \in \mathcal{E}_{N\times K}$, $\exists \hat{\mathbf{U}} \in \mathcal{M}_{N\times(K-1)}$, $\hat{u}_{ij} \in \{0,1\}$ s.t. $\mathbf{U} = \mathbf{F} + \hat{\mathbf{U}}\mathbf{G}$, where $\hat{\mathbf{U}} = \mathbf{U}(\mathbf{1}:\mathbf{N}, \mathbf{1}:(\mathbf{K}-\mathbf{1}))$. Note that the summation of each row of $\hat{\mathbf{U}}$ is 0 or 1.

Consider the minimization problem (8). By substituting (22), the objective function is

$$\begin{aligned} & \mathrm{trace}\left(\mathbf{S}\mathbf{U}^T\mathbf{Q}\mathbf{U} + 2\mathbf{C}\mathbf{U}\right) \\ = \; & \mathrm{trace}\left(\mathbf{S}(\mathbf{F}+\hat{\mathbf{U}}\mathbf{G})^T\mathbf{Q}(\mathbf{F}+\hat{\mathbf{U}}\mathbf{G}) + 2\mathbf{C}(\mathbf{F}+\hat{\mathbf{U}}\mathbf{G})\right) \\ = \; & \mathrm{trace}\Big(\mathbf{G}\mathbf{S}\mathbf{G}^T\hat{\mathbf{U}}^{\mathbf{T}}\mathbf{Q}\hat{\mathbf{U}} + \mathbf{G}\mathbf{S}\mathbf{F}^{\mathbf{T}}\mathbf{Q}\hat{\mathbf{U}} + \mathbf{Q}\mathbf{F}\mathbf{S}\mathbf{G}^{\mathbf{T}}\hat{\mathbf{U}}^{\mathbf{T}} \\ & \quad + \mathbf{G}\mathbf{C}\hat{\mathbf{U}} + \mathbf{C}^{\mathbf{T}}\mathbf{G}^{\mathbf{T}}\hat{\mathbf{U}}^{\mathbf{T}} + \mathbf{2}\mathbf{C}\mathbf{F} + \mathbf{S}\mathbf{F}^{\mathbf{T}}\mathbf{Q}\mathbf{F}\Big) \\ = \; & \mathrm{trace}\left(\hat{\mathcal{L}}\mathbf{W}_{\hat{\mathbf{U}}} + \mathbf{2}\mathbf{C}\mathbf{F} + \mathbf{S}\mathbf{F}^{\mathbf{T}}\mathbf{Q}\mathbf{F}\right), \end{aligned} \tag{23}$$

where

$$\begin{aligned}
\hat{\mathcal{L}} &= \begin{bmatrix} \mathbf{0} & \mathbf{QFSG}^T + \mathbf{C}^T\mathbf{G}^T & \mathbf{0} \\ \mathbf{GSF}^T\mathbf{Q} + \mathbf{GC} & \mathbf{0} & \frac{1}{2}\mathbf{Q} \\ \mathbf{0} & \frac{1}{2}\mathbf{Q} & \mathbf{0} \end{bmatrix}, \\
\mathbf{W}_{\hat{\mathbf{U}}} &= \begin{bmatrix} \mathbf{I} & \hat{\mathbf{U}}^{\mathbf{T}} & \hat{\mathbf{V}}^{\mathbf{T}} \\ \hat{\mathbf{U}} & \hat{\mathbf{U}}\hat{\mathbf{U}}^{\mathbf{T}} & \hat{\mathbf{U}}\hat{\mathbf{V}}^{\mathbf{T}} \\ \hat{\mathbf{V}} & \hat{\mathbf{V}}\hat{\mathbf{U}}^{\mathbf{T}} & \hat{\mathbf{V}}\hat{\mathbf{V}}^{\mathbf{T}} \end{bmatrix}, \\
\hat{\mathbf{V}} &= \hat{\mathbf{U}}\mathbf{GSG}^{\mathbf{T}}
\end{aligned} \tag{24}$$

Therefore, the decoding problem (6) can be written as

$$\begin{aligned}
\min \quad & \text{trace}\left(\hat{\mathcal{L}}\mathbf{W}_{\hat{\mathbf{U}}}\right) \\
s.t. \quad & \hat{\mathbf{U}} = \mathbf{U}(\mathbf{1}:\mathbf{N}, \mathbf{1}:(\mathbf{K}-\mathbf{1})); \quad \mathbf{U} \in \mathcal{E}_{\mathbf{N}\times\mathbf{K}} \\
& \hat{\mathbf{V}} = \hat{\mathbf{U}}\left(\mathbf{GSG}^{\mathbf{T}}\right)
\end{aligned} \tag{25}$$

Using the same procedure as before, we can show that the decoding problem is equivalent to the following reduced matrix-lifting semi-definite programming problem:

$$
\begin{aligned}
\min \quad & \operatorname{trace}\left(\hat{\mathcal{L}}\begin{bmatrix} \mathbf{I} & \hat{\mathbf{U}}^{\mathbf{T}} & \hat{\mathbf{V}}^{\mathbf{T}} \\ \hat{\mathbf{U}} & \hat{\mathbf{X}} & \hat{\mathbf{Y}} \\ \hat{\mathbf{V}} & \hat{\mathbf{Y}} & \hat{\mathbf{Z}} \end{bmatrix}\right) \\
s.t. \quad & \hat{\mathbf{U}}\mathbf{e}_{\mathbf{K-1}} \leq \mathbf{e}_{\mathbf{N}} \ ; \ \hat{\mathbf{U}} \geq \mathbf{0} \\
& \hat{\mathbf{V}} = \hat{\mathbf{U}}\left(\mathbf{G}\mathbf{S}\mathbf{G}^{\mathbf{T}}\right) \\
& \hat{X}_{ii}^2 = \hat{X}_{ii} \ \text{ for } \ i = 1, \cdots, N \\
& \operatorname{diag}(\hat{\mathbf{Y}}) = \hat{\mathbf{U}}\operatorname{diag}\left(\mathbf{G}\mathbf{S}\mathbf{G}^{\mathbf{T}}\right) \\
& \hat{\mathbf{Z}} = \left(\sum_{\mathbf{1}=\mathbf{i}}^{\mathbf{K-1}} (\mathbf{s_i} - \mathbf{s_K})^{\mathbf{2}}\right)\hat{\mathbf{Y}} \\
& \begin{bmatrix} \mathbf{I} & \hat{\mathbf{U}}^{\mathbf{T}} & \hat{\mathbf{V}}^{\mathbf{T}} \\ \hat{\mathbf{U}} & \hat{\mathbf{X}} & \hat{\mathbf{Y}} \\ \hat{\mathbf{V}} & \hat{\mathbf{Y}} & \hat{\mathbf{Z}} \end{bmatrix} \succeq 0 \\
& \hat{\mathbf{U}}, \hat{\mathbf{V}} \in \mathcal{M}_{\mathbf{N}\times(\mathbf{K-1})}, \hat{\mathbf{X}}, \hat{\mathbf{Y}}, \hat{\mathbf{Z}} \in \mathcal{S}^{\mathbf{N}}
\end{aligned} \tag{26}
$$

Note that the quadratic constraint $\hat{X}_{ii}^2 = \hat{X}_{ii}$ is equivalent to $\operatorname{diag}(\hat{\mathbf{X}}) = \hat{\mathbf{U}}\mathbf{e}_{\mathbf{K-1}}$.

B. Equivalent Formulation, but Different Results

The objective functions in the decoding problem (8) is $\operatorname{trace}\left(\mathbf{S}\mathbf{U}^T\mathbf{Q}\mathbf{U} + 2\mathbf{C}\mathbf{U}\right)$ which is equivalent to $\operatorname{trace}\left(\mathbf{Q}\mathbf{U}\mathbf{S}\mathbf{U}^T + 2\mathbf{U}\mathbf{C}\right)$. However, exchanging the role of $\mathbf{Q}$ and $\mathbf{S}$ results in two different formulation and bounds. Here, the auxiliary variable $\mathbf{V}$ is defined as $\mathbf{Q}\mathbf{U}$. Similar to previous notation, the auxiliary variables $\mathbf{X}, \mathbf{Y}$, and $\mathbf{Z}$ represents $\mathbf{U}^T\mathbf{U}, \mathbf{U}^T\mathbf{Q}\mathbf{U}$, and $\mathbf{U}^T\mathbf{Q}^2\mathbf{U}$, respectively. Therefore, it is easy to show that the equivalent minimization problem is

$$
\begin{aligned}
\min \quad & \operatorname{trace}\left(\begin{bmatrix} \mathbf{0} & \mathbf{C} & \mathbf{0} \\ \mathbf{C}^T & \mathbf{0} & \frac{1}{2}\mathbf{S} \\ \mathbf{0} & \frac{1}{2}\mathbf{S} & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{I} & \mathbf{U} & \mathbf{V} \\ \mathbf{U}^T & \mathbf{X} & \mathbf{Y} \\ \mathbf{V}^T & \mathbf{Y} & \mathbf{Z} \end{bmatrix}\right) \\
s.t. \quad & \mathbf{U}\mathbf{e}_K = \mathbf{e}_N \ \ ; \ \ \mathbf{U} \geq 0 \\
& \mathbf{V} = \mathbf{Q}\mathbf{U} \\
& \operatorname{diag}(\mathbf{X}) = \mathbf{U}^T\mathbf{e}_N \ \ ; \ \ X_{ij} = 0 \ i \neq j \\
& \mathbf{Y}\mathbf{e}_K = \mathbf{U}^T\mathbf{Q}\mathbf{e}_N \ \ ; \ \ \operatorname{trace}(\mathbf{Y}\mathbf{E}_K) = \operatorname{trace}(\mathbf{Q}\mathbf{E}_N) \\
& \mathbf{Z}\mathbf{e}_K = \mathbf{U}^T\mathbf{Q}^2\mathbf{e}_N \ \ ; \ \ \operatorname{trace}(\mathbf{Z}\mathbf{E}_K) = \operatorname{trace}(\mathbf{Q}^2\mathbf{E}_N) \\
& \begin{bmatrix} \mathbf{I} & \mathbf{U} & \mathbf{V} \\ \mathbf{U}^T & \mathbf{X} & \mathbf{Y} \\ \mathbf{V}^T & \mathbf{Y} & \mathbf{Z} \end{bmatrix} \succeq 0 \\
& \mathbf{U}, \mathbf{V} \in \mathcal{M}_{N\times K}, \mathbf{X}, \mathbf{Y}, \mathbf{Z} \in \mathcal{S}^K
\end{aligned} \tag{27}
$$

Note that the size of the variable matrix is $(2K+N)$.

V. Solving the SDP Problem

A. Common SDP Solvers

The relaxed decoding problems (19), (26), and (27) can be solved using Interior-Point Methods (IPMs), which are the most common methods for solving SDP problems of moderate sizes with polynomial computational complexities [17]. There are a large number of IPM-based solvers to handle SDP problems, e.g., DSDP [18], SeDuMi [19], SDPA [20], etc. In our numerical experiments, we use SeDuMi solver.

B. The augmented Lagrangian algorithm

Recently, Burer and Monteiro [21] proposed a new method for solving a full-rank SDP problem

$$\begin{aligned} \min \quad & \mathrm{trace}(\mathcal{L}\mathbf{W}) \\ s.t. \quad & \mathrm{trace}(\mathbf{A}_i\mathbf{W}) = b_i \ \text{ for } \ i = 1, \cdots, m \\ & \mathbf{W} \succeq 0. \end{aligned} \tag{28}$$

The algorithms distinguishing feature is a change of variables that replaces the symmetric, positive semi-definite variable $\mathbf{W} \in \mathcal{M}^n$ of (28) with a rectangular variable $\mathbf{R}$ according to the factorization $\mathbf{W} = \mathbf{R}\mathbf{R}^T$. In [22], [23], it is shown that, for a SDP problem (28) with m constraints, there exist an optimal solution with rank r such that $r(r+2) \leq m$. In [21], $\mathbf{R}$ is chosen in $\mathcal{M}_{n\times r}$. By using this formulation, the positive semi-definite constraint is removed since $\mathbf{W} = \mathbf{R}\mathbf{R}^T$ automatically enforces the constraint. Now the problem (28) can be reformulated as

$$\begin{aligned} \min_{\mathbf{R}\in\mathcal{M}_{n\times r}} \quad & \mathrm{trace}(\mathbf{R}^T\mathcal{L}\mathbf{R}) \\ s.t. \quad & \mathrm{trace}(\mathbf{R}^T\mathbf{A}_i\mathbf{R}) = b_i \ \text{ for } \ 0 \leq i \leq m \end{aligned} \tag{29}$$

In order to solve this non-linear problem, an *augmented Lagrangian* method is introduced in [21]. The augmented Lagrangian is defined as

$$\begin{aligned} L(\mathbf{R}, \boldsymbol{\lambda}, \sigma) \;=\; & \mathrm{trace}(\mathbf{R}^T\mathcal{L}\mathbf{R}) \\ & - \sum_{i=1}^{m} \lambda_i \left(\mathrm{trace}(\mathbf{R}^T\mathbf{A}_i\mathbf{R}) - b_i\right) \\ & + \frac{\sigma}{2}\sum_{i=1}^{m} \left(\mathrm{trace}(\mathbf{R}^T\mathbf{A}_i\mathbf{R}) - b_i\right)^2, \end{aligned} \tag{30}$$

where $\mathbf{R} \in \mathcal{M}_{n\times r}$, $\boldsymbol{\lambda} \in \mathbb{R}^m$, and $\sigma \in \mathbb{R}^+$. The last term is the penalty term indicating the Euclidean norm of the infeasibility of R with respect to r.

To minimize the augmented Lagrangian (30), this function is alternatively minimized with respect to $\mathbf{R}$ and with respect to $\boldsymbol{\lambda}$ and σ. The optimization of (30) with respect to

$\mathbf{R}$ can be achieved by a limited-memory BFGS algorithm, which uses the gradient of L:

$$\begin{aligned}\nabla_{\mathbf{R}} L(\mathbf{R}, \boldsymbol{\lambda}, \sigma) = & -2 \sum_{i=1}^{m} \left(\lambda_i - \sigma \left(\operatorname{trace}(\mathbf{R}^T \mathbf{A}_i \mathbf{R}) - b_i\right)\right) \mathbf{A}_i \mathbf{R} \\ & +2\mathcal{L}\mathbf{R}.\end{aligned} \tag{31}$$

This algorithm has the advantage of maintaining $O(nr)$ memory overhead, but also has the speed of a quasi-Newton method.

In order to optimize (30) with respect to $\boldsymbol{\lambda}$ and σ, the authors in [21] updates λ_i by $\lambda_i - \sigma \left(\operatorname{trace}(\mathbf{R}^T \mathbf{A}_i \mathbf{R}) - b_i\right)$ and σ by a multiplicative factor. They have shown that this technique always emprically converges to the globally optimal solution, although the quadratic programming problem is non-convex. Moreover, they have shown that this method is significantly faster than existing SDP solvers.

C. Decoding Problem as a Quadratic Non-Linear Problem

Define $\mathbf{R} = [\begin{array}{ccc} \mathbf{I} & \mathbf{U}^T & \mathbf{V}^T \end{array}]^T$. Therefore, the optimization problems (19), (26), and (27) can be reformulated as the problem in (29). Instead of applying SDP solvers, we can apply an augmented Lagrangian method directly to the mentioned decoding problems. However, despite the work in [21], our SDP problems have an explicit rank constraint. In other words, the size of the matrix $\mathbf{R}$ is precisely determined by the rank constraint.

The other difference is that we have some inequality constraints in our problem. The augmented Lagrangian method in [21] should be generalized to handle linear inequality constraints. We take the same approach as in [24] by solving directly for updates of the Lagrange multipliers for both equality and inequality constraints.

The updates of λ_i can be solved by treating a constraint $\operatorname{trace}(\mathbf{R}^T \mathbf{A}_i \mathbf{R}) \geq b_i$ as the constraint $\operatorname{trace}(\mathbf{R}^T \mathbf{A}_i \mathbf{R}) = b_i + s_i$ and $s_i \geq 0$. For inequality constraint, this results in

$$\lambda_i = \max(\lambda_i - \sigma(\operatorname{trace}(\mathbf{R}^T \mathbf{A}_i \mathbf{R}) - b_i), 0). \tag{32}$$

This update satisfies $\lambda_i \geq 0$ for Lagrange multipliers for inequality constraints. Computing the Lagrangian and the gradient of the Lagrangian are also straightforward for inequality constraints [24]. The second and third terms of the Lagrangian change: for each constraint

i, the Lagrangian term $-\lambda_i(\mathrm{trace}(\mathbf{R}^T\mathbf{A}_i\mathbf{R})-b_i)+\frac{\sigma}{2}(\mathrm{trace}(\mathbf{R}^T\mathbf{A}_i\mathbf{R})-b_i)^2$ is unchanged if $\sigma(\mathrm{trace}(\mathbf{R}^T\mathbf{A}_i\mathbf{R})-b_i)\leq\lambda_i$. Otherwise, this term becomes $-\lambda_i^2/2\sigma$. Similarly, in the computation of the gradient of Lagrangian, we only contribute the term $2(\lambda_i-\sigma(\mathrm{trace}(\mathbf{R}^T\mathbf{A}_i\mathbf{R})-b_i))\mathbf{A}_i\mathbf{R}$ if $\sigma(\mathrm{trace}(\mathbf{R}^T\mathbf{A}_i\mathbf{R})-b_i)\leq\lambda_i$. Otherwise, nothing is added to the gradient.

VI. Integer Solution - Matrix Nearness Problem

By solving the relaxed decoding problems, we can find a solution for matrix $\tilde{\mathbf{U}}$. In general, this matrix is not in $\mathcal{E}_{N\times K}$. The condition $\mathbf{U}\mathbf{e}_K=\mathbf{e}_N$ is satisfied. However, the elements are between 0 and 1. This matrix has to be converted to a 0-1 matrix by finding a matrix in $\mathcal{E}_{N\times K}$ which is nearest to this matrix.

A recurring problem in matrix theory is to find a structured matrix that best approximates a given matrix with respect to some distance measure. For example, it may be known a priori that a certain constraint ought to hold, and yet it fails on account of measurement errors or numerical roundoff. An attractive remedy is to replace the tainted matrix by the nearest matrix that does satisfy the constraint. Matrix approximation problems typically measure the distance between matrices with a norm. The Frobenius and spectral norms are pervasive choices because they are so analytically tractable.

In order to find the nearest solution in $\mathcal{E}_{N\times K}$ to $\tilde{\mathbf{U}}$, the solution of the relaxed problem, we solve the following minimization problem

$$\min_{\mathbf{U}\in\,\mathcal{E}_{N\times K}}\|\mathbf{U}-\tilde{\mathbf{U}}\|_{\mathbb{F}}^{\mathbf{2}},\tag{33}$$

where $\|\mathbf{A}\|_{\mathbb{F}}^2$ is the Frobenius norm of the matrix $\mathbf{A}$ and is defined as $\|\mathbf{A}\|_{\mathbb{F}}^2=\mathrm{trace}(\mathbf{A}\mathbf{A}^T)$. Therefore, the objective function can be reformulated as

$$\begin{aligned}\|\mathbf{U}-\tilde{\mathbf{U}}\|_{\mathbb{F}}^{\mathbf{2}} &= \mathrm{trace}\left((\mathbf{U}-\tilde{\mathbf{U}})(\mathbf{U}-\tilde{\mathbf{U}})^{\mathbf{T}}\right)\\ &= \mathrm{trace}\left(\mathbf{U}\mathbf{U}^T\right)-2\mathrm{trace}(\tilde{\mathbf{U}}\mathbf{U}^{\mathbf{T}})+\mathrm{trace}(\tilde{\mathbf{U}}\tilde{\mathbf{U}}^{\mathbf{T}})\\ &= N-2\mathrm{trace}(\tilde{\mathbf{U}}\mathbf{U}^{\mathbf{T}})+\mathrm{trace}(\tilde{\mathbf{U}}\tilde{\mathbf{U}}^{\mathbf{T}}).\end{aligned}\tag{34}$$

The last equality is due to the fact that for any $\mathbf{U} \in \mathcal{E}_{N\times K}$ we have $\mathrm{diag}(\mathbf{U}\mathbf{U}^T) = \mathbf{e}_N$, see (19). Therefore, after removing the constants, finding the integer solution is the solution of the following problem:

$$\max_{\mathbf{U}\in\,\mathcal{E}_{N\times K}} \ \mathrm{trace}(\tilde{\mathbf{U}}\mathbf{U}^{\mathbf{T}}) \tag{35}$$

Consider the maximization problem

$$\begin{aligned} \max \quad & \mathrm{trace}(\tilde{\mathbf{U}}\mathbf{U}^{\mathbf{T}}) \\ s.t. \quad & \mathbf{U}\mathbf{e}_K = \mathbf{e}_N \\ & 0 \leq \mathbf{U} \leq 1, \end{aligned} \tag{36}$$

where $\leq$ in the last constraint is element-wise. This problem is a linear programming problem with linear constraints and the optimum solution is a corner point meaning that constraint are satisfied with equality art the optimum point. In other words, at the optimum point, $\mathbf{U} \in \mathcal{E}_{N\times K}$. Therefore, in order to find the solution for (35), we can simply solve the linear problem (36), which is strongly polynomial time.

ACKNOWLEDGEMENT

The authors would like to thank Prof. H. Wolkowicz and Yichuan Ding for helpful discussions.

REFERENCES

[1] A. Mobasher, M. Taherzadeh, R. Sotirov, and A. K. Khandani, "A near maximum likelihood decoding algorithm for mimo systems based on semi-definite programming," *Submitted to IEEE Trans. on Info. Theory*, 2005.

[2] E. Telatar, "Capacity of multi-antenna gaussian channels," *European Trans. on Telecomm. ETT*, vol. 10, no. 6, pp. 585–596, November 1999.

[3] M. E. C. P. Schnorr, "Lattice basis reduction: Improved practical algorithms and solving subset sum problems," *Mathematical Programming*, no. 66, pp. 181–191, 1994.

[4] U. Fincke and M. Pohst, "Improved methods for calculating vectors of short length in a lattice, including a complexity analysis," *Mathematics of Computation*, vol. 44, pp. 463–471, Apr. 1985.

[5] M. O. Damen, A. Chkeif, and J.-C. Belfiore, "Lattice Code Decoder for Space-Time Codes," *IEEE Communications Letters*, vol. 4, no. 5, pp. 161–163, May 2000.

[6] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE Trans. on Signal Proc.*, vol. 53, no. 4, pp. 1474–1484, 2005.

[7] E. Agrell, T. Eriksson, A. Vardy, and K. Zeger, "Closest point search in lattices," *IEEE Transaction on Information Theory*, vol. 48, no. 8, pp. 2201–2214, Aug. 2002.

[8] B. Hassibi and H. Vikalo, "On Sphere Decoding Algorithm. Part I. Expected Complexity," 2003, submitted to IEEE Transaction on Signal Processing.

[9] B. Steingrimsson, T. Luo, and K. M. Wong, "Soft quasi-maximum-likelihood detection for multiple-antenna wireless channels," *IEEE Transactions on Signal Processing*, vol. 51, no. 11, pp. 2710– 2719, Nov. 2003.

[10] B. Steingrimsson, Z. Q. Luo, and K. M. Wong;, "Quasi-ML detectors with soft output and low complexity for PSK modulated MIMO channels," in *4th IEEE Workshop on Signal Processing Advances in Wireless Communications (SPAWC 2003)*, 15-18 June 2003, pp. 427 – 431.

[11] Z. Q. Luo, X. Luo, and M. Kisialiou, "An Efficient Quasi-Maximum Likelihood Decoder for PSK Signals," in *ICASSP '03*, 2003.

[12] Y. C. E. A. Wiesel and S. Shamai, "Semidefinite relaxation for detection of 16-QAM signaling in mimo channels," *IEEE Signal Processing Letters*, vol. 12, no. 9, Sept. 2005.

[13] Y. Ding and H. Wolkowicz, "A Matrix-lifting Semidefinite Relaxation for the Quadratic Assignment Problem," Department of Combinatorics & Optimization, University of Waterloo, Tech. Rep. CORR 06-22, 2006.

[14] A. Beck, "Quadratic matrix programming," *SIAM Journal on Optimization*, vol. 17, no. 4, pp. 1224–1238, 2007.

[15] A. Graham, *Kronecker Products and Matrix Calculus with Applications, Mathematics and its Applications*. Ellis Horwood Limited, Chichester, 1981.

[16] S. Hadley, F. Rendl, and H. Wolkowicz, "A new lower bound via projection for the quadratic assignment problem," *Math. Oper. Res.*, vol. 17, no. 3, pp. 727–739, 1992.

[17] F. Alizadeh, J.-P. A. Haeberly, and M. Overton, "Primal-dual interior-point methods for semidefinite programming: Stability, convergence, and numerical results," *SIAM Journal on Optimization*, vol. 8, no. 3, pp. 746–768, 1998.

[18] S. J. Benson and Y. Ye, "DSDP5 User Guide The Dual-Scaling Algorithm for Semidefinite Programming," Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, IL, Tech. Rep. ANL/MCS-TM-255, 2004, available via the WWW site at http://www.mcs.anl.gov/∼benson.

[19] J. Sturm, "Using sedumi 1.02, a matlab toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11-12, pp. 625–653, 1999.

[20] M. Kojima, K. Fujisawa, K. Nakata, and M. Yamashita, "SDPA(SemiDefinite Programming Algorithm) Users Mannual Version 6.00," Dept. of Mathematical and Computing Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguro-ku, Tokyo, 152-0033, Japan, Tech. Rep., 2002, available via the WWW site at http://sdpa.is.titech.ac.jp/SDPA/.

[21] S. Burer and R. D. C. Monteiro, "A nonlinear programming algorithm for solving semidefinite programs via low-rank factorization," *Mathematical Programming (Series B)*, vol. 95, pp. 329–357, 2003.

[22] A. Barvinok, "Problems of distance geometry and convex properties of quadratic maps," *Discrete Computational Geometry*, vol. 13, pp. 189–202, 1995.

[23] G. Pataki, "On the rank of extreme matrices in semidefinite programs and the multiplicity of optimal eigenvalue," *Mathematics of Operations Research*, vol. 23, pp. 339–358, 1998.

[24] B. Kulis, A. C. Surendran, and J. C. Platt, "Fast low-rank semidefinite programming for embedding and clustering," in *Eleventh International Conference on Artifical Intelligence and Statistics, AISTATS 2007*, March 2007.