On Equivalence of Semidefinite Relaxations for Quadratic Matrix Programming *

Yichuan Ding † Dongdong Ge ‡ Henry Wolkowicz §

April 28, 2010

Contents

1 Introduction **2**
1.1 Preliminaries 2
1.2 Outline 4

2 Quadratic Matrix Programming: Case I **4**
2.1 Lagrangian Relaxation 4
2.2 Equivalence of Vector and Matrix Lifting for QMP_1 5

3 Quadratic Matrix Programming: Case II **8**
3.1 Equivalence of Vector and Matrix Lifting for QMP_2 9
3.1.1 Proof of (Main) Theorem 3.1 12
3.1.2 Unbalanced Orthogonal Procrustes Problem 17
3.2 An Extension to QMP with Conic Constraints 18
3.2.1 Graph Partition Problem 18

4 Conclusion **19**

References **23**

Index **23**

Abstract

In this paper, we analyze two popular semidefinite programming $(\boldsymbol{SDP})$ relaxations for quadratically constrained quadratic programs $(\boldsymbol{QCQP})$ with matrix variables. These are based on *vector-lifting* and on *matrix lifting* and are of different size and expense. We prove, under mild assumptions, that these two relaxations provide equivalent bounds. Thus, our results provide a theoretical guideline for how to choose an inexpensive $\boldsymbol{SDP}$ relaxation and still obtain a strong bound. Our results also shed important insights on how to simplify large-scale $\boldsymbol{SDP}$ constraints by exploiting the particular sparsity pattern.

*Technical report CORR 2010-02; URL: orion.math.uwaterloo.ca/~hwolkowi/reports/ABSTRACTS.html
†Department of Management Science & Engineering, Stanford University. E-mail y7ding@stanford.edu
‡Antai College of Economics and Management, Shanghai Jiao Tong University. E-mail ddge@sjtu.edu.cn
§Research supported by Natural Sciences Engineering Research Council Canada. E-mail hwolkowicz@uwaterloo.ca

Keywords: Semidefinite Programming Relaxations, Quadratic Matrix Programming, Quadratic Constrained Quadratic Programming, Hard Combinatorial Problems.

1 Introduction

In this paper, we aim to provide theoretical insights on how to compare different semidefinite programming $(\boldsymbol{SDP})$ relaxations. In particular, we study a *vector lifting* relaxation and compare it to a significantly smaller *matrix lifting* relaxation and show that the resulting two bounds are equal.

Many hard combinatorial problems can be formulated as a quadratically constrained quadratic program $(\boldsymbol{QCQP})$ with matrix variables. If the resulting formulated problem is nonconvex, then $\boldsymbol{SDP}$ relaxations provide an efficient and successful approach for computing approximate solutions and strong bounds. Finding strong and inexpensive bounds is essential for branch and bound algorithms for solving large hard combinatorial problems. However, there can be many different $\boldsymbol{SDP}$ relaxations for the same problem, and it is usually not obvious which relaxation is overall *optimal* with regard to both computational efficiency and bound quality, e.g., [12].

For examples of using $\boldsymbol{SDP}$ relaxations for $\boldsymbol{QCQP}$ arising from hard problems, see e.g., quadratic assignment (QAP) [11, 12, 23, 26, 35], graph partitioning (GPP) [34], sensor network localization (SNL) [9, 10, 21], and the more general Euclidean distance matrix completions [1].

1.1 Preliminaries

The concept of *quadratic matrix programming* $(\boldsymbol{QMP})$ was introduced by Beck in [6], where it refers to a special instance of $\boldsymbol{QCQP}$ with matrix variables. Because we include the study of more general problems, we denote the model discussed in [6] as the *first case of quadratic matrix programming*, denoted $(\boldsymbol{QMP_1})$,

$$(\boldsymbol{QMP_1}) \qquad \begin{array}{rll} \mu^*_{P1} := & \min & \operatorname{trace}(X^TQ_0X) + 2\operatorname{trace}(C_0^TX) + \beta_0 \\ & \text{s.t.} & \operatorname{trace}(X^TQ_jX) + 2\operatorname{trace}(C_j^TX) + \beta_j \le 0, \quad j = 1,2,\ldots,m \\ & & X \in \mathcal{M}^{nr}, \end{array}$$

where $\mathcal{M}^{nr}$ denotes the set of n by r matrices, $Q_j \in \mathcal{S}^n, j = 0,1,\ldots,m$, $\mathcal{S}^n$ is the space of $n \times n$ symmetric matrices, and $C_j \in \mathcal{M}^{nr}$. Throughout this paper, we use the trace inner-product (dot product) $C \cdot X := \operatorname{trace} C^TX$.

The applicability of $\boldsymbol{QMP_1}$ is limited when compared to the more general class $\boldsymbol{QCQP}$. However, many applications use $\boldsymbol{QCQP}$ models in the form of $\boldsymbol{QMP_1}$, e.g., robust optimization [7] and sensor network localization [1]. In addition, many combinatorial problems are formulated with orthogonality constraints in one of the two forms

$$XX^T = I, \quad X^TX = I. \tag{1}$$

When X is square, the pair of constraints in (1) are equivalent to each other, *in theory*. However, relaxations that include both forms of the constraints rather than just one, can be expected to obtain stronger bounds. For example, Anstreicher and Wolkowicz [4] proved that strong duality holds for a certain relaxation of QAP when both forms of the orthogonality constraints in (1) are included; however, there can be a duality gap if only one of the forms is used. Motivated by this result, we extend our scope of problems so that the objective and constraint functions can include both forms of quadratic terms $X^T Q_j X$ and $X P_j X^T$. We now define the *second case of quadratic matrix programming* ($\boldsymbol{QMP_2}$) problems as

$$(\boldsymbol{QMP_2}) \quad \begin{array}{ll} \min & \operatorname{trace}(X^T Q_0 X) + \operatorname{trace}(X P_0 X^T) + 2\operatorname{trace}(C_0^T X) + \beta_0 \\ \text{s.t.} & \operatorname{trace}(X^T Q_j X) + \operatorname{trace}(X P_j X^T) + 2\operatorname{trace}(C_j^T X) + \beta_j \le 0, j = 1, \dots, m \\ & X \in \mathcal{M}^{nr}, \end{array} \tag{2}$$

where Q_j, P_j are symmetric matrices of appropriate sizes.

Both $\boldsymbol{QMP_1}$ and $\boldsymbol{QMP_2}$ can be vectorized into the $\boldsymbol{QCQP}$ form using

$$\operatorname{trace}(X^T Q X) = \operatorname{vec}(X)^T (I_r \otimes Q) \operatorname{vec}(X), \quad \operatorname{trace}(X P X^T) = \operatorname{vec}(X)^T (P \otimes I_n) \operatorname{vec}(X), \tag{3}$$

where $\otimes$ denotes the Kronecker product, and $\operatorname{vec}(X)$ vectorizes X by stacking columns of X on top of each other. The difference in the Kronecker products $(I_r \otimes Q), (P \otimes I_n)$ shows that there is a difference in the corresponding Lagrange multipliers and illustrates why the bounds from Lagrangian relaxation will be different for these two sets of constraints. The $\boldsymbol{SDP}$ relaxation for the vectorized $\boldsymbol{QCQP}$ is called the *vector-lifting semidefinite relaxation* ($\boldsymbol{VSDR}$). Under a constraint qualification assumption, $\boldsymbol{VSDR}$ for $\boldsymbol{QCQP}$ is equivalent to the dual of classical Lagrangian relaxation, see e.g., [24, 33, 5].

From (3), we get

$$\begin{array}{l} \operatorname{trace}(X^T Q X) = \operatorname{trace}(I_r \otimes Q) Y, \quad \text{if } \; Y = \operatorname{vec}(X) \operatorname{vec}(X)^T \\ \operatorname{trace}(X P X^T) = \operatorname{trace}(P \otimes I_n) Y, \quad \text{if } \; Y = \operatorname{vec}(X) \operatorname{vec}(X)^T. \end{array} \tag{4}$$

$\boldsymbol{VSDR}$ is derived using (4) with the relaxation $Y \succeq \operatorname{vec}(X) \operatorname{vec}(X)^T$. A Schur complement argument, e.g., [25, 22], implies the equivalence of this relaxation to the large matrix variable constraint $\begin{bmatrix} 1 & \operatorname{vec}(X^T) \\ \operatorname{vec}(X) & Y \end{bmatrix} \succeq 0$. A similar result holds for $\operatorname{trace}(X P X^T) = \operatorname{vec}(X)^T (P \otimes I_n) \operatorname{vec}(X)$.

Alternatively, from (3), we get the smaller system

$$\begin{array}{l} \operatorname{trace}(X^T Q X) = \operatorname{trace} Q Y, \quad \text{if } Y = X X^T \\ \operatorname{trace}(X P X^T) = \operatorname{trace} P Y, \quad \text{if } Y = X^T X. \end{array} \tag{5}$$

The *matrix-lifting semidefinite relaxation* $\boldsymbol{MSDR}$ is derived using (5) with the relaxation $Y \succeq X X^T$. A Schur complement argument now implies the equivalence of this relaxation to the smaller matrix variable constraint $\begin{bmatrix} I & X^T \\ X & Y \end{bmatrix} \succeq 0$. Again, a similar result holds for $\operatorname{trace}(X P X^T)$.

Intuitively, one expects that $\boldsymbol{VSDR}$ should provide stronger bounds than $\boldsymbol{MSDR}$. Beck [6] proved that $\boldsymbol{VSDR}$ is actually equivalent to $\boldsymbol{MSDR}$ for $\boldsymbol{QMP_1}$ if both $\boldsymbol{SDP}$ relaxations attain optimality and have a zero duality gap, e.g., when a constraint qualification, such as the Slater condition, holds for the dual program. In this paper we strengthen the above result by dropping the constraint qualification assumption. Then, we present our main contribution, i.e., we show the equivalence between $\boldsymbol{MSDR}$ and $\boldsymbol{VSDR}$ for the more general problem $\boldsymbol{QMP_2}$ under a constraint qualification. This result is of more interest because $\boldsymbol{QMP_2}$ does not possess the same nice structure (chordal pattern) as $\boldsymbol{QMP_1}$. Moreover, $\boldsymbol{QMP_2}$ encompasses a much richer class of problems and therefore has more significant applications.

1.2 Outline

In Section 2 we present the equivalence of the corresponding $\boldsymbol{VSDR}$ and $\boldsymbol{MSDR}$ formulations for $\boldsymbol{QMP_1}$ and prove Beck's result without the constraint qualification assumption, see Theorem 2.1. Section 3 proves the main result that $\boldsymbol{VSDR}$ and $\boldsymbol{MSDR}$ generate equivalent lower bounds for $\boldsymbol{QMP_2}$, under a constraint qualification assumption, see Theorem 3.1. Numerical tests are included. Section 4 provides concluding remarks.

2 Quadratic Matrix Programming: Case I

We first discuss the two relaxations for $\boldsymbol{QMP_1}$. We denote the matrices in the relaxations obtained from vector and matrix lifting by

$$\begin{array}{rcl} M(q_j^V(\bullet)) & := & \begin{bmatrix} \beta_j & \mathrm{vec}(C_j)^T \\ \mathrm{vec}(C_j) & I_r \otimes Q_j \end{bmatrix}, \\ M(q_j^M(\bullet)) & := & \begin{bmatrix} \frac{\beta_j}{r} I_r & C_j^T \\ C_j & Q_j \end{bmatrix}. \end{array}$$

We let

$$y = \begin{pmatrix} x_0 \\ \mathrm{vec}(X) \end{pmatrix} \in \mathbb{R}^{nr+1}, \quad Y = \begin{pmatrix} X_0 \\ X \end{pmatrix} \in \mathcal{M}^{(r+n)r},$$

and denote the quadratic and homogenized quadratic functions

$$q_j(X) := \mathrm{trace}(X^T Q_j X) + 2\,\mathrm{trace}(C_j^T X) + \beta_j$$

$$\begin{array}{rcl} q_j^V(X, x_0) & := & \mathrm{trace}(X^T Q_j X) + 2\,\mathrm{trace}(C_j^T X x_0) + \beta_j x_0^2 \\ & = & y^T M(q_j^V(\bullet)) y \end{array}$$

$$\begin{array}{rcl} q_j^M(X, X_0) & := & \mathrm{trace}(X^T Q_j X) + 2\,\mathrm{trace}(X_0^T C_j^T X) + \mathrm{trace}\, \frac{\beta_j}{r} X_0^T I_r X_0 \\ & = & \mathrm{trace}\, Y^T M(q_j^M(\bullet)) Y \end{array}$$

2.1 Lagrangian Relaxation

As mentioned above, under a constraint qualification, $\boldsymbol{VSDR}$ for $\boldsymbol{QCQP}$ is equivalent with the dual of classical Lagrangian relaxation. We include this result for completeness and to illustrate the role of a constraint qualification in the relaxation. We follow the approach in [24, Pg. 403] and use the strong duality of the trust region subproblem [31] to obtain the Lagrangian relaxation (or dual) for $\boldsymbol{QMP_1}$ as an $\boldsymbol{SDP}$.

$$
\begin{array}{rcl}
\mu_L^* & := & \max\limits_{\lambda\geq 0}\min\limits_{X} q_0(X)+\sum_{i=j}^m \lambda_j q_j(X)\\
& = & \max\limits_{\lambda\geq 0}\min\limits_{X,x_0^2=1} q_0^V(X,x_0)+\sum_{j=1}^m \lambda_i q_j^V(X,x_0)\\
& = & \max\limits_{\lambda\geq 0,t}\min\limits_{y} y^T\left(M(q_0^V(\bullet))+\sum_{j=1}^m \lambda_j M(q_j^V(\bullet))\right)y+t(1-x_0^2)\\
& = & \max\limits_{\lambda\geq 0,t}\min\limits_{y} \operatorname{trace}\left(M(q_0^V(\bullet))+\sum_{j=1}^m \lambda_j M(q_j^V(\bullet))\right)yy^T+t(1-x_0^2)\\
& = & (\boldsymbol{DVSDR_1})\left\{\begin{array}{ll} \max & t\\ \text{s.t.} & \begin{bmatrix} t & 0\\ 0 & 0\end{bmatrix}-\sum_{j=1}^m \lambda_j M(q_j^V(\bullet))\preceq M(q_0^V(\bullet))\\ & \lambda\in\mathbb{R}_+^m, t\in\mathbb{R}.\end{array}\right.
\end{array}
\tag{6}
$$

As illustrated in (6), Lagrangian relaxation is the dual program (denoted by $\boldsymbol{DVSDR_1}$) of the vector-lifting relaxation $\boldsymbol{VSDR_1}$ given below. Hence, under a constraint qualification, the Lagrangian relaxation is equivalent with the vector-lifting semidefinite relaxation. The usual constraint qualification is the Slater condition, i.e.,

$$
\exists\lambda\in\mathbb{R}_+^m,\ \text{s.t.}\quad M(q_0^V(\bullet))+\sum_{j=1}^m \lambda_j M(q_j^V(\bullet))\succ 0. \tag{7}
$$

2.2 Equivalence of Vector and Matrix Lifting for QMP$_1$

Recall that the dot product refers to the trace inner-product, $C\cdot X=\operatorname{trace} C^TX$. The vector-lifting relaxation is

$$
(\boldsymbol{VSDR_1})\qquad
\begin{array}{rll}
\mu_{V1}^* := & \min & M(q_0^V(\bullet))\cdot Z_V\\
& \text{s.t.} & M(q_j^V(\bullet))\cdot Z_V\leq 0,\ j=1,2,\ldots,m\\
& & (Z_V)_{1,1}=1\\
& & Z_V\succeq 0.
\end{array}
$$

Thus the constraint matrix is blocked as $Z_V=\begin{bmatrix} 1 & \operatorname{vec}(X)^T\\ \operatorname{vec}(X) & Y_V\end{bmatrix}$.

The matrix-lifting relaxation is

$$
(\boldsymbol{MSDR_1})\qquad
\begin{array}{rll}
\mu_{M1}^* := & \min & M(q_0^M(\bullet))\cdot Z_M\\
& \text{s.t.} & M(q_j^M(\bullet))\cdot Z_M\leq 0,\ j=1,2,\ldots,m\\
& & (Z_M)_{1:r,1:r}=I_r\\
& & Z_M\succeq 0.
\end{array}
$$

Thus the constraint matrix is blocked as $Z_M=\begin{bmatrix} I_r & X^T\\ X & Y_M\end{bmatrix}$.

$\boldsymbol{VSDR_1}$ is obtained by relaxing the quadratic equality constraint $Y_V=\operatorname{vec}(X)\operatorname{vec}(X)^T$ to $Y_V\succeq\operatorname{vec}(X)\operatorname{vec}(X)^T$, and then formulating this as $Z_V=\begin{bmatrix} 1 & \operatorname{vec}(X)^T\\ \operatorname{vec}(X) & Y_V\end{bmatrix}\succeq 0$. $\boldsymbol{MSDR_1}$ is obtained by relaxing the quadratic equality constraint $Y_M=XX^T$ to $Y_M\succeq XX^T$ and then reformulating this to the linear conic constraint $Z_M=\begin{bmatrix} I_r & X^T\\ X & Y_M\end{bmatrix}\succeq 0$. $\boldsymbol{VSDR_1}$ involves $O((nr)^2)$

variables and $O(m)$ constraints which is often at the complexity of $O(nr)$, whereas the smaller problem $\boldsymbol{MSDR_1}$ has only $O((n+r)^2)$ variables. The equivalence of relaxations using vector and matrix-liftings is proved in [6, Theorem 4.3] by assuming a constraint qualification for the dual programs. We now present our first main result and prove the above mentioned equivalence without any constraint qualification assumptions. The proof is of interest in itself in that we use the chordal property and matrix completions to connect the two relaxations.

Theorem 2.1. *As numbers in the extended real line* $[-\infty, +\infty]$, *the optimal values of the two relaxations obtained using vector and matrix-liftings are equal, i.e.,*

$$\mu_{V1}^* = \mu_{M1}^*.$$

Proof. The proof follows by showing that both $\boldsymbol{VSDR_1}$ and $\boldsymbol{MSDR_1}$ generate the same optimal values as the following program.

$$(\boldsymbol{VSDR_1'}) \qquad \begin{array}{rll} \mu_{V1'}^* := & \min & Q_0 \cdot \sum_{j=1}^r Y_{jj} + 2C_0 \cdot X + \beta_0 \\ & \text{s.t.} & Q_j \cdot \sum_{j=1}^r Y_{jj} + 2C_j \cdot X + \beta_j \leq 0, \ j = 1, 2, \ldots, m \\ & & Z_{jj} = \begin{bmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{bmatrix} \succeq 0, \ j = 1, 2, \ldots, r, \end{array}$$

where x_j, $j = 1, 2, \ldots, r$, are the columns of matrix X, and Y_{jj}, $j = 1, 2, \cdots, r$, *represent* the corresponding quadratic parts $x_j x_j^T$.

We first show that the optimal values of $\boldsymbol{VSDR_1}$ and $\boldsymbol{VSDR_1'}$ are equal, i.e., that

$$\mu_{V1}^* = \mu_{V1'}^*. \tag{8}$$

The equivalence of the two optimal values can be established by showing for each program, for each feasible solution, one can always construct a corresponding feasible solution to the other program with the same objective value.

First, suppose $\boldsymbol{VSDR_1'}$ has a feasible solution $Z_{jj} = \begin{bmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{bmatrix}$, $j = 1, 2, \ldots, r$. Construct the *partial symmetric matrix*

$$Z_V = \begin{bmatrix} 1 & x_1^T & x_2^T & \ldots & x_r^T \\ x_1 & Y_{11} & ? & ? & ? \\ x_2 & ? & Y_{22} & ? & ? \\ \vdots & ? & ? & \ddots & ? \\ x_r & ? & ? & ? & Y_{rr} \end{bmatrix},$$

where the entries denoted by '?' are unknown/unspecified. By observation, the unspecified entries of Z_V are not involved in the constraints or in the objective function of $\boldsymbol{VSDR_1}$. In other words, adding values to the unspecified positions will not change the constraint function values and the objective value. Therefore, any positive semidefinite completion of the partial matrix Z_V is feasible for $\boldsymbol{VSDR_1}$ and has the same objective value. The feasibility of Z_{jj} $(j = 1, 2, \ldots, r)$ for $\boldsymbol{VSDR_1'}$ implies $\begin{bmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{bmatrix} \succeq 0$ for each $j = 1, 2, \ldots, r$. So all the specified principal submatrices of Z_V are positive semidefinite, and hence Z_V is a *partial positive semidefinite matrix* (See references [2, 15, 16, 18, 32] for the specific definitions of partial positive semidefinite, chordal graph, semidefinite completion.) It is not difficult to verify the *chordal graph* property for the sparsity pattern of Z_V. Therefore Z_V has a positive semidefinite completion by the classical completion result [15, Theorem 7]. Thus we have constructed a feasible solution to $\boldsymbol{VSDR_1}$ with the same objective value as the feasible solution from $\boldsymbol{VSDR'_1}$. i.e., this shows that $\mu^*_{V1} \le \mu^*_{V1'}$.

Conversely, suppose $\boldsymbol{VSDR_1}$ has a feasible solution

$$Z_V = \begin{bmatrix} 1 & x_1^T & x_2^T & \dots & x_r^T \\ x_1 & Y_{11} & Y_{12} & \dots & Y_{1r} \\ x_2 & Y_{21} & Y_{22} & \dots & Y_{2r} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_r & Y_{r1} & Y_{r2} & \dots & Y_{rr} \end{bmatrix} \succeq 0.$$

Now we construct $Z_{jj} := \begin{bmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{bmatrix}$, $j = 1, 2, \dots, r$. Because each Z_{jj} is a principal submatrix of the positive semidefinite matrix Z_V, we have $Z_{jj} \succeq 0$. The feasibility of Z_V for $\boldsymbol{VSDR_1}$ also implies

$$M(q_i^V(\bullet)) \cdot Z_V \le 0, \qquad i = 1, 2, \dots, m. \tag{9}$$

It is easy to check that

$$Q_i \cdot \sum_{j=1}^{r} Y_{jj} + 2C_i \cdot X + \beta_i = M(q_i^V(\bullet)) \cdot Z_V \le 0, \qquad i = 1, 2, \dots, m, \tag{10}$$

where $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_r \end{bmatrix}$. Therefore, Z_{jj}, $j = 1, 2, \dots, r$, is feasible for $\boldsymbol{VSDR'_1}$, and also generates the same objective value for $\boldsymbol{VSDR'_1}$ as Z_V for $\boldsymbol{VSDR_1}$ by (10), i.e., this shows that $\mu^*_{V1} \ge \mu^*_{V1'}$. This completes the proof of (8).

Next we prove that the optimal values of $\boldsymbol{MSDR_1}$ and $\boldsymbol{VSDR'_1}$ are equal, i.e., that

$$\mu^*_{M1} = \mu^*_{V1'}. \tag{11}$$

The proof is similar to the one for (8). First suppose $\boldsymbol{VSDR'_1}$ has a feasible solution $Z_{jj} = \begin{bmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{bmatrix}$, $j = 1, 2, \cdots, m$. Let $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_r \end{bmatrix}$, and $Y_M = \sum_{j=1}^r Y_{jj}$. Now we construct $Z_M := \begin{bmatrix} I_r & X^T \\ X & Y_M \end{bmatrix}$. Then by $Z_{jj} \succeq 0$, $j = 1, 2, \dots, r$, we have $Y_M = \sum_{j=1}^r Y_{jj} \succeq \sum_{j=1}^r x_j x_j^T = XX^T$, which implies $Z_M^* \succeq 0$ by the Schur Complement [22, 25]. And, because

$$M(q_j^M(\bullet)) \cdot Z_M = Q_j \cdot \sum_{i=1}^{r} Y_{ii} + 2C_j \cdot X + \beta_j, \qquad j = 1, \dots, m, \tag{12}$$

we get Z_M is feasible for $\boldsymbol{MSDR}$ and it generates the same objective value as the one by Z_{jj}, $j = 1, 2, \cdots, m$, for $\boldsymbol{VSDR'_1}$, i.e., $\mu^*_{M1} \le \mu^*_{V1'}$.

Conversely, suppose $Z_M = \begin{bmatrix} I_r & X^T \\ X & Y_M \end{bmatrix} \succeq 0$ is feasible for $\boldsymbol{MSDR_1}$, and $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_r \end{bmatrix}$. Let $Y_{ii} = x_i (x_i)^T$ for $i = 1, 2, \dots, r-1$, and let $Y_{rr} = x_r x_r^T + (Y_M - XX^T)$. As a result, $Y_{ii} \succeq x_i x_i^T$ for $i = 1, 2, \dots, r$, and $\sum_{i=1}^r Y_{ii} = Y_M$. So, by constructing $Z_{jj} = \begin{bmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{bmatrix}$, $j = 1, 2, \cdots, r$, it is easy to show that Z_{jj} is feasible for $\boldsymbol{VSDR'_1}$ and generates an objective value equal to the objective value of $\boldsymbol{MSDR}$ with Z_M, i.e., $\mu^*_{M1} \ge \mu^*_{V1'}$. This completes the proof of (11). Combining this with (8) completes the proof of the Theorem. □

Though $\boldsymbol{MSDR_1}$ is significantly less expensive, Theorem 2.1 implies that the quality of the $\boldsymbol{MSDR_1}$ bound is no weaker than that from $\boldsymbol{VSDR_1}$. This tells us to always choose $\boldsymbol{MSDR_1}$ if the problem can be formulated as in $\boldsymbol{QMP_1}$.

Example 2.1 (Sensor Network Localization Problem)**.** *The Sensor Network Localization, SNL, problem is one of the most studied problems in Graph Realization, e.g., [20, 21, 30]. In this problem one is given a graph with m known points (anchors) $a_k \in R^d$, $k = 1, 2, \cdots, m$, and n unknown points (sensors) $x_j \in R^d$, $j = 1, 2, \cdots, n$, where d is the* embedding dimension. *A Euclidean distance $\hat{d}_{kj}$ between a_k and x_j or distance $\hat{d}_{ij}$ between x_i and x_j is also given for some pairs of two points. The goal is to seek estimates of the positions for all unknown points. One possible formulation of the problem is as follows.*

$$\begin{array}{ll} \min & 0 \\ s.t. & \operatorname{trace}(X^T(E_{ii} + E_{jj} - 2E_{ij})X) = d_{ij}, \qquad\qquad \forall\ (i,j) \in N_x \\ & \operatorname{trace}(X^T E_{ii} X) - 2\operatorname{trace}(\begin{bmatrix} a_j^T & 0 \end{bmatrix} X) + a_j^T a_j = d_{ij},\ \forall (i,j) \in N_a \\ & X \in \mathcal{M}^{nr}, \end{array} \tag{13}$$

where N_x, N_a refers to sets of known distances. This formulation is a $\boldsymbol{QMP_1}$, so we can develop both its $\boldsymbol{VSDR_1}$ and $\boldsymbol{MSDR_1}$ relaxations.

$$\begin{array}{ll} \min & 0 \\ s.t. & I \otimes (E_{ii} + E_{jj} - 2E_{ij}) \cdot Y = d_{ij},\ \forall\ (i,j) \in N_x \\ & I \otimes E_{ii} \cdot Y - 2a_j^T x_i + a_j^T a_j = d_{ij},\ \forall\ (i,j) \in N_a \\ & \begin{bmatrix} 1 & x^T \\ x & Y \end{bmatrix} \succeq 0 \end{array} \tag{14}$$

$$\begin{array}{ll} \min & 0 \\ s.t. & (E_{ii} + E_{jj} - 2E_{ij}) \cdot Y = d_{ij},\ \forall\ (i,j) \in N_x \\ & E_{ii} \cdot Y - 2\begin{bmatrix} a_j^T & 0 \end{bmatrix} \cdot X + a_j^T a_j = d_{ij},\ \forall (i,j) \in N_a \\ & \begin{bmatrix} I & X^T \\ X & Y \end{bmatrix} \succeq 0 \end{array} \tag{15}$$

Theorem 2.1 implies that the $\boldsymbol{MSDR_1}$ relaxation always provides the same lower bound as the $\boldsymbol{VSDR_1}$ one, while the number of variables for $\boldsymbol{MSDR_1}$ ($O((n+d)^2)$) is significantly smaller than the number for $\boldsymbol{VSDR_1}$ ($O(n^2d^2)$). The quality of the bounds combined with a lower computational complexity explains why $\boldsymbol{MSDR_1}$ is a favourite relaxation for researchers.

3 Quadratic Matrix Programming: Case II

In this section, we move to the main topic of our paper, i.e., the equivalence of the vector and matrix relaxations for the more general $\boldsymbol{QMP_2}$.

3.1 Equivalence of Vector and Matrix Lifting for QMP_2

We first propose the Vector-Lifting Semidefinite Relaxation $\boldsymbol{VSDR_2}$ for $\boldsymbol{QMP_2}$. From applying both equations in (4), we get the following.

$$(\boldsymbol{VSDR_2}) \qquad \begin{array}{rll} \mu^*_{V2} := & \min & \begin{bmatrix} \beta_0 & \operatorname{vec}(C_0)^T \\ \operatorname{vec}(C_0) & I_r \otimes Q_0 + P_0 \otimes I_n \end{bmatrix} \cdot Z_V \\ & \text{s.t.} & \begin{bmatrix} \beta_j & \operatorname{vec}(C_j)^T \\ \operatorname{vec}(C_j) & I_r \otimes Q_j + P_j \otimes I_n \end{bmatrix} \cdot Z_V \le 0, \;\; j = 1, 2, \ldots, m \\ & & (Z_V)_{1,1} = 1 \\ & & Z_V \in \mathcal{S}^{rn+1}_+ \qquad \left(Z_V = \begin{bmatrix} 1 & \operatorname{vec}(X)^T \\ \operatorname{vec}(X) & Y_V \end{bmatrix} \right) \end{array}$$

Matrix Y_V is $nr \times nr$ and can be partitioned into exactly r^2 block matrices Y_V^{ij}, $i, j = 1, 2, \ldots, r$, where each block is $n \times n$. From applying both equations in (5), we get the smaller Matrix-Lifting Semidefinite Relaxation $\boldsymbol{MSDR_2}$ for $\boldsymbol{QMP_2}$. (We add the additional constraint $\operatorname{trace} Y_1 = \operatorname{trace} Y_2$ since $\operatorname{trace} XX^T = \operatorname{trace} X^T X$.)

$$(\boldsymbol{MSDR_2}) \qquad \begin{array}{rll} \mu^*_{M2} := & \min & Q_0 \cdot Y_1 + P_0 \cdot Y_2 + 2C_0 \cdot X + \beta_0 \\ & \text{s.t.} & Q_j \cdot Y_1 + P_j \cdot Y_2 + 2C_j \cdot X + \beta_j \le 0, \;\; j = 1, 2, \ldots, m \\ & & Y_1 - XX^T \in \mathcal{S}^n_+ \qquad \left(Z_1 := \begin{bmatrix} I_r & X^T \\ X & Y_1 \end{bmatrix} \succeq 0 \right) \\ & & Y_2 - X^T X \in \mathcal{S}^r_+ \qquad \left(Z_2 := \begin{bmatrix} I_n & X \\ X^T & Y_2 \end{bmatrix} \succeq 0 \right) \\ & & \operatorname{trace} Y_1 = \operatorname{trace} Y_2. \end{array}$$

$\boldsymbol{VSDR_2}$ has $O((nr)^2)$ variables, whereas $\boldsymbol{MSDR_2}$ has only $O((n+r)^2)$ variables. The computational advantage of using the smaller problem $\boldsymbol{MSDR_2}$ motivates the comparison of the corresponding bounds. The main result is interesting and surprising, i.e., that $\boldsymbol{VSDR_2}$ and $\boldsymbol{MSDR_2}$ actually generate the same bound under a constraint qualification assumption. In general, the bound from $\boldsymbol{MSDR_2}$ is at least as strong as the bound from $\boldsymbol{VSDR_2}$.

Define the block-diag and block-offdiag transformations, respectively,

$$\mathrm{B}^0\mathrm{Diag}(Q) : \mathcal{S}^n \to \mathcal{S}^{rn+1}, \quad \mathrm{O}^0\mathrm{Diag}(P) : \mathcal{S}^r \to \mathcal{S}^{rn+1},$$

$$\mathrm{B}^0\mathrm{Diag}(Q) := \begin{bmatrix} 0 & 0 \\ 0 & I_r \otimes Q \end{bmatrix}, \quad \mathrm{O}^0\mathrm{Diag}(P) := \begin{bmatrix} 0 & 0 \\ 0 & P \otimes I_n \end{bmatrix}.$$

(See [35] for the $r = n$ case.) It is clear that $Q, P \succeq 0$ implies that both $\mathrm{B}^0\mathrm{Diag}(Q) \succeq 0, \mathrm{O}^0\mathrm{Diag}(P) \succeq 0$. The adjoints $\mathrm{b}^0\mathrm{diag}, \mathrm{o}^0\mathrm{diag}$ are, respectively,

$$\begin{array}{rcl} Y_1 & = & \mathrm{B}^0\mathrm{Diag}^*(Z_V) = \mathrm{b}^0\mathrm{diag}(Z_V) := \sum_{j=1}^r Y_V^{jj}, \\ Y_2 & = & \mathrm{O}^0\mathrm{Diag}^*(Z_V) = \mathrm{o}^0\mathrm{diag}(Z_V) := (\operatorname{trace} Y_V^{ij})_{i,j=1,2,\ldots,r}. \end{array} \tag{16}$$

Lemma 3.1. *Let $X \in \mathcal{M}^{nr}$ be given. Suppose that one of the following two conditions hold.*

1. *Let Y_v be given and Z_V defined as in $\boldsymbol{VSDR_2}$. Let the pair Z_1, Z_2 in $\boldsymbol{MSDR_2}$ be constructed using*

$$Y_1 = \mathrm{b}^0\mathrm{diag}(Z_V), Y_2 = \mathrm{o}^0\mathrm{diag}(Z_V). \tag{17}$$

2. *Let Y_1, Y_2 be given with* $\operatorname{trace} Y_1 = \operatorname{trace} Y_1$, *and let Z_1, Z_2 be defined as in* $\boldsymbol{MSDR_2}$. *Let Y_V, Z_V for* $\boldsymbol{VSDR_2}$ *be constructed from Y_1, Y_2 as follows.*

$$Y_V = \begin{bmatrix} V_1 & \frac{1}{n}(Y_2)_{12} I_n & \dots & \frac{1}{n}(Y_2)_{1r} I_n \\ \frac{1}{n}(Y_2)_{12} I_n & V_2 & \dots & \frac{1}{n}(Y_2)_{2r} I_n \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \frac{1}{n}(Y_2)_{(r-1)r} I_n \\ \dots & \dots & \dots & V_r \end{bmatrix}, \tag{18}$$

with

$$\sum_{i=1}^{r} V_i = Y_1, \quad \operatorname{trace} V_i = (Y_2)_{ii}, i = 1, \dots, r. \tag{19}$$

Then, Z_v satisfies the linear inequality constraints in $\boldsymbol{VSDR_2}$ *if, and only if, Z_1, Z_2 satisfies the linear inequality constraints in* $\boldsymbol{MSDR_2}$. *Moreover, the values of the objective functions with the corresponding variables are equal.*

Proof. 1. Note that

$$\begin{aligned} \begin{bmatrix} \beta & \operatorname{vec}(C)^T \\ \operatorname{vec}(C) & I_r \otimes Q + P \otimes I_n \end{bmatrix} \cdot Z_V &= \beta + 2C \cdot X + \left(\mathrm{B}^0\mathrm{Diag}(Q) + \mathrm{O}^0\mathrm{Diag}(P)\right) \cdot Z_V \\ &= \beta + 2C \cdot X + Q \cdot \mathrm{b}^0\mathrm{diag}(Z_V) + P \cdot \mathrm{o}^0\mathrm{diag}(Z_V) \\ &= \beta + 2C \cdot X + Q \cdot Y_1 + P \cdot Y_2, \quad \text{by (17).} \end{aligned} \tag{20}$$

2. Conversely, we note that $\operatorname{trace} Y_1 = \operatorname{trace} Y_2$ is a constraint in $\boldsymbol{MSDR_2}$. And, Z_V as constructed using (18) satisfies (17). In addition, the $n + r$ assignment type constraints in (19) on the rn variables in the diagonals of the $V_i, i = 1, \dots, r$, can always be solved. We can now apply the argument in (20) again.

□

Lemma 3.1 guarantees the equivalence of the feasible sets of the two relaxations with respect to the linear inequality constraints and the objective function. However, this ignores the semidefinite constraints. The following result partially addresses this deficiency.

Corollary 3.1. *If the feasible set of* $\boldsymbol{VSDR_2}$ *is nonempty, then the feasible set of* $\boldsymbol{MSDR_2}$ *is also nonempty and*

$$\mu^*_{M2} \le \mu^*_{V2}. \tag{21}$$

Proof. Suppose $Z_V = \begin{bmatrix} 1 & \operatorname{vec}(X)^T \\ \operatorname{vec}(X) & Y_V \end{bmatrix}$ is feasible for $\boldsymbol{VSDR_2}$. Recall that Matrix Y_V is $nr \times nr$ and can be partitioned into exactly r^2 block matrices Y_V^{ij}, $i, j = 1, 2, \dots, r$. As above, we set Y_1, Y_2 following (17), and we set $Z_1 = \begin{bmatrix} I_r & X^T \\ X & Y_1 \end{bmatrix}$, $Z_2 = \begin{bmatrix} I_n & X \\ X^T & Y_2 \end{bmatrix}$.

Denote the j-th column of X by $X_{:j}$, $j = 1, 2, \dots, r$. Now $Z_V \succeq 0$ implies $Y_V^{jj} - X_{:j} X_{:j}^T \succeq 0$. Therefore, $\sum_{j=1}^r Y_V^{jj} - \sum_{j=1}^r X_{:j} X_{:j}^T = Y_1 - X X^T \succeq 0$, i.e. $Z_1 \succeq 0$. Similarly, denote the k-th row of X by $X_{k:}$, $k = 1, 2, \dots, n$. Let $(Y_V^{ij})_{kk}$ denote the k-th diagonal entry of Y_V^{ij}, and define the $r \times r$ matrix $Y^k := ((Y_V^{ij})_{kk})_{i,j=1,2,\cdots,r}$. Then $Z_V \succeq 0$ implies $Y^k - X_{k:}^T X_{k:} \succeq 0$. Therefore, $\sum_{k=1}^n Y^k - \sum_{k=1}^n X_{k:}^T X_{k:} = Y_2 - X^T X \succeq 0$, i.e. $Z_2 \succeq 0$. The proof now follows from Lemma 3.1. □

Corollary 3.1 holds because $\boldsymbol{MSDR_2}$ only restricts the sum of some principal submatrices of Z_V (i.e., $\mathrm{b^0diag}(Z_V)$, $\mathrm{o^0diag}(Z_V)$) to be positive semidefinite; while $\boldsymbol{VSDR_2}$ restricts the whole matrix Z_V positive semidefinite. So the semidefinite constraints in $\boldsymbol{MSDR_2}$ are not as strong as in $\boldsymbol{VSDR_2}$. Moreover, the entries of Y_V involved in $\mathrm{b^0diag}(\bullet), \mathrm{o^0diag}(\bullet)$ form a partial semidefinite matrix which is not chordal and does not necessarily have a semidefinite completion. Therefore, the semidefinite completion technique we used to prove the equivalence between $\boldsymbol{VSDR_1}$ and $\boldsymbol{MSDR_1}$ is not applicable here. Instead, we will prove the equivalence of their dual programs. It is well known that the primal equals the dual when the generalized Slater condition holds [17, 28], and in this case we will then conclude that $\boldsymbol{VSDR_2}$ and $\boldsymbol{MSDR_2}$ generate the same bound.

Definition 3.1. *For $\lambda \in \mathbb{R}^m$, let:*

$$\beta_\lambda := \beta_0 + \sum_{j=1}^{m} \lambda_j \beta_j;$$

and let $C_\lambda, Q_\lambda, P_\lambda$ be defined similarly.

After substituting $\alpha \leftarrow \beta_\lambda - \alpha$, we see that the dual of $\boldsymbol{VSDR_2}$ is equivalent to

$$(\boldsymbol{DVSDR_2}) \qquad \begin{array}{ll} \max & \beta_\lambda - \alpha \\ \text{s.t.} & \begin{bmatrix} \alpha & \mathrm{vec}(C_\lambda)^T \\ \mathrm{vec}(C_\lambda) & I_r \otimes Q_\lambda + P_\lambda \otimes I_n \end{bmatrix} \succeq 0 \\ & \alpha \in \mathbb{R}, \lambda \in \mathbb{R}^m_+. \end{array}$$

And, the dual of $\boldsymbol{MSDR_2}$ is

$$(\boldsymbol{DMSDR_2}) \qquad \begin{array}{ll} \max & \beta_\lambda - \operatorname{trace} S_1 - \operatorname{trace} S_2 \\ \text{s.t.} & \begin{bmatrix} S_1 & R_1^T \\ R_1 & Q_\lambda - tI_n \end{bmatrix} \succeq 0 \\ & \begin{bmatrix} S_2 & R_2 \\ R_2^T & P_\lambda + tI_r \end{bmatrix} \succeq 0 \\ & R_1 + R_2 = C_\lambda \\ & \lambda \in \mathbb{R}^m_+, S_1 \in S^r, S_2 \in S^n, R_1, R_2 \in \mathcal{M}^{nr}, t \in R. \end{array}$$

The Slater condition for $\boldsymbol{DVSDR_2}$ is equivalent to the following:

$$\exists \lambda \in \mathbb{R}^m_+, \text{ s.t. } I_r \otimes Q_\lambda + P_\lambda \otimes I_n \succ 0. \tag{22}$$

The corresponding constraint qualification condition for $\boldsymbol{DMSDR_2}$ is

$$\exists t \in \mathbb{R}, \ \lambda \in \mathbb{R}^m_+, \text{ s.t. } Q_\lambda - tI_r \succ 0, P_\lambda + tI_n \succ 0. \tag{23}$$

These two conditions are equivalent due to the following lemma, which will also be used in our subsequent analysis.

Lemma 3.2. *Let $Q \in S^n$, $P \in S^r$. Then*

$$I_r \otimes Q + P \otimes I_n \succ 0, \quad (\textit{resp.} \ \succeq 0)$$

if, and only if,

$$\exists t \in \mathbb{R}, \ \textit{s.t.} \ Q - tI_n \succ 0, \ P + tI_r \succ 0, \quad (\textit{resp.} \ \succeq 0.)$$

Proof. Suppose the symmetric matrices Q, P have spectral decomposition

$$Q = V\Lambda_Q V^T, \quad P = U\Lambda_P U^T,$$

where V, U are orthogonal, and the eigenvalues are in the diagonal matrices

$$\Lambda_Q = \mathrm{Diag}\left(\begin{pmatrix}\lambda_1(Q) & \lambda_2(Q) & \dots & \lambda_n(Q)\end{pmatrix}\right), \quad \Lambda_P = \mathrm{Diag}\left(\begin{pmatrix}\lambda_1(P) & \lambda_2(P) & \dots & \lambda_r(P)\end{pmatrix}\right).$$

Therefore, we get the equivalences: $I_r \otimes Q + P \otimes I_n \succ 0$ if, and only if, $(V \otimes U)(I_r \otimes \Lambda_P + P \otimes I_n)(V \otimes U)^T \succ 0$ if, and only if, $I_r \otimes \Lambda_Q + \Lambda_P \otimes I_n \succ 0$ if, and only if, $\min_i \lambda_i(Q) + \min_j \lambda_j(P) > 0$ if, and only if,

$$\min_i \lambda_i(Q) - t > 0, \min_j \lambda_j(P) + t > 0, \quad \text{for some } t \in \mathbb{R}.$$

The equivalences hold if the strict inequalities, $\succ 0$ and $>$ are replaced by the inequalities $\succeq 0$ and $\geq$, respectively. $\square$

Now we state the main theorem of this paper on the equivalence of the two $\boldsymbol{SDP}$ relaxations for $\boldsymbol{QMP_2}$.

Theorem 3.1. *Suppose that $\boldsymbol{DVSDR_2}$ is strictly feasible. As numbers in the extended real line $(-\infty, +\infty]$, the optimal values of the two relaxations $\boldsymbol{VSDR_2}$, $\boldsymbol{MSDR_2}$, obtained using vector and matrix-liftings, are equal, i.e.,*

$$\mu^*_{V2} = \mu^*_{M2}.$$

3.1.1 Proof of (Main) Theorem 3.1

Since $\boldsymbol{DVSDR_2}$ is strictly feasible, Lemma 3.2 implies that both dual programs satisfy constraint qualifications. Therefore, both programs satisfy strong duality, see e.g., [28]. Therefore, both have zero duality gaps, i.e., the optimal values of $\boldsymbol{DVSDR_2}$, $\boldsymbol{DMSDR_2}$, are μ^*_{V2}, μ^*_{M2}, respectively.

Now assume that

$$\lambda \text{ is feasible for } \boldsymbol{DVSDR_2}. \tag{24}$$

Lemma 3.2 implies that λ is also feasible for $\boldsymbol{DMSDR_2}$, i.e., there exists $t \in \mathbb{R}$ such that

$$Q := Q_\lambda - tI_n \succeq 0, \quad P := P_\lambda + tI_r \succeq 0. \tag{25}$$

(To simplify notation, we use Q, P to denote these dual slack matrices.) The spectral decomposition of Q, P can be expressed as

$$Q = V\Lambda_Q V^T = [V_1 \ V_2] \begin{bmatrix} \Lambda_{Q+} & 0 \\ 0 & 0 \end{bmatrix} [V_1 \ V_2]^T, \quad P = U\Lambda_P U^T = [U_1 \ U_2] \begin{bmatrix} \Lambda_{P+} & 0 \\ 0 & 0 \end{bmatrix} [U_1 \ U_2]^T,$$

where the columns of the submatrices U_1, V_1 form an orthonormal basis that spans the range spaces $\mathcal{R}(P)$ and $\mathcal{R}(Q)$, respectively; while the columns of U_2, V_2 span the orthogonal complements $\mathcal{R}(P)^\perp$ and $\mathcal{R}(Q)^\perp$, respectively. Λ_{Q+} is a diagonal matrix where diagonal entries are nonzero eigenvalues of matrix Q, and Λ_{P+} is defined similarly. Let $\{\sigma_i\}$, $\{\theta_i\}$ denote the eigenvalues of P, Q, respectively.

We similarly simplify the notation

$$C := C_\lambda, \quad c := \mathrm{vec}(C), \quad \beta := \beta_\lambda. \tag{26}$$

Let $A^\dagger$ denote the *Moore-Penrose pseudoinverse* of matrix A. The following lemma allows us to express μ^*_{V2} as a function of Q, P, c and β.

Lemma 3.3. *Let* λ, P, Q, c, β *be defined as above in* (24), (25), (26)*. Let*

$$\alpha^* := c^T((I_r \otimes Q + P \otimes I_n)^\dagger)c. \tag{27}$$

Then, α^*, λ *is a feasible pair for* $\boldsymbol{DVSDR_2}$*. And, for any pair* α, λ *feasible to* $\boldsymbol{DVSDR_2}$*, we have*

$$-\alpha + \beta \leq -\alpha^* + \beta.$$

Proof. A general quadratic function $f(x) = x^T\bar{Q}x + 2\bar{c}^Tx + \bar{\beta}$ is nonnegative for any $x \in R^n$ if, and only if, the matrix $\begin{bmatrix} \bar{\beta} & \bar{c}^T \\ \bar{c} & \bar{Q} \end{bmatrix} \succeq 0$, e.g., [8, Pg. 163]. Therefore,

$$\begin{bmatrix} \alpha & c^T \\ c & I_r \otimes Q_\lambda + P_\lambda \otimes I_n \end{bmatrix} = \begin{bmatrix} \alpha & c^T \\ c & I_r \otimes Q + P \otimes I_n \end{bmatrix} \succeq 0 \tag{28}$$

if, and only if,

$$x^T(I_r \otimes Q + P \otimes I_n)x + 2c^Tx + \alpha \geq 0, \quad \forall x \in \mathbb{R}^{nr}.$$

For a fixed α, this is further equivalent to

$$\begin{array}{rcl} -\alpha & \leq & \min\limits_x \; x^T(I_r \otimes Q + P \otimes I_n)x + 2c^Tx \\ & = & -c^T((I_r \otimes Q + P \otimes I_n)^\dagger)c. \end{array}$$

Therefore, we can choose α^* as in (27). □

To further explore the structure of (27), we note that c can be decomposed as

$$c = (U_1 \otimes V_1)r_{11} + (U_1 \otimes V_2)r_{12} + (U_2 \otimes V_1)r_{21} + (U_2 \otimes V_2)r_{22}. \tag{29}$$

The validity of such an expression follows from the fact that the columns of $[U_1\ U_2] \otimes [V_1\ V_2]$ form an orthonormal basis of $\mathbb{R}^{nr}$. Furthermore, the dual feasibility of $\boldsymbol{DVSDR_2}$ includes the constraint $\begin{bmatrix} \alpha & c^T \\ c & I_r \otimes Q + P \otimes I_n \end{bmatrix} \succeq 0$, which implies $c \in \mathcal{R}(I_r \otimes Q + P \otimes I_n)$. This range space is spanned by the columns in the matrices $U_1 \otimes V_1$, $U_2 \otimes V_1$ and $U_1 \otimes V_2$, which implies that c has no component in $\mathcal{R}(U_2 \otimes V_2)$, i.e., $r_{22} = 0$ in (29).

The following lemma provides a key observation for the connections between the two dual programs. It deduces that if c is in $\mathcal{R}(U_1 \otimes V_1)$, then the α component of the objective value of $\boldsymbol{DVSDR_2}$ in Lemma 3.3 has a specific representation.

Lemma 3.4. *If* $c \in \mathcal{R}(U_1 \otimes V_1)$*, then the* α *component of the objective value of* $\boldsymbol{DVSDR_2}$ *in Lemma 3.3 satisfies*

$$\begin{array}{rcl} -\alpha & = & -c^T((I_r \otimes Q + P \otimes I_n)^\dagger)c \\ & = & \left\{ \begin{array}{ll} \max & -\operatorname{vec}(R_1)^T(I_r \otimes Q)^\dagger \operatorname{vec}(R_1) - \operatorname{vec}(R_2)^T(P \otimes I_n)^\dagger \operatorname{vec}(R_2) \\ s.t. & R_1 + R_2 = C \\ & R_1, R_2 \ \in \mathcal{M}^{nr} \end{array} \right. \end{array} \tag{30}$$

Proof. We can eliminate R_2 and express the maximization problem on the right hand side of the equality as $\max_{R_1} \phi(R_1)$, where

$$\begin{array}{rcl}\phi(R_1) & := & -\operatorname{vec}(R_1)^T((I_r \otimes Q)^\dagger + (P \otimes I_n)^\dagger)\operatorname{vec}(R_1) \\ & & +2c^T(P \otimes I_n)^\dagger \operatorname{vec}(R_1) - c^T(P \otimes I_n)^\dagger c.\end{array} \tag{31}$$

Since P, Q are both positive semidefinite, we know $(I_r \otimes Q)^\dagger + (P \otimes I_n)^\dagger \succeq 0$. Hence ϕ is concave. It is not difficult to verify that $(P \otimes I_n)^\dagger c \in \mathcal{R}((I_r \otimes Q)^\dagger + (P \otimes I_n)^\dagger)$. Therefore, the maximum of the quadratic concave function $\phi(R_1)$ is finite and attained at R_1^*,

$$\begin{array}{rcl}\operatorname{vec}(R_1^*) & = & ((I_r \otimes Q)^\dagger + (P \otimes I_n)^\dagger)^\dagger (P \otimes I_n)^\dagger c \\ & = & (P \otimes Q^\dagger + PP^\dagger \otimes I_n)^\dagger c;\end{array} \tag{32}$$

and this corresponds to a value

$$\begin{array}{rcl}\phi(R_1^*) & = & c^T((P \otimes Q^\dagger + P^\dagger P \otimes I_n)^\dagger (P \otimes I_n)^\dagger - (P \otimes I_n)^\dagger)c \\ & = & -c^T(U \otimes V)\hat{\Lambda}(U \otimes V)^T c,\end{array}$$

where

$$\hat{\Lambda} := \Lambda_P^\dagger \otimes I_n - (\Lambda_P^2 \otimes \Lambda_Q^\dagger + \Lambda_P \otimes I_n)^\dagger.$$

Matrix $\hat{\Lambda}$ is diagonal. Its diagonal entries can be calculated as

$$\hat{\Lambda}(i,j) = \begin{cases} \frac{1}{\sigma_i + \theta_j} & \text{if } \sigma_i > 0, \theta_j > 0 \\ 0 & \text{if } \sigma_i = 0 \text{ or } \theta_j = 0. \end{cases}$$

We now compare $\phi(R_1^*)$ with $-c^T(I_r \otimes Q + P \otimes I_n)^\dagger c$. Let

$$\begin{array}{rcl}\bar{\Lambda} & := & (I_r \otimes Q + P \otimes I_n)^\dagger = (U \otimes V)(I_r \otimes \Lambda_Q + \Lambda_P \otimes I_n)^\dagger (U \otimes V)^T \\ & = & (U \otimes V)((I_{\sigma_+} \otimes \Lambda_Q + \Lambda_P \otimes I_{\theta_+})^\dagger + I_{\sigma_0} \otimes \Lambda_Q^\dagger + \Lambda_P^\dagger \otimes \mathcal{I}_{\theta_0})(U \otimes V)^T,\end{array} \tag{33}$$

where matrix I_{σ_+} (resp. I_{σ_0}) is $r \times r$, diagonal, and zero, except for the i-th diagonal entries that are equal to one if $\sigma_i > 0$ (resp. $\sigma_i = 0$); and, matrix I_{θ_+} (resp. I_{θ_0}) is defined in the same way. Hence we know that matrix $\bar{\Lambda}$ is also diagonal. Its diagonal entries can be calculated as

$$\bar{\Lambda}(i,j) = \begin{cases} \frac{1}{\sigma_i + \theta_j} & \text{if } \sigma_i > 0, \theta_j > 0 \\ \frac{1}{\sigma_i} & \text{if } \sigma_i > 0, \theta_j = 0 \\ \frac{1}{\theta_j} & \text{if } \sigma_i = 0, \theta_j > 0 \\ 0 & \text{if } \sigma_i = 0, \theta_j = 0. \end{cases} \tag{34}$$

By assumption, $c = (U_1 \otimes V_1)r_{11}$, for some r_{11} of appropriate size. Note that $(U_1 \otimes V_1)r_{11}$ is orthogonal to the columns in $U_2 \otimes V_1$ and $U_1 \otimes V_2$. Thus only the part $(I_{\sigma_+} \otimes \Lambda_Q + \Lambda_P \otimes I_{\theta_+})^\dagger$ in the diagonal matrix is involved in computations, i.e.,

$$\begin{array}{rcl}-c^T(I_r \otimes Q + P \otimes I_n)^\dagger c & = & -r_{11}^T(U_1 \otimes V_1)^T(U \otimes V)\bar{\Lambda}(U \otimes V)^T(U_1 \otimes V_1)r_{11} \\ & = & -r_{11}^T(U_1 \otimes V_1)^T(U \otimes V)(I_{\sigma_+} \otimes \Lambda_Q + \Lambda_P \otimes I_{\theta_+})^\dagger(U \otimes V)^T(U_1 \otimes V_1)r_{11} \\ & = & -r_{11}^T(U_1 \otimes V_1)^T(U \otimes V)\hat{\Lambda}(U \otimes V)^T(U_1 \otimes V_1)r_{11} \\ & = & \phi(R_1^*).\end{array}$$

□

Now, for the given feasible α^*, λ of Lemma 3.3, we will construct a feasible solution for $\boldsymbol{DMSDR_2}$ that generates the same objective value. Using Lemma 3.2, we choose $t \in \mathbb{R}$ satisfying $Q = Q_\lambda - tI_n \succeq 0$, $P = P_\lambda + tI_r \succeq 0$.

We can now find a lower bound for the optimal value of $\boldsymbol{DMSDR_2}$.

Proposition 3.1. *Let $\lambda, t, P, Q, C, c, \beta$ be as above. Let R_1^* denote the maximizer of $\phi(R_1)$ in the proof of Lemma 3.4. Construct R_1 R_2 as follows:*

$$\begin{array}{rcl} \operatorname{vec}(R_1) & = & \operatorname{vec}(R_1^*) + (U_2 \otimes V_1) r_{21} \\ \operatorname{vec}(R_2) & = & \operatorname{vec}(R_2^*) + (U_1 \otimes V_2) r_{12}. \end{array} \tag{35}$$

Then we obtain a lower bound for the optimal value of $\boldsymbol{DMSDR_2}$.

$$\mu_{M2}^* \geq -\operatorname{vec}(R_1)^T (I_r \otimes Q)^\dagger \operatorname{vec}(R_1) - \operatorname{vec}(R_2)^T (P \otimes I_n)^\dagger \operatorname{vec}(R_2) + \beta. \tag{36}$$

Proof. Consider the subproblem that maximizes the objective with λ, t, R_1 and R_2 defined as above.

$$\begin{array}{ll} \max & \beta - \operatorname{trace} S_1 - \operatorname{trace} S_2 \\ \text{s.t.} & \begin{bmatrix} S_1 & R_1^T \\ R_1 & Q \end{bmatrix} \succeq 0 \\ & \begin{bmatrix} S_2 & R_2 \\ R_2^T & P \end{bmatrix} \succeq 0 \\ & S_1 \in S_r \ S_2 \in S_n, \end{array} \tag{37}$$

Because λ, t, R_1, and R_2 are all feasible for $\boldsymbol{DMSDR_2}$, this subproblem will generate a lower bound for μ_{M2}^*. We now invoke a result from [6], i.e., that there exists a symmetric matrix S such that $\operatorname{trace} S \leq \gamma$ and $\begin{bmatrix} S & C^T \\ C & Q \end{bmatrix} \succeq 0$ if, and only if, $f(X) = \operatorname{trace}(X^T Q X + 2C^T X) + \gamma \geq 0$ for any $X \in \mathcal{M}^{nr}$. This is equivalent to

$$-\gamma \leq \min_{X \in \mathcal{M}^{nr}} \operatorname{trace}(X^T Q X + 2C^T X) = -\operatorname{trace}(C^T Q^\dagger C).$$

Therefore, the subproblem (37) can be reformulated as

$$\begin{array}{ll} \max & \beta - \operatorname{trace} S_1 - \operatorname{trace} S_2 \\ \text{s.t.} & -\operatorname{trace} S_1 \leq -\operatorname{trace}(R_1^T Q^\dagger R_1) \\ & -\operatorname{trace} S_2 \leq -\operatorname{trace}(R_2 Q^\dagger R_2^T) \\ & S_1 \in S^r, S_2 \in S^n. \end{array} \tag{38}$$

Hence, the optimal value of (38) has an explicit expression and provides a lower bound for $\boldsymbol{DMSDR_2}$

$$\mu_{M2}^* \geq -\operatorname{vec}(R_1)^T (I_r \otimes Q)^\dagger \operatorname{vec}(R_1) - \operatorname{vec}(R_2)^T (P \otimes I_n)^\dagger \operatorname{vec}(R_2) + \beta.$$

□

With all the above preparations, we now complete the proof of (main) Theorem 3.1.

Proof. (of Theorem 3.1) Now we compare μ_{V2}^* from the expression (27) with μ_{M2}^* based on the lower bound expression in (36). By writing c in the form of (29), we get

$$\begin{aligned}\mu_{v2}^* \;=\; & -\left[r_{11}^T(U_1\otimes V_1)^T + r_{12}^T(U_1\otimes V_2)^T + r_{21}^T(U_2\otimes V_1)^T\right] \\ & \quad (I_r\otimes Q + P\otimes I_n)^\dagger \\ & \quad\quad [(U_1\otimes V_1)r_{11} + (U_1\otimes V_2)r_{12} + (U_2\otimes V_1)r_{21}] \\ & \quad\quad\quad +\beta.\end{aligned} \tag{39}$$

Consider the crossterm such as $r_{11}^T(U_1\otimes V_1)^T(I_r\otimes Q+P\otimes I_n)^\dagger(U_1\otimes V_2)r_{12}$. Since $(U_1\otimes V_1)r_{11}$ is orthogonal to $(U_1\otimes V_2)r_{12}$, and $(I_r\otimes Q+P\otimes I_n)^\dagger$ is diagonalizable by $[U_1\ U_2]\otimes[V_1\ V_2]$, this term is actually zero. Similarly, we can verify that the other crossterms equal zero. As a result, only the following sum of three quadratic terms remain, which we label using $C1, C2, C3$, respectively.

$$\begin{aligned}\mu_{v2}^* \;=\; & -r_{11}^T(U_1\otimes V_1)^T(I_r\otimes Q+P\otimes I_n)^\dagger(U_1\otimes V_1)r_{11} \\ & -r_{21}^T(U_2\otimes V_1)^T(I_r\otimes Q+P\otimes I_n)^\dagger(U_2\otimes V_1)r_{21} \\ & -r_{12}^T(U_1\otimes V_2)^T(I_r\otimes Q+P\otimes I_n)^\dagger(U_1\otimes V_2)r_{12}+\beta \\ =: \; & C1+C2+C3+\beta.\end{aligned} \tag{40}$$

We can also formulate the lower bound for μ_{M2}^* based on (36):

$$\begin{aligned}\mu_{M2}^* \;\geq\; & -\operatorname{vec}(R_1)^T(I_r\otimes Q)^\dagger\operatorname{vec}(R_1) - \operatorname{vec}(R_2)^T(P\otimes I_n)^\dagger\operatorname{vec}(R_2)+\beta \\ = \; & -(\operatorname{vec}(R_1^*)+(U_2\otimes V_1)r_{21})^T(I_r\otimes Q)^\dagger\operatorname{vec}(R_1^*+(U_2\otimes V_1)r_{21}) \\ & -(\operatorname{vec}(R_2^*)+(U_1\otimes V_2)r_{12})^T(P\otimes I_n)^\dagger(\operatorname{vec}(R_2^*)+(U_1\otimes V_2)r_{12})+\beta.\end{aligned} \tag{41}$$

Since $\operatorname{vec}(R_1^*)$ and $\operatorname{vec}(R_2^*)$ are both in $\mathcal{R}(U_1\otimes V_1)$, and this is orthogonal to both $(U_1\otimes V_2)r_{12}$ and $(U_2\otimes V_1)r_{21}$, and both matrices $(I_r\otimes Q)^\dagger$ and $(P\otimes I_n)^\dagger$ are diagonalizable by $[U_1\ U_2]\otimes[V_1\ V_2]$, we conclude that the crossterms, such as $\operatorname{vec}(R_1^*)^T(I_r\otimes Q)^\dagger(U_2\otimes V_1)r_{21}$, all equal zero. Therefore, the lower bound for μ_{M2}^* can be reformulated as

$$\begin{aligned}\mu_{M2}^* \;\geq\; & -(\operatorname{vec}(R_1^*)(I_r\otimes Q)^\dagger\operatorname{vec}(R_1^*)+\operatorname{vec}(R_2^*)(P\otimes I_n)^\dagger\operatorname{vec}(R_2^*)) \\ & -r_{21}^T(U_2\otimes V_1)^T(I_r\otimes Q)^\dagger(U_2\otimes V_1)r_{21} \\ & -r_{12}^T(U_1\otimes V_2)^T(P\otimes I_n)^\dagger(U_1\otimes V_2)r_{12}+\beta. \\ =: \; & T1+T2+T3+\beta.\end{aligned} \tag{42}$$

As above, denote the first three quadratic terms by $T1, T2, T3$, respectively.

We will show that term $C1$, $C2$ and $C3$ equal $T1$, $T2$ and $T3$, respectively. The equality between $C1$ and $T1$ follows from Lemma 3.4. For the other terms, consider $C2$ first. Write $(I_r\otimes Q+P\otimes I_n)^\dagger$ as the diagonal matrix $\bar{\Lambda}$ by (33). Note that $(U_2\otimes V_1)r_{21}$ is orthogonal with the columns in $U_1\otimes V_1$ and $U_1\otimes V_2$. Thus only the part $I_{\sigma_0}\otimes\Lambda_Q^\dagger$ in diagonal matrix $\bar{\Lambda}$ is involved in computing term $C2$, i.e.,

$$\begin{aligned} & -r_{21}^T(U_2\otimes V_1)^T(I_r\otimes Q+P\otimes I_n)^\dagger(U_2\otimes V_1)r_{21} \\ = \; & -r_{21}^T(U_2\otimes V_1)^T(U\otimes V)(I_{\sigma_0}\otimes\Lambda_Q^\dagger)(U\otimes V)^T(U_2\otimes V_1)r_{21}.\end{aligned} \tag{43}$$

Similarly, since $(U_2\otimes V_1)r_{21}$ is orthogonal with eigenvectors in $U_1\otimes V_2$, we have

$$\begin{aligned} & -r_{21}^T(U_2\otimes V_1)^T(I_r\otimes Q)^\dagger(U_2\otimes V_1)r_{21} \\ = \; & -r_{21}^T(U_2\otimes V_1)^T(U\otimes V)(I_{\sigma_+}\otimes\Lambda_Q^\dagger+I_{\sigma_0}\otimes\Lambda_Q^\dagger)(U\otimes V)^T(U_2\otimes V_1)r_{21} \\ = \; & -r_{21}^T(U_2\otimes V_1)^T(U\otimes V)(I_{\sigma_0}\otimes\Lambda_Q^\dagger)(U\otimes V)^T(U_2\otimes V_1)r_{21}.\end{aligned} \tag{44}$$

By (43) and (44), we conclude that that the term $C2$ equals $T2$. We may use the same argument to prove the equality of $C3$ and $T3$. Therefore, we conclude that

$$\begin{array}{rl} & -c^T((I_r \otimes Q + P \otimes I_n)^\dagger)c + \beta \\ = & -\operatorname{vec}(R_1)^T((I_r \otimes Q)^\dagger)\operatorname{vec}(R_1) - \operatorname{vec}(R_2)^T((P \otimes I_n)^\dagger)\operatorname{vec}(R_2) + \beta. \end{array} \tag{45}$$

Then by (27) and (36), we have established $\mu^*_{V2} \leq \mu^*_{M2}$. The other direction has been proved in Corollary 3.1. This completes the proof of the theorem. □

3.1.2 Unbalanced Orthogonal Procrustes Problem

Example 3.1. *In the Unbalanced Orthogonal Procrustes Problem [14] one seeks to solve the following minimizing problem.*

$$\begin{array}{ll} \min & \|AX - B\|_F^2 \\ s.t. & X^T X = I_r, \\ & X \in \mathcal{M}^{nr}, \end{array} \tag{46}$$

where $A \in \mathcal{M}^{nn}$, $B \in \mathcal{M}^{nr}$, and $n \geq r$.

The balanced case $n = r$ can be solved efficiently [29] and this special case also admits a $\boldsymbol{QMP_1}$ relaxation, [6]. Note that the unbalanced case is a typical $\boldsymbol{QMP_2}$. Its $\boldsymbol{VSDR_2}$ can be written as:

$$\begin{array}{ll} \min & \operatorname{trace}((I_r \otimes A^T A)Y) - \operatorname{trace}(2B^T AX) \\ s.t. & \operatorname{trace}((E_{ij} \otimes I_n)Y) = \delta_{i,j},\ i,j = 1,2,\cdots,r, \\ & \begin{bmatrix} 1 & x \\ x^T & Y \end{bmatrix} \succeq 0. \end{array} \tag{47}$$

It is easy to check that the $\boldsymbol{SDP}$ in (47) is feasible and its dual is strictly feasible, which implies the equivalence between $\boldsymbol{MSDR_2}$ and $\boldsymbol{VSDR_2}$. Thus we can obtain a nontrivial lower bound from its $\boldsymbol{MSDR_2}$ relaxation:

$$\begin{array}{ll} \min & \operatorname{trace}(A^T AY - 2B^T AX) \\ s.t. & \begin{bmatrix} I_r & X^T \\ X & Y \end{bmatrix} \succeq 0, \\ & \begin{bmatrix} I_n & X \\ X^T & I_r \end{bmatrix} \succeq 0, \\ & \operatorname{trace} Y = r. \end{array} \tag{48}$$

We also run preliminary computational experiments to compare the efficiency of the two $\boldsymbol{SDP}$ relaxations, see Table 3.1. All matrices in 5 instances are randomly generated and the problems are solved by SeDuMi 1.1 on a notebook computer(CPU: Intel Core 2 Duo, 2.53GHZ/Memory: 3GB). Table 3.1 exhibits the computational advantage of $\boldsymbol{MSDR_2}$ over $\boldsymbol{VSDR_2}$.

Table 1: The solution time(CPU seconds) of two $\boldsymbol{SDP}$ relaxations on the orthogonal Procrustes problem.

Problem Scale (n,r)	(15,5)	(20,10)	(30,10)	(30,15)	(40,20)
$\boldsymbol{VSDR_2}$ (CPU time)	2.14	23.03	65.01	196.70	954.70
$\boldsymbol{MSDR_2}$ (CPU time)	0.37	1.75	7.63	11.81	70.96

3.2 An Extension to QMP with Conic Constraints

Some $\boldsymbol{QMP}$ problems include conic constraints such as $X^TX \preceq (\succeq) S$, where S is a given positive semidefinite matrix. We can prove that the corresponding $\boldsymbol{MSDR_2}$ and $\boldsymbol{VSDR_2}$ are still equivalent for such problems.

Consider the following general form of $\boldsymbol{QMP_2}$ with conic constraints:

$$(\boldsymbol{QMP_3}) \quad \begin{array}{ll} \min & \operatorname{trace}(X^TQ_0X) + \operatorname{trace}(XP_0X^T) + 2\operatorname{trace}(C_0^TX) + \operatorname{trace}(H_0^TZ) \\ \text{s.t.} & \operatorname{trace}(X^TQ_jX) + \operatorname{trace}(XP_jX^T) + 2\operatorname{trace}(C_j^TX) + \operatorname{trace}(H_j^TZ) + \beta_j \le 0, \\ & \qquad\qquad j = 1, 2, \ldots, m \\ & X \in \mathcal{M}^{nr}, \; Z \in K \end{array}$$

where K can be the direct sum of convex cones (e.g., second-order cones, semidefinite cones). Note that the constraint $X^TX \preceq (\succeq) S$ can be formulated as

$$\begin{array}{rcl} \operatorname{trace}(X^TXE_{ij}) + (-)\operatorname{trace}(ZE_{ij}) & = & S_{ij} \\ Z & \succeq & 0. \end{array} \tag{49}$$

The formulations of $\boldsymbol{VSDR_2}$ and $\boldsymbol{MSDR_2}$ for $\boldsymbol{QMP_3}$ are the same as for $\boldsymbol{QMP_2}$ except for the additional term $H_j \cdot Z$ and the conic constraint $Z \in K$. Correspondingly, the dual programs $\boldsymbol{DVSDR_2}$ and $\boldsymbol{DMSDR_2}$ for $\boldsymbol{QMP_3}$ will both have an additional constraint

$$H_0 - \sum_{j=1}^{m} \lambda_j H_j \in K^*. \tag{50}$$

If a dual solution λ^* is feasible for $\boldsymbol{DVSDR_2}$, then it satisfies the constraint (50) in both $\boldsymbol{DVSDR_2}$ and $\boldsymbol{DMSDR_2}$. Therefore, we can follow the proof of Theorem 3.1, and construct a feasible solution for $\boldsymbol{DMSDR_2}$ with λ^* which generates the same objective value as μ^*_{V2}. This yields the following.

Corollary 3.2. *Assume $\boldsymbol{VSDR_2}$ for $\boldsymbol{QMP_3}$ is strictly feasible and its dual $\boldsymbol{DVSDR_2}$ is feasible. Then $\boldsymbol{DVSDR_2}$ and $\boldsymbol{DMSDR_2}$ both attain their optimum at the same λ and generate the same optimal value $\mu^*_{V2} = \mu^*_{M2}$.*

3.2.1 Graph Partition Problem

Example 3.2 (Graph Partition Problem, (GPP))**.** *GPP is an important combinatorial optimization problem with broad applications in network design and floor planning [3, 27]. Given a graph with n vertices, the problem is to find an r-partition $S_1, S_2, \ldots, S_r$ of the vertex set, such that $\|S_i\| = m_i$ with $m := (m_i)_{i=1,\ldots,r}$ given cardinalities of subsets, and the total number of edges across different subsets is minimized. Define matrix $X \in \mathcal{M}^{nr}$ to be the assignment of vertices, i.e., $X_{ij} = 1$ if vertex i is assigned to subset j; $X_{ij} = 0$ otherwise. With L the Laplacian matrix, the GPP can be formulated as an optimization problem:*

$$\begin{array}{rll} \mu^*_{GPP} = & \min & \frac{1}{2}\operatorname{trace}(X^TLX) \\ & s.t. & X^TX = \operatorname{Diag}(m) \\ & & \operatorname{diag}(XX^T) = e_n \\ & & X \ge 0 \end{array} \tag{51}$$

This formulation involves quadratic matrix constraints of both types $\mathrm{trace}(X^T E_{ii} X) = 1,\ i = 1,\dots,n$ *and* $\mathrm{trace}(X E_{ij} X^T = m_i \delta(i,j),\ i,j = 1,\dots,r$. *Thus it can be formulated as a* $\boldsymbol{QMP_2}$ *but not a* $\boldsymbol{QMP_1}$. *Anstreicher and Wolkowicz [5] proposed a semidefinite program relaxation with* $O(n^4)$ *variables and proved that its optimal value equals the so-called Donath-Hoffman lower bound [13]. This* $\boldsymbol{SDP}$ *formulation can be written in a more compact way as Povh [27] suggested:*

$$\begin{array}{rll} \mu^*_{DH} = & \min & \frac{1}{2}\mathrm{trace}((I_r \otimes L)V) \\ & s.t. & \sum_{i=1}^r \frac{1}{m_i} V^{ii} + W = I_n, \\ & & \mathrm{trace}(V^{ij}) = m_i \delta_{i,j},\ i,j = 1,\dots,r \\ & & \mathrm{trace}((I \otimes E_{ii})V) = 1,\ i = 1,\dots,n \\ & & V \in S^+_{rn},\ W \in S^+_n, \end{array} \tag{52}$$

where V *has been partitioned into* r^2 *square blocks, with each block size of* n *by* n, *and* V^{ij} *is the* (i,j)*-th block of* V. *Note that formulation* (52) *reduces the number of variables to* $O(n^2r^2)$.

An interesting application is the graph equipartition problem in which $m_i (= m_1)$*'s are all the same. Povh's* $\boldsymbol{SDP}$ *formulation is actually a* $\boldsymbol{VSDR_2}$ *for* $\boldsymbol{QMP_3}$ *:*

$$\begin{array}{ll} \min & \frac{1}{2}\mathrm{trace}(X^T L X) \\ s.t. & \mathrm{trace}(\frac{1}{m_1} X^T E_{ij} X + E_{ij} W) = \delta_{i,j},\ i,j = 1,\dots,n \\ & \mathrm{trace}(X E_{ij} X^T) = m_1 \delta_{i,j},\ i,j = 1,\dots,r \\ & \mathrm{trace}(X^T E_{ii} X) = 1,\ i = 1,\dots,n \\ & W \in S^+_n. \end{array} \tag{53}$$

It is easy to check that (52) *is feasible and its dual is strictly feasible. Hence by Corollary 3.2, the equivalence between* $\boldsymbol{MSDR_2}$ *and* $\boldsymbol{VSDR_2}$ *for* $\boldsymbol{QMP_3}$ *implies that the Donath-Hoffman bound can be computed by solving a small* $\boldsymbol{MSDR_2}$ *:*

$$\begin{array}{rll} \mu^*_{DH} = & \min & \frac{1}{2} L \cdot Y_1, \\ & s.t. & Y_1 \preceq m_1 I_n, \\ & & Y_2 = m_1 I_r, \\ & & \mathrm{diag}(Y_1) = e_n, \\ & & \begin{bmatrix} I_r & X^T \\ X & Y_1 \end{bmatrix} \succeq 0,\ \begin{bmatrix} I_n & X \\ X^T & Y_2 \end{bmatrix} \succeq 0 \end{array} \tag{54}$$

Because X *and* Y_2 *do not appear in the objective, formulation (54) can be reduced to a very simple form:*

$$\begin{array}{rll} \mu^*_{DH} = & \min & \frac{1}{2} L \cdot Y_1 \\ & s.t. & \mathrm{diag}(Y_1) = e_n \\ & & 0 \preceq Y_1 \preceq m_1 I_n. \end{array} \tag{55}$$

This $\boldsymbol{MSDR}$ *formulation has only* $O(n^2)$ *variables, which is a significant reduction from* $O(n^2r^2)$.

This result coincides with Karish and Rendl's result [19] for the graph equipartition. Their proof derives from the particular problem structure, while our result is based on the general equivalence of $\boldsymbol{VSDR_2}$ *and* $\boldsymbol{MSDR_2}$ *.*

4 Conclusion

This paper proves the equivalence of two $\boldsymbol{SDP}$ bounds for the hard $\boldsymbol{QCQP}$ in $\boldsymbol{QMP_2}$. Thus, it is clear that a user should use the smaller/inexpensive $\boldsymbol{MSDR}$ bound from matrix-lifting, rather than the more expensive $\boldsymbol{VSDR}$ bound from vector lifting. In particular, our results show that the large $\boldsymbol{VSDR_2}$ relaxation for the unbalanced orthogonal Procrustes problem can be replaced by the smaller $\boldsymbol{MSDR_2}$. And, with an extension of the main theorem, we proved the Karish and Rendl result [19] that the Donath-Hoffman bound for Graph Equipartition can be computed with a small $\boldsymbol{SDP}$.

The matrix completion and semidefinite inequality techniques used in our proofs are of independent interest.

Unfortunately, it is not clear how to formulate a general $\boldsymbol{QCQP}$ as a $\boldsymbol{MSDR}$. In particular, the objective function for the $\boldsymbol{QAP}$, trace $AXBX^T$, does not immediately admit an $\boldsymbol{MSDR}$ representation. (Though, Ding and Wolkowicz [12] provide a different matrix-lifting $\boldsymbol{SDP}$ relaxation, that generally has a strictly lower bound than the vectorized $\boldsymbol{SDP}$ relaxation proposed in [35].) The above motivates the need for finding efficient matrix-lifting representations for hard $\boldsymbol{QCQP}$ that allow efficient semidefinite relaxations.

References

[1] A. ALFAKIH, M.F. ANJOS, V. PICCIALLI, and H. WOLKOWICZ. Euclidean distance matrices, semidefinite programming, and sensor network localization. *Portugaliae Mathematica*, to appear. CORR 2009-05. 2

[2] A. ALFAKIH and H. WOLKOWICZ. Matrix completion problems. In *Handbook of semidefinite programming*, volume 27 of *Internat. Ser. Oper. Res. Management Sci.*, pages 533–545. Kluwer Acad. Publ., Boston, MA, 2000. 6

[3] C.J. ALPERT and A.B. KAHNG. Recent directions in netlist partitioning: a survey. *Integr. VLSI J.*, 19(1-2):1–81, 1995. 18

[4] K.M. ANSTREICHER, X. CHEN, H. WOLKOWICZ, and Y. YUAN. Strong duality for a trust-region type relaxation of the quadratic assignment problem. *Linear Algebra Appl.*, 301(1-3):121–136, 1999. 2

[5] K.M. ANSTREICHER and H. WOLKOWICZ. On Lagrangian relaxation of quadratic matrix constraints. *SIAM J. Matrix Anal. Appl.*, 22(1):41–55, 2000. 3, 19

[6] A. BECK. Quadratic matrix programming. *SIAM J. Optim.*, 17(4):1224–1238, 2006. 2, 3, 6, 15, 17

[7] A. BEN-TAL, L. El GHAOUI, and A. NEMIROVSKI. *Robust optimization.* Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2009. 2

[8] A. BEN-TAL and A.S. NEMIROVSKI. *Lectures on modern convex optimization.* MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2001. Analysis, algorithms, and engineering applications. 13

[9] P. BISWAS and Y. YE. Semidefinite programming for ad hoc wireless sensor network localization. In *Information Processing In Sensor Networks, Proceedings of the third international symposium on Information processing in sensor networks*, pages 46–54, Berkeley, Calif., 2004. 2

[10] M.W. CARTER, H.H. JIN, M.A. SAUNDERS, and Y. YE. SpaseLoc: an adaptive subproblem algorithm for scalable wireless sensor network localization. *SIAM J. Optim.*, 17(4):1102–1128, 2006. 2

[11] E. de KLERK and R. SOTIROV. Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem. *Math. Program.*, 122(2, Ser. A):225–246, 2010. 2

[12] Y. DING and H. WOLKOWICZ. A low-dimensional semidefinite relaxation for the quadratic assignment problem. *Math. Oper. Res.*, 34(4):1008–1022, 2009. 2, 20

[13] W.E. DONATH and A.J. HOFFMAN. Lower bounds for the partitioning of graphs. *IBM J. Res. Develop.*, 17:420–425, 1973. 19

[14] L. ELDÉN and H. PARK. A Procrustes problem on the Stiefel manifold. *Numer. Math.*, 82(4):599–619, 1999. 17

[15] B. GRONE, C.R. JOHNSON, E. MARQUES de SA, and H. WOLKOWICZ. Positive definite completions of partial Hermitian matrices. *Linear Algebra Appl.*, 58:109–124, 1984. 6, 7

[16] L. HOGBEN. Graph theoretic methods for matrix completion problems. *Linear Algebra Appl.*, 328(1-3):161–202, 2001. 6

[17] V. JEYAKUMAR and H. WOLKOWICZ. Generalizations of Slater's constraint qualification for infinite convex programs. *Math. Programming*, 57(1, Ser. B):85–101, 1992. 11

[18] C.R. JOHNSON. Matrix completion problems: a survey. *Proceedings of Symposium in Applied Mathematics*, 40:171–198, 1990. 6

[19] S.E. KARISCH and F. RENDL. Semidefinite programming and graph equipartition. In *Topics in Semidefinite and Interior-Point Methods*, volume 18 of *The Fields Institute for Research in Mathematical Sciences, Communications Series*, Providence, Rhode Island, 1998. American Mathematical Society. 19, 20

[20] N. KRISLOCK. *Efficient algorithms for large-scale Euclidean distance matrix problems with applications to wireless sensor network localization and molecular conformation.* PhD thesis, University of Waterloo, 2010. 8

[21] N. KRISLOCK and H. WOLKOWICZ. Explicit sensor network localization using semidefinite representations and clique reductions. Technical Report CORR 2009-04, University of Waterloo, Waterloo, Ontario, 2009. Available at URL: `www.optimization-online.org/DB_HTML/2009/05/2297.html`. 2, 8

[22] J. LIU. Eigenvalue and singular value inequalities of Schur complements. In F. ZHANG, editor, *The Schur complement and its applications*, volume 4 of *Numerical Methods and Algorithms*, pages xvi+295. Springer-Verlag, New York, 2005. 3, 7

[23] H.D. MITTELMANN and J. PENG. Estimating bounds for quadratic assignment problems associated with the hamming and manhattan distance matrices based on semidefinite programming. *SIAM J. Optim.*, page to appear, 2008. 2

[24] Y.E. NESTEROV, H. WOLKOWICZ, and Y. YE. Semidefinite programming relaxations of nonconvex quadratic optimization. In *Handbook of semidefinite programming*, volume 27 of *Internat. Ser. Oper. Res. Management Sci.*, pages 361–419. Kluwer Acad. Publ., Boston, MA, 2000. 3, 4

[25] D. OUELLETTE. Schur complements and statistics. *Linear Algebra Appl.*, 36:187–295, 1981. 3, 7

[26] J. PENG and H.D. MITTELMANN. Estimating bounds for quadratic assignment problems associated with the hamming and manhattan distance matrices based on semidefinite programming. *SIAM J. Optim.*, page to appear, 2008. 2

[27] J. POVH. Semidefinite approximations for quadratic programs over orthogonal matrices. In *Journal of Global Optimization*. Springer Netherlands, 2009. 18, 19

[28] R.T. ROCKAFELLAR. *Convex analysis*. Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1997. Reprint of the 1970 original, Princeton Paperbacks. 11, 12

[29] P. H. Schonemann. A generalized solution of the orthogonal procrustes problem. *Psychometrika*, 31(1):1–10, March 1966. 17

[30] A.M-C. SO and Y. YE. Theory of semidefinite programming for sensor network localization. *Math. Program.*, 109(2-3, Ser. B):367–384, 2007. 8

[31] R. STERN and H. WOLKOWICZ. Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. *SIAM J. Optim.*, 5(2):286–313, 1995. 4

[32] Z. WANG, S. ZHENG, S. BOYD, and Y. YE. Further relaxations of the semidefinite programming approach to sensor network localization. *SIAM J. Optim.*, 19(2):655–673, 2008. 6

[33] H. WOLKOWICZ. Semidefinite and Lagrangian relaxations for hard combinatorial problems. In M.J.D. Powell, editor, *Proceedings of 19th IFIP TC7 Conference on System Modelling and Optimization, July, 1999, Cambridge*, pages 269–309. Kluwer Academic Publishers, Boston, MA, 2000. 3

[34] H. WOLKOWICZ and Q. ZHAO. Semidefinite programming relaxations for the graph partitioning problem. *Discrete Appl. Math.*, 96/97:461–479, 1999. Selected for the special Editors' Choice, Edition 1999. 2

[35] Q. ZHAO, S.E. KARISCH, F. RENDL, and H. WOLKOWICZ. Semidefinite programming relaxations for the quadratic assignment problem. *J. Comb. Optim.*, 2(1):71–109, 1998. Semidefinite programming and interior-point approaches for combinatorial optimization problems (Fields Institute, Toronto, ON, 1996). 2, 9, 20

Index

$A^\dagger$, Moore-Penrose pseudoinverse, 12
C_λ, 11
P_λ, 11
Q_λ, 11
Y_V^{ij}, blocks of Y_V, 9
B^{0}Diag, Block-diag, 9
$\mathcal{M}^{nr}$, set of n by r matrices, 2
O^{0}Diag, Block-offdiag, 9
$\mathcal{R}(\cdot)$, range space, 13
$\mathcal{S}^n$, the space of $n \times n$ symmetric matrices, 2
β_λ, 11
b^{0}diag, block-diag adjoint, 9
vec(X), 3
o^{0}diag, block-offdiag adjoint, 9
$q_j^M(X)$, homogenized quadratic matrix function, 4
$q_j^V(X)$, homogenized quadratic vector function, 4
$q_j(X)$, quadratic function, 4
$\boldsymbol{MSDR}$, matrix-lifting semidefinite relaxation, 3
$\boldsymbol{MSDR_2}$, vector-lifting relaxation, 9
$\boldsymbol{QMP}$, quadratic matrix programming, 2
$\boldsymbol{QMP_1}$, first case of quadratic matrix programming, 2
$\boldsymbol{QMP_2}$, second case of quadratic matrix programming, 3, 8
$\boldsymbol{VSDR}$, vector-lifting semidefinite relaxation, 3
$\boldsymbol{VSDR_1}$, vector-lifting relaxation, 5
$\boldsymbol{VSDR_2}$, vector-lifting relaxation, 9

$\boldsymbol{MSDR_1}$, matrix-lifting relaxation, 5

block-diag adjoint, b^{0}diag, 9
Block-diag, B^{0}Diag, 9
block-offdiag adjoint, o^{0}diag, 9
Block-offdiag, O^{0}Diag, 9
blocks of Y_V, Y_V^{ij}, 9

chordal graph, 7

first case of quadratic matrix programming, $\boldsymbol{QMP_1}$, 2

homogenized quadratic matrix function, $q_j^M(X)$, 4
homogenized quadratic vector function, $q_j^V(X)$, 4

Kronecker product, 3

lifted matrices for $\boldsymbol{QMP_1}$, 4

matrix-lifting relaxation, $\boldsymbol{MSDR_1}$, 5
matrix-lifting semidefinite relaxation, $\boldsymbol{MSDR}$, 3
Moore-Penrose pseudoinverse, $A^\dagger$, 12

partial symmetric matrix, 6

quadratic function, $q_j(X)$, 4
quadratic matrix programming, $\boldsymbol{QMP}$, 2
quadratically constrained quadratic program, $\boldsymbol{QCQP}$, 2

range space, $\mathcal{R}(\cdot)$, 13

Schur complement, 3
second case of quadratic matrix programming, $\boldsymbol{QMP_2}$, 3, 8
sensor network localization problem, SNL, 8
set of n by r matrices, $\mathcal{M}^{nr}$, 2
SNL, sensor network localization problem, 8

the space of $n \times n$ symmetric matrices, $\mathcal{S}^n$, 2
trace dot product, 2, 5
trace inner-product, 2

vector-lifting relaxation, $\boldsymbol{MSDR_2}$, 9
vector-lifting relaxation, $\boldsymbol{VSDR_1}$, 5
vector-lifting relaxation, $\boldsymbol{VSDR_2}$, 9
vector-lifting semidefinite relaxation, $\boldsymbol{VSDR}$, 3