MATHEMATICS OF OPERATIONS RESEARCH
Vol. 36, No. 1, February 2011, pp. 88–104
ISSN 0364-765X | EISSN 1526-5471 | 11 | 3601 | 0088

DOI 10.1287/moor.1100.0473

On Equivalence of Semidefinite Relaxations for Quadratic Matrix Programming

Yichuan Ding
Department of Management Science and Engineering, Stanford University, Stanford, California 94305,
y7ding@standford.edu

Dongdong Ge
Antai College of Economics and Management, Shanghai Jiao Tong University, 200240 Shanghai, P. R. China,
ddge@sjtu.edu.cn

Henry Wolkowicz
Department of Combinatorics and Optimization, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada,
hwolkowicz@uwaterloo.ca

We analyze two popular semidefinite programming relaxations for quadratically constrained quadratic programs with matrix variables. These relaxations are based on *vector lifting* and on *matrix lifting*; they 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 a less expensive semidefinite programming relaxation and still obtain a strong bound. The main technique used to show the equivalence and that allows for the simplified constraints is the recognition of a class of nonchordal sparse patterns that admit a smaller representation of the positive semidefinite constraint.

Key words: semidefinite programming relaxations; quadratic matrix programming; quadratic constrained quadratic programming; hard combinatorial problems; sparsity patterns
MSC2000 subject classification: Primary: 90C20, 90C22; secondary: 90C26, 90C46
OR/MS subject classification: Primary: programming; secondary: quadratic
History: Received April 27, 2010; revised October 26, 2010. Published online in *Articles in Advance* February 2, 2011.

1. Introduction. We provide theoretical insights on how to compare different semidefinite programming (SDP) relaxations for quadratically constrained quadratic programs (QCQP) with matrix variables. In particular, we study a *vector lifting* relaxation and compare it to a significantly smaller *matrix lifting* relaxation to show that the resulting two bounds are equal.

Many hard combinatorial problems can be formulated as QCQPs with matrix variables. If the resulting formulated problem is nonconvex, then 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 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 (Ding and Wolkowicz [13]).

For examples of using SDP relaxations for QCQP arising from hard problems, see e.g., quadratic assignment (QAP) (de Klerk and Sotirov [12], Ding and Wolkowicz [13], Mittelmann and Peng [25], Zhao et al. [36]), graph partitioning (GPP) (Wolkowicz and Zhao [35]), sensor network localization (SNL) (Biswas and Ye [10], Carter et al. [11], Krislock and Wolkowicz [23]), and the more general Euclidean distance matrix completions (Alfakih et al. [2]).

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

$$
\begin{aligned}
(\mathrm{QMP}_1)\quad \mu^*_{P1} := \min\ & \operatorname{trace}(X^T Q_0 X) + 2\operatorname{trace}(C_0^T X) + \beta_0,\\
\text{s.t.}\ & \operatorname{trace}(X^T Q_j X) + 2\operatorname{trace}(C_j^T X) + \beta_j \le 0, \quad j = 1, 2, \ldots, m,\\
& X \in \mathbb{R}^{n\times r},
\end{aligned}
$$

where $\mathbb{R}^{n\times r}$ 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 $C_j \in \mathbb{R}^{n\times r}$. Throughout this paper, we use the trace inner product (dot product) $C \cdot X := \operatorname{trace} C^T X$.

The applicability of QMP_1 is limited when compared to the more general class QCQP. However, many applications use QCQP models in the form of QMP_1 c, e.g., robust optimization (Ben-Tal et al. [9]) and SNL

(Alfakih et al. [2]). In addition, many combinatorial problems are formulated with orthogonality constraints in one of the two forms

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

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 et al. [5] 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* QMP problems (QMP_2) as

$$\begin{aligned} (\text{QMP}_2) \quad \min \ & \operatorname{trace}(X^T Q_0 X) + \operatorname{trace}(X P_0 X^T) + 2\operatorname{trace}(C_0^T X) + \beta_0, \\ \text{s.t.} \ & \operatorname{trace}(X^T Q_j X) + \operatorname{trace}(X P_j X^T) + 2\operatorname{trace}(C_j^T X) + \beta_j \le 0, \quad j = 1, \ldots, m, \\ & X \in \mathbb{R}^{n\times r}, \end{aligned} \tag{2}$$

where Q_j and P_j are symmetric matrices of appropriate sizes.

Both QMP_1 and QMP_2 can be vectorized into the QCQP form using

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

where $\otimes$ denotes the Kronecker product (e.g., Graham [16]) 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 SDP relaxation for the vectorized QCQP is called the *vector-lifting semidefinite relaxation* (VSDR). Under a constraint qualification assumption, VSDR for QCQP is equivalent to the dual of classical Lagrangian relaxation (see e.g., Anstreicher and Wolkowicz [4], Nesterov et al. [26], Wolkowicz [34]).

From (3), we get

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

VSDR is derived using (4) with the relaxation $Y \succeq \operatorname{vec}(X)\operatorname{vec}(X)^T$. A Schur complement argument (e.g., Liu [24], Ouellette [27]) 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 $\operatorname{trace}(XPX^T) = \operatorname{vec}(X)^T (P \otimes I_n)\operatorname{vec}(X)$.

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

$$\begin{aligned} \operatorname{trace}(X^T Q X) &= \operatorname{trace} QY, \quad \text{if } Y = XX^T, \\ \operatorname{trace}(X P X^T) &= \operatorname{trace} PY, \quad \text{if } Y = X^T X. \end{aligned} \tag{5}$$

The *matrix-lifting semidefinite relaxation* (MSDR) is derived using (5) with the relaxation $Y \succeq XX^T$. 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 $\operatorname{trace}(XPX^T)$.

Intuitively, one expects that VSDR should provide stronger bounds than MSDR. Beck [6] proved that VSDR is actually equivalent to MSDR for QMP_1 if both 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 MSDR and VSDR for the more general problem QMP_2 under a constraint qualification. This result is of more interest because QMP_2 does not possess the same nice structure (chordal pattern) as QMP_1 c. Moreover, QMP_2 encompasses a much richer class of problems and therefore has more significant applications; for example, see the unbalanced orthogonal Procrustes problem (Eldén and Park [15]) discussed in §3.1.2 and the graph partition problem (Alpert and Kahng [3], Povh [28]) discussed in §3.2.1.

1.2. Outline. In §2 we present the equivalence of the corresponding VSDR and MSDR formulations for QMP_1 and prove Beck's result without the constraint qualification assumption (see Theorem 2.1). Section 3 proves the main result that VSDR and MSDR generate equivalent lower bounds for QMP_2, under a constraint qualification assumption (see Theorem 3.1). Numerical tests are included in §3.1.2. Section 4 provides concluding remarks.

2. Quadratic matrix programming: Case I. We first discuss the two relaxations for QMP_1 c. We denote the matrices in the relaxations obtained from vector and matrix lifting by

$$M(q_j^V(\cdot)) := \begin{bmatrix} \beta_j & \operatorname{vec}(C_j)^T \\ \operatorname{vec}(C_j) & I_r \otimes Q_j \end{bmatrix},$$
$$M(q_j^M(\cdot)) := \begin{bmatrix} \dfrac{\beta_j}{r} I_r & C_j^T \\ C_j & Q_j \end{bmatrix}.$$

We let

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

and we denote the quadratic and homogenized quadratic functions

$$\begin{aligned} q_j(X) &:= \operatorname{trace}(X^T Q_j X) + 2\operatorname{trace}(C_j^T X) + \beta_j, \\ q_j^V(X, x_0) &:= \operatorname{trace}(X^T Q_j X) + 2\operatorname{trace}(C_j^T X x_0) + \beta_j x_0^2 \\ &= y^T M(q_j^V(\cdot)) y, \\ q_j^M(X, X_0) &:= \operatorname{trace}(X^T Q_j X) + 2\operatorname{trace}(X_0^T C_j^T X) + \operatorname{trace}\left(\frac{\beta_j}{r} X_0^T I_r X_0\right) \\ &= \operatorname{trace}\big(Y^T M(q_j^M(\cdot)) Y\big). \end{aligned}$$

2.1. Lagrangian relaxation. As mentioned above, under a constraint qualification, VSDR for 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 Nesterov et al. [26, p. 403] and use the strong duality of the trust region subproblem (Stern and Wolkowicz [32]) to obtain the Lagrangian relaxation (or dual) for QMP_1 as an SDP.

$$\begin{aligned} \mu_L^* &:= \max_{\lambda \ge 0} \min_X q_0(X) + \sum_{i=j}^m \lambda_j q_j(X) \\ &= \max_{\lambda \ge 0} \min_{X,\, x_0^2 = 1} q_0^V(X, x_0) + \sum_{j=1}^m \lambda_i q_j^V(X, x_0) \\ &= \max_{\lambda \ge 0,\, t} \min_y y^T \left(M(q_0^V(\cdot)) + \sum_{j=1}^m \lambda_j M(q_j^V(\cdot)) \right) y + t(1 - x_0^2) \\ &= \max_{\lambda \ge 0,\, t} \min_y \operatorname{trace}\left(M(q_0^V(\cdot)) + \sum_{j=1}^m \lambda_j M(q_j^V(\cdot)) \right) y y^T + t(1 - x_0^2) \\ &= (\text{DVSDR}_1) \begin{cases} \max & t, \\ \text{s.t.} & \begin{bmatrix} t & 0 \\ 0 & 0 \end{bmatrix} - \sum_{j=1}^m \lambda_j M(q_j^V(\cdot)) \preceq M(q_0^V(\cdot)), \\ & \lambda \in \mathbb{R}_+^m, \quad t \in \mathbb{R}. \end{cases} \end{aligned} \tag{6}$$

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

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

2.2. Equivalence of vector and matrix lifting for QMP$_1$. Recall that the dot product refers to the trace inner product, $C \cdot X = \operatorname{trace} C^T X$. The vector-lifting relaxation is

$$
\begin{aligned}
(\mathrm{VSDR}_1) \quad \mu^*_{V1} := \min \quad & M(q_0^V(\cdot)) \cdot Z_V, \\
\text{s.t.} \quad & M(q_j^V(\cdot)) \cdot Z_V \le 0, \quad j = 1, 2, \dots, m, \\
& (Z_V)_{1,1} = 1, \\
& Z_V \succeq 0.
\end{aligned}
$$

Thus, the constraint matrix is blocked as $Z_V = \left[\begin{smallmatrix} 1 & \operatorname{vec}(X)^T \\ \operatorname{vec}(X) & Y_V \end{smallmatrix}\right]$.

The matrix-lifting relaxation is

$$
\begin{aligned}
(\mathrm{MSDR}_1) \quad \mu^*_{M1} := \min \quad & M(q_0^M(\cdot)) \cdot Z_M, \\
\text{s.t.} \quad & M(q_j^M(\cdot)) \cdot Z_M \le 0, \quad j = 1, 2, \dots, m, \\
& (Z_M)_{1:r,\,1:r} = I_r, \\
& Z_M \succeq 0.
\end{aligned}
$$

Thus, the constraint matrix is blocked as $Z_M = \left[\begin{smallmatrix} I_r & X^T \\ X & Y_M \end{smallmatrix}\right]$.

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 $Z_V = \left[\begin{smallmatrix} 1 & \operatorname{vec}(X)^T \\ \operatorname{vec}(X) & Y_V \end{smallmatrix}\right] \succeq 0$. 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 $Z_M = \left[\begin{smallmatrix} I_r & X^T \\ X & Y_M \end{smallmatrix}\right] \succeq 0$. VSDR$_1$ involves $O((nr)^2)$ variables and $O(m)$ constraints, which is often at the complexity of $O(nr)$, whereas the smaller problem MSDR$_1$ has only $O((n+r)^2)$ variables. The equivalence of relaxations using vector and matrix liftings is proved in Beck [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 itself is of interest 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 VSDR$_1$ and MSDR$_1$ generate the same optimal values as the following program.

$$
\begin{aligned}
(\mathrm{VSDR}'_1) \quad \mu^*_{V1'} := \min \quad & Q_0 \cdot \sum_{j=1}^{r} Y_{jj} + 2C_0 \cdot X + \beta_0, \\
\text{s.t.} \quad & Q_j \cdot \sum_{j=1}^{r} Y_{jj} + 2C_j \cdot X + \beta_j \le 0, \quad j = 1, 2, \dots, m, \\
& Z_{jj} = \begin{bmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{bmatrix} \succeq 0, \quad j = 1, 2, \dots, r,
\end{aligned}
$$

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

We first show that the optimal values of VSDR$_1$ and 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 that for each program, for each feasible solution, one can always construct a corresponding feasible solution with the same objective value.

First, suppose VSDR$'_1$ has a feasible solution $Z_{jj} = \left[\begin{smallmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{smallmatrix}\right]$, $j = 1, 2, \ldots, r$. 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 VSDR$_1$. In other words, giving 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 VSDR$_1$ and has the same objective value. The feasibility of Z_{jj} $(j = 1, 2, \ldots, r)$ for VSDR$'_1$ implies $\left[\begin{smallmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{smallmatrix}\right] \succeq 0$ for each $j = 1, 2, \ldots, r$. So all the specified principal submatrices of Z_V are positive semidefinite; hence, Z_V is a *partial positive semidefinite matrix* (see Alfakih and Wolkowicz [1], Grone et al. [17], Hogben [18], Johnson [20], Wang et al. [33] 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 (Grone et al. [17, Theorem 7]). Thus we have constructed a feasible solution to VSDR$_1$ with the same objective value as the feasible solution from VSDR$'_1$; i.e., this shows that $\mu^*_{V1} \le \mu^*_{V1'}$.

Conversely, suppose VSDR$_1$ has a feasible solution

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

Now we construct $Z_{jj} := \left[\begin{smallmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{smallmatrix}\right]$, $j = 1, 2, \ldots, r$. 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 VSDR$_1$ also implies

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

It is easy to check that

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

where $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_r \end{bmatrix}$. Therefore, Z_{jj}, $j = 1, 2, \ldots, r$, is feasible for VSDR$'_1$ and also generates the same objective value for VSDR$'_1$ as Z_V for 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 MSDR$_1$ and VSDR$'_1$ are equal, i.e., that

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

The proof is similar to the one for (8). First suppose VSDR$'_1$ has a feasible solution $Z_{jj} = \left[\begin{smallmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{smallmatrix}\right]$, $j = 1, 2, \ldots, m$. 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 $Z_M := \left[\begin{smallmatrix} I_r & X^T \\ X & Y_M \end{smallmatrix}\right]$. Then by $Z_{jj} \succeq 0$, $j = 1, 2, \ldots, r$, 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 (Liu [24], Ouellette [27]). Because

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

we get Z_M is feasible for MSDR and it generates the same objective value as the one by Z_{jj}, $j = 1, 2, \ldots, m$, for 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 MSDR$_1$, and $X = \begin{bmatrix} x_1 & x_2 & \cdots & x_r \end{bmatrix}$. Let $Y_{ii} = x_i(x_i)^T$ for $i = 1, 2, \ldots, r-1$, 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 $i = 1, 2, \ldots, r$, and $\sum_{i=1}^r Y_{ii} = Y_M$. So, by constructing $Z_{jj} = \begin{bmatrix} 1 & x_j^T \\ x_j & Y_{jj} \end{bmatrix}$, $j = 1, 2, \ldots, r$, it is easy to show that Z_{jj} is feasible for VSDR$_1'$ and generates an objective value equal to the objective value of 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. □

Remark 2.1. Though the MSDR$_1$ bound is significantly less expensive, Theorem 2.1 implies that the quality is no weaker than that from VSDR$_1$. Thus MSDR$_1$ is preferable as long as the problem can be formulated as a QMP$_1$ c. Moreover, a solution $Z_M = \begin{bmatrix} I_r & X^T \\ X & Y_M \end{bmatrix}$ to MSDR$_1$ can be used to construct the following corresponding solution to VSDR$_1$: $Z_V(2{:}\ nr+1, 1) = \text{vec}(X)$, and $Z_V(2{:}\ nr+1, 2{:}\ nr+1) = Y_V$, where Y_V is constructed by semidefinite completion, as in the proof of Theorem 2.1. In addition, the solution from MSDR$_1$ can also be used in a warm-start strategy applied to a vectorized semidefinite relaxation where additional constraints that do not allow a matrix lifting have been added.

Example 2.1 (SNL Problem). The SNL problem is one of the most studied problems in graph realization (e.g., Krislock [22], Krislock and Wolkowicz [23], So and Ye [31]). In this problem one is given a graph with m known points (anchors) $a_k \in R^d$, $k = 1, 2, \ldots, m$, and n unknown points (sensors) $x_j \in R^d$, $j = 1, 2, \ldots, n$, where d is the *embedding dimension*. A Euclidean distance d_{kj} between a_k and x_j or distance 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.

$$\begin{aligned} \min\ & 0, \\ \text{s.t.}\ & \text{trace}(X^T(E_{ii} + E_{jj} - 2E_{ij})X) = d_{ij}, \quad \forall (i,j) \in N_x, \\ & \text{trace}(X^T E_{ii} X) - 2\,\text{trace}\left(\begin{bmatrix} a_j^T & 0 \end{bmatrix} X\right) + a_j^T a_j = d_{ij}, \quad \forall (i,j) \in N_a, \\ & X \in \mathbb{R}^{n \times r}, \end{aligned} \tag{13}$$

where N_x, N_a refers to sets of known distances. This formulation is a QMP$_1$ c, so we can develop both its VSDR$_1$ and MSDR$_1$ relaxations.

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

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

Theorem 2.1 implies that the MSDR$_1$ relaxation always provides the same lower bound as the VSDR$_1$ one, although the number of variables for MSDR$_1$ ($O((n+d)^2)$) is significantly smaller than the number for VSDR$_1$ ($O(n^2d^2)$). The quality of the bounds combined with a lower computational complexity explains why 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 QMP$_2$.

3.1. Equivalence of vector and matrix lifting for QMP$_2$. We first propose the VSDR, VSDR$_2$ for QMP$_2$. From applying both equations in (4), we get the following:

$$(\text{VSDR}_2) \quad \mu^*_{V2} := \min \begin{bmatrix} \beta_0 & \text{vec}(C_0)^T \\ \text{vec}(C_0) & I_r \otimes Q_0 + P_0 \otimes I_n \end{bmatrix} \cdot Z_V,$$

$$\text{s.t.} \begin{bmatrix} \beta_j & \operatorname{vec}(C_j)^T \\ \operatorname{vec}(C_j) & I_r \otimes Q_j + P_j \otimes I_n \end{bmatrix} \cdot Z_V \le 0, \quad j = 1, 2, \ldots, m,$$
$$(Z_V)_{1,1} = 1,$$
$$Z_V \in \mathcal{S}_+^{rn+1} \quad \left(Z_V = \begin{bmatrix} 1 & \operatorname{vec}(X)^T \\ \operatorname{vec}(X) & Y_V \end{bmatrix} \right).$$

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 MSDR, MSDR_2 for QMP_2. (We add the additional constraint $\operatorname{trace} Y_1 = \operatorname{trace} Y_2$ because $\operatorname{trace} XX^T = \operatorname{trace} X^T X$.)

$$\begin{aligned} (\text{MSDR}_2) \quad \mu_{M2}^* := \min \quad & Q_0 \cdot Y_1 + P_0 \cdot Y_2 + 2C_0 \cdot X + \beta_0, \\ \text{s.t.} \quad & Q_j \cdot Y_1 + P_j \cdot Y_2 + 2C_j \cdot X + \beta_j \le 0, \quad j = 1, 2, \ldots, m, \\ & Y_1 - XX^T \in \mathcal{S}_+^n \quad \left(Z_1 := \begin{bmatrix} I_r & X^T \\ X & Y_1 \end{bmatrix} \succeq 0 \right), \\ & Y_2 - X^T X \in \mathcal{S}_+^r \quad \left(Z_2 := \begin{bmatrix} I_n & X \\ X^T & Y_2 \end{bmatrix} \succeq 0 \right), \\ & \operatorname{trace} Y_1 = \operatorname{trace} Y_2. \end{aligned}$$

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

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

$$\text{B}^0\text{Diag}(Q)\colon \mathcal{S}^n \to \mathcal{S}^{rn+1}, \qquad \text{O}^0\text{Diag}(P)\colon \mathcal{S}^r \to \mathcal{S}^{rn+1},$$
$$\text{B}^0\text{Diag}(Q) := \begin{bmatrix} 0 & 0 \\ 0 & I_r \otimes Q \end{bmatrix}, \qquad \text{O}^0\text{Diag}(P) := \begin{bmatrix} 0 & 0 \\ 0 & P \otimes I_n \end{bmatrix}.$$

(See Zhao et al. [36] for the $r = n$ case.) It is clear that $Q, P \succeq 0$ implies that both $\text{B}^0\text{Diag}(Q) \succeq 0$ and $\text{O}^0\text{Diag}(P) \succeq 0$. The adjoints b^0diag, o^0diag are, respectively,

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

Lemma 3.1. *Let $X \in \mathbb{R}^{n\times r}$ be given. Suppose that one of the following two conditions holds.*

(i) *Let Y_V be given and Z_V defined as in* VSDR_2. *Let the pair Z_1, Z_2 in* MSDR_2 *be constructed as in* (16).

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

$$Y_V = \begin{bmatrix} V_1 & \frac{1}{n}(Y_2)_{12} I_n & \ldots & \frac{1}{n}(Y_2)_{1r} I_n \\ \frac{1}{n}(Y_2)_{12} I_n & V_2 & \ldots & \frac{1}{n}(Y_2)_{2r} I_n \\ \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \frac{1}{n}(Y_2)_{(r-1)r} I_n \\ \ldots & \ldots & \ldots & V_r \end{bmatrix}, \tag{17}$$

with

$$\sum_{i=1}^r V_i = Y_1, \qquad \operatorname{trace} V_i = (Y_2)_{ii}, \quad i = 1, \ldots, r. \tag{18}$$

Then, Z_V satisfies the linear inequality constraints in VSDR_2 *if, and only if,* Z_1, Z_2 *satisfies the linear inequality constraints in* MSDR_2. *Moreover, the values of the objective functions with the corresponding variables are equal.*

PROOF. (i) Note that

$$\begin{aligned}\begin{bmatrix}\beta & \mathrm{vec}(C)^T\\ \mathrm{vec}(C) & I_r\otimes Q+P\otimes I_n\end{bmatrix}\cdot Z_V &= \beta+2C\cdot X+\big(\mathrm{B}^0\mathrm{Diag}(Q)+\mathrm{O}^0\mathrm{Diag}(P)\big)\cdot Z_V\\ &= \beta+2C\cdot X+Q\cdot \mathrm{b}^0\mathrm{diag}(Z_V)+P\cdot \mathrm{o}^0\mathrm{diag}(Z_V)\\ &= \beta+2C\cdot X+Q\cdot Y_1+P\cdot Y_2,\quad \text{by (16)}.\end{aligned}\tag{19}$$

(ii) Conversely, we note that $\operatorname{trace} Y_1=\operatorname{trace} Y_2$ is a constraint in MSDR_2. Z_V as constructed using (17) satisfies (16). In addition, the $n+r$ assignment type constraints in (18) 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 (19) 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* VSDR_2 *is nonempty, then the feasible set of* MSDR_2 *is also nonempty and*

$$\mu^*_{M2}\le\mu^*_{V2}.\tag{20}$$

PROOF. Suppose $Z_V=\left[\begin{smallmatrix}1 & \mathrm{vec}(X)^T\\ \mathrm{vec}(X) & Y_V\end{smallmatrix}\right]$ is feasible for 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 (16), and we set $Z_1=\left[\begin{smallmatrix}I_r & X^T\\ X & Y_1\end{smallmatrix}\right]$, $Z_2=\left[\begin{smallmatrix}I_n & X\\ X^T & Y_2\end{smallmatrix}\right]$.

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

Corollary 3.1 holds because MSDR_2 only restricts the sum of some principal submatrices of Z_V (i.e., $\mathrm{b}^0\mathrm{diag}(Z_V), \mathrm{o}^0\mathrm{diag}(Z_V)$) to be positive semidefinite, whereas VSDR_2 restricts the whole matrix Z_V positive semidefinite. So the semidefinite constraints in MSDR_2 are not as strong as in VSDR_2. Moreover, the entries of Y_V involved in $\mathrm{b}^0\mathrm{diag}(\cdot), \mathrm{o}^0\mathrm{diag}(\cdot)$ form a partial semidefinite matrix that is not chordal and does not necessarily have a semidefinite completion. Therefore, the semidefinite completion technique we used to prove the equivalence between VSDR_1 and 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 (Jeyakumar and Wolkowicz [19], Rockafellar [29]), and in this case we will then conclude that VSDR_2 and 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_λ, Q_λ, P_λ be defined similarly.

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

$$\begin{aligned}(\mathrm{DVSDR}_2)\quad \max\ & \beta_\lambda-\alpha,\\ \text{s.t.}\ & \begin{bmatrix}\alpha & \mathrm{vec}(C_\lambda)^T\\ \mathrm{vec}(C_\lambda) & I_r\otimes Q_\lambda+P_\lambda\otimes I_n\end{bmatrix}\succeq 0,\\ & \alpha\in\mathbb{R},\quad \lambda\in\mathbb{R}^m_+.\end{aligned}$$

The dual of MSDR_2 is

$$\begin{aligned}(\mathrm{DMSDR}_2)\quad \max\ & \beta_\lambda-\operatorname{trace} S_1-\operatorname{trace} S_2,\\ \text{s.t.}\ & \begin{bmatrix}S_1 & R_1^T\\ R_1 & Q_\lambda-tI_n\end{bmatrix}\succeq 0,\end{aligned}$$

$$\begin{bmatrix} S_2 & R_2 \\ R_2^T & P_\lambda + tI_r \end{bmatrix} \succeq 0,$$
$$R_1 + R_2 = C_\lambda,$$
$$\lambda \in \mathbb{R}^m_+,\quad S_1 \in S^r,\quad S_2 \in S^n,\quad R_1, R_2 \in \mathbb{R}^{n\times r},\quad t \in R.$$

The Slater condition for DVSDR_2 is equivalent to the following:

$$\exists \lambda \in \mathbb{R}^m_+, \quad \text{s.t.} \quad I_r \otimes Q_\lambda + P_\lambda \otimes I_n \succ 0. \tag{21}$$

The corresponding constraint qualification condition for DMSDR_2 is

$$\exists\, t \in \mathbb{R},\ \lambda \in \mathbb{R}^m_+, \quad \text{s.t.} \quad Q_\lambda - tI_r \succ 0, P_\lambda + tI_n \succ 0. \tag{22}$$

These two conditions are equivalent because of 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,

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

Proof. Assume $\{\lambda_i(Q)\}_{i=1,2,\dots,n}$ and $\{\lambda_j(P)\}_{j=1,2,\dots,r}$ are the sets of eigenvalues of Q and P, respectively. Thus, we get the equivalences $I_r \otimes Q + P \otimes I_n \succ 0$ if, and only if, $\lambda_i(Q) + \lambda_j(P) > 0$, $\forall\, i, j$ if, and only if, $\min_i \lambda_i(Q) + \min_j \lambda_j(P) > 0$ if, and only if,

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

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

Now we state the main theorem of this paper on the equivalence of the two SDP relaxations for QMP_2.

Theorem 3.1. *Suppose that* DVSDR_2 *is strictly feasible. As numbers in the extended real line* $(-\infty, +\infty]$, *the optimal values of the two relaxations* VSDR_2, 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. Because 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., Rockafellar [29]). Therefore, both have zero duality gaps; i.e., the optimal values of DVSDR_2, DMSDR_2, are μ^*_{V2}, μ^*_{M2}, respectively.

Now assume that

$$\lambda \text{ is feasible for } \text{DVSDR}_2. \tag{23}$$

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

$$Q := Q_\lambda - tI_n \succeq 0, \qquad P := P_\lambda + tI_r \succeq 0. \tag{24}$$

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

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

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, and 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

$$C := C_\lambda, \qquad c := \text{vec}(C), \qquad \beta := \beta_\lambda. \tag{25}$$

Let $A^\dagger$ denote the *Moore-Penrose pseudoinverse* of matrix A (e.g., Ben-Israel and Greville [7]). 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 (23), (24), (25). Let*

$$\alpha^* := c^T((I_r \otimes Q + P \otimes I_n)^\dagger)c. \tag{26}$$

Then α^*, λ is a feasible pair for DVSDR$_2$. *For any pair α, λ feasible to* DVSDR$_2$, *we have*

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

PROOF. A general quadratic function $f(x) = x^T \bar{Q} x + 2\bar{c}^T x + \bar{\beta}$ 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., Ben-Tal and Nemirovski [8, p. 163]). Therefore,

$$\begin{bmatrix} \alpha & c^T \\ c & I_r \otimes Q_\lambda + P_\lambda \otimes I_n \end{bmatrix} = \begin{bmatrix} \alpha & c^T \\ c & I_r \otimes Q + P \otimes I_n \end{bmatrix} \succeq 0 \tag{27}$$

if, and only if,

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

For a fixed α, this is further equivalent to

$$\begin{aligned} -\alpha &\le \min_x x^T(I_r \otimes Q + P \otimes I_n)x + 2c^T x \\ &= -c^T((I_r \otimes Q + P \otimes I_n)^\dagger)c. \end{aligned}$$

Therefore, we can choose α^* as in (26). □

To further explore the structure of (26), we note that c can be decomposed as

$$c = (U_1 \otimes V_1)r_{11} + (U_1 \otimes V_2)r_{12} + (U_2 \otimes V_1)r_{21} + (U_2 \otimes V_2)r_{22}. \tag{28}$$

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 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 (28).

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 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* DVSDR$_2$ *in Lemma 3.3 satisfies*

$$\begin{aligned} -\alpha &= -c^T((I_r \otimes Q + P \otimes I_n)^\dagger)c \\ &= \begin{cases} \max & -\operatorname{vec}(R_1)^T (I_r \otimes Q)^\dagger \operatorname{vec}(R_1) - \operatorname{vec}(R_2)^T (P \otimes I_n)^\dagger \operatorname{vec}(R_2), \\ \text{s.t.} & R_1 + R_2 = C, \\ & R_1, R_2 \in \mathbb{R}^{n \times r}. \end{cases} \end{aligned} \tag{29}$$

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

$$\begin{aligned} \phi(R_1) &:= -\operatorname{vec}(R_1)^T((I_r \otimes Q)^\dagger + (P \otimes I_n)^\dagger)\operatorname{vec}(R_1) \\ &\quad + 2c^T(P \otimes I_n)^\dagger \operatorname{vec}(R_1) - c^T(P \otimes I_n)^\dagger c. \end{aligned} \tag{30}$$

Because P and Q are both positive semidefinite, we get $I_r \otimes Q \succeq 0$, $P \otimes I_n \succeq 0$ and, therefore, $(I_r \otimes Q)^\dagger + (P \otimes I_n)^\dagger \succeq 0$. 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^*,

$$\begin{aligned} \operatorname{vec}(R_1^*) &= ((I_r \otimes Q)^\dagger + (P \otimes I_n)^\dagger)^\dagger (P \otimes I_n)^\dagger c \\ &= (P \otimes Q^\dagger + PP^\dagger \otimes I_n)^\dagger c; \end{aligned} \tag{31}$$

and this corresponds to a value

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

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

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

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

$$\begin{aligned} \bar{\Lambda} &:= (I_r \otimes Q + P \otimes I_n)^\dagger = (U \otimes V)(I_r \otimes \Lambda_Q + \Lambda_P \otimes I_n)^\dagger (U \otimes V)^T \\ &= (U \otimes V)((I_{\sigma_+} \otimes \Lambda_Q + \Lambda_P \otimes I_{\theta_+})^\dagger + I_{\sigma_0} \otimes \Lambda_Q^\dagger + \Lambda_P^\dagger \otimes \mathcal{I}_{\theta_0})(U \otimes V)^T, \end{aligned} \tag{32}$$

where matrix I_{σ_+} (resp. I_{σ_0}) is $r \times r$, diagonal, and zero, except for the ith 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

$$\bar{\lambda}_{i,j} = \begin{cases} \dfrac{1}{\sigma_i + \theta_j} & \text{if } \sigma_i > 0, \quad \theta_j > 0, \\ \dfrac{1}{\sigma_i} & \text{if } \sigma_i > 0, \quad \theta_j = 0, \\ \dfrac{1}{\theta_j} & \text{if } \sigma_i = 0, \quad \theta_j > 0, \\ 0 & \text{if } \sigma_i = 0, \quad \theta_j = 0. \end{cases} \tag{33}$$

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.,

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

For the given feasible α^*, λ of Lemma 3.3, we will construct a feasible solution for 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 DMSDR$_2$.

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

$$\begin{aligned} \text{vec}(R_1) &= \text{vec}(R_1^*) + (U_2 \otimes V_1)r_{21}, \\ \text{vec}(R_2) &= \text{vec}(R_2^*) + (U_1 \otimes V_2)r_{12}. \end{aligned} \tag{34}$$

Then we obtain a lower bound for the optimal value of DMSDR$_2$.

$$\mu_{M2}^* \geq -\text{vec}(R_1)^T(I_r \otimes Q)^\dagger \text{vec}(R_1) - \text{vec}(R_2)^T(P \otimes I_n)^\dagger \text{vec}(R_2) + \beta. \tag{35}$$

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

$$\begin{aligned} \max \quad & \beta - \text{trace}\, S_1 - \text{trace}\, S_2, \\ \text{s.t.} \quad & \begin{bmatrix} S_1 & R_1^T \\ R_1 & Q \end{bmatrix} \succeq 0, \\ & \begin{bmatrix} S_2 & R_2 \\ R_2^T & P \end{bmatrix} \succeq 0, \\ & S_1 \in S_r \quad S_2 \in S_n. \end{aligned} \tag{36}$$

Because λ, t, R_1, and R_2 are all feasible for DMSDR$_2$, this subproblem will generate a lower bound for μ^*_{M2}. We now invoke a result from Beck [6], i.e., that there exists a symmetric matrix S such that trace $S \le \gamma$ and $\begin{bmatrix} S & C^T \\ C & Q \end{bmatrix} \succeq 0$ if, and only if, $f(X) = \text{trace}(X^T QX + 2C^T X) + \gamma \ge 0$ for any $X \in \mathbb{R}^{n\times r}$. This is equivalent to

$$-\gamma \le \min_{X\in\mathbb{R}^{n\times r}} \text{trace}(X^T QX + 2C^T X) = -\text{trace}(C^T Q^\dagger C).$$

Therefore, the subproblem (36) can be reformulated as

$$\begin{aligned} \max \quad & \beta - \text{trace}\, S_1 - \text{trace}\, S_2, \\ \text{s.t.} \quad & -\text{trace}\, S_1 \le -\text{trace}(R_1^T Q^\dagger R_1), \\ & -\text{trace}\, S_2 \le -\text{trace}(R_2 Q^\dagger R_2^T), \\ & S_1 \in S^r, \quad S_2 \in S^n. \end{aligned} \tag{37}$$

Hence, the optimal value of (37) has an explicit expression and provides a lower bound for DMSDR$_2$

$$\mu^*_{M2} \ge -\text{vec}(R_1)^T (I_r \otimes Q)^\dagger \text{vec}(R_1) - \text{vec}(R_2)^T (P \otimes I_n)^\dagger \text{vec}(R_2) + \beta. \quad \Box$$

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

Proof of Theorem 3.1. We now compare μ^*_{V2} obtained by using α^* from the expression (26) with μ^*_{M2} based on the lower bound expression in (35). By writing c in the form of (28), we get

$$\begin{aligned} \mu^*_{V2} = & -[r_{11}^T (U_1 \otimes V_1)^T + r_{12}^T (U_1 \otimes V_2)^T + r_{21}^T (U_2 \otimes V_1)^T](I_r \otimes Q + P \otimes I_n)^\dagger \\ & \cdot [(U_1 \otimes V_1) r_{11} + (U_1 \otimes V_2) r_{12} + (U_2 \otimes V_1) r_{21}] + \beta. \end{aligned} \tag{38}$$

Consider the cross-term such as $r_{11}^T (U_1 \otimes V_1)^T (I_r \otimes Q + P \otimes I_n)^\dagger (U_1 \otimes V_2) r_{12}$. Because $(U_1 \otimes V_1) r_{11}$ 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 cross-terms equal zero. As a result, only the following sum of three quadratic terms remain, which we label using $C1$, $C2$, $C3$, respectively.

$$\begin{aligned} \mu^*_{V2} = & -r_{11}^T (U_1 \otimes V_1)^T (I_r \otimes Q + P \otimes I_n)^\dagger (U_1 \otimes V_1) r_{11}, \\ & - r_{21}^T (U_2 \otimes V_1)^T (I_r \otimes Q + P \otimes I_n)^\dagger (U_2 \otimes V_1) r_{21}, \\ & - r_{12}^T (U_1 \otimes V_2)^T (I_r \otimes Q + P \otimes I_n)^\dagger (U_1 \otimes V_2) r_{12} + \beta \\ =: & \; C1 + C2 + C3 + \beta. \end{aligned} \tag{39}$$

We can also formulate the lower bound for μ^*_{M2} based on (35):

$$\begin{aligned} \mu^*_{M2} & \ge -\text{vec}(R_1)^T (I_r \otimes Q)^\dagger \text{vec}(R_1) - \text{vec}(R_2)^T (P \otimes I_n)^\dagger \text{vec}(R_2) + \beta \\ & = -(\text{vec}(R_1^*) + (U_2 \otimes V_1) r_{21})^T (I_r \otimes Q)^\dagger \text{vec}(R_1^* + (U_2 \otimes V_1) r_{21}) \\ & \quad - (\text{vec}(R_2^*) + (U_1 \otimes V_2) r_{12})^T (P \otimes I_n)^\dagger (\text{vec}(R_2^*) + (U_1 \otimes V_2) r_{12}) + \beta. \end{aligned} \tag{40}$$

Because $\text{vec}(R_1^*)$ and $\text{vec}(R_2^*)$ 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 cross-terms, such as $\text{vec}(R_1^*)^T (I_r \otimes Q)^\dagger (U_2 \otimes V_1) r_{21}$, all equal zero. Therefore, the lower bound for μ^*_{M2} can be reformulated as

$$\begin{aligned} \mu^*_{M2} \ge & -(\text{vec}(R_1^*)(I_r \otimes Q)^\dagger \text{vec}(R_1^*) + \text{vec}(R_2^*)(P \otimes I_n)^\dagger \text{vec}(R_2^*)), \\ & - r_{21}^T (U_2 \otimes V_1)^T (I_r \otimes Q)^\dagger (U_2 \otimes V_1) r_{21}, \\ & - r_{12}^T (U_1 \otimes V_2)^T (P \otimes I_n)^\dagger (U_1 \otimes V_2) r_{12} + \beta \\ =: & \; T1 + T2 + T3 + \beta. \end{aligned} \tag{41}$$

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

We will show that terms $C1$, $C2$, 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 (32). 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 part $I_{\sigma_0} \otimes \Lambda_Q^\dagger$ in diagonal matrix $\bar{\Lambda}$ is involved in computing term $C2$, i.e.,

$$-r_{21}^T(U_2\otimes V_1)^T(I_r\otimes Q+P\otimes I_n)^\dagger(U_2\otimes V_1)r_{21}=-r_{21}^T(U_2\otimes V_1)^T(U\otimes V)(I_{\sigma_0}\otimes \Lambda_Q^\dagger)(U\otimes V)^T(U_2\otimes V_1)r_{21}. \tag{42}$$

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

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

By (42) and (43), 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

$$-c^T((I_r\otimes Q+P\otimes I_n)^\dagger)c+\beta=-\operatorname{vec}(R_1)^T((I_r\otimes Q)^\dagger)\operatorname{vec}(R_1)-\operatorname{vec}(R_2)^T((P\otimes I_n)^\dagger)\operatorname{vec}(R_2)+\beta. \tag{44}$$

Then by (26) and (35), we have established $\mu_{V2}^* \le \mu_{M2}^*$. The other direction has been proved in Corollary 3.1. □

Remark 3.1. Our result implies that the dual optimal solution λ^* to DVSDR$_2$ coincides with that from DMSDR$_2$. Therefore, if a QCQP contains constraints that cannot be formulated into DMSDR$_2$, we can first solve the corresponding DMSDR$_2$ without these constraints, and then we can use the dual solution in a warm-start strategy and continue solving the original QCQP.

3.1.2. Unbalanced orthogonal Procrustes problem

Example 3.1. In the unbalanced orthogonal Procrustes problem (Eldén and Park [15]) one seeks to solve the following minimizing problem:

$$\begin{aligned} \min\ & \|AX-B\|_F^2,\\ \text{s.t. } & X^TX=I_r,\\ & X\in\mathcal{M}^{nr}, \end{aligned}\tag{45}$$

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

The balanced case $n=r$ can be solved efficiently (Schönemann [30]), and this special case also admits a QMP$_1$ relaxation (Beck [6]). Note that the unbalanced case is a typical QMP$_2$. Its VSDR$_2$ can be written as

$$\begin{aligned} \min\ & \operatorname{trace}((I_r\otimes A^TA)Y)-\operatorname{trace}(2B^TAX),\\ \text{s.t. } & \operatorname{trace}((E_{ij}\otimes I_n)Y)=\delta_{i,j}, \quad i,j=1,2,\ldots,r,\\ & \begin{bmatrix}1 & x\\ x^T & Y\end{bmatrix}\succeq 0. \end{aligned}\tag{46}$$

It is easy to check that the SDP in (46) is feasible and its dual is strictly feasible, which implies the equivalence between MSDR$_2$ and VSDR$_2$. Thus we can obtain a nontrivial lower bound from its MSDR$_2$ relaxation:

$$\begin{aligned} \min\ & \operatorname{trace}(A^TAY-2B^TAX),\\ \text{s.t. } & \begin{bmatrix}I_r & X^T\\ X & Y\end{bmatrix}\succeq 0,\\ & \begin{bmatrix}I_n & X\\ X^T & I_r\end{bmatrix}\succeq 0,\\ & \operatorname{trace} Y = r. \end{aligned}\tag{47}$$

Preliminary computational experiments appear in Table 1. The matrices in the five instances are randomly generated, and they are solved using SeDuMi 1.1 and a 32-bit version of Matlab R2009a on a laptop running Windows XP, with a 2.53 GHz Intel Core 2 Duo processor and with 3 GB of RAM. Table 1 illustrates the computational advantage of MSDR$_2$ over VSDR$_2$.

TABLE 1. Solution times (CPU seconds) of two SDP relaxations on the orthogonal Procrustes problem.

Problem size (n, r)	(15, 5)	(20, 10)	(30, 10)	(30, 15)	(40, 20)
VSDR_2 (CPU sec)	2.14	23.03	65.01	196.70	954.70
MSDR_2 (CPU sec)	0.37	1.75	7.63	11.81	70.96

3.2. An extension to QMP with conic constraints. Some 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 MSDR_2 and VSDR_2 are still equivalent for such problems.

Consider the following general form of QMP_2 with conic constraints:

$$(\text{QMP}_3) \quad \begin{aligned} \min\ & \operatorname{trace}(X^TQ_0X) + \operatorname{trace}(XP_0X^T) + 2\operatorname{trace}(C_0^TX) + \operatorname{trace}(H_0^TZ), \\ \text{s.t.}\ & \operatorname{trace}(X^TQ_jX) + \operatorname{trace}(XP_jX^T) + 2\operatorname{trace}(C_j^TX) + \operatorname{trace}(H_j^TZ) + \beta_j \le 0, \quad j = 1, 2, \ldots, m, \\ & X \in \mathbb{R}^{n\times r}, \quad Z \in K, \end{aligned}$$

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

$$\begin{aligned} \operatorname{trace}(X^TXE_{ij}) + (-)\operatorname{trace}(ZE_{ij}) = S_{ij}, \\ Z \succeq 0. \end{aligned} \tag{48}$$

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

$$H_0 - \sum_{j=1}^m \lambda_j H_j \in K^*. \tag{49}$$

If a dual solution λ^* is feasible for DVSDR_2, then it satisfies the constraint (49) in both DVSDR_2 and DMSDR_2. Therefore, we can follow the proof of Theorem 3.1 and construct a feasible solution for DMSDR_2 with λ^*, which generates the same objective value as μ^*_{V2}. This yields the following.

COROLLARY 3.2. *Assume* VSDR_2 *for* QMP_3 *is strictly feasible and its dual* DVSDR_2 *is feasible. Then* DVSDR_2 *and* 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 (GPP). GPP is an important combinatorial optimization problem with broad applications in network design and floor planning (Alpert and Kahng [3], Povh [28]). 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 $X \in \mathbb{R}^{n\times r}$ 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:

$$\begin{aligned} \mu^*_{\text{GPP}} = \min\ & \tfrac{1}{2}\operatorname{trace}(X^TLX), \\ \text{s.t.}\ & X^TX = \operatorname{Diag}(m), \\ & \operatorname{diag}(XX^T) = e_n, \\ & X \ge 0. \end{aligned} \tag{50}$$

This formulation involves quadratic matrix constraints of both types $\operatorname{trace}(X^TE_{ii}X) = 1$, $i = 1, \ldots, n$ and $\operatorname{trace}(XE_{ij}X^T) = m_i\delta(i, j)$, $i, j = 1, \ldots, r$. Thus it can be formulated as a QMP_2 but not a QMP_1 c. Anstreicher and Wolkowicz [4] proposed a semidefinite program relaxation with $O(n^4)$ variables and proved that its optimal value equals the so-called Donath-Hoffman lower bound (Donath and Hoffman [14]). This SDP formulation can be written in a more compact way, as Povh [28] suggested:

$$
\begin{aligned}
\mu^*_{\mathrm{DH}} = \min \ & \tfrac{1}{2}\operatorname{trace}((I_r \otimes L)V),\\
\text{s.t.} \ & \sum_{i=1}^{r} \frac{1}{m_i} V^{ii} + W = I_n,\\
& \operatorname{trace}(V^{ij}) = m_i \delta_{i,j}, \quad i,j = 1,\ldots,r,\\
& \operatorname{trace}((I \otimes E_{ii})V) = 1, \quad i = 1,\ldots,n,\\
& V \in S^+_{rn}, \quad W \in S^+_n,
\end{aligned}
\tag{51}
$$

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 (51) 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 SDP formulation is actually a VSDR_2 for QMP_3:

$$
\begin{aligned}
\min \ & \tfrac{1}{2}\operatorname{trace}(X^T L X),\\
\text{s.t.} \ & \operatorname{trace}\left(\frac{1}{m_1} X^T E_{ij} X + E_{ij} W\right) = \delta_{i,j}, \quad i,j = 1,\ldots,n,\\
& \operatorname{trace}(X E_{ij} X^T) = m_1 \delta_{i,j}, \quad i,j = 1,\ldots,r,\\
& \operatorname{trace}(X^T E_{ii} X) = 1, \quad i = 1,\ldots,n,\\
& W \in S^+_n.
\end{aligned}
\tag{52}
$$

It is easy to check that (51) is feasible and its dual is strictly feasible. Hence, by Corollary 3.2, the equivalence between MSDR_2 and VSDR_2 for QMP_3 implies that the Donath-Hoffman bound can be computed by solving a small MSDR_2:

$$
\begin{aligned}
\mu^*_{\mathrm{DH}} = \min \ & \tfrac{1}{2} L \cdot Y_1,\\
\text{s.t.} \ & Y_1 \preceq m_1 I_n,\\
& Y_2 = m_1 I_r,\\
& \operatorname{diag}(Y_1) = e_n,\\
& \begin{bmatrix} I_r & X^T \\ X & Y_1 \end{bmatrix} \succeq 0, \qquad \begin{bmatrix} I_n & X \\ X^T & Y_2 \end{bmatrix} \succeq 0.
\end{aligned}
\tag{53}
$$

Because X and Y_2 do not appear in the objective, formulation (53) can be reduced to a very simple form:

$$
\begin{aligned}
\mu^*_{\mathrm{DH}} = \min \ & \tfrac{1}{2} L \cdot Y_1,\\
\text{s.t.} \ & \operatorname{diag}(Y_1) = e_n,\\
& 0 \preceq Y_1 \preceq m_1 I_n.
\end{aligned}
\tag{54}
$$

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

This result coincides with Karisch and Rendl's result (Karisch and Rendl [21]) for the graph equipartition. Their proof derives from the particular problem structure, while our result is based on the general equivalence of VSDR_2 and MSDR_2.

4. Conclusion. This paper proves the equivalence of two SDP bounds for the hard QCQP in QMP_2. Thus, it is clear that a user should use the smaller/inexpensive MSDR bound from matrix lifting, rather than the more expensive VSDR bound from vector lifting. In particular, our results show that the large VSDR_2 relaxation for the unbalanced orthogonal Procrustes problem can be replaced by the smaller MSDR_2. And, with an extension of the main theorem, we proved the Karisch and Rendl result (Karisch and Rendl [21]) that the Donath-Hoffman bound for graph equipartition can be computed with a small SDP.

The key idea of the paper is to simplify the semidefinite constraint using a sparse completion technique. Most existing literature on this topic requires the matrix to have a chordal structure (Grone et al. [17], Beck [6], Wang et al. [33]); whereas in our case, the dual matrix of QMP_2 is not chordal, but it can be decomposed as a sum of two matrices, each of which admits a chordal structure. This idea, we hope, will lead to further studies on identifying other nonchordal sparse patterns that can be used to simplify the semidefinite constraints. The sparse matrix completion results and semidefinite inequality techniques used in our proofs are of independent interest.

Unfortunately, it is not clear how to formulate a general QCQP as an MSDR. For example, the objective function for the QAP, trace $AXBX^T$, does not immediately admit an MSDR representation (though a relaxed MSDR is presented in Ding and Wolkowicz [13] that generally has a strictly lower bound than the vectorized SDP relaxation proposed in Zhao et al. [36]). The above motivates the need for finding efficient matrix-lifting representations for hard QCQP problems.

Acknowledgments. Research by the first and third author was supported by Natural Sciences Engineering Research Council Canada and a grant from AFOSR. Research by the second author was supported by the National Natural Science Foundation of China under Grants 71001062. The authors express their gratitude to the editor and anonymous referees for their careful reading of the manuscript.

References

[1] Alfakih, A., H. Wolkowicz. 2000. Matrix completion problems. H. Wolkowicz, R. Saigal, L. Vandenberghe, eds. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, *International Series in Operations Research and Management Science*, Vol. 27. Kluwer Academic Publishers, Boston, 533–545.

[2] Alfakih, A., M. F. Anjos, V. Piccialli, H. Wolkowicz. 2010. Euclidean distance matrices, semidefinite programming, and sensor network localization. *Portugal Math.* Forthcoming.

[3] Alpert, C. J., A. B. Kahng. 1995. Recent directions in netlist partitioning: A survey. *Integrated VLSI J.* **19**(1–2) 1–81.

[4] Anstreicher, K. M., H. Wolkowicz. 2000. On Lagrangian relaxation of quadratic matrix constraints. *SIAM J. Matrix Anal. Appl.* **22**(1) 41–55.

[5] Anstreicher, K. M., X. Chen, H. Wolkowicz, Y. Yuan. 1999. Strong duality for a trust-region type relaxation of the quadratic assignment problem. *Linear Algebra Appl.* **301**(1–3) 121–136.

[6] Beck, A. 2006. Quadratic matrix programming. *SIAM J. Optim.* **17**(4) 1224–1238.

[7] Ben-Israel, A., T. N. E. Greville. 1974. *Generalized Inverses: Theory and Applications*. Wiley-Interscience, New York.

[8] Ben-Tal, A., A. S. Nemirovski. 2001. *Lectures on Modern Convex Optimization*. MPS/SIAM Series on Optimization. Society for Industrial and Applied Mathematics (SIAM), Philadelphia.

[9] Ben-Tal, A., L. El Ghaoui, A. Nemirovski. 2009. *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ.

[10] Biswas, P., Y. Ye. 2004. Semidefinite programming for ad hoc wireless sensor network localization. *Proc. Third Internat. Sympos. Inform. Processing in Sensor Networks*, Berkeley, CA, 46–54.

[11] Carter, M. W., H. H. Jin, M. A. Saunders, Y. Ye. 2006. SpaseLoc: An adaptive subproblem algorithm for scalable wireless sensor network localization. *SIAM J. Optim.* **17**(4) 1102–1128.

[12] de Klerk, E., R. Sotirov. 2010. Exploiting group symmetry in semidefinite programming relaxations of the quadratic assignment problem. *Math. Programming* **122**(2, Ser. A) 225–246.

[13] Ding, Y., H. Wolkowicz. 2009. A low-dimensional semidefinite relaxation for the quadratic assignment problem. *Math. Oper. Res.* **34**(4) 1008–1022.

[14] Donath, W. E., A. J. Hoffman. 1973. Lower bounds for the partitioning of graphs. *IBM J. Res. Development* **17**(5) 420–425.

[15] Eldén, L., H. Park. 1999. A Procrustes problem on the Stiefel manifold. *Numer. Math.* **82**(4) 599–619.

[16] Graham, A. 1981. *Kronecker Products and Matrix Calculus: With Applications*. Halsted Press, Toronto.

[17] Grone, B., C. R. Johnson, E. Marques de Sa, H. Wolkowicz. 1984. Positive definite completions of partial Hermitian matrices. *Linear Algebra Appl.* **58** 109–124.

[18] Hogben, L. 2001. Graph theoretic methods for matrix completion problems. *Linear Algebra Appl.* **328**(1–3) 161–202.

[19] Jeyakumar, V., H. Wolkowicz. 1992. Generalizations of Slater's constraint qualification for infinite convex programs. *Math. Programming* **57**(1, Ser. B) 85–101.

[20] Johnson, C. R. 1990. Matrix completion problems: A survey. *Proc. Sympos. Appl. Math.*, Vol. 40. American Mathematical Society, Providence, RI, 171–198.

[21] Karisch, S. E., F. Rendl. 1998. Semidefinite programming and graph equipartition. P. M. Pardalos, H. Wolkowicz, eds. *Topics in Semidefinite and Interior-Point Methods, Fields Institute Communications Series*, Vol. 18. American Mathematical Society, Providence, RI, 77–96.

[22] Krislock, N. 2010. Semidefinite facial reduction for low-rank euclidean distance matrix completion. Doctoral dissertation, University of Waterloo, Waterloo, Ontario.

[23] Krislock, N., H. Wolkowicz. 2010. Explicit sensor network localization using semidefinite representations and facial reductions. *SIAM J. Optim.* **20**(5) 2679–2708.

[24] Liu, J. 2005. Eigenvalue and singular value inequalities of Schur complements. C. Brezinski, F. Zhang, eds. *The Schur Complement and Its Applications*, *Numerical Methods and Algorithms*, Vol. 4. Springer-Verlag, New York, 47–82.

[25] Mittelmann, H. D., J. Peng. 2010. Estimating bounds for quadratic assignment problems associated with the hamming and Manhattan distance matrices based on semidefinite programming. *SIAM J. Optim.* **20**(6) 3408–3426.

[26] Nesterov, Y. E., H. Wolkowicz, Y. Ye. 2000. Semidefinite programming relaxations of nonconvex quadratic optimization. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications*, *International Series in Operations Research and Management Science*, Vol. 27. Kluwer Academic Publishers, Boston, 361–419.

[27] Ouellette, D. V. 1981. Schur complements and statistics. *Linear Algebra Appl.* **36** 187–295.
[28] Povh, J. 2010. Semidefinite approximations for quadratic programs over orthogonal matrices. *J. Global Optim.* **48**(3) 447–463.
[29] Rockafellar, R. T. 1997. *Convex Analysis. Princeton Landmarks in Mathematics*. Princeton University Press, Princeton, NJ.
[30] Schönemann, P. H. 1966. A generalized solution of the orthogonal Procrustes problem. *Psychometrika* **31**(1) 1–10.
[31] So, A. M.-C., Y. Ye. 2007. Theory of semidefinite programming for sensor network localization. *Math. Programming* **109**(2–3, Ser. B) 367–384.
[32] Stern, R., H. Wolkowicz. 1995. Indefinite trust region subproblems and nonsymmetric eigenvalue perturbations. *SIAM J. Optim.* **5**(2) 286–313.
[33] Wang, Z., S. Zheng, S. Boyd, Y. Ye. 2008. Further relaxations of the semidefinite programming approach to sensor network localization. *SIAM J. Optim.* **19**(2) 655–673.
[34] Wolkowicz, H. 2000. Semidefinite and Lagrangian relaxations for hard combinatorial problems. M. J. D. Powell, S. Scholtes, eds. *Proc. 19th IFIP TC7 Conf. System Modelling Optim.* Kluwer Academic Publishers, Boston, 269–309.
[35] Wolkowicz, H., Q. Zhao. 1999. Semidefinite programming relaxations for the graph partitioning problem. *J. Discrete Appl. Math.* **96/97**(1) 461–479.
[36] Zhao, Q., S. E. Karisch, F. Rendl, H. Wolkowicz. 1998. Semidefinite programming relaxations for the quadratic assignment problem. *J. Combin. Optim.* **2**(1) 71–109.