Tighter yet more tractable relaxations
and nontrivial instance generation
for sparse standard quadratic optimization

Immanuel Bomze[*]   Bo Peng[†]   Yuzhou Qiu[‡]   E. Alper Yıldırım[§]

June 3, 2024

Abstract

The Standard Quadratic optimization Problem (StQP), arguably the simplest among all classes of NP-hard optimization problems, consists of extremizing a quadratic form (the simplest nonlinear polynomial) over the standard simplex (the simplest polytope/compact feasible set). As a problem class, StQPs may be nonconvex with an exponential number of inefficient local solutions. StQPs arise in a multitude of applications, among them mathematical finance, machine learning (clustering), and modeling in biosciences (e.g., selection and ecology). This paper deals with such StQPs under an additional sparsity or cardinality constraint, which, even for convex objectives, renders NP-hard problems. One motivation to study StQPs under such sparsity restrictions is the high-dimensional portfolio selection problem with too many assets to handle, in particular, in the presence of transaction costs. Here, relying on modern conic optimization techniques, we present tractable convex relaxations for this relevant but difficult problem. We propose novel equivalent reformulations of these relaxations with significant dimensional reduction, which is essential for the tractability of these relaxations when the problem size grows. Moreover, we propose an instance generation procedure which systematically avoids too easy instances. Our extensive computational results illustrate the high quality of the relaxation bounds in a significant number of instances. Furthermore, in contrast with exact mixed-integer quadratic programming models, the solution time of the relaxations is very robust to the choices of the problem parameters. In particular, the reduced formulations achieve significant improvements in terms of the solution time over their counterparts.

[*]VCOR and Research Network Data Science, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria. ORCID ID: 0000-0002-6288-9226 E-mail: `immanuel.bomze@univie.ac.at`

[†]VGSCO, University of Vienna, Oskar-Morgenstern-Platz 1, 1090 Wien, Austria. ORCID ID: 0000-0002-2650-0295 E-mail: `bo.peng@univie.ac.at`

[‡]School of Mathematics, Peter Guthrie Tait Road, The University of Edinburgh, Edinburgh, EH9 3FD, United Kingdom. E-mail: `y.qiu-16@sms.ed.ac.uk`

[§]School of Mathematics, Peter Guthrie Tait Road, The University of Edinburgh, Edinburgh, EH9 3FD, United Kingdom. ORCID ID: 0000-0003-4141-3189 E-mail: `E.A.Yildirim@ed.ac.uk`

1   The problem — introduction and some motivation

In many applications, one is interested in sparse solutions to optimization problems. In this paper, we consider Standard Quadratic optimization Problems (StQPs) under a hard sparsity constraint.

An StQP consists of minimizing a (not necessarily convex) quadratic form over the standard simplex:

$$\ell_n(\mathbf{Q}) := \min_{\mathbf{x}\in\mathbb{R}^n}\left\{\mathbf{x}^\top\mathbf{Q}\mathbf{x} : \mathbf{x}\in F_n\right\}, \qquad\qquad \text{(StQP)}$$

where

$$F_n := \left\{\mathbf{x}\in\mathbb{R}^n_+ : \mathbf{e}^\top\mathbf{x} = 1\right\} \qquad\qquad (1)$$

is the standard simplex, the simplest polytope in the n-dimensional Euclidean space $\mathbb{R}^n$. Here, $\mathbf{e}\in\mathbb{R}^n$ denotes the vector of all ones.

Despite its simplicity, this problem class serves to model manifold real-world applications, ranging from the analysis of social networks (community detection via dominant-set-clustering) [10, 34] through biology, game theory to economy and finance [8, 31], to cite just a few. For a survey on mixed-integer convex quadratic optimization approaches to portfolio selection, see, e.g., [32].

One motivation to introduce sparsity constraints comes from high-dimensional portfolio selection. Recently, the significant benefits of controlling the cardinality of a portfolio have been widely recognized in the literature, e.g., by [22] who observed that maintaining a well-diversified portfolio on S&P 500 datasets only required approximately $10\sim 30$ assets. By fixing the number of holding stocks in advance, they bypass estimation of the covariance matrix Σ, focusing solely on the cardinality constraints. Even earlier, [28] already found empirically that commonly employed shrinkage methods significantly underperform in constructing portfolios. This effect, namely that using a mathematically well-behaved surrogate instead of the true sparsity term (which is discontinuous and thus non-convex) can be grossly misleading, is by now widely known in other application domains, e.g., in signal reconstruction.

Therefore, we propose to study the StQP with an exact, hard sparsity constraint, with the aim to develop tight yet tractable relaxations. Such a problem can be expressed as follows:

$$\ell_\rho(\mathbf{Q}) := \min_{\mathbf{x}\in\mathbb{R}^n}\left\{\mathbf{x}^\top\mathbf{Q}\mathbf{x} : \mathbf{x}\in F_\rho\right\}, \qquad\qquad \text{(StQP(}\rho\text{))}$$

where

$$F_\rho := \{\mathbf{x}\in F_n : \|\mathbf{x}\|_0 \le \rho\} = \left\{\mathbf{x}\in\mathbb{R}^n_+ : \mathbf{e}^\top\mathbf{x} = 1, \quad \|\mathbf{x}\|_0 \le \rho\right\}. \qquad (2)$$

Here, $\|\mathbf{x}\|_0$ denotes the number of nonzero components of a vector $\mathbf{x}$ and $\rho\in[1\!:\!n] := \{1,\dots,n\}$ is the sparsity parameter.

Before we start our discussion, let us stress that the primary aim of this study is to provide rigorous and tight bounds on an extremely difficult problem. In practice, this will most likely be complemented by fast, high-performance heuristics. Combining both methods will, as we hope, provide practical solutions of reasonable, guaranteed quality.

2

As an aside, let us mention that we address here any StQP of any finite size n, not just the average case as $n \to \infty$. As shown in [18, 19], asymptotically (for large n) with high probability, an StQP instance already has a global solution in F_2, so very sparse. However, it is unclear, in general, for which value of n this effect kicks in (if at all) and what happens in highly structured (e.g., convex) StQPs as in the portfolio selection case, which may be somehow hidden in possibly lower-dimensional regions in the space of all instances; see, e.g., [12].

1.1 Literature review and contributions

The sparse StQP problem arises in a wide range of applications. A prominent class is comprised of portfolio selection problems with an upper bound on the number of assets [5, 37, 44]. The relevance of this problem is well illustrated by the fact that the most recent release of `MATLAB` [42] now offers an iterative mixed-integer linear programming (MILP) algorithmic framework to solve the mixed-integer quadratic programming (MIQP) formulation of (convex) portfolio selection problems with lower and upper bounds on the cardinality, as well as the possibility of using semi-continuous variables (taking either zero or positive values above a certain fixed threshold). However, a quick test of this toolbox reveals that even for $n = 30$ and extreme sparsity $\rho = 1$, the solution time was around 40 seconds (note that the solution is trivial for $\rho = 1$ by extracting the smallest diagonal element of the data matrix Q). For $\rho = 2$, no solution was found within two hours. Clearly, solution times are dominated by all models discussed here, cf. Figure 2a below.

Further applications include selecting subsets of regression variables in computational algorithms [33], and sparsity in the information rate of compressive sampling [17].

There are various studies on the sparse StQP problem. Those studies can be separated into two main categories. The first group of studies indirectly handles the sparsity constraint by regularization via the ℓ_1-norm or the ℓ_p-pseudonorms with $p \in (0, 1)$, see for instance [14, 19]. The second category of studies considers the cardinality constraints in the sparse StQP problem directly with different approaches. A semidefinite programming based reformulation of the sparse StQP problem that helps generating perspective cuts has been studied in [23, 45]; see also [27]. In [15], the authors study a nonconvex reformulation of the sparse StQP problem with continuous variables and apply a regularization method for solving the reformulation. A semidefinite programming based heuristic method that can give an upper bound on the sparse StQP problem is introduced in [13]. A tight semidefinite relaxation of the sparse StQP problem with strong computational behavior under the assumption of convexity of the objective function is studied in [43]. Based on the mixed-integer formulation of the sparse StQP problem, there are also many algorithms using discrete optimization techniques to either solve the problem to optimality, or to find an approximate solution (see, for instance, [3, 5, 20, 35, 38, 41, 46]).

This work builds upon the recent work of the same set of authors on the sparse StQP problem [6], where the focus was on classical convex relaxations given by the reformulation-linearization technique (RLT), Shor relaxation (SDP), and their combination (SDP-RLT) arising from $(\text{P1}(\rho))$, one of the mixed-integer quadratic formulations considered this paper. They established several structural properties of these relaxations in relation to the corresponding relaxations of StQPs without any sparsity constraints, and utilized these relations to obtain several results about the quality of the lower bounds arising from different relaxations. In contrast, in this paper, we consider convex relaxations arising from exact copositive reformulations of two different MIQP formulations, which are provably tighter than each of the aforementioned classical relaxations. Furthermore, our focus is on finding tight but tractable relaxations that scale well with the problem dimension.

Our contributions are as follows:

(i) We present exact copositive reformulations of two mixed integer quadratic optimization models $(\text{P1}(\rho))$ and $(\text{P2}(\rho))$.

(ii) We propose novel equivalent reformulations of doubly nonnegative relaxations of each of the two copositive reformulations in significantly smaller dimensions.

(iii) We establish that one of the two relaxations is provably tighter than the other one.

(iv) We propose an instance generation procedure that avoids instances of $(\text{StQP}(\rho))$ for which the cardinality constraint is redundant.

1.2 Notation and organisation of the paper

We use $\mathbb{R}^n$, $\mathbb{R}^n_+$, $\mathbb{R}^{m \times n}$, and $\mathcal{S}^n$ to denote the n-dimensional Euclidean space, nonnegative orthant, the space of $m \times n$ real matrices, and the space of $n \times n$ real symmetric matrices, respectively. Throughout the paper, vectors and matrices are denoted by bold lowercase and bold uppercase letters, respectively (e.g. $\mathbf{x}$ and X). In particular, we reserve $\mathbf{e}$ (resp., E) and $\mathbf{o}$ (resp., O) for the vector (resp. matrix) of all ones and all zeroes of appropriate dimensions, respectively. The dimension should be clear from the context. Scalars are denoted by regular lowercase Roman or Greek letters. We employ subscripts to indicate a specific element of vectors or matrices. For instance, we denote the ith component of the vector $\mathbf{x}$ by x_i, and the (i,j)-entry of the matrix X by X_{ij}. For $\mathsf{X} \in \mathbb{R}^{n \times n}$ and two index sets $\mathbb{A} \subset \{1,...,n\}$, $\mathbb{B} \subset \{1,...,n\}$, $\mathsf{X}_{\mathbb{A}\mathbb{B}}$ is the submatrix of X given by the rows and columns indexed by $\mathbb{A}$ and $\mathbb{B}$, respectively. Similarly, $\mathbf{x}_{\mathbb{A}}$ denotes the subvector of $\mathbf{x} \in \mathbb{R}^n$ given by the components indexed by $\mathbb{A}$. For $\mathsf{X} \in \mathbb{R}^{n \times n}$, we denote the column vector formed by the diagonal entries of X by diag (X). Inequalities on vectors and matrices are understood to be componentwise scalar inequalities on each of the corresponding components. For $\mathsf{X} \in \mathcal{S}^n$, we use $\mathsf{X} \succeq \mathsf{O}$ (resp., $\mathsf{X} \succ \mathsf{O}$) to denote that X is positive semidefinite (resp., positive definite). The trace inner product of $\mathsf{X} \in \mathbb{R}^{m \times n}$ and $\mathsf{Y} \in \mathbb{R}^{m \times n}$ is denoted by $\langle \mathsf{X}, \mathsf{Y} \rangle = \text{trace} \left(\mathsf{X}^\top \mathsf{Y} \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} X_{ij} Y_{ij}$.

We define the following closed, convex cones in $\mathcal{S}^n$:

$$\mathcal{CP}_n = \left\{ \mathsf{M} \in \mathcal{S}^n : \mathsf{M} = \mathsf{A}\mathsf{A}^\top \quad \text{for some } \mathsf{A} \in \mathbb{R}^{n \times k} \text{ such that } \mathsf{A} \geq \mathsf{O} \right\}, \tag{3}$$

$$\mathcal{D}_n = \left\{ \mathsf{M} \in \mathcal{S}^n : \mathsf{M} \succeq \mathsf{O}, \ \mathsf{M} \geq \mathsf{O} \right\}, \tag{4}$$

$$\mathcal{SPN}_n = \left\{ \mathsf{M} \in \mathcal{S}^n : \mathsf{M} = \mathsf{P} + \mathsf{N}, \ \text{for some } \mathsf{P} \succeq \mathsf{O}, \ \mathsf{N} \geq \mathsf{O} \right\}, \tag{5}$$

$$\mathcal{COP}_n = \left\{ \mathsf{M} \in \mathcal{S}^n : \mathbf{u}^\top \mathsf{M}\mathbf{u} \geq 0, \quad \text{for all } \mathbf{u} \in \mathbb{R}^n_+ \right\}, \tag{6}$$

i.e., $\mathcal{CP}_n$ is the cone of completely positive matrices, $\mathcal{D}_n$ is the cone of doubly nonnegative matrices, $\mathcal{SPN}^n$ is the cone of SPN matrices (those which can be decomposed into the sum of a positive semidefinite (PSD) and a componentwise nonnegative matrix), and $\mathcal{COP}^n$ is the cone of copositive matrices. It is well-known and easy to check that

$$\mathcal{CP}_n \subseteq \mathcal{D}_n \subseteq \mathcal{SPN}_n \subseteq \mathcal{COP}_n. \tag{7}$$

Furthermore, $\mathcal{CP}_n = \mathcal{D}_n$ and $\mathcal{SPN}_n = \mathcal{COP}_n$ if and only if $n \leq 4$ (see [21]).

The boundary of $\mathcal{COP}_n$, denoted by $\mathbf{bd} \ \mathcal{COP}_n$, is given by

$$\mathbf{bd} \ \mathcal{COP}_n = \left\{ \mathsf{M} \in \mathcal{COP}_n : \exists \ \mathbf{u} \in F_n \ \text{s.t.} \ \mathbf{u}^\top \mathsf{M}\mathbf{u} = 0 \right\}, \tag{8}$$

4

where F_n is given by (1).

For $\mathsf{M} \in \mathbf{bd}\ \mathcal{COP}_n$, the set of zeroes of M is given by

$$\mathbf{V}^{\mathsf{M}} = \left\{\mathbf{u} \in F_n : \mathbf{u}^\top \mathsf{M}\mathbf{u} = 0\right\}. \tag{9}$$

The paper is organized as follows: Section 2 presents two exact conic reformulations, one by using binary variables directly (Section 2.1) and the other involving complementarity constraints (Section 2.2). In Section 3, we discuss their doubly nonnegative relaxations and propose smaller equivalents. Section 3.3 compares the tractable lower bounds generated by them. Section 4 is devoted to careful generation procedures which systematically avoid too easy instances and provide testbeds of different complexity scale: PSD instances, then SPN instances, and finally COP instances which provably cannot be reduced to an instance in the previously mentioned classes. In Section 5, we describe our computational experiment in detail (Section 5.1), discuss solution times in Section 5.2, and the quality of the lower bounds achieved in Section 5.3. We conclude in Section 6.

2 Two exact completely positive formulations

In this section, we present two exact formulations of the sparse StQP problem (StQP(ρ)) as mixed-binary quadratic optimization problems, a direct one covered by Section 2.1, whereas Section 2.2 is devoted to a formulation involving a complementarity constraint.

2.1 Conic formulation of direct MIQP

By introducing binary variables, the sparse StQP can be reformulated as the following mixed-binary quadratic optimization problem:

$$
\begin{aligned}
\ell_\rho(\mathsf{Q}) = \min_{(\mathbf{x},\mathbf{u}) \in \mathbb{R}^n \times \mathbb{R}^n} \quad & \mathbf{x}^\top \mathsf{Q}\mathbf{x} \\
\text{s.t.} \quad & \mathbf{e}^\top \mathbf{x} = 1 \\
& \mathbf{e}^\top \mathbf{u} = \rho \\
& \mathbf{x} \leq \mathbf{u} \\
& \mathbf{u} \in \{0,1\}^n \\
& \mathbf{x} \geq \mathbf{o}.
\end{aligned}
\tag{P1(ρ)}
$$

By introducing the redundant constraints $\mathbf{u} \leq \mathbf{e}$, slack variables $\mathbf{y} \geq \mathbf{o}$ and $\mathbf{v} \geq \mathbf{o}$ so that $\mathbf{x} + \mathbf{y} = \mathbf{u}$ and $\mathbf{u} + \mathbf{v} = \mathbf{e}$, we arrive at the following reformulation of (P1(ρ)):

$$
\begin{aligned}
\ell_\rho(\mathsf{Q}) = \min_{(\mathbf{x},\mathbf{u},\mathbf{v},\mathbf{y}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n} \quad & \mathbf{x}^\top \mathsf{Q}\mathbf{x} \\
\text{s.t.} \quad & \mathbf{e}^\top \mathbf{x} = 1 \\
& \mathbf{e}^\top \mathbf{u} = \rho \\
& \mathbf{x} + \mathbf{y} = \mathbf{u} \\
& \mathbf{u} + \mathbf{v} = \mathbf{e} \\
& \mathbf{u} \in \{0,1\}^n \\
& \mathbf{x} \geq \mathbf{o} \\
& \mathbf{y} \geq \mathbf{o} \\
& \mathbf{v} \geq \mathbf{o}.
\end{aligned}
\tag{P1A(ρ)}
$$

By [16], (P1A(ρ)) admits an equivalent reformulation as the following conic optimization problem:

$$
\begin{aligned}
\ell_\rho(\mathsf{Q}) = \min_{\mathsf{Z} \in \mathcal{S}^{4n+1}} \quad & \langle \mathsf{Q}, \mathsf{Z}^{\mathbf{xx}} \rangle \\
\text{s.t.} \qquad\qquad \mathbf{e}^\top \mathbf{x} &= 1 \\
\mathbf{e}^\top \mathbf{u} &= \rho \\
\mathbf{x} + \mathbf{y} &= \mathbf{u} \\
\mathbf{u} + \mathbf{v} &= \mathbf{e} \\
\langle \mathsf{E}, \mathsf{Z}^{\mathbf{xx}} \rangle &= 1 \\
\langle \mathsf{E}, \mathsf{Z}^{\mathbf{uu}} \rangle &= \rho^2 \\
\operatorname{diag}\left(\mathsf{Z}^{\mathbf{uu}}\right) &= \mathbf{u} \\
\operatorname{diag}\left(\mathsf{Z}^{\mathbf{xx}}\right) + \operatorname{diag}\left(\mathsf{Z}^{\mathbf{yy}}\right) + \operatorname{diag}\left(\mathsf{Z}^{\mathbf{uu}}\right) + 2\left[\operatorname{diag}\left(\mathsf{Z}^{\mathbf{xy}}\right) - \operatorname{diag}\left(\mathsf{Z}^{\mathbf{xu}}\right) - \operatorname{diag}\left(\mathsf{Z}^{\mathbf{uy}}\right)\right] &= \mathbf{o} \\
\operatorname{diag}\left(\mathsf{Z}^{\mathbf{uu}}\right) + 2\operatorname{diag}\left(\mathsf{Z}^{\mathbf{uv}}\right) + \operatorname{diag}\left(\mathsf{Z}^{\mathbf{vv}}\right) &= \mathbf{e} \\
\mathsf{Z} := \begin{pmatrix}
1 & \mathbf{x}^\top & \mathbf{u}^\top & \mathbf{v}^\top & \mathbf{y}^\top \\
\mathbf{x} & \mathsf{Z}^{\mathbf{xx}} & \mathsf{Z}^{\mathbf{xu}} & \mathsf{Z}^{\mathbf{xv}} & \mathsf{Z}^{\mathbf{xy}} \\
\mathbf{u} & (\mathsf{Z}^{\mathbf{xu}})^\top & \mathsf{Z}^{\mathbf{uu}} & \mathsf{Z}^{\mathbf{uv}} & \mathsf{Z}^{\mathbf{uy}} \\
\mathbf{v} & (\mathsf{Z}^{\mathbf{xv}})^\top & (\mathsf{Z}^{\mathbf{uv}})^\top & \mathsf{Z}^{\mathbf{vv}} & \mathsf{Z}^{\mathbf{vy}} \\
\mathbf{y} & (\mathsf{Z}^{\mathbf{xy}})^\top & (\mathsf{Z}^{\mathbf{uy}})^\top & (\mathsf{Z}^{\mathbf{vy}})^\top & \mathsf{Z}^{\mathbf{yy}}
\end{pmatrix} &\in \quad \mathcal{CP}_{4n+1}.
\end{aligned}
$$
$$\text{(CP1A(ρ))}$$

By [16], (P1A(ρ)) and (CP1A(ρ)) are equivalent in the following sense: Both of their optimal values are equal to $\ell_\rho(\mathsf{Q})$ and for any optimal solution $\mathsf{Z} \in \mathcal{S}^{4n+1}$ of (CP1A(ρ)), $(\mathbf{x}, \mathbf{u}, \mathbf{y}, \mathbf{v}) \in \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n$ is in the convex hull of the set of optimal solutions of (P1A(ρ)). Therefore, (CP1A(ρ)) is an exact convex reformulation of the nonconvex optimization problem (P1A(ρ)).

2.2 Conic formulation of MIQP with complementarity constraint

In (P1(ρ)), sparsity constraints are handled by big-M constraints. Alternatively, sparsity can be enforced by dropping the big-M constraints $\mathbf{x} \le \mathbf{u}$ and replacing them with the complementarity constraints $\mathbf{x}_j(1 - \mathbf{u}_j) = 0$, $j = 1, \ldots, n$ (see, e.g., [15]). Defining $\mathbf{v} = \mathbf{e} - \mathbf{u} \in \{0,1\}^n$, we obtain the following quadratic optimization problem with complementarity constraints:

$$
\begin{aligned}
\ell_\rho(Q) = \min_{(\mathbf{x}, \mathbf{v}) \in \mathbb{R}^n \times \mathbb{R}^n} \quad & \mathbf{x}^\top \mathsf{Q} \mathbf{x} \\
\text{s.t.} \quad \mathbf{e}^\top \mathbf{x} &= 1 \\
\mathbf{e}^\top \mathbf{v} &= n - \rho \\
\mathbf{x}^\top \mathbf{v} &= 0 \\
\mathbf{v} &\in \{0,1\}^n \\
\mathbf{x} &\ge \mathbf{o}.
\end{aligned}
$$
$$\text{(P2(ρ))}$$

Similarly, adding the redundant constraints $\mathbf{v} \le \mathbf{e}$ and reintroducing the slack variables $\mathbf{u} \ge \mathbf{o}$ so that

6

$\mathbf{u} + \mathbf{v} = \mathbf{e}$, (P2($\rho$)) can be reformulated as follows:

$$
\begin{aligned}
\ell_\rho(\mathsf{Q}) = \min_{(\mathbf{x},\mathbf{u},\mathbf{v})\in\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n} \quad & \mathbf{x}^\top \mathsf{Q}\mathbf{x} \\
\text{s.t.} \quad \mathbf{e}^\top\mathbf{x} &= 1 \\
\mathbf{e}^\top\mathbf{u} &= \rho \\
\mathbf{u}+\mathbf{v} &= \mathbf{e} \\
\mathbf{x}^\top\mathbf{v} &= 0 \qquad\qquad\text{(P2A(ρ))} \\
\mathbf{u} &\in \{0,1\}^n \\
\mathbf{x} &\geq \mathbf{o} \\
\mathbf{v} &\geq \mathbf{o} \\
\mathbf{u} &\geq \mathbf{o}.
\end{aligned}
$$

By [11, 16], (P2A(ρ)) admits an equivalent reformulation as the following copositive optimization problem:

$$
\begin{aligned}
\ell_\rho(\mathsf{Q}) = \min_{\mathsf{Y}\in\mathcal{S}^{3n+1}} \quad & \langle \mathsf{Q}, \mathsf{Y}^{\mathbf{xx}}\rangle \\
\text{s.t.} \quad \mathbf{e}^\top\mathbf{x} &= 1 \\
\mathbf{e}^\top\mathbf{u} &= \rho \\
\mathbf{u}+\mathbf{v} &= \mathbf{e} \\
\langle \mathsf{E}, \mathsf{Y}^{\mathbf{xx}}\rangle &= 1 \\
\langle \mathsf{E}, \mathsf{Y}^{\mathbf{uu}}\rangle &= \rho^2 \\
\mathrm{diag}\,(\mathsf{Y}^{\mathbf{uu}}) &= \mathbf{u} \qquad\qquad\text{(CP2A(ρ))} \\
\mathrm{diag}\,(\mathsf{Y}^{\mathbf{uu}}) + 2\,\mathrm{diag}\,(\mathsf{Y}^{\mathbf{uv}}) + \mathrm{diag}\,(\mathsf{Y}^{\mathbf{vv}}) &= \mathbf{e} \\
\mathbf{e}^\top\mathrm{diag}\,(\mathsf{Y}^{\mathbf{xv}}) &= 0 \\
\mathsf{Y} := \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top & \mathbf{v}^\top \\ \mathbf{x} & \mathsf{Y}^{\mathbf{xx}} & \mathsf{Y}^{\mathbf{xu}} & \mathsf{Y}^{\mathbf{xv}} \\ \mathbf{u} & (\mathsf{Y}^{\mathbf{xu}})^\top & \mathsf{Y}^{\mathbf{uu}} & \mathsf{Y}^{\mathbf{uv}} \\ \mathbf{v} & (\mathsf{Y}^{\mathbf{xv}})^\top & (\mathsf{Y}^{\mathbf{uv}})^\top & \mathsf{Y}^{\mathbf{vv}} \end{pmatrix} &\in \mathcal{CP}_{3n+1}.
\end{aligned}
$$

Once again, the equivalence between (P2A(ρ)) and (CP2A(ρ)) is understood similarly to that between (P1A(ρ)) and (CP1A(ρ)).

3 Two convex relaxations and their smaller equivalents

In this section, we consider two convex relaxations arising from the conic formulations (CP1A(ρ)) and (CP2A(ρ)).

Despite the fact that (CP1A(ρ)) and (CP2A(ρ)) are convex optimization problems, they are, in general, NP-hard [36]. A well-known tractable convex relaxation of completely positive conic problems is obtained by replacing the intractable cone $\mathcal{CP}_d$ with the larger cone $\mathcal{D}_d$ of doubly nonnegative (DNN) matrices, i.e., the convex cone of symmetric $d \times d$ matrices that are both PSD and componentwise nonnegative. It follows that the optimal value of the DNN relaxation is a lower bound on the optimal value of the completely positive conic optimization problem. Here, we present the DNN relaxations arising from conic optimization problems (CP1A(ρ)) and (CP2A(ρ)). Furthermore, we establish that each of the two DNN relaxations can, in fact, be reformulated in smaller dimensions without sacrificing the strength of the corresponding lower bound. Finally, we present a theoretical comparison of the strength of the two DNN relaxations.

3.1 DNN relaxations of the direct MIQP formulation

In this section, we consider the DNN relaxation of the problem $(\mathrm{CP1A}(\rho))$. By replacing the intractable conic constraint $\mathsf{Z} \in \mathcal{CP}_{4n+1}$ in $(\mathrm{CP1A}(\rho))$ by $\mathsf{Z} \in \mathcal{D}_{4n+1}$, where $\mathcal{D}_{4n+1}$ denotes the cone of doubly nonnegative (DNN) matrices in $\mathcal{S}^{4n+1}$, we obtain the following tractable convex relaxation of $(\mathrm{P1A}(\rho))$:

$$\nu(\mathrm{D1A}(\rho)) := \min_{\mathsf{Z} \in \mathcal{S}^{4n+1}} \quad \langle \mathsf{Q}, \mathsf{Z^{xx}} \rangle$$

$$\text{s.t.} \qquad \mathbf{e}^\top \mathbf{x} = 1$$
$$\mathbf{e}^\top \mathbf{u} = \rho$$
$$\mathbf{x} + \mathbf{y} = \mathbf{u}$$
$$\mathbf{u} + \mathbf{v} = \mathbf{e}$$
$$\langle \mathsf{E}, \mathsf{Z^{xx}} \rangle = 1$$
$$\langle \mathsf{E}, \mathsf{Z^{uu}} \rangle = \rho^2$$
$$\mathrm{diag}\,(\mathsf{Z^{uu}}) = \mathbf{u}$$
$$\mathrm{diag}\,(\mathsf{Z^{xx}}) + \mathrm{diag}\,(\mathsf{Z^{yy}}) + \mathrm{diag}\,(\mathsf{Z^{uu}}) + 2\,[\mathrm{diag}\,(\mathsf{Z^{xy}}) - \mathrm{diag}\,(\mathsf{Z^{xu}}) - \mathrm{diag}\,(\mathsf{Z^{uy}})] = \mathbf{o}$$
$$\mathrm{diag}\,(\mathsf{Z^{uu}}) + 2\,\mathrm{diag}\,(\mathsf{Z^{uv}}) + \mathrm{diag}\,(\mathsf{Z^{vv}}) = \mathbf{e}$$

$$\mathsf{Z} := \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top & \mathbf{v}^\top & \mathbf{y}^\top \\ \mathbf{x} & \mathsf{Z^{xx}} & \mathsf{Z^{xu}} & \mathsf{Z^{xv}} & \mathsf{Z^{xy}} \\ \mathbf{u} & (\mathsf{Z^{xu}})^\top & \mathsf{Z^{uu}} & \mathsf{Z^{uv}} & \mathsf{Z^{uy}} \\ \mathbf{v} & (\mathsf{Z^{xv}})^\top & (\mathsf{Z^{uv}})^\top & \mathsf{Z^{vv}} & \mathsf{Z^{vy}} \\ \mathbf{y} & (\mathsf{Z^{xy}})^\top & (\mathsf{Z^{uy}})^\top & (\mathsf{Z^{vy}})^\top & \mathsf{Z^{yy}} \end{pmatrix} \in \mathcal{D}_{4n+1}.$$

$$(\mathrm{D1A}(\rho))$$

Note that $(\mathrm{D1A}(\rho))$ consists of $5n + 4$ linear equality constraints and a doubly nonnegative constraint in $\mathcal{S}^{4n+1}$. First, we make several observations about feasible solutions of $(\mathrm{D1A}(\rho))$.

Lemma 3.1. *Let* $\mathsf{Z} \in \mathcal{S}^{4n+1}$ *be* $(D1A(\rho))$*-feasible. Then, the following relations hold:*

$$\mathbf{x}\mathbf{e}^\top - \mathsf{Z^{xu}} = \mathsf{Z^{xv}} \geq \mathsf{O}, \tag{10}$$

$$-\mathsf{Z^{xx}} + \mathsf{Z^{xu}} = \mathsf{Z^{xy}} \geq \mathsf{O}, \tag{11}$$

$$\mathsf{Z^{xx}} - \mathsf{Z^{xu}} - (\mathsf{Z^{xu}})^\top + \mathsf{Z^{uu}} = \mathsf{Z^{yy}} \geq \mathsf{O}, \tag{12}$$

$$\mathbf{e}\mathbf{e}^\top - \mathbf{e}\mathbf{u}^\top - \mathbf{u}\mathbf{e}^\top + \mathsf{Z^{uu}} = \mathsf{Z^{vv}} \geq \mathsf{O}, \tag{13}$$

$$-\mathbf{e}\mathbf{x}^\top + (\mathsf{Z^{xu}})^\top + \mathbf{e}\mathbf{u}^\top - \mathsf{Z^{uu}} = \mathsf{Z^{vy}} \geq \mathsf{O}, \tag{14}$$

$$\mathit{diag}\,(\mathsf{Z^{xu}}) = \mathbf{x}, \tag{15}$$

$$\mathit{diag}\,(\mathsf{Z^{xv}}) = \mathbf{o}, \tag{16}$$

$$\mathit{diag}\,(\mathsf{Z^{vy}}) = \mathbf{o}, \tag{17}$$

$$\mathit{diag}\,(\mathsf{Z^{vv}}) = \mathbf{v}, \tag{18}$$

$$\mathit{diag}\,(\mathsf{Z^{uv}}) = \mathbf{o}. \tag{19}$$

Proof. Let $\mathsf{Z} \in \mathcal{S}^{4n+1}$ be (D1A(ρ))-feasible. By using the Schur complementation, we obtain

$$
\begin{pmatrix}
\mathsf{Z^{xx}} & \mathsf{Z^{xu}} & \mathsf{Z^{xv}} & \mathsf{Z^{xy}} \\
(\mathsf{Z^{xu}})^\top & \mathsf{Z^{uu}} & \mathsf{Z^{uv}} & \mathsf{Z^{uy}} \\
(\mathsf{Z^{xv}})^\top & (\mathsf{Z^{uv}})^\top & \mathsf{Z^{vv}} & \mathsf{Z^{vy}} \\
(\mathsf{Z^{xy}})^\top & (\mathsf{Z^{uy}})^\top & (\mathsf{Z^{vy}})^\top & \mathsf{Z^{yy}}
\end{pmatrix}
=
\begin{pmatrix} \mathbf{x} \\ \mathbf{u} \\ \mathbf{v} \\ \mathbf{y} \end{pmatrix}
\begin{pmatrix} \mathbf{x} \\ \mathbf{u} \\ \mathbf{v} \\ \mathbf{y} \end{pmatrix}^\top
+ \sum_{k \in K}
\begin{pmatrix} \mathbf{a}^k \\ \mathbf{b}^k \\ \mathbf{c}^k \\ \mathbf{d}^k \end{pmatrix}
\begin{pmatrix} \mathbf{a}^k \\ \mathbf{b}^k \\ \mathbf{c}^k \\ \mathbf{d}^k \end{pmatrix}^\top ,
$$

where K is a finite set and $\{\mathbf{a}^k, \mathbf{b}^k, \mathbf{c}^k, \mathbf{d}^k\} \subset \mathbb{R}^n$ for each $k \in K$. Using $\mathrm{diag}\,(\mathsf{Z^{uu}}) + 2\,\mathrm{diag}\,(\mathsf{Z^{uv}}) + \mathrm{diag}\,(\mathsf{Z^{vv}}) = \mathbf{e}$, we obtain

$$
\begin{aligned}
u_i^2 + \sum_{k \in K} \left(b_i^k\right)^2 + 2\left[u_i v_i + \sum_{k \in K} \left(b_i^k\right)\left(c_i^k\right)\right] + v_i^2 + \sum_{k \in K} \left(c_i^k\right)^2 &= (u_i + v_i)^2 + \sum_{k \in K}\left(b_i^k + c_i^k\right)^2 \\
&= 1 + \sum_{k \in K}\left(b_i^k + c_i^k\right)^2 = 1
\end{aligned}
$$

for each $i \in \{1, \ldots, n\}$, where we used $\mathbf{u} + \mathbf{v} = \mathbf{e}$ in the second equality. We conclude that

$$
\mathbf{b}^k = -\mathbf{c}^k, \quad \text{for all } k \in K . \tag{20}
$$

In a similar manner, by using $\mathrm{diag}\,(\mathsf{Z^{xx}}) + \mathrm{diag}\,(\mathsf{Z^{yy}}) + \mathrm{diag}\,(\mathsf{Z^{uu}}) + 2\left[\mathrm{diag}\,(\mathsf{Z^{xy}}) - \mathrm{diag}\,(\mathsf{Z^{xu}}) - \mathrm{diag}\,(\mathsf{Z^{uy}})\right] = \mathbf{o}$, we arrive at

$$
(x_i + y_i - u_i)^2 + \sum_{k \in K}\left(a_i^k + d_i^k - b_i^k\right)^2 = 0
$$

for each $i \in \{1, \ldots, n\}$, which, together with $\mathbf{x} + \mathbf{y} = \mathbf{u}$, implies that

$$
\mathbf{a}^k + \mathbf{d}^k = \mathbf{b}^k, \quad \text{for all } k \in K . \tag{21}
$$

Therefore,

$$
\mathbf{x}\mathbf{e}^\top - \mathsf{Z^{xu}} = \mathbf{x}\mathbf{e}^\top - \mathbf{x}\mathbf{u}^\top - \sum_{k \in K} \mathbf{a}^k\left(\mathbf{b}^k\right)^\top = \mathbf{x}\mathbf{v}^\top + \sum_{k \in K} \mathbf{a}^k\left(\mathbf{c}^k\right)^\top = \mathsf{Z^{xv}} \geq \mathsf{O},
$$

where we used $\mathbf{u} + \mathbf{v} = \mathbf{e}$, (20), and $\mathsf{Z} \in \mathcal{D}_{4n+1}$. This establishes (10). Arguing similarly,

$$
-\mathsf{Z^{xx}} + \mathsf{Z^{xu}} = -\mathbf{x}\mathbf{x}^\top - \sum_{k \in K} \mathbf{a}^k\left(\mathbf{a}^k\right)^\top + \mathbf{x}\mathbf{u}^\top + \sum_{k \in K} \mathbf{a}^k\left(\mathbf{b}^k\right)^\top = \mathbf{x}\mathbf{y}^\top + \sum_{k \in K} \mathbf{a}^k\left(\mathbf{d}^k\right)^\top = \mathsf{Z^{xy}} \geq \mathsf{O},
$$

where we used $\mathbf{x} + \mathbf{y} = \mathbf{u}$, (21), and $\mathsf{Z} \in \mathcal{D}_{4n+1}$. This establishes (11). Next,

$$
\begin{aligned}
\mathsf{Z^{xx}} - \mathsf{Z^{xu}} - (\mathsf{Z^{xu}})^\top + \mathsf{Z^{uu}} &= \mathbf{x}\mathbf{x}^\top + \sum_{k \in K} \mathbf{a}^k\left(\mathbf{a}^k\right)^\top - \mathbf{x}\mathbf{u}^\top - \sum_{k \in K} \mathbf{a}^k\left(\mathbf{b}^k\right)^\top \\
&\quad - \mathbf{u}\mathbf{x}^\top - \sum_{k \in K} \mathbf{b}^k\left(\mathbf{a}^k\right)^\top + \mathbf{u}\mathbf{u}^\top + \sum_{k \in K} \mathbf{b}^k\left(\mathbf{b}^k\right)^\top \\
&= \mathbf{y}\mathbf{y}^T + \sum_{k \in K} \mathbf{d}^k\left(\mathbf{d}^k\right)^\top \\
&= \mathsf{Z^{yy}} \geq \mathsf{O},
\end{aligned}
$$

where we used $\mathbf{x} + \mathbf{y} = \mathbf{u}$, (21), and $\mathsf{Z} \in \mathcal{D}_{4n+1}$. Similarly,

$$
\mathbf{e}\mathbf{e}^\top - \mathbf{e}\mathbf{u}^\top - \mathbf{u}\mathbf{e}^\top + \mathsf{Z^{uu}} = \mathbf{e}\mathbf{e}^\top - \mathbf{e}\mathbf{u}^\top - \mathbf{u}\mathbf{e}^\top + \mathbf{u}\mathbf{u}^\top + \sum_{k \in K} \mathbf{b}^k\left(\mathbf{b}^k\right)^\top = \mathbf{v}\mathbf{v}^\top + \sum_{k \in K} \mathbf{c}^k\left(\mathbf{c}^k\right)^\top = \mathsf{Z^{vv}} \geq \mathsf{O},
$$

where we used $\mathbf{u} + \mathbf{v} = \mathbf{e}$, (20), and $\mathsf{Z} \in \mathcal{D}_{4n+1}$, establishing (13). Next,

$$
\begin{aligned}
-\mathbf{e}\mathbf{x}^\top + (\mathsf{Z^{xu}})^\top + \mathbf{e}\mathbf{u}^\top - \mathsf{Z^{uu}} &= -\mathbf{e}\mathbf{x}^\top + \mathbf{u}\mathbf{x}^\top + \sum_{k \in K} \mathbf{b}^k\left(\mathbf{a}^k\right)^\top + \mathbf{e}\mathbf{u}^\top - \mathbf{u}\mathbf{u}^\top - \sum_{k \in K} \mathbf{b}^k\left(\mathbf{b}^k\right)^\top \\
&= -\mathbf{v}\mathbf{x}^\top + \mathbf{v}\mathbf{u}^\top + \sum_{k \in K} \mathbf{c}^k\left(\mathbf{d}^k\right)^\top \\
&= \mathsf{Z^{vy}} \geq \mathsf{O},
\end{aligned}
$$

where we used $\mathbf{u} + \mathbf{v} = \mathbf{e}$, $\mathbf{x} + \mathbf{y} = \mathbf{u}$, (20), (21), and $\mathsf{Z} \in \mathcal{D}_{4n+1}$, establishing (14).

Using diag $(\mathsf{Z^{uu}}) = \mathbf{u}$, (10), and (14),

$$\mathbf{o} \leq \mathrm{diag}\left(\mathbf{x}\mathbf{e}^\top - \mathsf{Z^{xu}}\right) = \mathbf{x} - \mathrm{diag}\left(\mathsf{Z^{xu}}\right) \quad \text{and}$$

$$\mathbf{o} \leq \mathrm{diag}\left(-\mathbf{e}\mathbf{x}^\top + (\mathsf{Z^{xu}})^\top + \mathbf{e}\mathbf{u}^\top - \mathsf{Z^{uu}}\right) = -\mathbf{x} + \mathrm{diag}\left(\mathsf{Z^{xu}}\right) + \mathbf{u} - \mathrm{diag}\left(\mathsf{Z^{uu}}\right) = -\mathbf{x} + \mathrm{diag}\left(\mathsf{Z^{xu}}\right),$$

which together yield (15). By (15) and diag $(\mathsf{Z^{uu}}) = \mathbf{u}$, it is easy to see that (16) and (17) follow from (10) and (14), respectively. Using (13), diag $(\mathsf{Z^{uu}}) = \mathbf{u}$ and $\mathbf{u} + \mathbf{v} = \mathbf{e}$, we obtain

$$\mathrm{diag}\left(\mathbf{e}\mathbf{e}^\top - \mathbf{e}\mathbf{u}^\top - \mathbf{u}\mathbf{e}^\top + \mathsf{Z^{uu}}\right) = \mathbf{e} - 2\,\mathbf{u} + \mathbf{u} = \mathbf{v} = \mathrm{diag}\left(\mathsf{Z^{vv}}\right),$$

which establishes (18). Finally, (19) follows from (18), diag $(\mathsf{Z^{uu}}) = \mathbf{u}$, diag $(\mathsf{Z^{uu}}) + 2\,\mathrm{diag}\,(\mathsf{Z^{uv}}) + \mathrm{diag}\,(\mathsf{Z^{vv}}) = \mathbf{e}$, and $\mathbf{u} + \mathbf{v} = \mathbf{e}$. This completes the proof. $\qquad\square$

In view of Lemma 3.1, let us now consider the following optimization problem:

$$
\begin{aligned}
\nu(\mathrm{D1B}(\rho)) := \min_{\mathsf{W} \in \mathcal{S}^{2n+1}} \quad & \langle \mathsf{Q}, \mathsf{W^{xx}} \rangle \\
\text{s.t.} \quad & \mathbf{e}^\top \mathbf{x} = 1 \\
& \mathbf{e}^\top \mathbf{u} = \rho \\
& \langle \mathsf{E}, \mathsf{W^{xx}} \rangle = 1 \\
& \langle \mathsf{E}, \mathsf{W^{uu}} \rangle = \rho^2 \\
& \mathrm{diag}\,(\mathsf{W^{uu}}) = \mathbf{u} \\
& \mathbf{x}\mathbf{e}^\top - \mathsf{W^{xu}} \geq \mathsf{O} \\
& -\mathsf{W^{xx}} + \mathsf{W^{xu}} \geq \mathsf{O} \\
& \mathsf{W^{xx}} - \mathsf{W^{xu}} - (\mathsf{W^{xu}})^\top + \mathsf{W^{uu}} \geq \mathsf{O} \\
& \mathbf{e}\mathbf{e}^\top - \mathbf{e}\mathbf{u}^\top - \mathbf{u}\mathbf{e}^\top + \mathsf{W^{uu}} \geq \mathsf{O} \\
& -\mathbf{e}\mathbf{x}^\top + (\mathsf{W^{xu}})^\top + \mathbf{e}\mathbf{u}^\top - \mathsf{W^{uu}} \geq \mathsf{O} \\
& \mathsf{W^{xx}} \geq \mathsf{O} \\
& \mathsf{W} := \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top \\ \mathbf{x} & \mathsf{W^{xx}} & \mathsf{W^{xu}} \\ \mathbf{u} & (\mathsf{W^{xu}})^\top & \mathsf{W^{uu}} \end{pmatrix} \succeq \mathsf{O}.
\end{aligned}
\qquad (\mathrm{D1B}(\rho))
$$

It is easy to see that $(\mathrm{D1B}(\rho))$ is a convex relaxation of $(\mathrm{P1}(\rho))$ since, for any feasible solution $(\mathbf{x}, \mathbf{u}) \in \mathbb{R}^n \times \mathbb{R}^n$ of $(\mathrm{P1}(\rho))$,

$$\mathsf{W} = \begin{pmatrix} 1 \\ \mathbf{x} \\ \mathbf{u} \end{pmatrix} \begin{pmatrix} 1 \\ \mathbf{x} \\ \mathbf{u} \end{pmatrix}^\top \in \mathcal{S}^{2n+1}$$

is a feasible solution of $(\mathrm{D1B}(\rho))$ with the same objective function value. Note that $(\mathrm{D1B}(\rho))$ has $n + 4$ linear equality constraints, $(9/2)n^2 + (3/2)n$ linear inequality constraints, and a PSD constraint in $\mathcal{S}^{2n+1}$. It is worth noticing that the inequality constraints of $(\mathrm{D1B}(\rho))$ are precisely given by all the RLT (reformulation-linearization technique) constraints obtained from the inequalities $\mathbf{x} \geq \mathbf{o}$, $\mathbf{x} \leq \mathbf{u}$, and $\mathbf{u} \leq \mathbf{e}$ (see, e.g., [40]).

First, we present some properties of feasible solutions of $(\mathrm{D1B}(\rho))$.

Lemma 3.2. *Let* $\mathsf{W} \in \mathcal{S}^{2n+1}$ *be* $(D1B(\rho))$*-feasible. Then,*

$$\mathsf{W} \quad \in \quad \mathcal{D}_{2n+1}, \tag{22}$$

$$-(\mathsf{W}^{\mathbf{xu}})^\top + \mathsf{W}^{\mathbf{uu}} \quad \geq \quad \mathsf{O}, \tag{23}$$

$$\mathbf{ue}^\top - \mathsf{W}^{\mathbf{uu}} \quad \geq \quad \mathsf{O}, \tag{24}$$

$$\mathbf{u} \quad \leq \quad \mathbf{e}, \tag{25}$$

$$\mathbf{x} \quad \leq \quad \mathbf{u}. \tag{26}$$

Proof. Let $\mathsf{W} \in \mathcal{S}^{2n+1}$ be $(D1B(\rho))$-feasible. Since $\mathsf{W}^{\mathbf{xx}} \geq \mathsf{O}$ and $-\mathsf{W}^{\mathbf{xx}} + \mathsf{W}^{\mathbf{xu}} \geq \mathsf{O}$, we obtain $\mathsf{W}^{\mathbf{xu}} \geq \mathsf{O}$. By $\mathbf{xe}^\top - \mathsf{W}^{\mathbf{xu}} \geq \mathsf{O}$, we arrive at $\mathbf{x} \geq \mathbf{o}$. Since $\mathrm{diag}\,(\mathsf{W}^{\mathbf{uu}}) = \mathbf{u}$ and $\mathsf{W} \succeq \mathsf{O}$, we have $\mathbf{u} \geq \mathbf{o}$. Then, adding the inequalities $-\mathsf{W}^{\mathbf{xx}} + \mathsf{W}^{\mathbf{xu}} \geq \mathsf{O}$ and $\mathsf{W}^{\mathbf{xx}} - \mathsf{W}^{\mathbf{xu}} - (\mathsf{W}^{\mathbf{xu}})^\top + \mathsf{W}^{\mathbf{uu}} \geq \mathsf{O}$ yields (23), which, in turn, establishes (22) since $\mathsf{W}^{\mathbf{uu}} \geq (\mathsf{W}^{\mathbf{xu}})^\top \geq \mathsf{O}$. Similarly, adding the inequalities $\left(\mathbf{xe}^\top - \mathsf{W}^{\mathbf{xu}}\right)^\top = \mathbf{ex}^\top - (\mathsf{W}^{\mathbf{xu}})^\top \geq \mathsf{O}$ and $-\mathbf{ex}^\top + (\mathsf{W}^{\mathbf{xu}})^\top + \mathbf{eu}^\top - \mathsf{W}^{\mathbf{uu}} \geq \mathsf{O}$, we arrive at (24). Adding (24) to $\mathbf{ee}^\top - \mathbf{eu}^\top - \mathbf{ue}^\top + \mathsf{W}^{\mathbf{uu}} \geq \mathsf{O}$, we get $\mathbf{e}(\mathbf{e} - \mathbf{u})^\top \geq \mathsf{O}$, which yields (25). Similarly, using $\mathrm{diag}\,(\mathsf{W}^{\mathbf{uu}}) = \mathbf{u}$,

$$\mathbf{o} \leq \mathrm{diag}\,\left(\mathbf{xe}^\top - \mathsf{W}^{\mathbf{xu}}\right) \quad = \quad \mathbf{x} - \mathrm{diag}\,(\mathsf{W}^{\mathbf{xu}}) \quad \text{and}$$

$$\mathbf{o} \leq \mathrm{diag}\,\left(-\mathbf{ex}^\top + (\mathsf{W}^{\mathbf{xu}})^\top + \mathbf{eu}^\top - \mathsf{W}^{\mathbf{uu}}\right) \quad = \quad -\mathbf{x} + \mathrm{diag}\,(\mathsf{W}^{\mathbf{xu}}) + \mathbf{u} - \mathrm{diag}\,(\mathsf{W}^{\mathbf{uu}}) = -\mathbf{x} + \mathrm{diag}\,(\mathsf{W}^{\mathbf{xu}}),$$

which together yield $\mathrm{diag}\,(\mathsf{W}^{\mathbf{xu}}) = \mathbf{x}$. Combining this with (23), we get

$$\mathbf{o} \leq \mathrm{diag}\,\left(-(\mathsf{W}^{\mathbf{xu}})^\top + \mathsf{W}^{\mathbf{uu}}\right) = -\mathrm{diag}\,(\mathsf{W}^{\mathbf{xu}}) + \mathrm{diag}\,(\mathsf{W}^{\mathbf{uu}}) = -\mathbf{x} + \mathbf{u},$$

where we used $\mathrm{diag}\,(\mathsf{W}^{\mathbf{uu}}) = \mathbf{u}$. This establishes (26) and completes the proof. $\square$

Our next result shows that $(D1B(\rho))$ is equivalent to the DNN relaxation $(D1A(\rho))$ of $(CP1A(\rho))$.

Proposition 3.1. $(D1A(\rho))$ *and* $(D1B(\rho))$ *are equivalent to each other. Therefore,* $\nu(D1A(\rho)) = \nu(D1B(\rho))$.

Proof. Let $\mathsf{W} \in \mathcal{S}^{2n+1}$ be $(D1B(\rho))$-feasible. Let us define $\mathsf{Z} \in \mathcal{S}^{4n+1}$ as follows:

$$\mathsf{Z} = \begin{pmatrix} 1 & \mathbf{o}^\top & \mathbf{o}^\top \\ \mathbf{o} & \mathsf{I} & \mathbf{o} \\ \mathbf{o} & \mathsf{O} & \mathsf{I} \\ \mathbf{e} & \mathsf{O} & -\mathsf{I} \\ \mathbf{o} & -\mathsf{I} & \mathsf{I} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top \\ \mathbf{x} & \mathsf{W}^{\mathbf{xx}} & \mathsf{W}^{\mathbf{xu}} \\ \mathbf{u} & (\mathsf{W}^{\mathbf{xu}})^\top & \mathsf{W}^{\mathbf{uu}} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{o}^\top & \mathbf{o}^\top \\ \mathbf{o} & \mathsf{I} & \mathbf{o} \\ \mathbf{o} & \mathsf{O} & \mathsf{I} \\ \mathbf{e} & \mathsf{O} & -\mathsf{I} \\ \mathbf{o} & -\mathsf{I} & \mathsf{I} \end{pmatrix}^\top. \tag{27}$$

Therefore,

$$\mathsf{Z} = \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top & \mathbf{e}^\top - \mathbf{u}^\top & -\mathbf{x}^\top + \mathbf{u}^\top \\ \mathbf{x} & \mathsf{W}^{\mathbf{xx}} & \mathsf{W}^{\mathbf{xu}} & \mathbf{xe}^\top - \mathsf{W}^{\mathbf{xu}} & -\mathsf{W}^{\mathbf{xx}} + \mathsf{W}^{\mathbf{xu}} \\ \mathbf{u} & (\mathsf{W}^{\mathbf{xu}})^\top & \mathsf{W}^{\mathbf{uu}} & \mathbf{ue}^\top - \mathsf{W}^{\mathbf{uu}} & -(\mathsf{W}^{\mathbf{xu}})^\top + \mathsf{W}^{\mathbf{uu}} \\ \mathbf{e} - \mathbf{u} & \mathbf{ex}^\top - (\mathsf{W}^{\mathbf{xu}})^\top & \mathbf{eu}^\top - \mathsf{W}^{\mathbf{uu}} & \mathbf{ee}^\top - \mathbf{eu}^\top - \mathbf{ue}^\top + \mathsf{W}^{\mathbf{uu}} & -\mathbf{ex}^\top + (\mathsf{W}^{\mathbf{xu}})^\top + \mathbf{eu}^\top - \mathsf{W}^{\mathbf{uu}} \\ -\mathbf{x} + \mathbf{u} & -\mathsf{W}^{\mathbf{xx}} + (\mathsf{W}^{\mathbf{xu}})^\top & -\mathsf{W}^{\mathbf{xu}} + \mathsf{W}^{\mathbf{uu}} & -\mathbf{xe}^\top + \mathsf{W}^{\mathbf{xu}} + \mathbf{ue}^\top - \mathsf{W}^{\mathbf{uu}} & \mathsf{W}^{\mathbf{xx}} - \mathsf{W}^{\mathbf{xu}} - (\mathsf{W}^{\mathbf{xu}})^\top + \mathsf{W}^{\mathbf{uu}} \end{pmatrix}. \tag{28}$$

We claim that $\mathsf{Z} \in \mathcal{S}^{4n+1}$ is $(D1A(\rho))$-feasible. By Lemma 3.2, (27), and (28), we conclude that $\mathsf{Z} \in \mathcal{D}_{4n+1}$. Since $\mathbf{v} = \mathbf{e} - \mathbf{u}$, $\mathbf{y} = \mathbf{u} - \mathbf{x}$, $\mathsf{Z}^{\mathbf{xx}} = \mathsf{W}^{\mathbf{xx}}$, and $\mathsf{Z}^{\mathbf{uu}} = \mathsf{W}^{\mathbf{uu}}$, the first seven sets of equality constraints of $(D1A(\rho))$

are satisfied. Considering the eighth set of equality constraints of (D1A(ρ)) and substituting the corresponding submatrices in (28), we obtain

$$
\begin{aligned}
\mathrm{diag}\,(\mathsf{Z^{xx}}) &= \mathrm{diag}\,(\mathsf{W^{xx}}) \\
\mathrm{diag}\,(\mathsf{Z^{yy}}) &= \mathrm{diag}\,(\mathsf{W^{xx}}) - 2\,\mathrm{diag}\,(\mathsf{W^{xu}}) + \mathrm{diag}\,(\mathsf{W^{uu}}) \\
\mathrm{diag}\,(\mathsf{Z^{uu}}) &= \mathrm{diag}\,(\mathsf{W^{uu}}) \\
\mathrm{diag}\,(\mathsf{Z^{xy}}) &= -\mathrm{diag}\,(\mathsf{W^{xx}}) + \mathrm{diag}\,(\mathsf{W^{xu}}) \\
\mathrm{diag}\,(\mathsf{Z^{xu}}) &= \mathrm{diag}\,(\mathsf{W^{xu}}) \\
\mathrm{diag}\,(\mathsf{Z^{uy}}) &= -\mathrm{diag}\,(\mathsf{W^{xu}}) + \mathrm{diag}\,(\mathsf{W^{uu}}),
\end{aligned}
$$

which implies that

$$
\mathrm{diag}\,(\mathsf{Z^{xx}}) + \mathrm{diag}\,(\mathsf{Z^{yy}}) + \mathrm{diag}\,(\mathsf{Z^{uu}}) + 2\left[\mathrm{diag}\,(\mathsf{Z^{xy}}) - \mathrm{diag}\,(\mathsf{Z^{xu}}) - \mathrm{diag}\,(\mathsf{Z^{uy}})\right] = \mathbf{o}\,.
$$

Finally, consider the last set of equality constraints of (D1A(ρ)):

$$
\mathrm{diag}\,(\mathsf{Z^{uu}}) + 2\,\mathrm{diag}\,(\mathsf{Z^{uv}}) + \mathrm{diag}\,(\mathsf{Z^{vv}}) = \mathrm{diag}\,(\mathsf{W^{uu}}) + 2\,(\mathbf{u} - \mathrm{diag}\,(\mathsf{W^{uu}})) + \mathbf{e} - 2\,\mathbf{u} + \mathrm{diag}\,(\mathsf{W^{uu}}) = \mathbf{e}\,,
$$

which implies that $\mathsf{Z} \in \mathcal{S}^{4n+1}$ is (D1A(ρ))-feasible and achieves the same objective function value as W in (D1B(ρ)).

Conversely, let $\mathsf{Z} \in \mathcal{S}^{4n+1}$ be (D1A(ρ))-feasible. Let $\mathsf{W} \in \mathcal{S}^{2n+1}$ be given by the top left 3×3-block of Z, i.e.,

$$
\mathsf{W} = \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top \\ \mathbf{x} & \mathsf{Z^{xx}} & \mathsf{Z^{xu}} \\ \mathbf{u} & (\mathsf{Z^{xu}})^\top & \mathsf{Z^{uu}} \end{pmatrix}\,. \tag{29}
$$

Clearly, $\mathsf{W} \in \mathcal{D}_{2n+1}$ and satisfies all linear equality constraints and $\mathsf{W^{xx}} = \mathsf{Z^{xx}} \geq \mathsf{O}$ in (D1B(ρ)). By Lemma 3.1, we conclude that W satisfies each of the remaining linear inequality constraints and achieves the same objective function value as Z in (D1A(ρ)). It follows that $\nu(\mathrm{D1A}(\rho)) = \nu(\mathrm{D1B}(\rho))$. $\qquad\square$

Recall that (D1A(ρ)) consists of $5n+4$ linear equality constraints and **a doubly nonnegative constraint** in $\mathcal{S}^{4n+1}$. In contrast, (D1B(ρ)) has $n+4$ linear equality constraints, $(9/2)n^2 + (3/2)n$ linear inequality constraints, and **a PSD constraint** in $\mathcal{S}^{2n+1}$. By Proposition 3.1, we conclude that the lower bound arising from the DNN relaxation of (CP1A(ρ)) can be computed by solving a conic optimization problem in a much smaller dimension.

3.2   DNN relaxations of the complementarity constraint formulation

In this section, we focus on the DNN relaxation of the copositive optimization problem (CP2A(ρ)). By replacing the intractable conic constraint $\mathsf{Z} \in \mathcal{CP}_{3n+1}$ in (CP2A(ρ)) by $\mathsf{Z} \in \mathcal{D}_{3n+1}$, we obtain the following convex optimization problem:

$$\nu(\mathrm{D2A}(\rho)) := \min_{\mathsf{Y}\in\mathcal{S}^{3n+1}} \quad \langle \mathsf{Q}, \mathsf{Y^{xx}} \rangle$$

$$\begin{aligned}
\text{s.t.} \qquad \mathbf{e}^\top \mathbf{x} &= 1 \\
\mathbf{e}^\top \mathbf{u} &= \rho \\
\mathbf{u} + \mathbf{v} &= \mathbf{e} \\
\langle \mathsf{E}, \mathsf{Y^{xx}} \rangle &= 1 \\
\langle \mathsf{E}, \mathsf{Y^{uu}} \rangle &= \rho^2 \\
\mathrm{diag}\,(\mathsf{Y^{uu}}) &= \mathbf{u} \qquad\qquad (\mathrm{D2A}(\rho))\\
\mathbf{e}^\top \mathrm{diag}\,(\mathsf{Y^{xv}}) &= 0 \\
\mathrm{diag}\,(\mathsf{Y^{uu}}) + 2\,\mathrm{diag}\,(\mathsf{Y^{uv}}) + \mathrm{diag}\,(\mathsf{Y^{vv}}) &= \mathbf{e}
\end{aligned}$$

$$\mathsf{Y} := \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top & \mathbf{v}^\top \\ \mathbf{x} & \mathsf{Y^{xx}} & \mathsf{Y^{xu}} & \mathsf{Y^{xv}} \\ \mathbf{u} & (\mathsf{Y^{xu}})^\top & \mathsf{Y^{uu}} & \mathsf{Y^{uv}} \\ \mathbf{v} & (\mathsf{Y^{xv}})^\top & (\mathsf{Y^{uv}})^\top & \mathsf{Y^{vv}} \end{pmatrix} \in \mathcal{D}_{3n+1}.$$

Once again, it is easy to verify that $(\mathrm{D2A}(\rho))$ is a tractable convex relaxation of $(\mathrm{P2A}(\rho))$. Note that $(\mathrm{D2A}(\rho))$ consists of $3n+5$ linear equality constraints and a doubly nonnegative constraint in $\mathcal{S}^{3n+1}$.

First, we establish several properties of feasible solutions of $(\mathrm{D2A}(\rho))$.

Lemma 3.3. *Let* $\mathsf{Y} \in \mathcal{S}^{3n+1}$ *be* $(\mathrm{D2A}(\rho))$-*feasible. Then, the following relations hold:*

$$\begin{aligned}
\mathbf{x}\mathbf{e}^\top - \mathsf{Y^{xu}} &= \mathsf{Y^{xv}} \geq \mathsf{O}, & (30) \\
\mathbf{e}\mathbf{e}^\top - \mathbf{e}\mathbf{u}^\top - \mathbf{u}\mathbf{e}^\top + \mathsf{Y^{uu}} &= \mathsf{Y^{vv}} \geq \mathsf{O}, & (31) \\
\mathbf{u}\mathbf{e}^\top - \mathsf{Y^{uu}} &= \mathsf{Y^{uv}} \geq \mathsf{O}, & (32) \\
\mathit{diag}\,(\mathsf{Y^{xv}}) &= \mathbf{o}, & (33) \\
\mathit{diag}\,(\mathsf{Y^{xu}}) &= \mathbf{x}, & (34) \\
\mathit{diag}\,(\mathsf{Y^{vv}}) &= \mathbf{v}, & (35) \\
\mathit{diag}\,(\mathsf{Y^{uv}}) &= \mathbf{o}. & (36)
\end{aligned}$$

Proof. Let $\mathsf{Y} \in \mathcal{S}^{3n+1}$ be $(\mathrm{D2A}(\rho))$-feasible. By using the Schur complementation, we obtain

$$\begin{pmatrix} \mathsf{Y^{xx}} & \mathsf{Y^{xu}} & \mathsf{Y^{xv}} \\ (\mathsf{Y^{xu}})^\top & \mathsf{Y^{uu}} & \mathsf{Y^{uv}} \\ (\mathsf{Y^{xv}})^\top & (\mathsf{Y^{uv}})^\top & \mathsf{Y^{vv}} \end{pmatrix} = \begin{pmatrix} \mathbf{x} \\ \mathbf{u} \\ \mathbf{v} \end{pmatrix}\begin{pmatrix} \mathbf{x} \\ \mathbf{u} \\ \mathbf{v} \end{pmatrix}^\top + \sum_{k\in K}\begin{pmatrix} \mathbf{a}^k \\ \mathbf{b}^k \\ \mathbf{c}^k \end{pmatrix}\begin{pmatrix} \mathbf{a}^k \\ \mathbf{b}^k \\ \mathbf{c}^k \end{pmatrix}^\top,$$

where K is a finite set and $\{\mathbf{a}^k, \mathbf{b}^k, \mathbf{c}^k\} \subset \mathbb{R}^n$ for each $k \in K$. Arguing similarly to the proof of Lemma 3.1, we obtain $\mathbf{b}^k = -\mathbf{c}^k$ for each $k \in K$. The relations (30), (31), and (32) can be established in a very similar manner. Since $\mathbf{e}^\top \mathrm{diag}\,(\mathsf{Y^{xv}}) = 0$ and $\mathsf{Y} \in \mathcal{D}_{3n+1}$, we obtain (33). By using $\mathbf{u} + \mathbf{v} = \mathbf{e}$, $\mathbf{b}^k = -\mathbf{c}^k$ for each $k \in K$, and the decomposition above, it is easy to verify that $\mathrm{diag}\,(\mathsf{Y^{xu}}) + \mathrm{diag}\,(\mathsf{Y^{xv}}) = \mathbf{x}$, which establishes (34) due to (33). Similarly, we obtain $\mathrm{diag}\,(\mathsf{Y^{vv}}) = \mathbf{e} - 2\,\mathbf{u} + \mathrm{diag}\,(\mathsf{Y^{uu}}) = \mathbf{e} - \mathbf{u} = \mathbf{v}$ since $\mathrm{diag}\,(\mathsf{Y^{uu}}) = \mathbf{u}$, which yields (35). Finally, we obtain (36) from $\mathrm{diag}\,(\mathsf{Y^{uu}}) = \mathbf{u}$, $\mathrm{diag}\,(\mathsf{Y^{uu}}) + 2\,\mathrm{diag}\,(\mathsf{Y^{uv}}) + \mathrm{diag}\,(\mathsf{Y^{vv}}) = \mathbf{e}$, $\mathbf{u} + \mathbf{v} = \mathbf{e}$, and (35). This completes the proof. $\square$

Once again, we can utilize Lemma 3.3 to obtain the following optimization problem:

$$\nu(\text{D2B}(\rho)) := \min_{\mathsf{S} \in \mathcal{S}^{2n+1}} \qquad \langle \mathsf{Q}, \mathsf{S}^{\mathbf{xx}} \rangle$$

$$\begin{aligned}
\text{s.t.} \qquad \mathbf{e}^\top \mathbf{x} &= 1 \\
\mathbf{e}^\top \mathbf{u} &= \rho \\
\langle \mathsf{E}, \mathsf{S}^{\mathbf{xx}} \rangle &= 1 \\
\langle \mathsf{E}, \mathsf{S}^{\mathbf{uu}} \rangle &= \rho^2 \\
\text{diag}\,(\mathsf{S}^{\mathbf{uu}}) &= \mathbf{u} \\
\text{diag}\,(\mathsf{S}^{\mathbf{xu}}) &= \mathbf{x} \qquad\qquad (\text{D2B}(\rho)) \\
\mathbf{x}\mathbf{e}^\top - \mathsf{S}^{\mathbf{xu}} &\geq \mathsf{O} \\
\mathbf{e}\mathbf{e}^\top - \mathbf{e}\mathbf{u}^\top - \mathbf{u}\mathbf{e}^\top + \mathsf{S}^{\mathbf{uu}} &\geq \mathsf{O} \\
\mathbf{u}\mathbf{e}^\top - \mathsf{S}^{\mathbf{uu}} &\geq \mathsf{O} \\
\mathsf{S} := \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top \\ \mathbf{x} & \mathsf{S}^{\mathbf{xx}} & \mathsf{S}^{\mathbf{xu}} \\ \mathbf{u} & (\mathsf{S}^{\mathbf{xu}})^\top & \mathsf{S}^{\mathbf{uu}} \end{pmatrix} &\in \quad \mathcal{D}_{2n+1}.
\end{aligned}$$

Note that $(\text{D2B}(\rho))$ has $2n+4$ linear equality constraints, $(5/2)n^2 + (1/2)n$ linear inequality constraints, and a doubly nonnegative constraint in $\mathcal{S}^{2n+1}$. Once again, it is worth noticing that $(\text{D2B}(\rho))$ contains all the RLT constraints obtained from the inequalities $\mathbf{x} \geq \mathbf{o}$, $\mathbf{u} \geq \mathbf{o}$, and $\mathbf{u} \leq \mathbf{e}$.

Our next result establishes the equivalence between $(\text{D2A}(\rho))$ and $(\text{D2B}(\rho))$.

Proposition 3.2. $(\text{D2A}(\rho))$ and $(\text{D2B}(\rho))$ are equivalent to each other. Therefore, $\nu(\text{D2A}(\rho)) = \nu(\text{D2B}(\rho))$.

Proof. The proof is very similar to that of Proposition 3.1. Let $\mathsf{S} \in \mathcal{S}^{2n+1}$ be $(\text{D2B}(\rho))$-feasible. Let us define $\mathsf{Y} \in \mathcal{S}^{3n+1}$ as follows:

$$\mathsf{Y} = \begin{pmatrix} 1 & \mathbf{o}^\top & \mathbf{o}^\top \\ \mathbf{o} & \mathsf{I} & \mathbf{o} \\ \mathbf{o} & \mathsf{O} & \mathsf{I} \\ \mathbf{e} & \mathsf{O} & -\mathsf{I} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top \\ \mathbf{x} & \mathsf{S}^{\mathbf{xx}} & \mathsf{S}^{\mathbf{xu}} \\ \mathbf{u} & (\mathsf{S}^{\mathbf{xu}})^\top & \mathsf{S}^{\mathbf{uu}} \end{pmatrix} \begin{pmatrix} 1 & \mathbf{o}^\top & \mathbf{o}^\top \\ \mathbf{o} & \mathsf{I} & \mathbf{o} \\ \mathbf{o} & \mathsf{O} & \mathsf{I} \\ \mathbf{e} & \mathsf{O} & -\mathsf{I} \end{pmatrix}^\top. \tag{37}$$

Therefore,

$$\mathsf{Y} = \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top & \mathbf{e}^\top - \mathbf{u}^\top \\ \mathbf{x} & \mathsf{S}^{\mathbf{xx}} & \mathsf{S}^{\mathbf{xu}} & \mathbf{x}\mathbf{e}^\top - \mathsf{S}^{\mathbf{xu}} \\ \mathbf{u} & (\mathsf{S}^{\mathbf{xu}})^\top & \mathsf{S}^{\mathbf{uu}} & \mathbf{u}\mathbf{e}^\top - \mathsf{S}^{\mathbf{uu}} \\ \mathbf{e} - \mathbf{u} & \mathbf{e}\mathbf{x}^\top - (\mathsf{S}^{\mathbf{xu}})^\top & \mathbf{e}\mathbf{u}^\top - \mathsf{S}^{\mathbf{uu}} & \mathbf{e}\mathbf{e}^\top - \mathbf{e}\mathbf{u}^\top - \mathbf{u}\mathbf{e}^\top + \mathsf{S}^{\mathbf{uu}} \end{pmatrix}. \tag{38}$$

We claim that $\mathsf{Y} \in \mathcal{S}^{3n+1}$ is $(\text{D2A}(\rho))$-feasible. Since $\text{diag}\,(\mathsf{S}^{\mathbf{uu}}) = \mathbf{u}$, and $\mathsf{S}^{\mathbf{uu}} - \mathbf{u}\mathbf{u}^\top \succeq 0$, we obtain $\mathbf{u} \leq \mathbf{e}$. Therefore, by (37) and (38), we conclude that $\mathsf{Y} \in \mathcal{D}_{4n+1}$. Since $\mathbf{v} = \mathbf{e} - \mathbf{u}$, $\mathsf{Y}^{\mathbf{xx}} = \mathsf{S}^{\mathbf{xx}}$, $\mathsf{Y}^{\mathbf{xu}} = \mathsf{S}^{\mathbf{xu}}$, and $\mathsf{Y}^{\mathbf{uu}} = \mathsf{S}^{\mathbf{uu}}$, the first six sets of equality constraints of $(\text{D2A}(\rho))$ are satisfied. Consider the seventh set of equality constraints of $(\text{D2A}(\rho))$:

$$\mathbf{e}^\top \text{diag}\,(\mathsf{Y}^{\mathbf{xv}}) = \mathbf{e}^\top \text{diag}\,(\mathbf{x}\mathbf{e}^\top - \mathsf{S}^{\mathbf{xu}}) = \mathbf{e}^\top\,[\mathbf{x} - \text{diag}\,(\mathsf{S}^{\mathbf{xu}})] = 0,$$

where we used $\text{diag}\,(\mathsf{S}^{\mathbf{xu}}) = \mathbf{x}$. Finally, the last set of equality constraints of $(\text{D2A}(\rho))$ can be directly verified using (38). Therefore, $\mathsf{Y} \in \mathcal{S}^{3n+1}$ is $(\text{D2A}(\rho))$-feasible and achieves the same objective function value as S in $(\text{D2B}(\rho))$.

Conversely, let $\mathsf{Y} \in \mathcal{S}^{3n+1}$ be (D2A(ρ))-feasible. Let $\mathsf{S} \in \mathcal{S}^{2n+1}$ be given by the top left 3×3-block of Z, i.e.,

$$\mathsf{S} = \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{u}^\top \\ \mathbf{x} & \mathsf{Y^{xx}} & \mathsf{Y^{xu}} \\ \mathbf{u} & (\mathsf{Y^{xu}})^\top & \mathsf{Y^{uu}} \end{pmatrix}. \tag{39}$$

Clearly, $\mathsf{S} \in \mathcal{D}_{2n+1}$ and satisfies all linear equality and inequality constraints in (D2B(ρ)) by Lemma 3.3. Furthermore, S has the same objective function value in (D2B(ρ)) as that of Y in (D2A(ρ)). It follows that $\nu(\text{D2A}(\rho)) = \nu(\text{D2B}(\rho))$. $\qquad\square$

A comparison of (D2A(ρ)) and (D2B(ρ)) reveals that the former consists of $3n+5$ linear equality constraints and **a doubly nonnegative constraint** in $\mathcal{S}^{3n+1}$ whereas the latter has $2n+4$ linear equality constraints, $(5/2)n^2+(1/2)n$ linear inequality constraints, and **a doubly nonnegative constraint** in $\mathcal{S}^{2n+1}$. Proposition 3.2 implies that the lower bound arising from the DNN relaxation of (CP2A(ρ)) can be computed by solving a conic optimization problem in a smaller dimension.

3.3 Comparing the two DNN relaxations

Now we compare the lower bounds arising from the two DNN relaxations of (CP1A(ρ)) and (CP2A(ρ)).

Proposition 3.3. (D1A(ρ)) *(or equivalently, (D1B(ρ))) is at least as tight as (D2A(ρ)) (or equivalently, (D2B(ρ))), i.e.,*

$$\nu(D2A(\rho)) = \nu(D2B(\rho)) \leq \nu(D1A(\rho)) = \nu(D1B(\rho)) \leq \ell_\rho(\mathsf{Q}).$$

Proof. Let $\mathsf{Z} \in \mathcal{S}^{4n+1}$ be (D1A(ρ))-feasible. Let $\mathsf{Y} \in \mathcal{S}^{3n+1}$ be given by the top left 4×4-block of Z, i.e.,

$$\mathsf{Y} = \begin{pmatrix} 1 & \mathbf{x}^\top & \mathbf{v}^\top & \mathbf{u}^\top \\ \mathbf{x} & \mathsf{Z^{xx}} & \mathsf{Z^{xv}} & \mathsf{Z^{xu}} \\ \mathbf{u} & (\mathsf{Z^{xu}})^\top & \mathsf{Z^{uu}} & \mathsf{Z^{uv}} \\ \mathbf{v} & (\mathsf{Z^{xv}})^\top & (\mathsf{Z^{uv}})^\top & \mathsf{Z^{vv}} \end{pmatrix}.$$

We claim that $\mathsf{Y} \in \mathcal{S}^{3n+1}$ is (D2A(ρ))-feasible. Clearly, $\mathsf{Y} \in \mathcal{D}_{3n+1}$ and all constraints of (D2A(ρ)) are satisfied by Lemma 3.1. Therefore, for every (D1A(ρ))-feasible solution, there exists a corresponding (D2A(ρ))-feasible solution with the same objective function value. By Propositions 3.1 and 3.2, we conclude that $\nu(\text{D2A}(\rho)) = \nu(\text{D2B}(\rho)) \leq \nu(\text{D1A}(\rho)) = \nu(\text{D1B}(\rho)) \leq \ell_\rho(\mathsf{Q})$. $\qquad\square$

Our next example illustrates that we can have $\nu(\text{D2A}(\rho)) < \nu(\text{D1A}(\rho))$. i.e., (D2A($\rho$)) can be strictly weaker than (D1A(ρ)).

Example 3.1. *Consider the following instance, where $n = 6$ and $\rho = 3$:*

$$\mathsf{Q} = \begin{pmatrix} 2.6947 & -0.2028 & -1.1144 & -2.4230 & -2.1633 & 0.7710 \\ -0.2028 & 5.1998 & 0.5005 & -2.0941 & 0.7828 & -3.3611 \\ -1.1144 & 0.5005 & 5.2918 & 1.4119 & -1.1526 & -1.9723 \\ -2.4230 & -2.0941 & 1.4119 & 4.1140 & -0.2025 & 0.3132 \\ -2.1633 & 0.7828 & -1.1526 & -0.2025 & 7.6645 & -0.0170 \\ 0.7710 & -3.3611 & -1.9723 & 0.3132 & -0.0170 & 3.3040 \end{pmatrix}.$$

For this instance, we have $\nu(D2A(\rho)) = \nu(D2B(\rho)) \approx 0.1320 < \nu(D1A(\rho)) = \nu(D1B(\rho)) \approx 0.1333$. In addition, we observe that the computed optimal solution S^* of (D2B(ρ)) is not feasible for (D1B(ρ)). For instance, $(S_{31}^{\mathbf{xu}})^* - (S_{31}^{\mathbf{xx}})^* \approx -0.00137 < 0$.

4 Generating nontrivial instances of sparse StQPs

In this section, we discuss procedures for generating nontrivial instances of the sparse StQP problem (StQP(ρ)). To this end, we say that an instance of (StQP(ρ)) is *nontrivial* if the sparsity constraint is violated by each optimal solution of the corresponding instance of StQP given by (StQP) (or, equivalently, (StQP(ρ)) with $\rho = n$). For a nontrivial instance of the sparse StQP, it follows that the sparsity constraint cuts off all optimal solutions of the corresponding StQP instance without the sparsity constraints.

We aim to construct a nontrivial instance of (StQP(ρ)) in two steps. First, we discuss how to construct an instance of (StQP) with a prespecified *unique* optimal solution. We then choose the sparsity parameter ρ to be strictly smaller than the sparsity of the designated unique optimal solution of (StQP). It follows that the resulting instance of (StQP(ρ)) is nontrivial since the unique optimal solution of (StQP) is cut off by the sparsity constraint.

In Section 4.1, we consider generating an instance of (StQP) with a unique optimal solution. We then present simple algorithms for generating nontrivial instances of (StQP(ρ)) in Section 4.2.

4.1 Generating StQP instances with a unique optimal solution

Let us next focus on generating an instance of (StQP) with a unique optimal solution. To that end, we first recall that an StQP can be reformulated as the following copositive optimization problem [9]:

$$\ell_n(\mathsf{Q}) = \min_{\mathbf{x} \in \mathbb{R}^n} \left\{ \mathbf{x}^\top \mathsf{Q} \mathbf{x} : \mathbf{x} \in F_n \right\} = \min_{\mathsf{X} \in \mathcal{S}^n} \left\{ \langle \mathsf{Q}, \mathsf{X} \rangle : \langle \mathsf{E}, \mathsf{X} \rangle = 1, \quad \mathsf{X} \in \mathcal{CP}_n \right\},$$

where F_n is given by (1) and $\mathsf{E} = \mathbf{e}\mathbf{e}^\top \in \mathcal{S}^n$.

The doubly nonnegative (DNN) relaxation of an StQP is therefore given by

$$\mu(\mathsf{Q}) = \min_{\mathsf{X} \in \mathcal{S}^n} \left\{ \langle \mathsf{Q}, \mathsf{X} \rangle : \langle \mathsf{E}, \mathsf{X} \rangle = 1, \quad \mathsf{X} \in \mathcal{D}_n \right\}. \tag{40}$$

Note that $\mu(\mathsf{Q}) \leq \ell_n(\mathsf{Q})$. For a given instance of StQP, we say that its DNN relaxation is *exact* if $\mu(\mathsf{Q}) = \ell_n(\mathsf{Q})$ and *inexact* otherwise.

We will review the following useful results, which will play a fundamental role in our instance construction procedures.

Theorem 4.1. *Let $\mathbf{x} \in F_n$, where F_n is given by (1).*

(i) *$\mathbf{x}$ is a globally optimal solution of (StQP) if and only if there exist $\mathsf{M} \in \mathbf{bd} \, \mathcal{COP}_n$ and $\lambda \in \mathbb{R}$ such that $\mathbf{x} \in \mathbf{V}^\mathsf{M}$ and $\mathsf{Q} = \mathsf{M} + \lambda \mathsf{E}$, where $\mathbf{bd} \, \mathcal{COP}_n$ and $\mathbf{V}^\mathsf{M}$ are given by (8) and (9), respectively. Furthermore, in this case, λ is the optimal value of (StQP), and the set of optimal solutions of (StQP) is given by $\mathbf{V}^\mathsf{M}$.*

(ii) *$\mathbf{x}$ is a globally optimal solution of (StQP) and its doubly nonnegative relaxation is exact if and only if there exist $\mathsf{R} \succeq \mathsf{O}$, $\mathsf{N} \geq \mathsf{O}$, and $\lambda \in \mathbb{R}$ such that $\mathbf{x}^\top \mathsf{N} \mathbf{x} = 0$ and $\mathsf{Q} = (\mathsf{I} - \mathbf{e}\mathbf{x}^\top)\mathsf{R}(\mathsf{I} - \mathbf{x}\mathbf{e}^\top) + \mathsf{N} + \lambda \mathsf{E}$.*

(iii) $\mathbf{x}$ *is a globally optimal solution of* (StQP) *and its doubly nonnegative relaxation is inexact if and only if there exist* $\mathsf{M} \in \mathbf{bd}\, \mathcal{COP}_n \backslash \mathcal{SPN}_n$ *and* $\lambda \in \mathbb{R}$ *such that* $\mathbf{x} \in \mathbf{V}^\mathsf{M}$ *and* $\mathsf{Q} = \mathsf{M} + \lambda\, \mathsf{E}$, *where* $\mathbf{bd}\, \mathcal{COP}_n$ *and* $\mathbf{V}^\mathsf{M}$ *are given by* (8) *and* (9), *respectively.*

Proof. The reader is referred to [7] for (i), and to [24] for (ii) and (iii). □

In view of Theorem 4.1, we first give a simple recipe to construct an instance of (StQP) such that it has a unique optimal solution and its doubly nonnegative relaxation is exact.

Proposition 4.1. *Let* $\mathbf{x} \in F_n$, *where* F_n *is given by* (1). *Let* $\mathbb{A} = \{j \in \{1, \ldots, n\} : \mathbf{x}_j > 0\}$ *and* $\mathbb{B} = \{j \in \{1, \ldots, n\} : \mathbf{x}_j = 0\}$. *Then, for any* $\mathsf{R} \succ \mathsf{O}$, *any* $\mathsf{N} \in \mathcal{S}^n$ *is such that* $\mathsf{N}_{\mathbb{A}\mathbb{A}} = \mathsf{O}$, $\mathsf{N}_{\mathbb{A}\mathbb{B}} \geq \mathsf{O}$, *and* $\mathsf{N}_{\mathbb{B}\mathbb{B}} \geq \mathsf{O}$, *and any* $\lambda \in \mathbb{R}$, *if* $\mathsf{Q} = (\mathsf{I} - \mathbf{e}\mathbf{x}^\top)\mathsf{R}(\mathsf{I} - \mathbf{x}\mathbf{e}^\top) + \mathsf{N} + \lambda\, \mathsf{E}$, *then* $\mathbf{x}$ *is the unique globally optimal solution of* (StQP) *and its DNN relaxation is exact.*

Proof. It follows from the hypothesis and Theorem 4.1(ii) that $\mathbf{x}$ is a globally optimal solution of (StQP) and its DNN relaxation is exact. We next show the uniqueness. For any $\mathbf{y} \in F_n$, where F_n is given by (1), we have

$$\mathbf{y}^\top \mathsf{Q}\mathbf{y} = (\mathbf{y} - \mathbf{x})^\top \mathsf{R}(\mathbf{y} - \mathbf{x}) + \mathbf{y}^\top \mathsf{N}\mathbf{y} + \lambda \geq \lambda = \mathbf{x}^\top \mathsf{Q}\mathbf{x},$$

where we used $\mathbf{x} \in F_n$, $\mathbf{y} \in F_n$, $\mathsf{R} \succ \mathsf{O}$, $\mathbf{y} \geq \mathbf{o}$, and $\mathsf{N} \geq \mathsf{O}$. Since $\mathsf{R} \succ \mathsf{O}$, the inequality above holds with equality if and only if $\mathbf{y} = \mathbf{x}$. We conclude that $\mathbf{x}$ is the unique globally optimal solution of (StQP). □

In Proposition 4.1, if one chooses $\mathsf{N} = \mathsf{O}$ and $\lambda \geq \mathbf{o}$, we obtain $\mathsf{Q} \succeq \mathsf{O}$, which implies that the resulting instance of (StQP) is a convex optimization problem. On the other hand, by choosing N and λ in such a way that $\mathsf{Q} \not\succeq \mathsf{O}$, Proposition 4.1 allows us to construct a nonconvex instance of StQP that admits an exact DNN relaxation. We will utilize both of these observations in our computational experiments.

Next, we consider constructing an instance of StQP with a unique optimal solution but an inexact DNN relaxation. In contrast with the previous case, such a construction is more involved and requires additional care. Our next result presents such a procedure.

Proposition 4.2. *Let* $\mathbf{x} \in F_n$, *where* F_n *is given by* (1). *Let* $\mathbb{A} = \{j \in \{1, \ldots, n\} : \mathbf{x}_j > 0\}$ *and* $\mathbb{B} = \{j \in \{1, \ldots, n\} : \mathbf{x}_j = 0\}$. *Let* $\mathsf{Q} = (\mathsf{I} - \mathbf{e}\mathbf{x}^\top)\mathsf{R}(\mathsf{I} - \mathbf{x}\mathbf{e}^\top) + \mathsf{N} + \lambda\, \mathsf{E}$, *where* $\lambda \in \mathbb{R}$, $\mathsf{R} \in \mathcal{S}^n$ *is such that* $\mathsf{R}_{\mathbb{A}\mathbb{A}} \succeq \mathsf{O}$, $\mathsf{R}_{\mathbb{B}\mathbb{B}} \in \mathcal{COP}_{|\mathbb{B}|}$, $\mathsf{R}_{\mathbb{A}\mathbb{B}} = \mathsf{O}$, *and* $\mathsf{N} \in \mathcal{S}^n$ *is such that* $\mathsf{N}_{\mathbb{A}\mathbb{A}} = \mathsf{O}$, $\mathsf{N}_{\mathbb{A}\mathbb{B}} \geq \mathsf{O}$, *and* $\mathsf{N}_{\mathbb{B}\mathbb{B}} \geq \mathsf{O}$.

  (i) $\mathbf{x}$ *is a globally optimal solution of* (StQP). *Furthermore, if* $\mathsf{R}_{\mathbb{A}\mathbb{A}} \succ \mathsf{O}$, *then* $\mathbf{x}$ *is the unique globally optimal solution.*

  (ii) *If* $\mathsf{R}_{\mathbb{B}\mathbb{B}} \in \mathcal{COP}_{|\mathbb{B}|} \backslash \mathcal{SPN}_{|\mathbb{B}|}$ *and* $\mathsf{N} = \mathsf{O}$, *there exists an* $\epsilon > 0$ *such that, for all* $\mathsf{R}_{\mathbb{A}\mathbb{A}} \succ \mathsf{O}$ *with* $\|\mathsf{R}_{\mathbb{A}\mathbb{A}}\| < \epsilon$, *where* $\|\cdot\|$ *denotes the operator norm,* $\mathbf{x}$ *is the unique globally optimal solution of* (StQP) *and its doubly nonnegative relaxation is inexact.*

Proof. (i) Let $\mathsf{M} = (\mathsf{I} - \mathbf{e}\mathbf{x}^\top)\mathsf{R}(\mathsf{I} - \mathbf{x}\mathbf{e}^\top) + \mathsf{N}$. By Theorem 4.1(i), it suffices to show that $\mathsf{M} \in \mathbf{bd}\, \mathcal{COP}_n$ and $\mathbf{x} \in \mathbf{V}^\mathsf{M}$, where $\mathbf{V}^\mathsf{M}$ is given by (9). For any $\mathbf{y} \in F_n$,

$$\mathbf{y}^\top \mathsf{M}\mathbf{y} = (\mathbf{y} - \mathbf{x})^\top \mathsf{R}(\mathbf{y} - \mathbf{x}) + \mathbf{y}^\top \mathsf{N}\mathbf{y} = (\mathbf{y}_\mathbb{A} - \mathbf{x}_\mathbb{A})^\top \mathsf{R}_{\mathbb{A}\mathbb{A}}(\mathbf{y}_\mathbb{A} - \mathbf{x}_\mathbb{A}) + \mathbf{y}_\mathbb{B}^\top \mathsf{R}_{\mathbb{B}\mathbb{B}}\mathbf{y}_\mathbb{B} + \mathbf{y}^\top \mathsf{N}\mathbf{y} \geq 0 = \mathbf{x}^\top \mathsf{M}\mathbf{x}, \quad (41)$$

where we used $\mathbf{y} \in F_n$ in the first equality, $\mathsf{R}_{\mathbb{A}\mathbb{B}} = \mathsf{O}$ and $\mathbf{x}_\mathbb{B} = \mathbf{o}$ in the second equality, $\mathsf{R}_{\mathbb{A}\mathbb{A}} \succeq \mathsf{O}$, $\mathsf{R}_{\mathbb{B}\mathbb{B}} \in \mathcal{COP}_{|\mathbb{B}|}$, $\mathsf{N} \geq \mathsf{O}$, and $\mathbf{y} \geq \mathbf{o}$ to derive the inequality, and $\mathbf{x} \in F_n$ in the last equality. We conclude

17

that $M \in \mathbf{bd}\,\mathcal{COP}_n$ and $\mathbf{x} \in \mathbf{V}^M$. By Theorem 4.1(i), $\mathbf{x}$ is a globally optimal solution of (StQP). If, in addition, $R_{\mathbb{AA}} \succ O$, then it follows from (41) that $\mathbf{y}^\top M \mathbf{y} = 0$ if and only if $(\mathbf{y}_{\mathbb{A}} - \mathbf{x}_{\mathbb{A}})^\top R_{\mathbb{AA}} (\mathbf{y}_{\mathbb{A}} - \mathbf{x}_{\mathbb{A}}) = \mathbf{y}_{\mathbb{B}}^\top R_{\mathbb{BB}} \mathbf{y}_{\mathbb{B}} = \mathbf{y}^\top N \mathbf{y} = 0$. Since $R_{\mathbb{AA}} \succ O$, we conclude that $\mathbf{y}_{\mathbb{A}} = \mathbf{x}_{\mathbb{A}}$. Since $\mathbf{x} \in F_n$, $\mathbf{x}_{\mathbb{B}} = \mathbf{o}$, and $\mathbf{y}_{\mathbb{A}} = \mathbf{x}_{\mathbb{A}}$, we obtain $1 = \mathbf{e}^\top \mathbf{x} = \mathbf{e}_{\mathbb{A}}^\top \mathbf{x}_{\mathbb{A}} + \mathbf{e}_{\mathbb{B}}^\top \mathbf{x}_{\mathbb{B}} = \mathbf{e}_{\mathbb{A}}^\top \mathbf{x}_{\mathbb{A}} = \mathbf{e}_{\mathbb{A}}^\top \mathbf{y}_{\mathbb{A}}$, which implies that $\mathbf{y}_{\mathbb{B}} = \mathbf{o}$ since $\mathbf{y} \in F_n$. It follows that $\mathbf{y} = \mathbf{x}$. Therefore, $\mathbf{V}^M = \{\mathbf{x}\}$. By Theorem 4.1(i), $\mathbf{x}$ is the unique globally optimal solution of (StQP).

(ii) Let $M = (I - \mathbf{e}\mathbf{x}^\top) R (I - \mathbf{x}\mathbf{e}^\top)$. By Theorem 4.1(ii), it suffices to show that $M \in \mathbf{bd}\,\mathcal{COP}_n \backslash \mathcal{SPN}_n$ and $\mathbf{x} \in \mathbf{V}^M$. By (i), $M \in \mathbf{bd}\,\mathcal{COP}_n$ and $\mathbf{x} \in \mathbf{V}^M$. Since $R_{\mathbb{BB}} \in \mathcal{COP}_{|\mathbb{B}|} \backslash \mathcal{SPN}_{|\mathbb{B}|}$, $\mathcal{CP}_{|\mathbb{B}|}$ and $\mathcal{D}_{|\mathbb{B}|}$ are the dual cones of $\mathcal{COP}_{|\mathbb{B}|}$ and $\mathcal{SPN}_{|\mathbb{B}|}$, respectively, it follows from (7) that there exists $T \in \mathcal{D}_{|\mathbb{B}|} \backslash \mathcal{CP}_{|\mathbb{B}|}$ such that $\langle R_{\mathbb{BB}}, T \rangle = -\delta < 0$. Let $U \in \mathcal{S}^n$ be such that $U_{\mathbb{AA}} = O$, $U_{\mathbb{AB}} = O$, and $U_{\mathbb{BB}} = T$. Clearly, $U \in \mathcal{D}_n$. Furthermore,

$$
\begin{aligned}
\langle M, U \rangle &= \langle M_{\mathbb{BB}}, T \rangle \\
&= \left\langle \left(\mathbf{x}_{\mathbb{A}}^\top R_{\mathbb{AA}} \mathbf{x}_{\mathbb{A}} \right) \mathbf{e}_{\mathbb{B}} \mathbf{e}_{\mathbb{B}}^\top + R_{\mathbb{BB}}, T \right\rangle \\
&= \left(\mathbf{x}_{\mathbb{A}}^\top R_{\mathbb{AA}} \mathbf{x}_{\mathbb{A}} \right) \left(\mathbf{e}_{\mathbb{B}}^\top T \mathbf{e}_{\mathbb{B}} \right) - \delta \\
&\leq \|R_{\mathbb{AA}}\| \mu - \delta,
\end{aligned}
$$

where $\mu = \mathbf{e}_{\mathbb{B}}^\top T \mathbf{e}_{\mathbb{B}} > 0$ and we used $\|\mathbf{x}_{\mathbb{A}}\| \leq 1$ to derive the last inequality. We conclude that $\langle M, U \rangle < 0$ provided that $\|R_{\mathbb{AA}}\| < \epsilon$, where $\epsilon = \delta/\mu > 0$. Therefore, for any $R_{\mathbb{AA}} \succeq O$ such that $\|R_{\mathbb{AA}}\| < \epsilon$, we obtain $M \in \mathbf{bd}\,\mathcal{COP}_n \backslash \mathcal{SPN}_n$ since $U \in \mathcal{D}_n$. By Theorem 4.1(ii), it follows that the DNN relaxation of (StQP) is inexact. If, in addition, $R_{\mathbb{AA}} \succ O$, then $\mathbf{x}$ is the unique globally optimal solution of (StQP) by part (i). This concludes the proof. $\qquad\square$

4.2 Two algorithms for generating nontrivial sparse StQP instances

In this section, we present two algorithms for generating nontrivial instances of (StQP(ρ)). Given $\mathbf{x} \in F_n$ with $\|\mathbf{x}\|_0 \geq 2$, both algorithms initially generate an instance of (StQP) such that $\mathbf{x}$ is the unique optimal solution. Then, a nontrivial instance of (StQP(ρ)) is constructed by choosing $\rho \in \{1, \ldots, \|\mathbf{x}\|_0 - 1\}$.

The details of our algorithms are presented in Algorithm 1 and Algorithm 2. The two algorithms differ only in terms of whether the StQP instance generated in the first stage admits an exact or inexact DNN relaxation.

Algorithm 1 Nontrivial instance of (StQP(ρ)) with (StQP) admitting an exact DNN relaxation
---
Require: $n \geq 2$, $\mathbf{x} \in F_n$ such that $\|\mathbf{x}\|_0 \geq 2$, $\lambda \in \mathbb{R}$

Ensure: A nontrivial instance of (StQP(ρ))

1: $\mathbb{A} \leftarrow \{j \in \{1, \ldots, n\} : \mathbf{x}_j > 0\}$, $\mathbb{B} \leftarrow \{j \in \{1, \ldots, n\} : \mathbf{x}_j = 0\}$

2: Choose an arbitrary $R \in \mathcal{S}^n$ such that $R \succ O$.

3: Choose an arbitrary $N \in \mathcal{S}^n$ such that $N_{\mathbb{AA}} = O$, $N_{\mathbb{AB}} \geq O$, and $N_{\mathbb{BB}} \geq O$.

4: $Q \leftarrow (I - \mathbf{e}\mathbf{x}^\top) R (I - \mathbf{x}\mathbf{e}^\top) + N + \lambda E$

5: Choose an arbitrary $\rho \in \{1, \ldots, \|\mathbf{x}\|_0 - 1\}$.

Our first algorithm, given by Algorithm 1, requires as input $n \geq 2$, $\mathbf{x} \in F_n$ such that $\|\mathbf{x}\|_0 \geq 2$, and $\lambda \in \mathbb{R}$, and generates an instance of (StQP(ρ)). By Proposition 4.1, the instance generated by Algorithm 1 is a

nontrivial instance of (StQP(ρ)) such that $\mathbf{x}$ is the unique optimal solution of (StQP), which admits an exact DNN relaxation. We remark that each step of Algorithm 1 can be implemented in a fairly straightforward way.

Algorithm 2 Nontrivial instance of (StQP(ρ)) with (StQP) admitting an inexact DNN relaxation

---

Require: $n \geq 7$, $\mathbf{x} \in F_n$ such that $2 \leq \|\mathbf{x}\|_0 \leq n - 5$, $\lambda \in \mathbb{R}$

Ensure: A nontrivial instance of (StQP(ρ))

1: $\mathbb{A} \leftarrow \{j \in \{1, \ldots, n\} : \mathbf{x}_j > 0\}$, $\mathbb{B} \leftarrow \{j \in \{1, \ldots, n\} : \mathbf{x}_j = 0\}$

2: Construct $\mathsf{R} \in \mathcal{S}^n$ using the following steps.

3: $\mathsf{R}_{\mathbb{A}\mathbb{B}} \leftarrow \mathsf{O}$

4: Choose an arbitrary $\mathsf{R}_{\mathbb{B}\mathbb{B}} \in \mathcal{COP}_{|\mathbb{B}|} \backslash \mathcal{SPN}_{|\mathbb{B}|}$.

5: Choose an arbitrary $\mathsf{T} \in \mathcal{D}_{|\mathbb{B}|} \backslash \mathcal{CP}_{|\mathbb{B}|}$ such that $\langle \mathsf{R}_{\mathbb{B}\mathbb{B}}, \mathsf{T} \rangle < 0$.

6: $\delta \leftarrow -\langle \mathsf{R}_{\mathbb{B}\mathbb{B}}, \mathsf{T} \rangle$, $\mu \leftarrow \mathbf{e}_{\mathbb{B}}^\top \mathsf{T} \mathbf{e}_{\mathbb{B}}$, $\epsilon \leftarrow \frac{\delta}{\mu}$

7: Choose any $\mathsf{R}_{\mathbb{A}\mathbb{A}} \succ \mathsf{O}$ such that $\|\mathsf{R}_{\mathbb{A}\mathbb{A}}\| < \epsilon$.

8: $\mathsf{Q} \leftarrow (\mathsf{I} - \mathbf{e}\mathbf{x}^\top) \mathsf{R} (\mathsf{I} - \mathbf{x}\mathbf{e}^\top) + \lambda \mathsf{E}$

9: Choose an arbitrary $\rho \in \{1, \ldots, \|\mathbf{x}\|_0 - 1\}$.

Our second algorithm, given by Algorithm 2, takes as input $n \geq 7$, $\mathbf{x} \in F_n$ such that $2 \leq \|\mathbf{x}\|_0 \leq n - 5$, and $\lambda \in \mathbb{R}$, and outputs an instance of (StQP(ρ)). Similarly, it follows from Proposition 4.2 that the instance generated by Algorithm 2 is a nontrivial instance of (StQP(ρ)) such that $\mathbf{x}$ is the unique optimal solution of (StQP), which admits an inexact DNN relaxation.

In contrast with Algorithm 1, the practical implementation of Algorithm 2 merits further discussion. Step 4 of Algorithm 2 requires $\mathsf{R}_{\mathbb{B}\mathbb{B}} \in \mathcal{COP}_{|\mathbb{B}|} \backslash \mathcal{SPN}_{|\mathbb{B}|}$. Such matrices are called *exceptional*. Recall that $\mathcal{SPN}_n \subseteq \mathcal{COP}_n$, and $\mathcal{SPN}_n = \mathcal{COP}_n$ if and only if $n \leq 4$ (see [21]). Therefore, such an exceptional matrix exists if and only if $|\mathbb{B}| \geq 5$, which is ensured by the choices of the input parameters of Algorithm 2. A well-known exceptional 5×5 matrix is the Horn matrix given by

$$
\mathsf{H} = \begin{pmatrix} 1 & -1 & 1 & 1 & -1 \\ -1 & 1 & -1 & 1 & 1 \\ 1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & 1 \end{pmatrix} \in \mathcal{COP}_5 \backslash \mathcal{SPN}_5. \tag{42}
$$

Furthermore, H is an extreme ray of $\mathcal{COP}_5$ [26].

We next present a practical procedure for constructing an exceptional matrix in higher dimensions. Our procedure is inspired by a similar procedure outlined in [24, Section 6.2].

Let $n \geq 7$ and let $\mathbf{x} \in F_n$ satisfy $\|\mathbf{x}\|_0 \leq n - 5$, which ensures that $|\mathbb{B}| \geq 5$. We now construct $\mathsf{R}_{\mathbb{B}\mathbb{B}} \in \mathcal{COP}_{|\mathbb{B}|} \backslash \mathcal{SPN}_{|\mathbb{B}|}$ as follows. Choose an arbitrary $\mathsf{B} \in \mathcal{COP}_{|\mathbb{B}|-5}$, which can, for instance, be ensured by choosing $\mathsf{B} \in \mathcal{SPN}_{|\mathbb{B}|-5}$ by (7) (or even $\mathsf{B} \geq \mathsf{O}$ or $\mathsf{B} \succeq \mathsf{O}$), and an arbitrary $\mathsf{C} \in \mathbb{R}^{(|\mathbb{B}|-5) \times |\mathbb{B}|}$ such that $\mathsf{C} \geq \mathsf{O}$. Let us define

$$
\mathsf{R}_{\mathbb{B}\mathbb{B}} = \begin{bmatrix} \mathsf{B} & \mathsf{C} \\ \mathsf{C}^\top & \mathsf{H} \end{bmatrix} \in \mathcal{S}^{|\mathbb{B}|}. \tag{43}
$$

It is easy to see that $\mathsf{R}_{\mathbb{B}\mathbb{B}} \in \mathcal{COP}_{|\mathbb{B}|}$ (see also [39, Lemma 3.4 (a)]). Since $\mathsf{H} \in \mathcal{COP}_5 \backslash \mathcal{SPN}_5$, there exists

$\mathsf{F} \in \mathcal{D}_5 \backslash \mathcal{CP}_5$ such that $\langle \mathsf{H}, \mathsf{F} \rangle = -\delta < 0$. Let

$$\mathsf{T} = \begin{pmatrix} \mathsf{O} & \mathsf{O} \\ \mathsf{O} & \mathsf{F} \end{pmatrix}. \tag{44}$$

Note that $\mathsf{T} \in \mathcal{D}_{|\mathbb{B}|}$ and $\langle \mathsf{R}_{\mathbb{BB}}, \mathsf{T} \rangle = \langle \mathsf{H}, \mathsf{F} \rangle = -\delta < 0$, which implies that $\mathsf{R}_{\mathbb{BB}} \notin \mathcal{SPN}_{|\mathbb{B}|}$. We conclude that $\mathsf{R}_{\mathbb{BB}} \in \mathcal{COP}_{|\mathbb{B}|} \backslash \mathcal{SPN}_{|\mathbb{B}|}$. We then define $\mu = \mathbf{e}_{\mathbb{B}}^\top \mathsf{T} \mathbf{e}_{\mathbb{B}} > 0$ and $\epsilon = \delta/\mu$.

While any choice of $\mathsf{F} \in \mathcal{D}_5 \backslash \mathcal{CP}_5$ such that $\langle \mathsf{H}, \mathsf{F} \rangle = -\delta < 0$ works in the above construction, it may be preferable in practice to choose F such that the corresponding value of ϵ is as large as possible. This can be easily achieved in practice by solving the following small semidefinite programming (SDP) problem:

$$\epsilon := \max_{\mathsf{F} \in \mathcal{D}_5 \backslash \{\mathsf{O}\}} \frac{-\langle \mathsf{H}, \mathsf{F} \rangle}{\mathbf{e}_{\mathbb{B}}^\top \mathsf{F} \mathbf{e}_{\mathbb{B}}} = \max \{ -\langle \mathsf{H}, \mathsf{F} \rangle : \mathbf{e}_{\mathbb{B}}^\top \mathsf{F} \mathbf{e}_{\mathbb{B}} = 1, \, \mathsf{F} \in \mathcal{D}_5 \}, \tag{45}$$

where the second equality follows from the homogeneity of the objective function. An approximate optimal solution of (45) is given by

$$\mathsf{F} = \frac{1}{78.2} \begin{pmatrix} 7 & 4.32 & 0 & 0 & 4.32 \\ 4.32 & 7 & 4.32 & 0 & 0 \\ 0 & 4.32 & 7 & 4.32 & 0 \\ 0 & 0 & 4.32 & 7 & 4.32 \\ 4.32 & 0 & 0 & 4.32 & 7 \end{pmatrix}, \tag{46}$$

which yields $\epsilon \approx 0.1049$.

Remark 1. *The above construction works for any exceptional $\mathsf{R}_{\mathbb{BB}}$ once we can control T, the normal separating $\mathsf{R}_{\mathbb{BB}}$ (corresponding to H above) from $\mathcal{SPN}_{|\mathbb{B}|}$, in the sense to solve the (effectively small) SDP problem*

$$\epsilon := \max_{\mathsf{T} \in \mathcal{D}_{|\mathbb{B}|} \backslash \{\mathsf{O}\}} \frac{-\langle \mathsf{R}_{\mathbb{BB}}, \mathsf{T} \rangle}{\mathbf{e}_{\mathbb{B}}^\top \mathsf{T} \mathbf{e}_{\mathbb{B}}} = \max \{ -\langle \mathsf{R}_{\mathbb{BB}}, \mathsf{T} \rangle : \mathbf{e}_{\mathbb{B}}^\top \mathsf{T} \mathbf{e}_{\mathbb{B}} = 1, \, \mathsf{T} \in \mathcal{D}_{|\mathbb{B}|} \}. \tag{47}$$

Observe that any extremal ray of $\mathcal{COP}_{|\mathbb{B}|}$ different from a rank-one matrix or a symmetric componentwise non-negative matrix with exactly two positive entries – which constitute the extremal rays of the cone of PSD matrices and the cone of componentwise nonnegative matrices – is necessarily exceptional, by extremality. While for larger $|\mathbb{B}|$, these extremal rays are (yet) unknown, we can imitate the above extension strategy and build upon the known ones for $|\mathbb{B}| \in \{5, 6\}$, e.g., using Hildebrand matrices [1, 29]. All these used in $\mathsf{R}_{\mathbb{BB}}$ would render (47) feasible (recommended only for extremal rays of moderate order to keep this effort low), and proceeding as above with H and F, this approach provides more nontrivial examples. Finally, note that extremality is used in this context only to ensure exceptionality; any non-extremal but exceptional matrix would work as well.

4.3 Set of instances

To assess the impact of the choice of the instance set on the solution time of each exact model and each convex relaxation as well as on the quality of the lower bound, we conduct extensive experiments on a variety of carefully constructed instances of $(\text{StQP}(\rho))$. Let us now describe the set of instances in detail.

Using Algorithm 1 and Algorithm 2, we generated three sets of nontrivial instances of $(\text{StQP}(\rho))$ with different characteristics.

(i) *PSD Instances:* This set of instances of $(\text{StQP}(\rho))$ is constructed using Algorithm 1 in such a way that $\mathsf{Q} \succeq \mathsf{O}$, i.e., the corresponding instance of (StQP) is a convex optimization problem. For instance, such

instances may arise in portfolio optimization problems. We remark that the resulting instances of $(\mathrm{StQP}(\rho))$ are, in general, still NP-hard. In Step 2 of Algorithm 1, we construct $\mathsf{R} \succ \mathsf{O}$ so that all of its eigenvalues are uniformly distributed in $(0, 3)$. By choosing $\mathsf{N} = \mathsf{O}$ and $\lambda = 0$, we ensure that $\mathsf{Q} = (\mathsf{I} - \mathbf{e}\mathbf{x}^\top)\mathsf{R}(\mathsf{I} - \mathbf{x}\mathbf{e}^\top) \succeq \mathsf{O}$. Since $\mathbf{x}^\top \mathsf{Q} \mathbf{x} = 0$, Q is PSD but not positive definite. By Proposition 4.1, each PSD instance satisfies

$$\ell_\rho(\mathsf{Q}) > \ell_n(\mathsf{Q}) = \mu(\mathsf{Q}) = 0, \tag{48}$$

where $\mu(\mathsf{Q})$ is defined as in (40).

(ii) *SPN Instances:* This set of instances of $(\mathrm{StQP}(\rho))$ is generated using Algorithm 1 in a similar manner to PSD instances, i.e., we set $\lambda = 0$ and use the same procedure to construct R. However, instead of choosing $\mathsf{N} = \mathsf{O}$ as in PSD instances, each entry of $\mathsf{N}_{\mathbb{A}\mathbb{B}} \in \mathbb{R}^{|\mathbb{A}| \times |\mathbb{B}|}$ and $\mathsf{N}_{\mathbb{B}\mathbb{B}} \in \mathcal{S}^{|\mathbb{B}|}$ is uniformly generated in $(0, 3)$. In contrast with PSD instances, $\mathsf{Q} = (\mathsf{I} - \mathbf{e}\mathbf{x}^\top)\mathsf{R}(\mathsf{I} - \mathbf{x}\mathbf{e}^\top) + \mathsf{N}$ is, in general, not PSD. In fact, we numerically verified that Q was an indefinite matrix for each SPN instance. By Proposition 4.1, each SPN instance satisfies

$$\ell_\rho(\mathsf{Q}) > \ell_n(\mathsf{Q}) = \mu(\mathsf{Q}) = 0. \tag{49}$$

(iii) *COP Instances:* This set of instances of $(\mathrm{StQP}(\rho))$ is generated by Algorithm 2. In Step 2, we construct $\mathsf{R} \in \mathcal{S}^n$ using the procedure outlined in Section 4.2. In particular, $\mathsf{R}_{\mathbb{B}\mathbb{B}}$ is constructed as in (43), where $\mathsf{B} \succ \mathsf{O}$ is generated with eigenvalues uniformly distributed in $(0, 3)$, and each entry of $\mathsf{C} \geq \mathsf{O}$ is uniformly distributed in $(0, 1)$. Our choices are, in part, motivated by the observation that the eigenvalues of the Horn matrix H given by (42) lie in $[-1.236, 3.236]$ so that the eigenvalues of $\mathsf{B} \succ \mathsf{O}$ have similar magnitude. The separating matrix T is constructed as in (44), where F is given by (46), which yields $\epsilon \approx 0.1049$. Finally, $\mathsf{R}_{\mathbb{A}\mathbb{A}} \succ \mathsf{O}$ is constructed with eigenvalues uniformly distributed in $(0, 0.99\epsilon)$, ensuring that $\|\mathsf{R}_{\mathbb{A}\mathbb{A}}\| < \epsilon$. We then choose $\lambda = 0$ and set $\mathsf{Q} = (\mathsf{I} - \mathbf{e}\mathbf{x}^\top)\mathsf{R}(\mathsf{I} - \mathbf{x}\mathbf{e}^\top)$. By Proposition 4.2, this procedure ensures that $\mathsf{Q} \not\succeq \mathsf{O}$ and each COP instance satisfies

$$\ell_\rho(\mathsf{Q}) > \ell_n(\mathsf{Q}) = 0 > \mu(\mathsf{Q}). \tag{50}$$

In summary, each of the three sets of instances of $(\mathrm{StQP}(\rho))$ is comprised of nontrivial instances that exhibit different characteristics. While each PSD instance and each SPN instance has a corresponding StQP instance that admits an exact doubly nonnegative (DNN) relaxation, they differ in terms of the convexity of the objective function. On the other hand, each COP instance has a nonconvex objective function with the additional property that the DNN relaxation of the corresponding StQP instance is inexact.

5   Computational results

In this section, we report and discuss the results of our computational experiments. After describing the experimental setup in Section 5.1, we report our results in detail in Sections 5.2 and 5.3.

5.1   Experimental setup

The set of instances was already described in Section 4.3. Here, we outline the details of our experimental setup.

In our experiments, we chose $n \in \{25, 50\}$. For each choice of n, we identified three sparsity levels for the designated optimal solution $\mathbf{x} \in F_n$ of (StQP), denoted by $\rho_0 = \|\mathbf{x}\|_0$. In particular, we considered $\rho_0 \in$

$\{\lfloor 0.25n \rceil, \lfloor 0.5n \rceil, \lfloor 0.75n \rceil\}$, where $\lfloor \cdot \rceil$ denotes the nearest integer function. Finally, for each choice of ρ_0, we chose $\rho \in \{\lfloor 0.25\rho_0 \rceil, \lfloor 0.5\rho_0 \rceil, \lfloor 0.75\rho_0 \rceil\}$. The parameters are summarized in Table 1.

n	ρ_0	ρ
	6	{2,3,4}
25	12	{3,6,9}
	19	{5,10,14}
	12	{3,6,9}
50	25	{6,12,19}
	38	{10,19,28}

Table 1: Choices of the parameters n, ρ_0, and ρ

For each choice of n, ρ_0, and ρ presented in Table 1, we generated 25 instances from each of the three sets of instances described in Section 4.3, which gave rise to a total of 1,350 nontrivial instances of $(\text{StQP}(\rho))$ comprising of 450 PSD instances, 450 SPN instances, and 450 COP instances. For each instance of $(\text{StQP}(\rho))$, we solved $(\text{P1}(\rho))$, $(\text{P2}(\rho))$, $(\text{D1A}(\rho))$, $(\text{D1B}(\rho))$, $(\text{D2A}(\rho))$, and $(\text{D2B}(\rho))$.

Algorithm 1 and Algorithm 2 were implemented in Julia version 1.8.5 [4]. Each instance of $(\text{P1}(\rho))$ and $(\text{P2}(\rho))$ was solved by `Gurobi` version 11.0.0 [25] via the modeling language `JuMP` v1.19.0 [30]. We solved $(\text{D1A}(\rho))$, $(\text{D1B}(\rho))$, $(\text{D2A}(\rho))$, and $(\text{D2B}(\rho))$ by `MOSEK` version 10.1.24 [2] via `JuMP`. Our computational experiments were carried out on a 64-bit HP workstation with 24 threads (2 sockets, 6 cores per socket, 2 threads per core) running `Ubuntu Linux` with 96 GB of RAM and Intel Xeon CPU E5-2667 processors with a clock speed of 2.90 GHz. We imposed a time limit of 600 seconds on each solve. The default settings were used for all of the other parameters of `Gurobi` and `MOSEK`.

5.2 Solution time

Now we focus on the solution time of each of the six models $(\text{P1}(\rho))$, $(\text{P2}(\rho))$, $(\text{D1A}(\rho))$, $(\text{D1B}(\rho))$, $(\text{D2A}(\rho))$, and $(\text{D2B}(\rho))$.

Figure 1 depicts the average solution times of each of the six models on each of the three instance sets. The results for PSD instances, SPN instances, and COP instances are presented in the first, second, and third rows of Figure 1, respectively. In each row, the results for $n = 25$ and $n = 50$ are given in the first and second column, respectively. In each graph, the horizontal axis denotes the average solution time (in seconds) presented in logarithmic scale. We used identical axis limits in each graph to facilitate a better comparison across different graphs. The vertical axis is comprised of nine sets of bar charts, each of which represents a particular choice of the tuple (ρ_0, ρ) corresponding to the choice of n as outlined in Table 1. Finally, for each choice of (ρ_0, ρ), the corresponding bar chart represents the average solution times of the six models $(\text{P1}(\rho))$, $(\text{P2}(\rho))$, $(\text{D1A}(\rho))$, $(\text{D1B}(\rho))$, $(\text{D2A}(\rho))$, and $(\text{D2B}(\rho))$ over 25 instances generated using the procedure in Section 4.3. Note that the solution times of instances that were terminated due to the time limit of 600 seconds were also included in the average solution times. Therefore, each bar chart represents the average computational requirement of the corresponding model. We recall that $(\text{P1}(\rho))$ and $(\text{P2}(\rho))$ were solved by `Gurobi` whereas $(\text{D1A}(\rho))$, $(\text{D1B}(\rho))$,
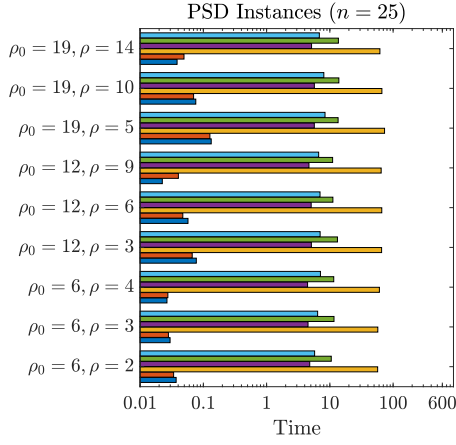
(a) PSD Instances ($n = 25$)
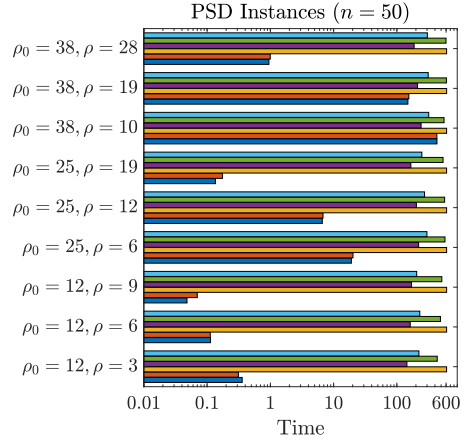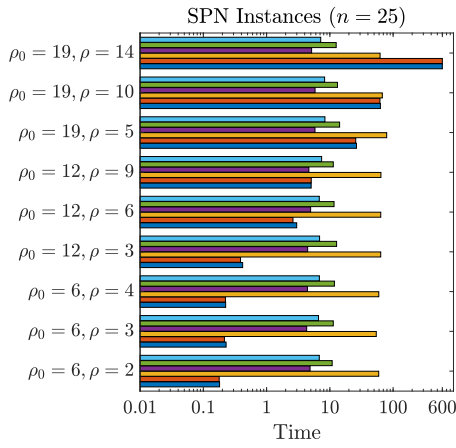
(b) PSD Instances ($n = 50$)
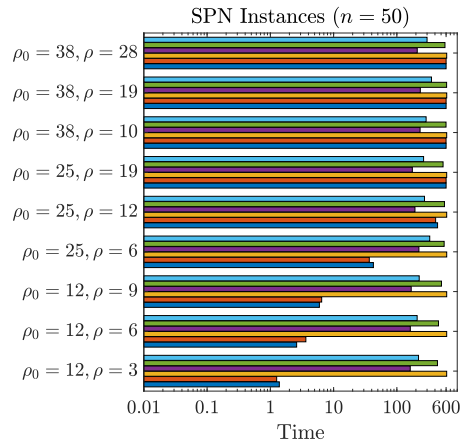
(c) SPN Instances ($n = 25$)
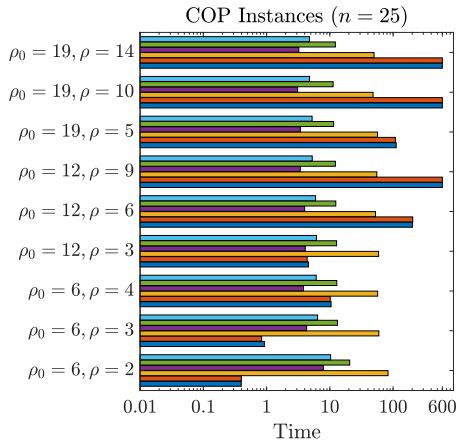
(d) SPN Instances ($n = 50$)
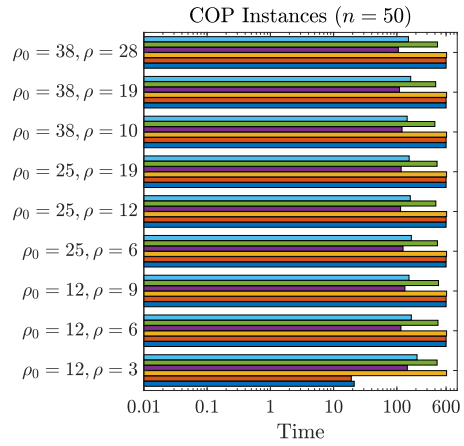
(e) COP Instances ($n = 25$)

(f) COP Instances ($n = 50$)

Figure 1: Average solution times (in seconds) of (P1(ρ)), (P2(ρ)), (D1A(ρ)), (D1B(ρ)), (D2A(ρ)), and (D2B(ρ)) for PSD, SPN, and COP instances

(D2A(ρ)), and (D2B(ρ)) by `MOSEK`.

Figure 1 reveals several interesting relations about average solution times. We first outline our observations

concerning the exact MIQP models $(P1(\rho))$ and $(P2(\rho))$:

(i) Average solution times of the two exact models exhibit very similar behavior for each fixed choice of the instance set and the triple (n, ρ_0, ρ). However, there seems to be a very strong correlation between the average solution time and the choices of (n, ρ_0, ρ) as well as the instance set.

(ii) Considering each row of Figure 1 separately, the average solution time of each of the two exact models across all choices of the tuple (ρ_0, ρ) increases as n increases. However, the rate of increase seems to be highly dependent on the choice of the instance set.

We next present the corresponding observations for the average solution times of the convex relaxations $(D1A(\rho))$, $(D1B(\rho))$, $(D2A(\rho))$, and $(D2B(\rho))$:

(i) In contrast with the exact models, for each choice of n, Figure 1 reveals that the average solution times of the convex relaxations exhibit very similar behavior across all three sets of instances and all choices of (ρ_0, ρ).

(ii) The average solution time of our reduced formulation $(D1B(\rho))$ consistently achieves the lowest average solution times, followed by the reduced formulation $(D2B(\rho))$, which, in turn, is followed by $(D2A(\rho))$ and $(D1A(\rho))$, respectively.

(iii) The average solution time of our reduced formulation $(D1B(\rho))$ exhibits at least an order of magnitude improvement over that of $(D1A(\rho))$. On the other hand, the corresponding improvement of $(D2B(\rho))$ over $(D2A(\rho))$ seems to be less pronounced, especially in smaller dimensions.

(iv) As expected, the average solution time of each relaxation increases with n. While the rate of increase does not seem to be influenced by the specific instance set, we observe that it does depend on the specific relaxation.

Since Figure 1 is based only on average solution times, it cannot capture the variability in the data. In an attempt to shed more light on the distribution of the solution times across all instance sets, we present the empirical cumulative distribution functions of all six models in Figure 2, which is organized similarly to Figure 1. The three rows display the results for PSD, SPN, and COP instances, respectively, whereas the two columns are devoted to the results for $n = 25$ and $n = 50$, respectively. Each of the six graphs presents six plots corresponding to the empirical cumulative distribution functions of solution times of each of the six models on all 225 instances for the corresponding choice of the instance set and n (see Table 1). The horizontal axis denotes the solution time (in seconds) on a logarithmic scale, and the vertical axis represents the fraction of the instances solved. The markers represent the data points for every 25th instance. Once again, we employed identical axis limits in each of the six graphs.

Figure 2 clearly reveals the difference between the distributions of solution times of the exact models $(P1(\rho))$ and $(P2(\rho))$ and those of the convex relaxations $(D1A(\rho))$, $(D1B(\rho))$, $(D2A(\rho))$, and $(D2B(\rho))$. We outline our observations below:

(i) The solution times of the two exact models, in general, exhibit a very similar distribution but a significant variability across different instance sets.

(ii) In contrast, we observe that the solution times of the convex relaxations tend to have a much smaller variance and exhibit very similar distributions across different instance sets.
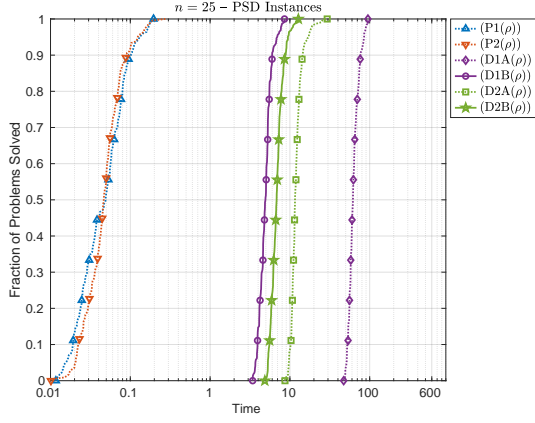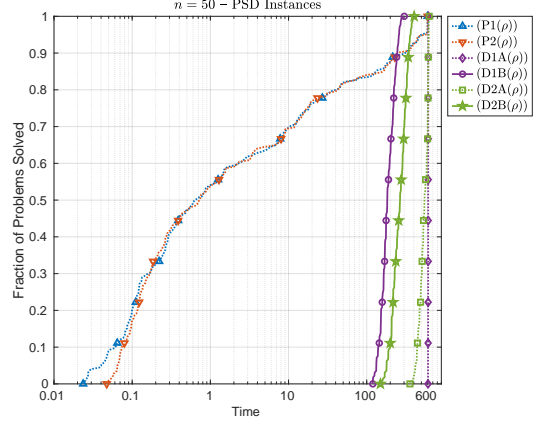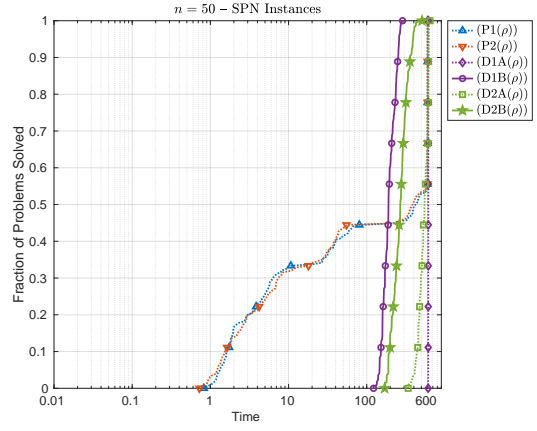
24

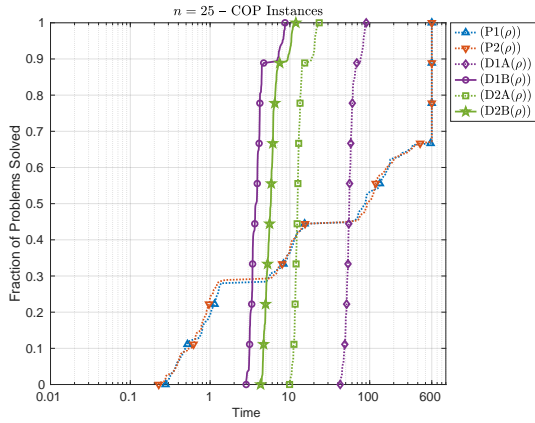(a) PSD Instances ($n = 25$)
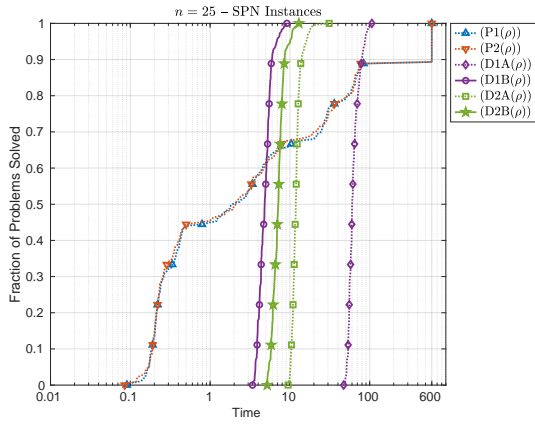
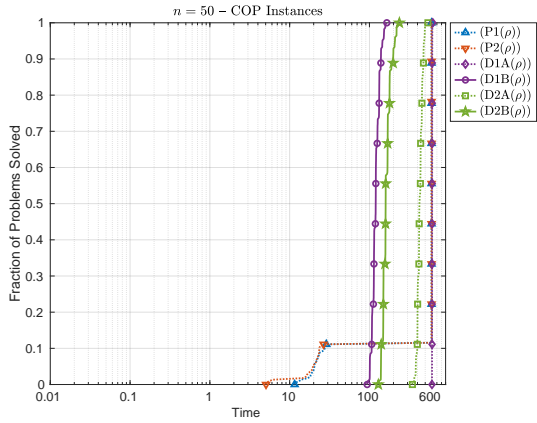(b) PSD Instances ($n = 50$)

(c) SPN Instances ($n = 25$)

(d) SPN Instances ($n = 50$)

(e) COP Instances ($n = 25$)

(f) COP Instances ($n = 50$)

Figure 2: Empirical cumulative distribution functions of solution times of $(\text{P1}(\rho))$, $(\text{P2}(\rho))$, $(\text{D1A}(\rho))$, $(\text{D1B}(\rho))$, $(\text{D2A}(\rho))$, and $(\text{D2B}(\rho))$

Therefore, in contrast with the exact models $(\text{P1}(\rho))$ and $(\text{P2}(\rho))$, we conclude that the solution time of each convex relaxation seems to be very robust with respect to the choice of the instance set and the choice of (ρ_0, ρ)

in our setting.

In the sequel, we present a more detailed discussion of the behavior of the solution times of the exact models and convex relaxations.

5.2.1 Exact MIQP models

As illustrated by Figures 1 and 2, the solution time of the exact models is highly dependent on the choices of the instance set and on the triple (n, ρ_0, ρ). Here, we aim to shed more light on this dependence.

In Figures 3 and 4, we present the empirical cumulative distribution functions of solution times of $(P1(\rho))$ and $(P2(\rho))$ for all instances with $n = 25$ and $n = 50$, respectively. Each graph in each figure consists of six plots, each of which corresponds to the solution times of $(P1(\rho))$ and $(P2(\rho))$ on each of PSD, SPN, and COP instances for a fixed choice of (n, ρ_0, ρ). In each column of each figure, we fix the tuple (n, ρ_0) and present the empirical cumulative distributions of the solution times of 25 instances corresponding to the three different choices of ρ as presented in Table 1. On the other hand, each row corresponds to a different ratio of $\rho/\rho_0 \in \{0.25, 0.5, 0.75\}$. Again, we use a logarithmic scale for the solution time and ensure that all axis limits are identical for a meaningful comparison across different graphs. Finally, we note that the markers represent the data points at every 5th instance.

A close examination of Figures 3 and 4 reveal the following observations about the two exact models $(P1(\rho))$ and $(P2(\rho))$:

(i) The two models, in general, exhibit very similar distributions across all parameter choices, except for some small differences for a few choices of (n, ρ_0, ρ) on PSD instances, where $(P1(\rho))$ seems to be solved slightly faster than $(P2(\rho))$.

(ii) We clearly observe an empirical first-order stochastic dominance among PSD, SPN, and COP instances. For each choice of (n, ρ_0, ρ), PSD instances tend to achieve the lowest solution times, followed by SPN instances, whereas COP instances exhibit the highest solution times.

(iii) For each fixed (n, ρ_0), we observe that the solution time tends to decrease as ρ increases for PSD instances, especially for $n = 50$, whereas it increases for each of SPN and COP instances.

(iv) If we fix n and the ratio $\rho/\rho_0 \in \{0.25, 0.5, 0.75\}$, the solution time increases as ρ_0 increases across all three sets of instances. For fixed n, we therefore conclude that the solution time of the exact models seems to be also highly influenced by the choice of the ratio ρ/ρ_0.

(v) For $n = 25$, both exact models can be solved very quickly for all PSD instances. In contrast, for SPN and COP instances, each of $(P1(\rho))$ and $(P2(\rho))$ is terminated due to the time limit of 600 seconds on certain subsets of the instances. Therefore, we observe that SPN and COP instances can be particularly challenging for $(P1(\rho))$ and $(P2(\rho))$, even for $n = 25$ with particular choices of (ρ_0, ρ). For $n = 50$, we observe that even some PSD instances were terminated due to the time limit. Recall that each instance in this set has a convex objective function. For each choice of (n, ρ_0, ρ), a comparison of PSD, SPN, and COP instances reveals that the number of instances terminated due to the time limit exhibits a monotonically increasing behavior.
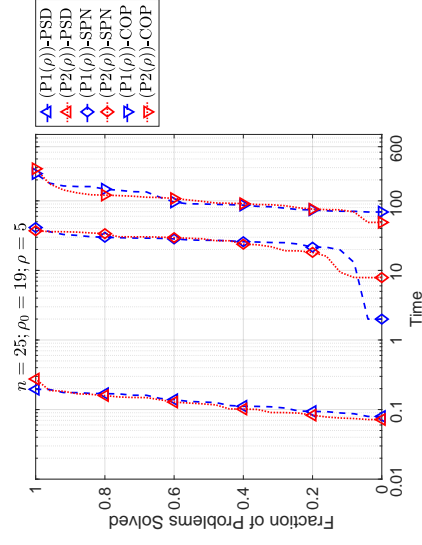
In summary, the solution time of exact models $(P1(\rho))$ and $(P2(\rho))$ seem to be highly dependent on each of the choices of the instance set, (n, ρ_0, ρ), as well as on the ratio ρ/ρ_0. COP instances tend to be the most
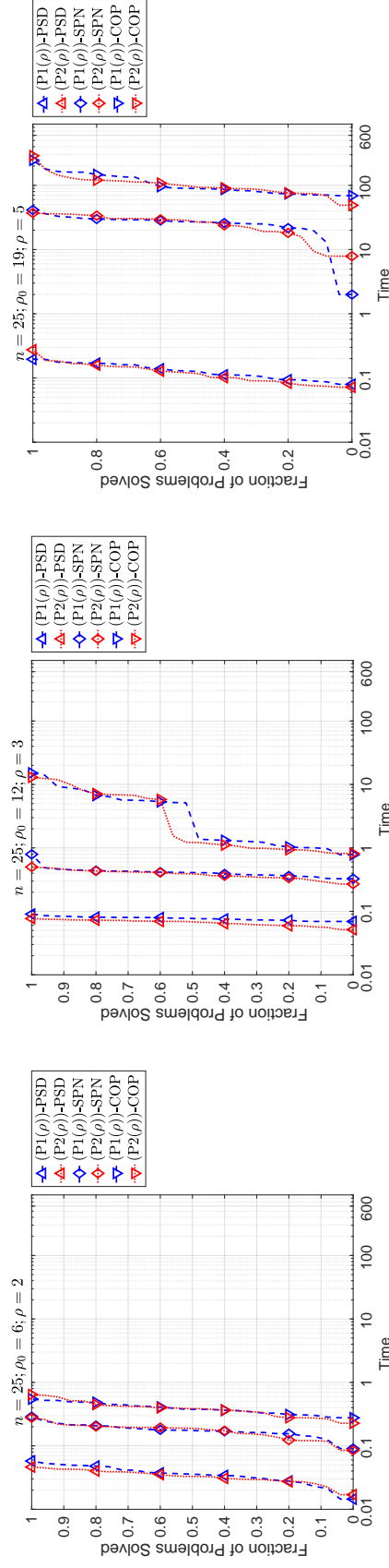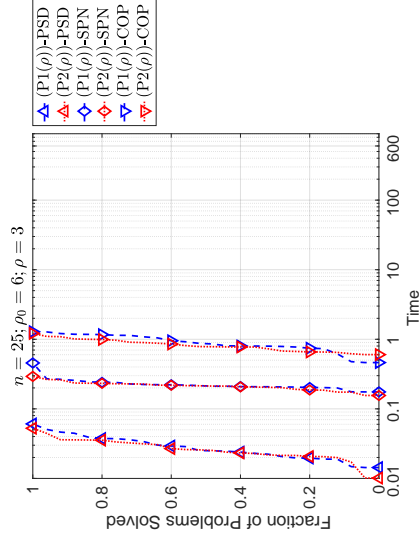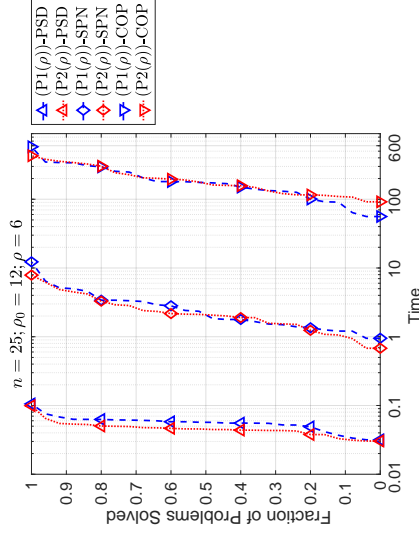
(a) $n = 25, \rho_0 = 6, \rho = 2$
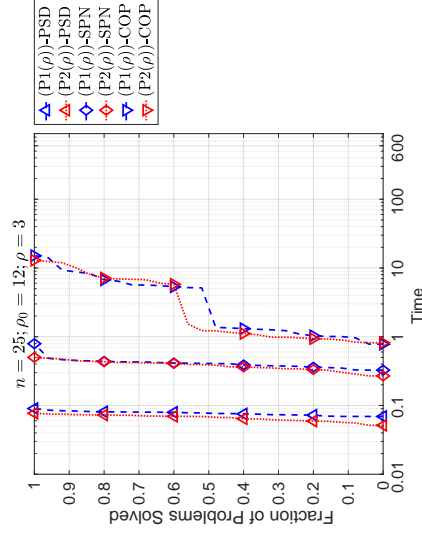
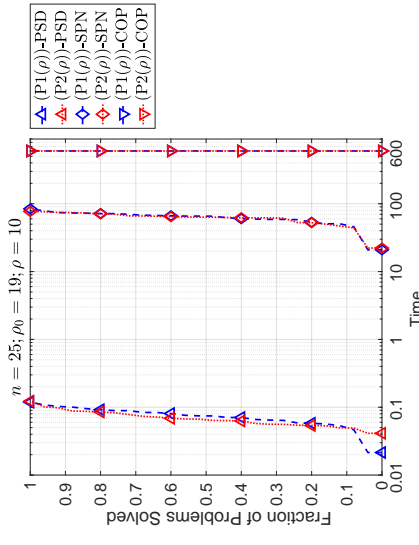(b) $n = 25, \rho_0 = 12, \rho = 3$

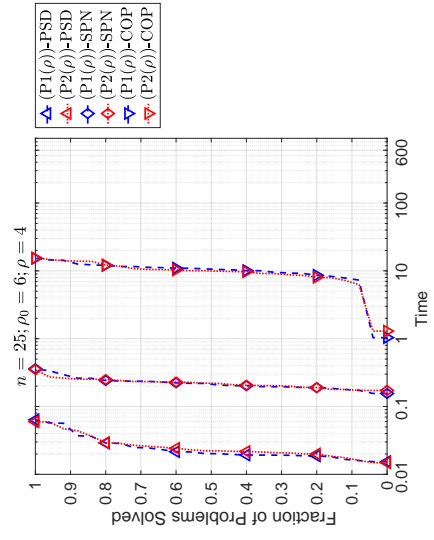(c) $n = 25, \rho_0 = 19, \rho = 5$
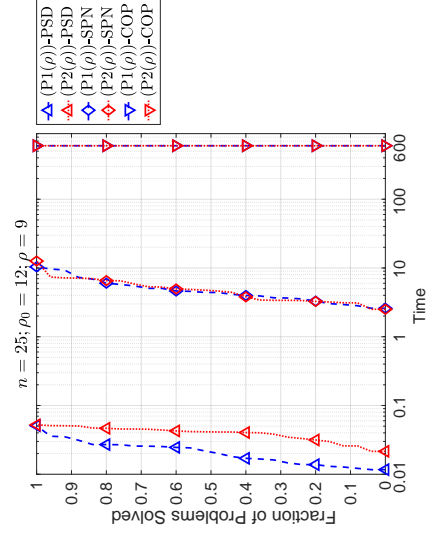
(d) $n = 25, \rho_0 = 6, \rho = 3$
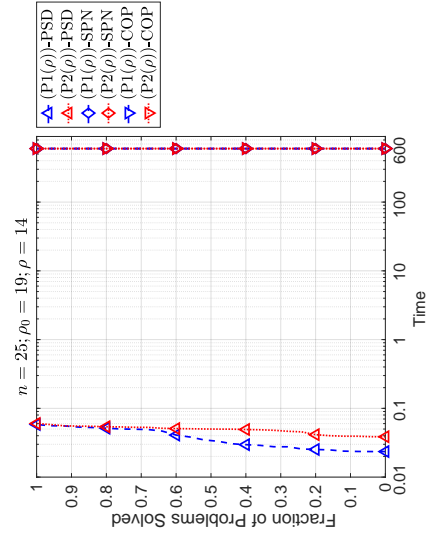
(e) $n = 25, \rho_0 = 12, \rho = 6$

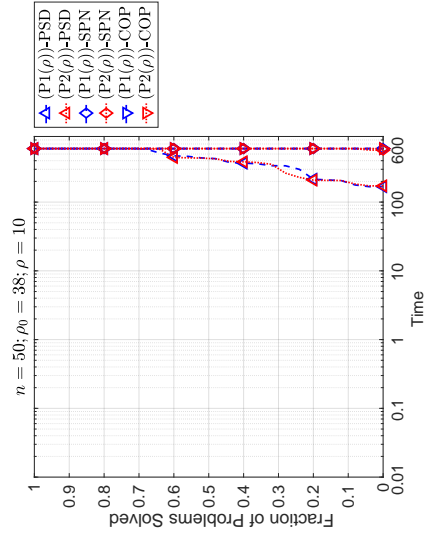(f) $n = 25, \rho_0 = 19, \rho = 10$

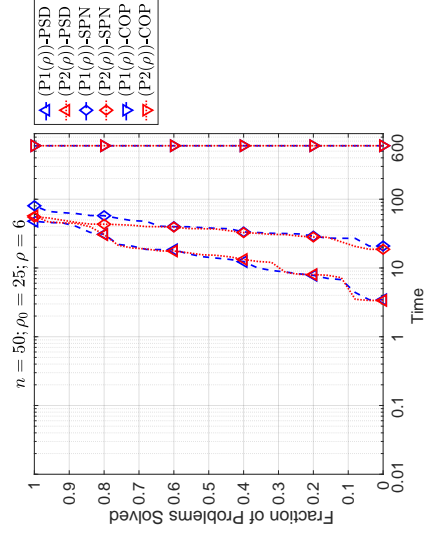(g) $n = 25, \rho_0 = 6, \rho = 4$

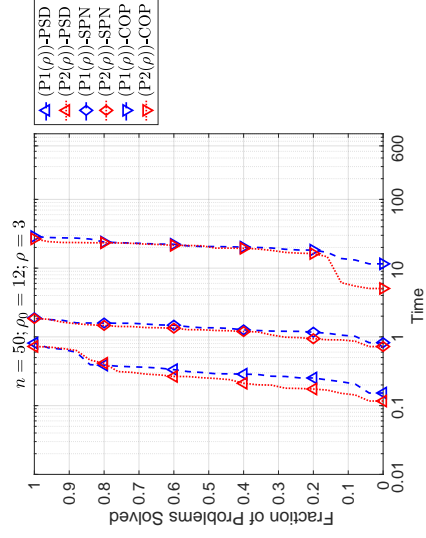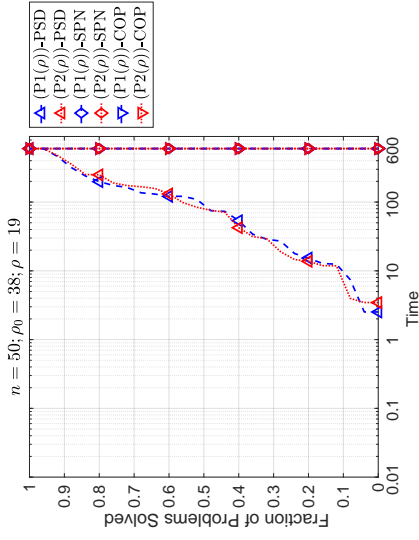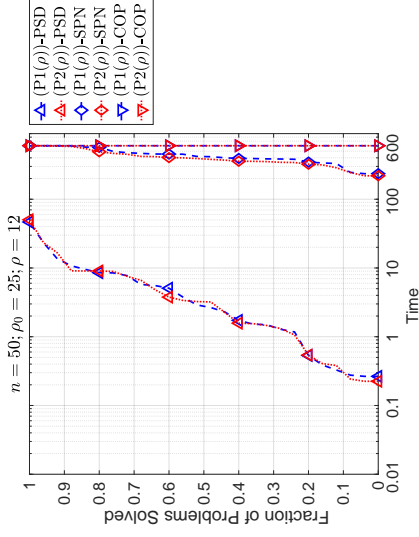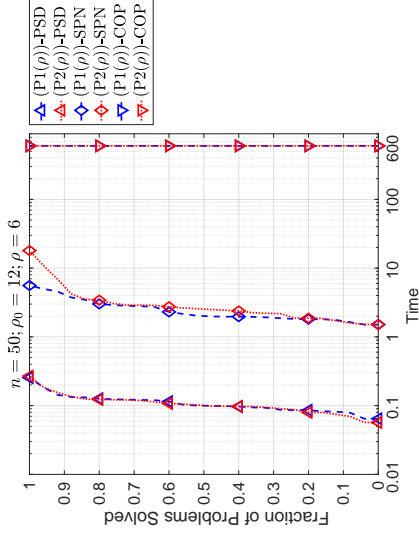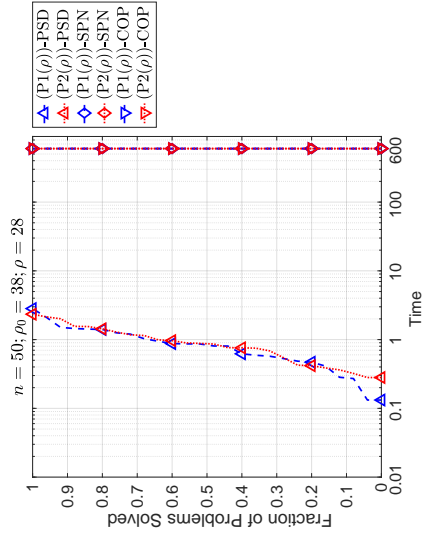(h) $n = 25, \rho_0 = 12, \rho = 9$

(i) $n = 25, \rho_0 = 19, \rho = 14$

Figure 3: Empirical cumulative distribution functions of solution times of (P1(ρ)) and (P2(ρ)) for all instances with $n = 25$

(a) $n = 50, \rho_0 = 12, \rho = 3$

(b) $n = 50, \rho_0 = 25, \rho = 6$

(c) $n = 50, \rho_0 = 38, \rho = 10$

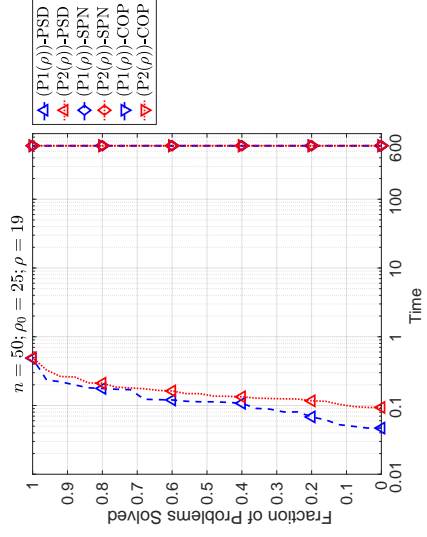(d) $n = 50, \rho_0 = 12, \rho = 6$

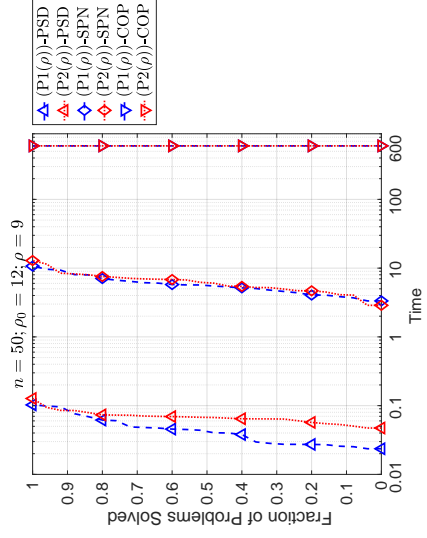(e) $n = 50, \rho_0 = 25, \rho = 12$

(f) $n = 50, \rho_0 = 38, \rho = 19$

(g) $n = 50, \rho_0 = 12, \rho = 9$

(h) $n = 50, \rho_0 = 25, \rho = 19$

(i) $n = 50, \rho_0 = 38, \rho = 28$

Figure 4: Empirical cumulative distribution functions of solution times of (P1(ρ)) and (P2(ρ)) for all instances with $n = 50$

challenging for the exact models, followed by SPN instances, which, in turn, are followed by PSD instances. These results illustrate that Algorithms 1 and 2 are capable of generating instances of $(\text{StQP}(\rho))$ that are particularly challenging for Gurobi, even in relatively small dimensions.

We close this section with a brief discussion of why we choose not to report optimality gaps for $(\text{P1}(\rho))$ and $(\text{P2}(\rho))$ on instances that were terminated by the time limit. Gurobi computes the optimality gap using the formula $|z_P - z_D|/|z_P|$, where z_P and z_D denote the best objective function value and the best lower bound, respectively. In our settings, the optimal value is usually very close to zero, which leads to very large optimality gaps due to the division by a very small number. Therefore, optimality gaps reported by Gurobi tend to be extremely large and do not seem to reflect the quality of the solution accurately. We will, however, discuss the quality of the solutions relying directly on the lower bounds in Section 5.3.

5.2.2 Convex DNN relaxations

Now let us focus on solution times of the four convex relaxations $(\text{D1A}(\rho))$, $(\text{D1B}(\rho))$, $(\text{D2A}(\rho))$, and $(\text{D2B}(\rho))$.

In contrast with the exact models, Figure 2 illustrates that the solution times of convex relaxations do not exhibit a strong correlation with the choices of the instance set and the tuple (ρ_0, ρ). As such, we do not present the counterparts of Figures 3 and 4 for the convex relaxations due to space considerations.

Based on Figure 2, we make the following observations about the convex relaxations $(\text{D1A}(\rho))$, $(\text{D1B}(\rho))$, $(\text{D2A}(\rho))$, and $(\text{D2B}(\rho))$:

(i) The distributions of the solution times clearly illustrate the computational advantages of the reduced formulations $(\text{D2A}(\rho))$ and $(\text{D2B}(\rho))$ over their counterparts.

(ii) We observe an empirical first-order stochastic dominance among the distributions of the solution times of $(\text{D1A}(\rho))$, $(\text{D1B}(\rho))$, $(\text{D2A}(\rho))$, and $(\text{D2B}(\rho))$ on each of the six graphs. In particular, we observe that the reduced formulation $(\text{D1B}(\rho))$ consistently achieves the lowest solution times, followed by the reduced formulation $(\text{D2B}(\rho))$, which, in turn, is followed by $(\text{D2A}(\rho))$ and $(\text{D1A}(\rho))$, respectively. In addition to being the theoretically tightest relaxation, it is worth noticing that $(\text{D1B}(\rho))$ also outperforms $(\text{D2B}(\rho))$ in terms of the solution time.

(iii) For $n = 25$, each of the four relaxations can be solved to optimality within the time limit for each instance set and each choice of ρ_0 and ρ. For $n = 50$, we observe that $(\text{D1A}(\rho))$ was terminated due to the time limit on every instance whereas $(\text{D2A}(\rho))$ was terminated due to the time limit on some subsets of the PSD and SPN instances. In contrast, each of the two reduced formulations $(\text{D1B}(\rho))$ and $(\text{D2B}(\rho))$ was solved to optimality within the time limit on all instances. These observations clearly demonstrate the computational advantages of the reduced formulations $(\text{D2A}(\rho))$ and $(\text{D2B}(\rho))$ over their counterparts, especially in higher dimensions.

(iv) For PSD instances with $n = 25$, the distributions of the solution times of each of the two exact models exhibit a first-order stochastic dominance on the corresponding distribution of each of the four convex relaxations. In contrast, we observe that the solution times of the convex relaxations achieve better performance on a larger subset of instances for each instance set with a larger n. For fixed n, a comparison of PSD, SPN, and COP instances reveals an increasingly better performance of the convex relaxations in comparison with the exact models.

29

Table 2 reports the number of instances on which (D2A(ρ)) is terminated due to the time limit. Note that we omit (D1A(ρ)) since all instances with $n = 50$ were terminated due to the time limit. It is worth noticing that (D2A(ρ)) was solved to optimality within the time limit on all COP instances with $n = 50$. In contrast with $n = 25$, we therefore conclude that the solution time of (D2A(ρ)) for $n = 50$ seems to be somewhat sensitive to the set of instances and the choices of the parameters (ρ_0, ρ).

Instance Set	n	ρ_0	ρ	(D2A(ρ))
PSD	50	12	9	1
		25	6	14
		25	12	10
		25	19	3
		38	10	8
		38	19	22
		38	28	18
SPN	50	12	3	1
		25	6	12
		25	12	8
		25	19	2
		38	10	20
		38	19	24
		38	28	10
COP	50	–	–	–

Table 2: Number of instances (out of 25) terminated due to the time limit (excluding (D1A(ρ)) that was terminated due to the time limit on *all* instances with $n = 50$)

Finally, we recall that optimality gaps are not reported on instances terminated due to the time limit since they are not particularly meaningful in our setting.

In conclusion, the solution time of each relaxation seems to be very robust with respect to the choice of the instance set and the choice of (ρ_0, ρ) in our setting. The reduced formulations (D2A(ρ)) and (D2B(ρ)) yield significant computational advantages over their counterparts, especially in higher dimensions.

5.3   Quality of lower bounds

In this section, we discuss the quality of the lower bounds arising from the convex relaxations. To that end, we compare the absolute gaps of the exact models with those of the convex relaxations. For an instance, we define the MIQP gap as the difference between the best upper bound and the best lower bound obtained from (P1(ρ)) and (P2(ρ)). Similarly, we define the DNN gap to be the difference between the best upper bound obtained from the two exact models, (P1(ρ)) and (P2(ρ)), and the best lower bound obtained from the reduced formulations (D2A(ρ)) and (D2B(ρ)). Note that we use the lower bounds from the reduced formulations since all of them can

be solved to optimality on all instances.

If at least one of $(P1(\rho))$ and $(P2(\rho))$ can solve an instance to optimality, then the MIQP gap will be very close to zero. On such an instance, we deem that our relaxation is *exact* if the DNN gap is of similar magnitude to that of the MIQP gap. If each of $(P1(\rho))$ and $(P2(\rho))$ is terminated due to the limit on an instance, then the MIQP gap will be sufficiently away from zero. In this case, if the DNN gap is smaller than or equal to the MIQP gap, we say that our relaxations are better than the exact modes. Otherwise, our relaxations are *worse* than the exact models.

In Figure 5, which is organized similarly to Figure 2, we plot the DNN and MIQP gaps. To ease reading, we have accumulated all instances of the same dimensions from each instance set across all different parameter constellations (ρ_0, ρ), yielding 225 such instances in total for every graph (a)-(f), depicting the situation in the different hardness classes of the generated instances. In each graph, the horizontal axis represents the instances ordered in nondecreasing MIQP gaps, which are represented by the blue curve, and the vertical axis denotes the absolute gap. Finally, the markers indicate the data points at every 25th instance.

We outline our observations:

(i) For PSD instances, our relaxations are exact on the vast majority of all instances solved to optimality by an exact model and are better than the exact models on all instances with a positive MIQP gap.

(ii) For SPN instances, our relaxations are exact on a relatively smaller proportion of all instances solved to optimality by an exact model and are better than the exact models on all instances with a positive MIQP gap. Furthermore, it is worth noticing that our relaxations continue to be exact on several instances with a positive MIQP gap.

(iii) On COP instances, our relaxations are worse than the exact models on the vast majority of instances. However, we remark that the DNN gap is smaller on a small subset of instances with $n = 50$ that admit a positive MIQP gap.

The results show clearly that the DNN relaxations provide tight lower bounds on a large number of PSD and SPN instances and that there is a significant number of instances where the MIQP gap is larger than the gaps achieved by the DNN relaxations, sometimes drastically so (one or even two orders of magnitude). The dominance of MIQP models on the extremely difficult COP instances in graphs (e) and (f) may be explained by the fact that additional cuts may be needed to tighten the DNN gaps without resorting to higher-order relaxations of the (intractable) exact conic reformulations.

We close this part by reporting that we have not observed a significant difference between the quality of the lower bounds arising from the provably tighter relaxation $(D1B(\rho))$ and the weaker $(D2B(\rho))$. However, we still recommend using the tighter $(D1B(\rho))$ since it also outperforms $(D2B(\rho))$ in terms of the solution time on our instance sets.

6 Discussion, conclusions and outlook

In this paper, we studied the sparse StQP. By utilizing exact copositive reformulations of two mixed integer quadratic optimization models, we proposed two tractable convex relaxations. For both relaxations, we established equivalent tractable reformulations in significantly smaller dimensions. We also presented a theoretical comparison of the two relaxations. Finally, we proposed an instance generation scheme that is guaranteed to
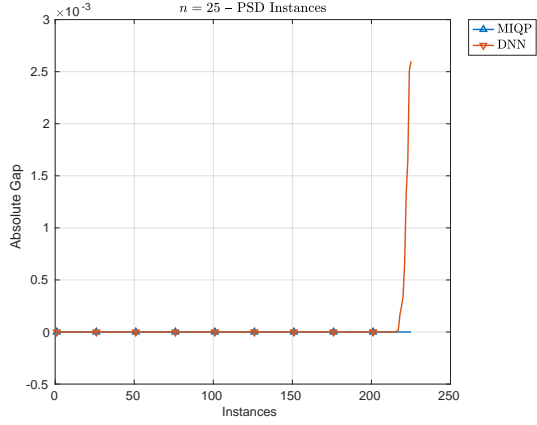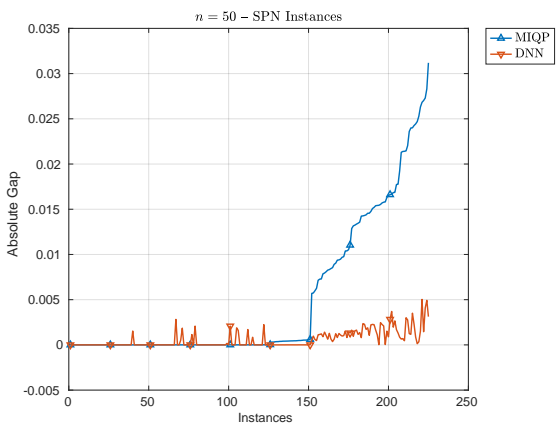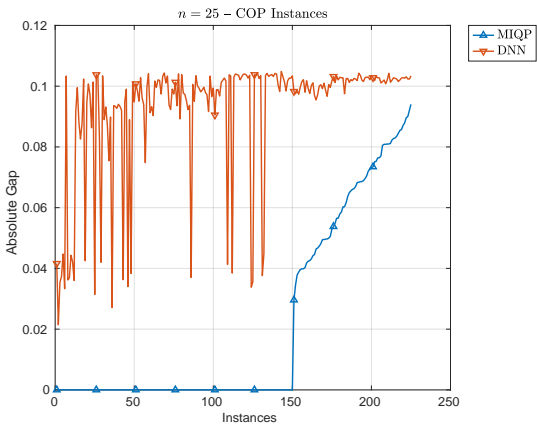
(a) PSD Instances ($n = 25$)
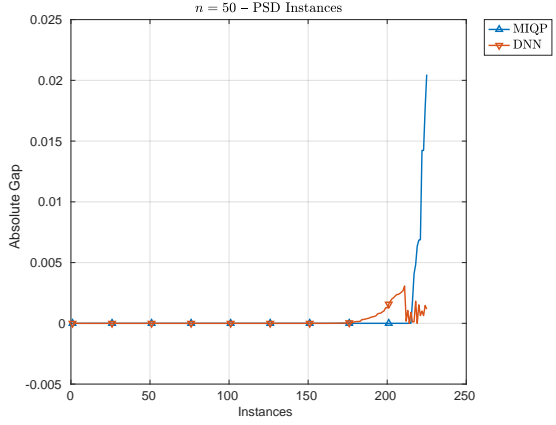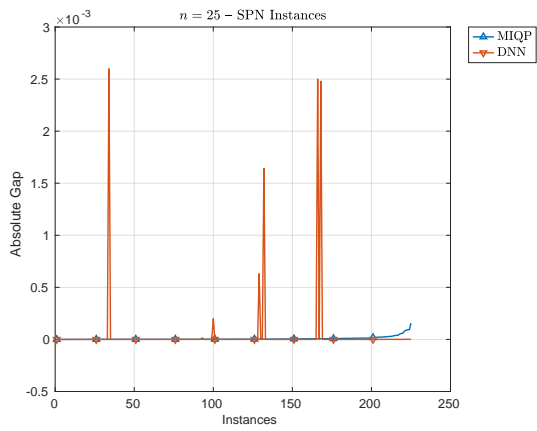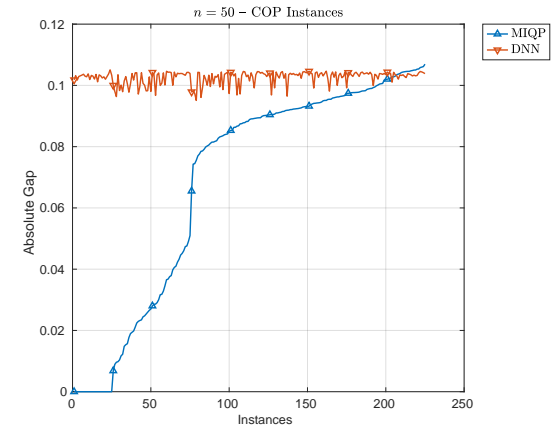
(b) PSD Instances ($n = 50$)

(c) SPN Instances ($n = 25$)

(d) SPN Instances ($n = 50$)

(e) COP Instances ($n = 25$)

(f) COP Instances ($n = 50$)

Figure 5: Absolute gaps of the best of $(P1(\rho))$ and $(P2(\rho))$ versus the best of $(D1B(\rho))$ and $(D2B(\rho))$

construct nontrivial instances. Our computational results clearly illustrate the computational advantages of our reduced relaxations as well as the quality of the relaxation bounds.

In the near future, we aim to focus on the description of the set of instances of $(StQP(\rho))$ that admit

exact relaxations. Motivated by our computational results, we intend to investigate how our relaxations can be strengthened for hard instances such as COP instances.

Statements and Declarations

Funding

Competing Interests

The authors declare that they have no conflict of interest.

Author Contributions

All authors contributed significantly to the study conception and design. The first draft of the manuscript was written with contributions from all authors. The experiments were designed and graphically represented mostly by E. Alper Yıldırım. All authors commented on all previous versions of the manuscript, and have read and approved the final manuscript.

Data Availability

The datasets generated and analyzed during the current study are available in the `github` repository.

Code Availability

The codes created during the current study are available in the `github` repository.

References

[1] A. Afonin, R. Hildebrand, and P. J. Dickinson. The extreme rays of the 6×6 copositive cone. *Journal of Global Optimization*, 79(1):153–190, 2021.

[2] M. ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 10.1.*, 2024.

[3] D. Bertsimas and R. Shioda. Algorithm for cardinality-constrained quadratic optimization. *Computational Optimization and Applications*, 43(1):1–22, 2009.

[4] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. Julia: A fresh approach to numerical computing. *SIAM Rev.*, 59(1):65–98, 2017.

[5] D. Bienstock. Computational study of a family of mixed-integer quadratic programming problems. *Mathematical Programming*, 74:121–140, 1996.

[6] I. Bomze, B. Peng, Y. Qiu, and E. A. Yıldırım. On tractable convex relaxations of standard quadratic optimization problems under sparsity constraints. *arXiv preprint arXiv:2310.04340*, 2023.

[7] I. M. Bomze. Evolution towards the maximum clique. *Journal of Global Optimization*, 10(2):143–164, Mar 1997.

[8] I. M. Bomze. Regularity versus degeneracy in dynamics, games, and optimization: A unified approach to different aspects. *SIAM Review*, 44(3):394–414, 2002.

[9] I. M. Bomze, M. Dür, E. de Klerk, C. Roos, A. J. Quist, and T. Terlaky. On copositive programming and standard quadratic optimization problems. *J. Glob. Optim.*, 18(4):301–320, 2000.

[10] I. M. Bomze, M. Kahr, and M. Leitner. Trust your data or not – StQP remains StQP: Community detection via robust Standard Quadratic Optimization. *Math. Oper. Res.*, 46:301–316, 2021.

[11] I. M. Bomze and B. Peng. Conic formulation of QPCCs applied to truly sparse QPs. *Computational Optimization and Applications*, 84:703–735, 2023.

[12] I. M. Bomze, W. Schachinger, and R. Ullrich. The complexity of simple models – a study of worst and typical hard cases for the standard quadratic optimization problem. *Mathematics of Operations Research*, 43(2):651–674, 2018.

[13] S. Braun and J. E. Mitchell. A semidefinite programming heuristic for quadratic programming problems with complementarity constraints. *Computational Optimization and Applications*, 31(1):5–29, 2005.

[14] J. Brodie, I. Daubechies, C. De Mol, D. Giannone, and I. Loris. Sparse and stable Markowitz portfolios. *Proceedings of the National Academy of Sciences*, 106(30):12267–12272, 2009.

[15] O. P. Burdakov, C. Kanzow, and A. Schwartz. Mathematical programs with cardinality constraints: Reformulation by complementarity-type conditions and a regularization method. *SIAM Journal on Optimization*, 26(1):397–425, 2016.

[16] S. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Math. Program.*, 120(2):479–495, 2009.

[17] E. J. Candès and M. B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008.

[18] X. Chen and J. Peng. New analysis on sparse solutions to random standard quadratic optimization problems and extensions. *Mathematics of Operations Research*, 40(3):725–738, 2015.

[19] X. Chen, J. Peng, and S. Zhang. Sparse solutions to random standard quadratic optimization problems. *Mathematical Programming*, 141:273–293, 2013.

[20] D. Di Lorenzo, G. Liuzzi, F. Rinaldi, F. Schoen, and M. Sciandrone. A concave optimization-based approach for sparse portfolio selection. *Optimization Methods and Software*, 27(6):983–1000, 2012.

[21] P. H. Diananda. On non-negative forms in real variables some or all of which are non-negative. *Mathematical Proceedings of the Cambridge Philosophical Society*, 58(1):17–25, 1962.

[22] J.-H. Du, Y. Guo, and X. Wang. High-dimensional portfolio selection with cardinality constraints. *Journal of the American Statistical Association*, pages 1–13, 2022.

[23] A. Frangioni and C. Gentile. SDP diagonalizations and perspective cuts for a class of nonseparable MIQP. *Operations Research Letters*, 35(2):181–185, 2007.

[24] Y. G. Gökmen and E. A. Yıldırım. On standard quadratic programs with exact and inexact doubly nonnegative relaxations. *Mathematical Programming*, 193(Ser. A):365–403, 2022.

[25] Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2023.

[26] M. Hall and M. Newman. Copositive and completely positive quadratic forms. *Mathematical Proceedings of the Cambridge Philosophical Society*, 59(2):329–339, 1963.

[27] S. Han, A. Gómez, and A. Atamtürk. The equivalence of optimal perspective formulation and Shor's SDP for quadratic programs with indicator variables. *Operations Research Letters*, 50(2):195–198, 2022.

[28] N. Hautsch and S. Voigt. Large-scale portfolio allocation under transaction costs and model uncertainty. *Journal of Econometrics*, 212(1):221–240, 2019.

[29] R. Hildebrand. The extreme rays of the 5×5 copositive cone. *Linear Algebra and its Applications*, 437(7):1538–1547, 2012.

[30] M. Lubin, O. Dowson, J. Dias Garcia, J. Huchette, B. Legat, and J. P. Vielma. JuMP 1.0: Recent improvements to a modeling language for mathematical optimization. *Mathematical Programming Computation*, 15:581–589, 2023.

[31] H. Markowitz. Portfolio selection. *The Journal of Finance*, 7(1):77–91, 1952.

[32] L. Mencarelli and C. D'Ambrosio. Complex portfolio selection via convex mixed-integer quadratic programming: A survey. *International Transactions in Operational Research*, 26(2):389–414, 2019.

[33] A. J. Miller. Selection of subsets of regression variables. *Journal of the Royal Statistical Society Series A*, 147(3):389–410, 1984.

[34] T. S. Motzkin and E. G. Straus. Maxima for graphs and a new proof of a theorem of Turán. *Canadian Journal of Mathematics*, 17:533–540, 1965.

[35] W. Murray and H. Shek. A local relaxation method for the cardinality constrained portfolio optimization problem. *Computational Optimization and Applications*, 53:681–709, 2012.

[36] K. G. Murty and S. N. Kabadi. Some NP-complete problems in quadratic and nonlinear programming. *Math. Program.*, 39(2):117–129, 1987.

[37] A. F. Perold. Large-scale portfolio optimization. *Management Science*, 30(10):1143–1160, 1984.

[38] R. Ruiz-Torrubiano, S. García-Moratilla, and A. Suárez. Optimization problems with cardinality constraints. In *Computational Intelligence in Optimization: Applications and Implementations*, pages 105–130. Springer, 2010.

[39] N. Shaked-Monderer. SPN graphs: when copositive = SPN. *Linear Algebra and its Applications*, 509:82–113, 2016.

[40] H. D. Sherali and W. P. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*. Springer US, Boston, MA, 1999.

[41] X. Sun, X. Zheng, and D. Li. Recent advances in mathematical programming with semi-continuous variables and cardinality constraint. *Journal of the Operations Research Society of China*, 1(1):55–77, 2013.

[42] The MathWorks Inc. `Matlab` Optimization Toolbox version 24.1 (R2024a), 2024.

[43] A. Wiegele and S. Zhao. Tight SDP relaxations for cardinality-constrained problems. In N. Trautmann and M. Gnägi, editors, *Operations Research Proceedings 2021*, Lecture Notes in Operations Research, pages 167–172. Springer, 2021.

[44] W. Xu, J. Tang, K. F. C. Yiu, and J. W. Peng. An efficient global optimal method for cardinality constrained portfolio optimization. *INFORMS Journal on Computing*, 36(2):690–704, 2024.

[45] X. Zheng, X. Sun, and D. Li. Improving the performance of MIQP solvers for quadratic programs with cardinality and minimum threshold constraints: A semidefinite program approach. *INFORMS Journal on Computing*, 26(4):690–703, 2014.

[46] X. Zheng, X. Sun, D. Li, and J. Sun. Successive convex approximations to cardinality-constrained convex programs: A piecewise-linear DC approach. *Computational Optimization and Applications*, 59(1):379–397, 2014.