Trust your data or not - StQP remains StQP:
Community Detection via Robust Standard Quadratic Optimization*

Immanuel M. Bomze, Michael Kahr and Markus Leitner
University of Vienna, Department of Statistics and Operations Research, Austria.
{immanuel.bomze|m.kahr|markus.leitner}@univie.ac.at
July 31, 2018

Abstract

We consider the Robust Standard Quadratic Optimization Problem (RStQP), in which an uncertain (possibly indefinite) quadratic form is extremized over the standard simplex. Following most approaches, we model the uncertainty sets by ellipsoids, polyhedra, or spectrahedra, more precisely, by intersections of sub-cones of the copositive matrix cone. We show that the copositive relaxation gap of the RStQP equals the minimax gap under some mild assumptions on the curvature of uncertainty sets, and present conditions under which the RStQP reduces to a single Standard Quadratic Optimization Problem. These conditions also ensure that the copositive relaxation of an RStQP is exact. The theoretical findings are accompanied by the results of computational experiments for a specific application from the domain of graph clustering, more precisely, community detection in (social) networks. The results indicate that the cardinality of communities tend to increase for ellipsoidal uncertainty sets and to decrease for spectrahedral uncertainty sets.

1 Introduction

1.1 Motivation, problem definition, and outline

Robust optimization allows to guarantee feasibility of solutions when uncertainty affects problem parameters. This approach has gained remarkable attention during the last decades, most likely triggered by the influential papers of Ben-Tal and Nemirovski [2, 3, 4]. Another reason for the attractiveness is that robust approaches (see, e.g., [8, 15, 17, 18] for comprehensive surveys) do not presume knowledge about the probability distribution of the uncertain data realizations, but only about bounds on them that define so-called uncertainty sets. However, there is typically a trade-off between the chosen uncertainty sets and the computational tractability of optimization.

*Supported by the Vienna Science and Technology Fund (project ICT15-014), and the Austrian Science Fund (project P 26755-N26).
problems in which uncertain data realizations are considered [2]. For instance, in linear optimization computational tractability translates from the uncertainty set to the robust problem. By contrast, this is not true for conic programs in which even simple uncertainty sets typically increase the computational effort needed to solve robust optimization problems significantly [5, p.152ff].

It is frequently assumed that uncertainty only affects constraints [5, 18], sometimes exploiting the epigraph formulation for the objective function, but there are, however, important cases in which one can safely trust the constraints. In particular, this is the case for the Standard Quadratic Optimization Problem (StQP) [9] in which a (possibly indefinite) quadratic form is extremized over the standard simplex

$$\Delta^n := \{ x \in \mathbb{R}_+^n : e^T x = 1 \},$$

in the $n$-dimensional Euclidean space $\mathbb{R}^n$. Indeed, thinking of probability distributions over $[1 : n]$, the integer interval $\{1, \ldots, n\}$, it is widely accepted that probabilities are nonnegative and sum up to one. Here we already used some notation: $\mathbb{R}_+^n := \{ x \in \mathbb{R}^n : x_i \geq 0 \text{ for all } i \in [1 : n] \}$ denotes the nonnegative orthant in $\mathbb{R}^n$, the vector of all-ones is denoted by $e := [1, \ldots, 1]^T \in \mathbb{R}_+^n$, and $(\cdot)^T$ denotes transposition. Moreover, we will use $e_i$ to denote the $i$-th column vector of the identity matrix $I$. Notice that throughout this article we reserve small bold-faced letters for column vectors (e.g., the all-zeros vector $o$) with $x_i$ referring to the $i$-th coordinate of a vector $x$; and bold-faced, capital letters for matrices (e.g., the all-zeros matrix $O$). The set of symmetric matrices of order $n$ is denoted by $S^n := \{ A \in \mathbb{R}^{n \times n} : A = A^T \}$. For some $Q \in S^n$, the StQP is defined by

$$\max_{x \in \Delta^n} x^T Q x .$$

Despite its simplicity, the StQP is quite versatile and has numerous applications from different domains [12], e.g., finance (Markowitz portfolio selection), economics (evolutionary game-theoretic algorithms), ecology (replicator dynamics), machine learning (image analysis), and graph theory (graph clustering). In all these fields, uncertain input data may arise quite naturally.

As a seemingly novel application, Section 5 discusses an application in social network community detection based on the similarity among community members which can only be roughly quantified and is therefore naturally subject to uncertainty. The proposed method allows (i) to extract communities without partitioning the whole network, and (ii) for multiple community memberships, i.e., individuals can belong to several different communities which is naturally the case in social networks; cf. [16, 23].

Since the StQP is NP-hard [19], adding significant complexity (due to uncertain input data) may lead to computational intractability even for very small problem instances. In this article we show how to (possibly) overcome these computational issues and derive conditions under which the problem complexity of the StQP does not increase when its objective function parameters are subject to uncertainty; see Definition 1.

**Definition 1. (Robust Standard Quadratic Optimization Problem (RStQP)).**

Let the nominal problem data be given by deterministic $Q \in S^n$, which is possibly affected by uncertain additive perturbations $U$ ranging within an uncertainty set $U \subset S^n$. Then the RStQP is defined by

$$\max_{x \in \Delta^n} \min_{U \in U} x^T (Q + U) x .$$
Remark 1. While we will discuss the RStQP in its max-min form in this article, all results can be easily transformed to the analogous min-max form. Alternatively, each StQP instance in maximization form can be solved in minimization form by appropriate manipulations of $Q$, see, e.g., [14]. Indeed this transformation also applies for the RStQP by manipulating set $\mathcal{U}$ accordingly. Moreover, we will desist from discussing general quadratic problems of the form $x^T Q x + c^T x$ over $\Delta^n$, since they can always be homogenized by considering rank-two update $Q' = Q + ec^T + ce^T$ in which case the objective values of the aforementioned general quadratic problem and the one of the StQP coincide [13] by replacing $Q$ with $Q'$ in (1). We can also solve general robust quadratic problems over $\Delta^n$ with the RStQP by manipulating set $U$ accordingly in (2) such that the uncertainty set for $Q'$ is compatible with the uncertainty model for $(Q, c)$. However, as in our main application, i.e., community detection in (social) networks, we deal with the homogeneous case in the first place, we will not pursue this issue further in this paper.

Contribution and outline. In the remainder of this section, we introduce further notation and terminology, as well as the completely positive (CP) relaxation of the RStQP. In Section 2 we show that the gap between the RStQP and its corresponding CP relaxation is equal to the minimax gap of Sion’s theorem [29], and show that this gap is closed under some reasonable assumptions on the uncertainty sets. In Section 3 we show that the RStQP reduces to an StQP for several (commonly used) uncertainty sets, while in Section 4 we derive a relaxation of the RStQP that avoids the minimax problem structure for the case of polyhedral uncertainty sets. In Section 5 we apply our findings to the Dominant Set Clustering Problem [22, 24] which aims to identify homogeneous clusters in networks such as communities formed by similar individuals in social networks. Thereby, we provide new insights for the deterministic and robust problem variants. The results of our computational experiments that accompany our theoretical findings are discussed in Section 6.

1.2 Further notation and terminology; lifting

We denote the Frobenius norm of a square matrix $A \in \mathbb{R}^{n \times n}$ by $\|A\|_F := (\sum_i \sum_j a_{ij}^2)^{1/2}$, where $a_{ij}$ is the entry in the $i$-th row and $j$-th column of $A$. The Euclidean norm of a vector $x$ is denoted by $\|x\|_2 := (\sum_i x_i^2)^{1/2}$.

We proceed with a lifting for problem (2) into a higher dimensional space in which the objective function is linear, and which goes back to Shor [28]. The aim hereby is to obtain the CP relaxation of (2) as will be seen later. Let the trace of matrix $A \in \mathbb{R}^{n \times n}$ be denoted by $\text{Tr}(A) := \sum_i a_{ii}$. Moreover, $\langle A, B \rangle := \text{Tr}(A^T B) = \sum_i \sum_j a_{ij} b_{ij}$ denotes the Frobenius inner product of matrices $A$ and $B$. Then observe that the objective function of (2) can be rearranged as

$$x^T Q x = \text{Tr}(x^T Q x) = \text{Tr}(Q x x^T) = \langle Q, x x^T \rangle = \langle Q, X \rangle,$$

where we define the matrix variable $X := x x^T$ and consider it as a point in $\mathbb{R}^d$ with $d = \binom{n+1}{2}$.

For an arbitrary cone $\mathcal{K} \subseteq \mathbb{R}^d$ we use ‘$\preceq_{\mathcal{K}}$’ to denote the Löwner ordering with respect to $\mathcal{K}$, meaning $A \preceq_{\mathcal{K}} B$ if and only if $(B - A) \in \mathcal{K}$. Moreover, we denote a set which is induced by such a conic ordering by

$$[A, B]_\mathcal{K} := \{ U \in \mathbb{R}^d : A \preceq_{\mathcal{K}} U \preceq_{\mathcal{K}} B \}.$$
The dual cone of a cone is indicated by an asterisk, that is, 
\[ K^* := \{ y : \langle x, y \rangle \geq 0 \text{ for all } x \in K \} . \]

To introduce the most prominent example of a pair of dual cones in this paper, let \( \text{conv}\{\cdot\} \) denote the convex hull of a set, i.e., the smallest convex set containing that set. Then, the cone of \textit{completely positive matrices} (see, e.g., [7]) is defined by
\[ C_n := \text{conv}\{xx^T : x \in \mathbb{R}_+^n \} = \{ X \in S^n : X = \sum_{i=1}^{k} y_i y_i^T \text{ for some } y_i \in \mathbb{R}_+^n \} . \]

The minimum integer \( k \) for which the condition of the latter definition of \( C_n \) is satisfied is the so-called \textit{cp-rank} of \( X \), which can be bounded by \[ \max\{\left(\frac{n+1}{2}\right) - 4, n\} \text{, see [27].} \]

Its dual cone is
\[ C_n^* := \{ Q \in S^n : \langle Q, X \rangle \geq 0 \text{ for all } X \in C_n \} = \{ Q \in S^n : x^T Q x \geq 0 \text{ for all } x \in \mathbb{R}_+^n \} , \]
i.e., the cone of \textit{copositive matrices}. Another familiar example, this time of a self-dual cone, is the cone of \textit{positive semidefinite matrices} obtained by dropping the nonnegativity requirement in the definition of \( C_n \), that is,
\[ P_n := \text{conv}\{xx^T : x \in \mathbb{R}^n \} = \{ Q \in S^n : x^T Q x \geq 0 \text{ for all } x \in \mathbb{R}^n \} . \]

A further self-dual cone is the set of all symmetric matrices with no negative entries, \( N_n := S^n \cap \mathbb{R}_+^{n \times n} \). Combinations of the latter two self-dual cones will give approximations of \( C_n \) and \( C_n^* \) respectively which are dual to each other: the \textit{doubly nonnegative (DNN)} cone \( P_n \cap N_n \supset C_n \) and the \textit{nonnegative decomposable (NND)} cone \( P_n + N_n \subset C_n^* \).

By denoting the matrix of all-ones \( E := ee^T \in S^n \) we define the \textit{lifted standard simplex} (marked in bold-faced letters) by
\[ \Delta^n := \{ X \in C_n : \langle E, X \rangle = 1 \} , \]
which is a compact, convex set with infinitely many extremal points \( \{xx^T : x \in \Delta^n \} \). Finally, using Equation (3), we can formulate the completely positive relaxation of the RStQP (in the lifted space) as a bilinear minimax problem over the set \( U \times \Delta^n \) by
\[ \max_{X \in \Delta^n} \min_{U \in U} \langle Q + U, X \rangle . \tag{4} \]

Notice, however, that although the objective function of (4) is bilinear, the problem difficulty did not vanish but is “hidden” in the constraint \( X \in \Delta^n \subseteq C_n \). The deterministic case (which can be embedded by putting \( U = \{Q\} \) in the current setting) satisfies
\[ \max_{X \in \Delta^n} \langle Q, X \rangle = \max_{x \in \Delta^n} x^T Q x . \tag{5} \]

Thus, the deterministic counterpart of (4) is a conic reformulation of (1), a completely positive optimization problem. The conic dual is a \textit{copositive (COP)} optimization problem of the form
\[ \min \{ y \in \mathbb{R} : y E - S = Q , S \in C_n^* \} , \tag{6} \]
which corresponds to identifying the smallest \( y \) such that \( y E - Q \) is copositive. For this problem class the conic duality and the relaxation gap is zero, as shown by Bomze et al. [13]. By contrast, in general there is a relaxation gap which exactly corresponds to the gap in the minimax inequality, even if the conic duality gap is zero (e.g., because of a Slater condition for the conic problems). We will deal with this and further related results in Section 2.
2 CP relaxation, minimax gap and Sion’s theorem

First we establish the fact that the completely positive relaxation (4) gives an upper bound on (2) and therefore deserves its name. Furthermore, we will, under mild assumptions on the uncertainty set \( \mathcal{U} \), establish an alternative interpretation of the relaxation value, and specify two conditions ensuring zero gap.

2.1 CP relaxation gap is minimax gap

Theorem 1. (CP relaxation gap is minimax gap).

(a) For general \( \mathcal{U} \), we have
\[
\max_{\mathbf{x} \in \Delta^n} \min_{\mathbf{u} \in \mathcal{U}} \mathbf{x}^T(\mathbf{Q} + \mathbf{u})\mathbf{x} \leq \max_{\mathbf{x} \in \Delta^n} \min_{\mathbf{u} \in \mathcal{U}} \langle \mathbf{Q} + \mathbf{u}, \mathbf{x} \rangle.
\]
(b) Suppose \( \mathcal{U} \) is closed and convex. Then
\[
\max_{\mathbf{x} \in \Delta^n} \min_{\mathbf{u} \in \mathcal{U}} \langle \mathbf{Q} + \mathbf{u}, \mathbf{x} \rangle = \min_{\mathbf{u} \in \mathcal{U}} \max_{\mathbf{x} \in \Delta^n} \mathbf{x}^T(\mathbf{Q} + \mathbf{u})\mathbf{x},
\]
so that the CP relaxation gap is exactly the gap in the minimax inequality.

Proof. (a) follows directly by lifting \( \mathbf{x}^T(\mathbf{Q} + \mathbf{u})\mathbf{x} = \langle \mathbf{Q} + \mathbf{u}, \mathbf{x}\mathbf{x}^T \rangle \) since \( \{\mathbf{x}\mathbf{x}^T : \mathbf{x} \in \Delta^n\} \subset \Delta^n \). To establish (b), we apply the already mentioned fact that for deterministic StQP, the CP relaxation is exact, using Sion’s minimax-theorem [29] for the lifted (linear) formulation:
\[
\min_{\mathbf{u} \in \mathcal{U}} \max_{\mathbf{x} \in \Delta^n} \mathbf{x}^T(\mathbf{Q} + \mathbf{u})\mathbf{x} = \max_{\mathbf{u} \in \mathcal{U}} \min_{\mathbf{x} \in \Delta^n} \langle \mathbf{Q} + \mathbf{u}, \mathbf{x} \rangle.
\]
Hence, the result follows.

By now, the possible non-concavity of the objective function \( \mathbf{x}^T(\mathbf{Q} + \mathbf{u})\mathbf{x} \) in \( \mathbf{x} \) suggests that relaxation (4) is in general not an exact reformulation of the RStQP (2), unlike the traditional StQP (with \( \mathcal{U} = \{\mathbf{0}\} \)). Somehow surprisingly, it turns out that some simple conditions on \( \mathcal{U} \) in fact ensure that the CP relaxation is exact and, moreover, that Sion’s theorem holds (i.e., there is no minimax gap) for the original problem, despite the possible non-concavity.

2.2 Linear dependence closes gap

In this subsection we discuss the RStQP for arbitrary uncertainty sets, however, assuming that the worst-case data realizations and the deterministic data are linearly dependent. Notice that such uncertainty sets occur naturally when the uncertain data realizations are estimated to lie within a relative tolerance with respect to the deterministic data, e.g., within a tolerance of \( \pm 5\% \).

Theorem 2. (Linear dependence closes gap).

Let \( \mathcal{L} = \{\lambda \mathbf{Q} : \lambda \in \mathbb{R}\} \) and \( \mathcal{U} \) be an arbitrary uncertainty set such that a suitable \( \mathcal{L} \in \mathcal{U} \cap \mathcal{L} \) satisfies
\[
\min_{\mathbf{u} \in \mathcal{U}} \langle \mathbf{Q} + \mathbf{u}, \mathbf{x} \rangle = \langle \mathbf{Q} + \mathbf{L}, \mathbf{x} \rangle \quad \text{for all} \; \mathbf{x} \in \Delta^n.
\]
Then the RStQP (2) reduces to an StQP equivalent to the nominal one (with data \( \mathbf{Q} \)). Hence, its CP relaxation is exact and Sion’s theorem holds.
Proof. By assumption there exists a $\lambda \in \mathbb{R}$ such that $L = \lambda Q$. Applying (7) to $X = xx^T \in \Delta^n$ for arbitrary $x \in \Delta^n$, we get
\[
\max_{x \in \Delta^n} \min_{U \in \mathcal{U}} \langle Q + U, xx^T \rangle = \max_{x \in \Delta^n} \langle Q + \lambda Q, xx^T \rangle = (1 + \lambda) \max_{x \in \Delta^n} x^T Q x ,
\]
a single StQP which has the same solutions as $\max_{x \in \Delta^n} x^T Q x$. Using the general minimax inequality, we arrive at
\[
\min_{U \in \mathcal{U}} \max_{x \in \Delta^n} x^T (Q + U)x \geq \max_{x \in \Delta^n} \min_{U \in \mathcal{U}} x^T (Q + U)x \geq \min_{U \in \mathcal{U}} \max_{x \in \Delta^n} x^T (Q + U)x .
\]
So all inequalities above are in fact equalities, and the proof is complete. \qed

Remark 2. Note that we impose no curvature assumptions on $U - L$ for Sion’s theorem to hold. By contrast, we will now assume that for some cone $K$, the relation $(U - L) \in K$ holds for all $U \in \mathcal{U}$, which relaxes concavity in $x$ and still ensures zero minimax gap.

2.3 COP-minimality closes gap

Lemma 3. Let $\mathcal{K} \subseteq \mathcal{C}_n^*$ be a sub-cone of the cone of copositive matrices. Let $\{L, U\} \subset \mathcal{S}^n$ be given matrices such that $L \preceq_{\mathcal{K}} U$. Then $\langle L, X \rangle \leq \langle U, X \rangle$ for all $X \in \mathcal{C}_n$.

Proof. Indeed $L \preceq_{\mathcal{K}} U$ is equivalent to $(U - L) \in \mathcal{K} \subseteq \mathcal{C}_n^*$ and therefore $\langle U - L, X \rangle \geq 0$ for all $X \in \mathcal{C}_n$ if and only if $\langle U, X \rangle \geq \langle L, X \rangle$. \qed

This elementary observation has the following consequences [30]:

Theorem 4. (COP-minimality closes gap).

Let $\mathcal{K} \subseteq \mathcal{C}_n^*$ be a sub-cone of the cone of copositive matrices. Suppose that $\mathcal{U}$ contains a matrix $L \in \mathcal{U}$ such that $L \preceq_{\mathcal{K}} U$ for all $U \in \mathcal{U}$. Then Sion’s theorem holds for the RStQP (2) and its CP relaxation (4) is exact.

Proof. Observe that the inner minimization problem of (4) only depends on the realizations $U$ for any given matrices $X \in \Delta^n$. Using Lemma 3 it follows by $\Delta^n \subseteq \mathcal{C}_n$ that
\[
\min_{U \in \mathcal{U}} \langle Q + U, X \rangle = \langle Q + L, X \rangle .
\]
Repeating the argument of Theorem 2 completes the proof. \qed

Remark 3. Observe that uncertainty sets constructed according to Theorem 4 also implies that the dual of the CP relaxation, i.e., its copositive relaxation (COP relaxation), is exact since strong duality holds for the StQP [11] (and its CP/COP relaxations). Hence,
\[
\max_{x \in \Delta^n} \min_{U \in \mathcal{U}} \langle Q + U, X \rangle = \min \{ y \in \mathbb{R} : yE - S = Q + L, S \in \mathcal{C}_n^* \} .
\]
3 Trust your data or not – StQP remains StQP

In this section we will use the results from Section 2 to derive conditions under which the RStQP reduces to a single (deterministic) StQP. In particular we discuss uncertainty sets \( \mathcal{U} \) that allow to solve the inner minimization problem of (2) in closed form and, hence, do not increase the problem’s complexity.

3.1 Conic uncertainty

We proceed with discussing uncertainty sets composed by the intersection of cones, which include the frequently investigated box and spectrahedral uncertainty sets (see, e.g., [6]), respectively, as special cases.

**Corollary 1.** (Conic uncertainty with zero CP relaxation gap).

Let set \( \mathcal{K} \subseteq \mathbb{C}^n \) be a sub-cone of the cone of copositive matrices and \( L, B \in \mathcal{S}^n \) be given matrices. Assume that \( \mathcal{U} = [L, B]_\mathcal{K} \) with \( L \in \mathcal{U} \). Then the CP relaxation (4) is an exact reformulation of (2) and the RStQP reduces to an StQP (1) with problem data \( Q + L \).

**Proof.** Follows directly from Theorem 4.

The results from above allow to construct (infinitely) many conic uncertainty sets while ensuring that the resulting RStQP has the same complexity as the underlying deterministic StQP. These include the most important cones in conic optimization, and the aforementioned, frequently considered uncertainty sets (i.e., box-uncertainty or spectrahedral uncertainty) as special cases. To only list a few of such conic uncertainty sets with the key property \( L \in \mathcal{U} \), take one of the following options for \( \mathcal{U} \), along with their self-explanatory names:

1. \( [L, B]_{\mathbb{C}_n} \) CP-uncertainty,
2. \( [L, B]_{\mathbb{N}_n} \) box-uncertainty,
3. \( [L, B]_{\mathbb{P}_n} \) spectrahedral uncertainty,
4. \( [L, B]_{\mathbb{N}_n \cap \mathbb{P}_n} \) DNN-uncertainty,
5. \( [L, B]_{\mathbb{N}_n + \mathbb{P}_n} \) NND-uncertainty, and
6. \( [L, B]_{\mathbb{C}_n^*} \) COP-uncertainty.

**Remark 4.** Despite of the fact that there is no efficient separation oracle for uncertainty sets (i) and (vi), we nonetheless end up with a robust problem that has the same complexity as the deterministic problem. We further remark that set \( \mathcal{U} = [L, B]_\mathcal{K} \) need not even be convex if \( \mathcal{K} \) is not convex.
3.2 Ellipsoidal uncertainty

We will now discuss ellipsoidal uncertainty sets which have also received much attention in robust optimization.

**Theorem 5. (Ellipsoidal uncertainty).**

Let $C$ be a nonsingular $n \times n$ matrix and define, for some scalar $\rho > 0$, the uncertainty set $U = \{ U \in S^n : \|C^TUC\|_F \leq \rho \}$. Then

$$\max_{x \in \Delta^n} \min_{U \in U} \langle Q + U, X \rangle = \max_{x \in \Delta^n} \langle Q - \rho(C^T)^{-1}, X \rangle,$$

i.e., the RStQP reduces to an instance of an StQP.

**Proof.** Note that the inner minimization problem of (8) depends only on the realizations of $U$ for any given matrices $Q$ and $X$. Since $C$ is nonsingular by definition, the relation $y = C^{-1}x$ is equivalent to $x = Cy$. Rearranging the variable term of the inner problem of (8) amounts to

$$\langle U, X \rangle = x^T U x = x^T [(C^T)^{-1}C^TUC]^{-1} x = y^T (C^T)^{-1}C^TUC = \langle C^TUC, yy^T \rangle,$$

which is bounded from below via the Cauchy-Schwarz inequality by

$$\langle C^TUC, yy^T \rangle \geq -\|C^TUC\|_F \|yy^T\|_F \geq -\|yy^T\|_F.$$

Note that the Cauchy-Schwarz inequality holds with equality if and only if $U = \lambda X$:

$$\langle C^TUC, yy^T \rangle = -\|C^TUC\|_F \|yy^T\|_F \iff C^TUC = \lambda yy^T,$$

for some scalar $\lambda$ bounded by $0 \leq \lambda \leq -\|yy^T\|_F$ due to the right-most inequality of (9). Consequently, we can solve the inner minimization problem by

$$\min_{U \in U} \langle C^TUC, yy^T \rangle = \min_{\lambda \leq 0} \langle \lambda yy^T, yy^T \rangle = -\frac{\rho}{\|yy^T\|_F^2} \|yy^T\|_F^2,$$

or equivalently,

$$\min_{U \in U} \langle U, X \rangle = -\rho \|yy^T\|_F^2 = -\rho \|C^{-1}x\|_2^2 = -\rho \langle C^{-1}x, C^{-1}x \rangle = -\rho \langle (CC^T)^{-1}, X \rangle.$$

Plugging the latter minimum over $U \in U$ in (8) and collecting $X$, the result follows. \qed

**Remark 5.** If $C$ is an orthogonal matrix, that is, $CC^T = I$, the ellipsoidal uncertainty set corresponds to a Frobenius-ball, the simplest ellipsoid. Moreover, notice that by defining $L := -\rho(CC^T)^{-1}$, the statement is equivalent to the result derived in Theorem 4, i.e.,

$$\max_{x \in \Delta^n} \min_{U \in U} x^T(Q + U)x = \max_{x \in \Delta^n} x^T(Q + L)x.$$

However, in case of ellipsoidal uncertainty, this $L$ need not lie in $U$, in contrast to the setting of Theorem 4. Indeed, $\|C^TLC\|_F = \rho \sqrt{n} > \rho$ in the orthogonal case.
4 Polyhedral uncertainty

As discussed in Section 2, the gap of the CP relaxation (4) is equal to the minimax gap. This limits the practicability of conic reformulations of the RStQP for arbitrary uncertainty sets. Thus, the computational complexity of the RStQP may be significantly higher than the one of a single StQP in these cases.

However, as will be demonstrated below one can obtain an approximation that avoids the minimax structure of the RStQP when considering polyhedral uncertainty sets. The cost incurred by tractable approximations comes from two gaps, the minimax and the approximation gap. The latter originates when replacing $C_n$ with a tractable cone, e.g. the DNN cone $\mathcal{P}_n \cap \mathcal{N}_n$. But first we reformulate the RStQP as a quadratic optimization problem under nonconvex, quadratic and linear constraints:

**Theorem 6.** Let $\mathcal{U} = \{U \in S^n : \langle A_i, U \rangle \leq b_i, \text{ for all } i \in [1 : m]\}$ be a nonempty polyhedral uncertainty set. Then

$$\max_{x \in \Delta_n} \min_{U \in \mathcal{U}} \langle Q + U, xx^T \rangle = \max_{(x,y) \in \Delta_n \times \mathbb{R}^m_+} \left\{ x^T Q x - b^T y : -\sum_{i=1}^m y_i A_i = xx^T \right\}. \quad (10)$$

**Proof.** By assumption, the inner minimization problem of the RStQP is a linear program, hence, strong duality holds. If $y \in \mathbb{R}^m_+$ are dual variables associated with the constraints in the definition of $\mathcal{U}$, dualizing yields (10). \qed

As such, Theorem 6 enables applying all methods for treating (nonconvex) all-quadratic problems, but it also opens an avenue of getting tractable bounds:

**Theorem 7.** Let $\mathcal{U}$ be as in Theorem 6. Then we have the following upper bounds for the RStQP in terms of linear conic optimization problems:

$$\max_{x \in \Delta_n} \min_{U \in \mathcal{U}} \langle Q + U, xx^T \rangle$$

$$\leq \max_{(X,y) \in \Delta_n \times \mathbb{R}^m_+} \left\{ \langle Q, X \rangle - b^T y : -\sum_{i=1}^m y_i A_i = X \right\} \quad (11)$$

$$\leq \max_{(X,y) \in \mathcal{N}_n \cap \mathcal{P}_n \times \mathbb{R}^m_+} \left\{ \langle Q, X \rangle - b^T y : -\sum_{i=1}^m y_i A_i = X, \langle E, X \rangle = 1 \right\}. \quad (12)$$

**Proof.** The first inequality is obvious by (10) via Shor lifting, and the second by the inclusion $C_n \subseteq \mathcal{P}_n \cap \mathcal{N}_n$. \qed

Inequality (11) corresponds exactly to the CP relaxation gap. Indeed, note that dualizing the lifted formulation in (11) by associating dual variables $U \in S^n$ to the constraint yields

$$\max_{x \in \Delta_n} \min_{U \in S^n} \left\{ \langle Q + U, X \rangle : \langle A_i, U \rangle \leq b_i, \text{ for all } i \in [1 : m] \right\},$$

and a similar statement holds for the tractable approximation (12).
5 Robust dominant set clustering

In this section we address the Robust Dominant Set Clustering Problem which has been proposed in its deterministic form by Pavan [21] who established relations between dominant sets and extrema of an StQP similar to the well-known theorem of Motzkin and Straus [20] for maximal cliques in simple undirected graphs $G = (V, E)$ with finite vertex-set $V$ and edge-set $E \subseteq V \times V$. A clique is a vertex-set $S \subseteq V$ such that its induced subgraph $G_S$ is complete. A clique $S$ is maximal if there exists no other clique $S'$ such that $S \subset S'$ and a maximum clique has the largest cardinality among all (maximal) cliques.

5.1 Description and definition of the deterministic problem

The Dominant Set Clustering Problem aims to cluster similar vertices of an undirected, weighted graph $G = (V, E, w)$ with a given weight function $w : E \mapsto \mathbb{R}_+$. An edge weight $w(i, j)$ reflects the similarity between vertices $i$ and $j$ for each $\{i, j\} \in E$. These edge weights are also used to define the weighted adjacency matrix $A \in \mathcal{N}_{|V|}$ in which positive values are assigned to all pairs of adjacent vertices and zero otherwise, i.e., $a_{ij} = w(i, j)$ for all $\{i, j\} \in E$, $a_{ij} = 0$ for all $\{i, j\} \notin E$, and $a_{ii} = 0$ for all $i \in V$.

Dominant sets can be informally defined as sets containing entities that are internally alike but externally different. Thus, elements of a dominant set are dissimilar to elements outside that set. The following formal definition of dominant sets follows [21, 22]. For a non-empty subset $S \subseteq V$ let the average weighted degree of vertex $i \in S$ with respect to $S$ be defined by $awdeg_S(i) = \frac{1}{|S|} \sum_{j \in S} a_{ij}$. For all $j \notin S$ let $\phi_S(i, j) = a_{ij} - awdeg_S(i)$. Intuitively, $\phi_S(i, j)$ measures the similarity of $i$ and $j$ with respect to the average similarity of $i$ to its neighbors within $S$. The main idea behind dominant set clustering is to associate weights (similarity values with respect to sets) recursively defined via

$$w_S(i) = \begin{cases} 1 & \text{if } |S| = 1, \\ \frac{1}{\sum_{j \in S \setminus \{i\}} \phi_{S \setminus \{i\}}(j, i) \cdot w_{S \setminus \{i\}}(j)} & \text{otherwise}, \end{cases}$$

to vertices $i \in S \subseteq V$. The total weight of set $S$ is given by $W(S) = \sum_{i \in S} w_S(i)$ and dominant sets are (intuitively) composed of vertices with similar weights; see Definition 2 for a formal definition.

Definition 2. (Dominant sets).

A non-empty set of vertices $S \subseteq V$ is a dominant set if

(i) $W(T) > 0$ for all $T \subseteq S$, $T \neq \emptyset$,

(ii) $w_S(i) > 0$ for all $i \in S$, and

(iii) $w_{S \cup \{i\}}(i) < 0$ for all $i \notin S$.

Similar to the theorem of Motzkin and Straus [20], Pavan [21] showed that dominant sets correspond to strict local maximizers $x'$ of $\max_{x \in \Delta^{|V|}} x^T A x$. More precisely, if $x'$ is a strict local maximizer, then $S = \sigma(x') := \{i \in V : x'_i > 0\}$ is a dominant set, if $w_{S \cup \{i\}}(i) \neq 0$ for all $i \in V \setminus S$. Conversely, any dominant set $S$ coincides with the support $\sigma(x')$ of a strict local maximizer $x'$. We
Figure 1: Dominant sets in a complete graph. Vertices connected with bold edges (of the line style) belong to a dominant set.

refer to [24] for a recent, comprehensive review on dominant set clustering, solution algorithms, and applications from several domains including bioinformatics, computer vision, climatology, and medical data analysis.

Some intuition (based on [24]) is provided in Figure 1 in which the concept of dominant sets is sketched for instances based on a complete graph. In Figure 1a, a dominant set is formed by vertices \{i, j, l\}, since they are equally similar to each other, while vertex k which is quite different due to its weak connection to all other vertices. Conversely, in Figure 1b the comparably strong similarity between k and its neighbors induces a dominant set of cardinality four despite the dissimilarity of i, j, and l. Figure 1c shows that dominant sets need not be disjoint since vertex j belongs to dominant sets \{i, j, k\} and \{j, l, m\}. Finally, Figure 1d depicts a similar situation as Figure 1a. Notice that although set \{i, j, l, m\} is “more similar” (i.e., stronger connected) than set \{i, j, l\} in Figure 1a it is not dominant. This illustrates that similarity is a relative measure that is immune against scaling but sensitive to translation, i.e., additive perturbations on the instance matrix A. Observe that insensitivity against scaling is also implied by the fact that we can always scale matrix A by a scalar \(\gamma > 0\) without changing the maximizers of an StQP; cf. Theorem 2.

The following results relate the notion of a dominant set to the structure of the underlying graph.

**Proposition 8. (Connectedness).**

Let \(x'\) be a locally optimal solution to \(\max_{x \in \Delta^{|V|}} x^T A x\) with \(A \in \mathcal{N}_n \setminus \{O\}\), and \(\sigma(x') = \{i \in V : x'_i > 0\}\). Then the graph induced by node set \(\sigma(x')\) is connected.

**Proof.** Suppose we are given an instance,

\[
A = \begin{bmatrix}
A_1 & O & * \\
O & A_2 & * \\
* & * & *
\end{bmatrix} \in \mathcal{N}_n \setminus \{O\}
\]

in which \(A_1 \in \mathcal{N}_q\), \(A_2 \in \mathcal{N}_p\), and asterisks * denote arbitrary block matrices of appropriate dimensions, where the graph induced by \(\sigma(x')\) contains at least two connected components, if \(x' = [x_1, x_2, o]^T\) is a local maximizer of \(f(x) = x^T A x\) over \(\Delta\) such that \(x_1 \in \mathbb{R}_+^q \setminus \{o\}\) and
\(x_2 \in \mathbb{R}_+^n \setminus \{0\}\) are both non-zero vectors. From \(e^T x' = 1\) we gather that \(x'\) is a convex combination

\[x' = (1 - \lambda)u + \lambda v\]

with \(\{u, v\} = \left\{ \begin{bmatrix} \frac{x_1}{e^T x_1} & o \\ o & o \end{bmatrix}, \begin{bmatrix} o & \frac{x_2}{e^T x_2} \\ o & o \end{bmatrix} \right\} \subseteq \Delta^n\)

where \(\lambda = e^T x_2 \in (0, 1)\) locally maximizes \(g(t) = f((1 - t)u + tv)\) in \(t = \lambda\). Therefore we must have \(\dot{g}(\lambda) = 0\) and \(\ddot{g}(\lambda) = (v - u)^T A(v - u) \leq 0\). On the other hand, the assumed structure of \(A\), \(u\) and \(v\) yields

\[\dot{g} = 0\] hence, \(\ddot{g} = 0\) so \(g\) must be affine-linear in \(t\), hence constant, hence zero because of \(g(0) = f(u) = f(v) = g(1) = 0\), which follows from (13) since we have established equality to zero there. Local optimality of \(x\) over \(\Delta^n\) implies \(0 \leq f(y) \leq f(x') = 0\) for all \(y \in \Delta^n\) close to \(x\) and then, since \(f\) is quadratic, for all \(y \in \Delta^n\), which is impossible if \(A \neq O\).

**Proposition 9. (Clique condition).**

Every dominant set of a graph \(G\) forms a clique.

**Proof.** Consider the weighted adjacency matrix \(A\) of \(G\) and let \(x'\) be a strict local maximizer of \(x^T Ax\). Then \(\{i, j\} \subseteq \sigma(x')\) implies \(x_i' x_j' > 0\). Let \(t\) with \(|t| > 0\) be small enough and consider a slight perturbation of the maximizer \(x'\) of the form \(x(t) = x' + t(e_j - e_i)\). Arguing as in the proof of Proposition 8, strict local optimality of \(x'\) implies the second-order condition \((e_i - e_j)^T A(e_i - e_j) = \ddot{g}(0) < 0\) for the quadratic function \(g(t) = f(x(t))\), i.e., \(a_{ii} + a_{jj} - 2a_{ij} < 0\). Thus, \(a_{ij} > 0\) (which follows from \(a_{ii} \geq 0\) and \(a_{jj} \geq 0\)) holds and, hence, \(G\) contains an edge \(\{i, j\} \in E\). \(\square\)

### 5.2 Robust dominant set clustering

Adapting the notation of the RStQP slightly we define the **Robust Dominant Set Clustering Problem** by

\[
\max_{x \in \Delta^{|V|}} \min_{U \in \mathcal{U}} x^T (A + U)x, \tag{14}
\]

in which \(A \in N_{|V|}\) is the weighted adjacency matrix of graph \(G = (V, E, w)\), and edge weights are uncertain. Such uncertainty may arise if the appropriateness of a certain similarity measure is fuzzy. This is, for instance, the case when vertex similarities are used in social network analysis to identify communities. The choice of a particular similarity measure may be a further source of uncertainty. To this end, observe that there exists a large body of literature dedicated to the investigation of such measures; see, e.g., [16, 26].

Now, if the minimax theorem holds (e.g., if the uncertainty sets are constructed as described in Section 3) and if the maximizer \(x'\) of \(x^T (Q + L)x\) is strict, then \(S = \sigma(x')\) is a clique for the graph with adjacency matrix \(A = Q + L\). However, it need not be a clique for the graph underlying the nominal instance \(Q\); see, e.g., Figure 3d. Whenever \(a_{ij} > 0\) implies \(q_{ij} > 0\) we can use Proposition 8 to show connectedness under \(Q\). This will be the case if uncertainty sets \(\mathcal{U}\) are chosen such that they respect the graph structure of the nominal instance. One such possibility is discussed in Section 2.2.
6 Computational experiments

This section discusses the results of our computational study for which the considered algorithm has been implemented in MATLAB 2016b. Each experiment has been performed on a single core of an Intel Xeon E5-2670v2 machine with 2.5 GHz and 8 GB random-access memory.

Besides reporting usual quantitative results such as numbers of identified dominant sets, their sizes, and corresponding runtimes, we also investigate the influence of different uncertainty sets and their sizes on (the structure of) the obtained solutions. All results will be compared to the deterministic case, whose results will be provided in Section 6.3 after describing our benchmark instances and the used solution algorithm in Sections 6.1 and 6.2, respectively. The uncertain case is discussed in Section 6.4. Thereby, results for ellipsoidal uncertainty sets are given relative to box uncertainty sets \( U = [L, B]_{\mathbb{R}^n} \) for the sake of comparability.

6.1 Instance description

We use benchmark instances from the 10th DIMACS Implementation Challenge (on graph partitioning and graph clustering) [1] and elementary (comparable) similarity measures, i.e., weight-functions \( w^{(k)} : E \rightarrow [0, 1]|E| \), to construct \( K \) similarity graphs \( G^{(k)} = (V, E, w^{(k)}) \) and weighted adjacency matrices \( A^{(k)} \) for all \( k \in [1 : K] \). Thereby, we assume that none of these similarity measures reflects the “true” similarity between two vertices \( i, j \in V \) but use them to construct conic uncertainty sets as follows. From each similarity measure, i.e., weight-functions \( w^{(1)}, \ldots, w^{(K)} \), we construct weighted adjacency matrices \( A^{(1)}, \ldots, A^{(K)} \) and derive worst-case data realizations \( A^{\text{min}} \) by

\[
A^{\text{min}}_{ij} = \min_{k \in [1 : K]} A^{(k)}_{ij}, \quad \text{for all } i, j \in V,
\]

i.e., by taking the entry-wise minimum data realizations over all similarity measures \( K \). Moreover, we construct the deterministic data by

\[
A = \frac{1}{K} \sum_{k=1}^{K} A^{(k)},
\]

i.e., by averaging weight-functions \( w^{(k)} \) across all \( k \in [1 : K] \). The conic uncertainty sets \( U = [L, B]_{\mathbb{R}^n} \) which we use as reference for ellipsoidal uncertainty sets are constructed such that \( L := A^{\text{min}} - A \). Therefore, \( A^{\text{min}} = A + L \) and \( A^{\text{min}} \leq A + U \) holds entrywise for all \( U \in \mathcal{U} \). Notice that we could define matrix \( B \) similarly to (15) by taking the entry-wise maximum. This is, however, not necessary for the considered uncertainty sets which we detail in Section 6.4.

Exemplary, we use three well-known (vertex-)similarity measures, i.e., \( K = 3 \), based on the common neighbors of vertices \( i, j \in V \) defined by sets \( |N(i) \cap N(j)| \). Thereby, for each \( i \in V \), \( N(i) = \{ j \in V \mid \{i, j\} \in E \} \) is its set of adjacent nodes and the vertex similarities are defined as

\[
a^{(k)}_{ij} = \frac{|N(i) \cap N(j)|}{d^{(k)}}, \quad \text{for all } i, j \in V, \quad k \in [1 : 3].
\]

Denominators \( d^{(k)} \) are defined by \( d^{(1)} = |N(i) \cup N(j)| \) for the Jaccard-index, \( d^{(2)} = \sqrt{|N(i)||N(j)|} \) for the Salton-index, and \( d^{(3)} = \min\{|N(i)|, |N(j)|\} \) for the hub-promoted-index; see, e.g., [31]. Notice that all these similarity values have been scaled by a factor 100.
6.2 Description of the solution algorithm

*Infection Immunization Dynamics (INImDyn)* [25] used for all computations is based on evolutionary game theory and derives solutions which satisfy conditions (17) and (18).

\[
\begin{align*}
[Ax]_i &= x^TAx, \text{ for all } i \in \sigma(x), \\
[Ax]_i &\leq x^TAx, \text{ for all } i \notin \sigma(x).
\end{align*}
\]  

(17)

Conditions (17) refer (in a game-theoretic sense) to a *Nash strategy* \(x \in \Delta^n\) for a given payoff matrix \(A\). To guarantee that Nash strategy \(x \in \Delta^n\) realizes a strict (local) maximum of the RStQP it has to be *evolutionary stable*, that is, it satisfies the condition

\[
y^TAx = x^TAx \quad \text{and} \quad y \in \Delta^n \quad \implies \quad y^TAy < x^TAy.
\]

(18)

A detailed description of INImDyn that is sketched in Algorithm 1 can be found in [25]. Given a starting point \(x \in \Delta^n\) the algorithm chooses a direction \(y \in \Delta^n\) by selection function \(S_{pure}(x)\), thereby, maximizing the improvement step in each iteration based on *pure strategies* \(e_i\) or *co-strategies* \(\bar{e}_x^i\) with respect to vector \(x\) for all \(i \in V\). Hence, \(y \leftarrow e_i\) or \(y \leftarrow \bar{e}_x^i = \frac{1}{x_i} (x_i - x) + x\). As long as such improving directions exist, that is, the condition in Algorithm 1 is satisfied, vector \(x\) is updated by “injecting” a fraction \(\alpha\) of the improving direction \(y\). The algorithm terminates once the stopping tolerance \(\tau\) is reached. Rota Bulò et al. [25] show that given a starting point \(x^{(0)} \in \Delta^n\) that does not satisfy conditions (17), the objective value of the StQP (and hence the RStQP) strictly increases along any non-constant trajectory in each iteration step. They also prove that generally across instances \(A\), all iteration trajectories converge to a strict local solution \(x\) satisfying (18) such that \(\sigma(x)\) is a dominant set; cf. [10].

**Data:** data matrix \(A \in \mathcal{N}_n\), worst-case perturbations \(L \in \mathcal{U}\), starting point \(x \in \Delta^n\), stopping tolerance \(\tau\).

**Result:** (local) optimum \(x \in \Delta^n\)

\[
A' \leftarrow A + L
\]

while \(\epsilon(x) > \tau\) do

\[
\begin{array}{l}
\quad y \leftarrow S_{pure}(x) \\
\quad \alpha \leftarrow 1 \\
\quad \pi \leftarrow (y - x)^TA'y - (y - x)^TA'x \\
\quad \text{if } \pi < 0 \text{ then} \\
\quad \quad \alpha \leftarrow \min \left[\frac{(y-x)^TA'x}{\pi}, 1\right] \\
\quad \quad x \leftarrow \alpha(y - x) + x \\
\quad \epsilon(x) \leftarrow \sum_i \min\{x_i, x^TA'x - [A'x]_i\}^2
\end{array}
\]

**Algorithm 1:** Robust Infection Immunization Dynamics Algorithm

6.3 Results for deterministic instances

Computational results for the benchmark instances of the deterministic case, i.e., solving \(\max_{x \in \Delta^n} x^TAx\) are given in Table 1. Thereby, we used 1000 different starting points (indicated
Table 1: Results of the deterministic problem instances described by the numbers of vertices (\(|V|\)) and numbers of edges (\(|E|\)) based on 1000 multi-starts of \textsc{InImDyn} per instance. Numbers of identified different solutions (\(\# \sigma(x)\)) and the corresponding average cardinalities of dominant sets (\(|\sigma(x)|\)) as well as average objective function values (\(f(x)\)) and average runtimes in seconds (\(t[s]\)) of each multi-start iteration are reported.

| Instance name                  | \(|V|\) | \(|E|\) | \(\# \sigma(x)\) | \(|\sigma(x)|\) | \(f(x)\) | \(t[s]\) |
|-------------------------------|--------|--------|-----------------|-----------------|---------|--------|
| AG-Monien_3elt                | 4720   | 13722  | 928             | 3.02            | 20.78   | 0.72   |
| Arenas_celegans_metabolic     | 453    | 2007   | 100             | 3.44            | 37.83   | 0.42   |
| Arenas_email                  | 1133   | 4229   | 145             | 3.78            | 29.36   | 0.12   |
| Arenas_jazz                   | 198    | 2734   | 17              | 6.76            | 74.86   | 0.18   |
| DIMACS10_chesapeake           | 39     | 163    | 9               | 3.11            | 43.84   | 0.01   |
| DIMACS10_data                 | 2851   | 15081  | 786             | 3.93            | 35.77   | 0.69   |
| DIMACS10_delaunay_n10         | 1024   | 3056   | 373             | 3.14            | 23.77   | 0.07   |
| DIMACS10_delaunay_n11         | 2048   | 6127   | 554             | 3.09            | 23.86   | 0.16   |
| DIMACS10_delaunay_n12         | 4096   | 12264  | 727             | 3.12            | 23.73   | 0.55   |
| Hamm_add20                    | 2395   | 7096   | 566             | 3.28            | 27.89   | 0.22   |
| Hamm_add32                    | 4960   | 8782   | 706             | 3.14            | 24.57   | 0.74   |
| Newman_adjnoun                | 112    | 308    | 6               | 3.83            | 28.69   | 0.06   |
| Newman_celegansneural         | 297    | 2029   | 59              | 3.54            | 37.15   | 0.22   |
| Newman_dolphins               | 62     | 121    | 7               | 3.71            | 40.51   | 0.01   |
| Newman_football               | 115    | 517    | 18              | 6.17            | 49.57   | 0.06   |
| Newman_karate                 | 34     | 67     | 5               | 3.40            | 38.50   | 0.01   |
| Newman_lesmis                 | 77     | 232    | 9               | 5.00            | 66.51   | 0.02   |
| Newman_netscience             | 1589   | 2521   | 274             | 3.98            | 44.70   | 0.12   |
| Newman_polblogs               | 1490   | 16029  | 36              | 4.06            | 46.84   | 0.36   |
| Newman_polbooks               | 105    | 423    | 21              | 4.48            | 39.44   | 0.05   |

6.4 Results for robust instances

This section considers conic, ellipsoidal, and ball uncertainty and compares the obtained results for different sizes of these uncertainty sets. Thereby, conic uncertainty sets \(U = \delta L\) are considered for

by superscript \((0)\) for each multi-start of \textsc{InImDyn} for each instance, i.e., \(x_m^{(0)} \in \Delta^n\) for all \(m \in [1 : 1000]\), such that

\[
x_m^{(0)} = \begin{cases} \frac{1}{|V|} e & \text{if } m = 1, \\ e_m & \text{if } 1 < m \leq |V|, \\ \tilde{x} & \text{otherwise}, \end{cases}
\]

where \(e_m\) is the \(m\)–th column of the identity matrix, and \(\tilde{x}\) denotes a vector chosen uniformly at random. As a stopping criterion we set \(\tau = 10^{-11}\), based on prior tests showing that even using \(\tau = 10^{-8}\) would not change the number of identified dominant sets in the deterministic case. The results in Table 1 build the reference points for all computations considering uncertain data realizations. More precisely, the average across all multi-starts and instances of numbers of different solutions found (\(\# \sigma(x)\)), average cardinalities of dominant sets (\(|\sigma(x)|\)), as well as average objective function values (\(f(x)\)), correspond to points \((0, 1)\) in Figure 2.
different sizes $\|\delta L\|_1 = \sum_i \sum_j |\delta |_{ij}$ induced by scalar $\delta \in [0, 1]$. Worst-case data realization $L$ are calculated as described in Section 6.1 and we consider $\delta = \ell/20$ for all $\ell \in [0; 20]$ in our experiments.

As described in Section 3.2, the worst-case solutions for considering ellipsoidal uncertainty is given by

$$\max_{x \in \Delta |V|} \min_{U \in \ell} x^T(A + U)x = \max_{x \in \Delta |V|} x^T(A + \rho(C^T)^{-1})x,$$

for some nonsingular matrix $C$ which controls the shape of the ellipsoid, and some scalar $\rho > 0$. Ellipsoidal worst-case solutions are computed by shifting matrix $A$ by the rank-one update $A' = A + \gamma ee^T$ with $\gamma = \rho \min_{i,j \in V}[((C^T)^{-1})_{ij}$ such that $A' \geq 0$, and adapt scalar $\rho$ such that $\|\delta L\|_1 = \|A' - A\|_1$. In other words, we compare uncertainty sets based on the changes in absolute values they impose on the deterministic data matrix $A$ and take scalar $\delta$ (which controls the size of the conic uncertainty set) as reference value. Different values for $\rho$ are obtained according to the 21 values considered for $\delta$; see above. A corresponding rank-one downgrade is considered for a fair comparison of objective function values. We consider two different ellipsoidal uncertainty sets, one defined by $C = I$ that corresponds to a ball, and another one such that $(C^T)^{-1} = D = \text{Diag}^r(\tilde{r})$. The latter set is defined by a vector with coordinates $\tilde{r}_i \in [1/2, 3/2]$ chosen independently at random across all $i \in V$. Note that $D$ is nonsingular by construction, hence, diagonal matrices $CC^T$ and also $C$ exist with entries $C_{ii} = \sqrt{1/\tilde{r}_i}$ for all $i \in V$.

Aggregated results are given in Figure 2, and we again remark that the points $(0,1)$ correspond to the average values across all deterministic instances; see Table 1. Figure 2a depicts averages over the mean objective function values of all different solutions that have been identified by 1000 multi-starts of INMdyn. We observe that the worst-case results for ellipsoidal uncertainty sets results on average in larger objective function values than for conic and ball uncertainty sets. From Figure 2b we observe that the number of different solutions found typically decreases with increasing size of an uncertainty set when considering ellipsoidal uncertainty. The opposite trend can be observed from Figure 2c with respect to the average cardinality of found dominant sets. This observation may stem from the fact that ellipsoidal uncertainty sets contribute to flatten out the objective function and, hence, tend to increase the similarity of vertices (until the whole graph is one dominant set). We also observe, however, that the opposite is true for conic uncertainty sets for which increasing the size of the uncertainty set results in a larger number of identified dominant sets of lower cardinality.

### 6.5 Community detection in social networks.

This section demonstrates how (robust) dominant set clustering can be used to identify communities in social networks based among their members. In particular we discuss the instance Newman Lesmis in which vertices correspond to the characters of Victor Hugo’s novel Les Misérables, while each edge weight corresponds to the number of simultaneous appearances of two characters in a chapter. The modifications described in Section 6.1 have been applied to the original instance. The computational results which are visualized in Figure 3 confirm that dominant sets correspond to communities formed by similar characters. The community (1) of Figures 3a-3d corresponds, for example, to the rebels in the novel. We also observe that the cardinality of dominants sets decreases for conic uncertainty sets and increases for ball and ellipsoidal uncertainty when compared to the deterministic case. The reverse effect can be observed regarding the number of identified different dominant sets ($\#\sigma(x)$). Finally, observe that all dominant sets form cliques
Figure 2: (a) Mean objective function values; (b) number of different solutions found; (c) corresponding dominant set cardinalities, relative to the deterministic solutions for each instance with 1000 multi-starts of INIMdyn. Results are averaged across all instances.

(cf. Proposition 9) in the original graph except one in Figure 3d, which is, however, connected (cf. Proposition 8) as discussed in Section 5.2 and forms a clique in the perturbed instance.

7 Conclusions

In this article, we have introduced the Robust Standard Quadratic Optimization Problem and have shown that its CP relaxation gap is equal to the minimax gap. For several, frequently considered uncertainty sets we also have shown that this CP relaxation is an exact reformulation and that Sion’s minimax theorem holds without the usual convexity assumptions. A direct consequence of these results is that the RStQP reduces to an StQP in such cases which therefore retains the computational complexity of the deterministic problem variant. We have also investigated the robust variant of the dominant set clustering problem. For its deterministic case, we have shown that dominant sets form cliques in the similarity graph. The results of our computational study indicate that considering ellipsoidal uncertainty sets tend to identify a smaller number of dominant sets but with higher cardinality and objective function values compared to the results of considering conic uncertainty sets. Promising research directions for future work include the consideration of less conservative robust optimization approaches such as, e.g., budgeted uncertainty or distributionally robust optimization.

Acknowledgments.

The authors thank Juan Vera Lizcano for pointing out a generalization of our initial argument in Theorem 4.
Figure 3: (Robust) dominant sets (indicated by -) before the character names) of the instance Newman_lesmis: (a) deterministic case, (b) conic uncertainty, (c) ball uncertainty, and (d) ellipsoidal uncertainty.

References


