Math. Program., Ser. B 107, 5–36 (2006)

Digital Object Identifier (DOI) 10.1007/s10107-005-0677-1

Dimitris Bertsimas · Melvyn Sim

Tractable Approximations to Robust Conic Optimization Problems

Received: April 6, 2004 / Accepted: July 7, 2005
Published online: December 30, 2005 –

Abstract. In earlier proposals, the robust counterpart of conic optimization problems exhibits a lateral increase in complexity, i.e., robust linear programming problems (LPs) become second order cone problems (SOCPs), robust SOCPs become semidefinite programming problems (SDPs), and robust SDPs become NP-hard. We propose a relaxed robust counterpart for general conic optimization problems that **(a)** preserves the computational tractability of the nominal problem; specifically the robust conic optimization problem retains its original structure, i.e., robust LPs remain LPs, robust SOCPs remain SOCPs and robust SDPs remain SDPs, and **(b)** allows us to provide a guarantee on the probability that the robust solution is feasible when the uncertain coefficients obey independent and identically distributed normal distributions.

Key words. Robust Optimization – Conic Optimization – Stochastic Optimization

1. Introduction

The general optimization problem under parameter uncertainty is as follows:

$$\begin{aligned} \max\ & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.}\ & f_i(\boldsymbol{x}, \tilde{\boldsymbol{D}}_i) \geq 0, \qquad i \in I, \\ & \boldsymbol{x} \in X, \end{aligned} \tag{1}$$

where $f_i(\boldsymbol{x}, \tilde{\boldsymbol{D}}_i), i \in I$ are given functions, X is a given set and $\tilde{\boldsymbol{D}}_i,\ i \in I$ is the vector of random coefficients. Without loss of generality, we can move the objective function to the constraints and hence, assume that the objective is linear with deterministic coefficients.

We define the nominal problem to be Problem (1) when the random coefficients $\tilde{\boldsymbol{D}}_i$ take values equal to their expected values $\boldsymbol{D}_i^0$. In order to protect the solution against infeasibility of Problem (1), we may formulate the problem using chance constraints as follows:

$$\begin{aligned} \max\ & \boldsymbol{c}'\boldsymbol{x} \\ \text{s.t.}\ & \mathrm{P}(f_i(\boldsymbol{x}, \tilde{\boldsymbol{D}}_i) \geq 0) \geq 1 - \epsilon_i, \qquad i \in I, \\ & \boldsymbol{x} \in X. \end{aligned} \tag{2}$$

D. Bertsimas: Boeing Professor of Operations Research, Sloan School of Management and Operations Research Center, Massachusetts Institute of Technology, E53-363, Cambridge, MA 02139, USA. The research of the author was partially supported by the Singapore-MIT alliance. e-mail: dbertsim@mit.edu

M. Sim: NUS Business School, National University of Singapore, Singapore. The research of the author is supported by NUS academic research grant R-314-000-066-122 and the Singapore-MIT alliance. e-mail: dscsimm@nus.edu.sg

Unfortunately, it is well known that such chance constraints are non-convex and generally intractable. However, we would like to solve a tractable problem and obtain a "robust" solution that is feasible to the chance constraints Problem (2) when ϵ_i is very small and without having to reduce the objective function excessively. In order to address Problem (1) Ben-Tal and Nemirovski [1, 3] and independently by El Ghaoui et al. [12, 13] propose to solve the following robust optimization problem

$$\begin{aligned} &\max\ \boldsymbol{c}'\boldsymbol{x} \\ &\text{s.t.}\ \min_{\boldsymbol{D}_i \in \mathcal{U}_i} f_i(\boldsymbol{x}, \boldsymbol{D}_i) \geq 0, \qquad i \in I, \\ &\qquad \boldsymbol{x} \in X, \end{aligned} \tag{3}$$

where $\mathcal{U}_i$, $i \in I$ are given uncertainty sets. The motivation for solving Problem (3) is to find a solution $\boldsymbol{x}^* \in X$ that "immunizes" Problem (1) against parameter uncertainty. That is, by selecting appropriate sets $\mathcal{U}_i$, $i \in I$, we can find solutions $\boldsymbol{x}^*$ to Problem (3) that give guarantees ϵ_i in Problem (2). However, this is done at the expense of decreasing the achievable objective. It is important to note that we describe uncertainty in Problem (3) (using the sets $\mathcal{U}_i$, $i \in I$) in a deterministic manner. In selecting uncertainty sets $\mathcal{U}_i$, $i \in I$ we feel that two criteria are important:

(a) Preserving the computational tractability both theoretically and most importantly practically of the nominal problem. From a theoretical perspective it is desirable that if the nominal problem is solvable in polynomial time, then the robust problem is also polynomially solvable. More specifically, it is desirable that robust conic optimization problems retain their original structure, i.e., robust linear programming problems (LPs) remain LPs, robust second order cone problems (SOCPs) remain SOCPs and robust semidefinite programming problems (SDPs) remain SDPs.
(b) Being able to find a guarantee on the probability that the robust solution is feasible, when the uncertain coefficients obey some natural probability distributions.

Let us examine whether the state of the art in robust optimization has the two properties mentioned above:

1. **Linear Programming:** A uncertain LP constraint is of the form $\tilde{\boldsymbol{a}}'\boldsymbol{x} \geq \tilde{b}$, for which $\tilde{\boldsymbol{a}}$ and $\tilde{b}$ are subject to uncertainty. When the corresponding uncertainty set $\mathcal{U}$ is a polyhedron, then the robust counterpart is also an LP (see Ben-Tal and Nemirovski [3, 4] and Bertsimas and Sim [9, 10]). When $\mathcal{U}$ is ellipsoidal, then the robust counterpart becomes an SOCP. For linear programming there are probabilistic guarantees for feasibility available ([3, 4] and [9, 10]) under reasonable probabilistic assumptions on data variation.
2. **Quadratic Constrained Quadratic Programming (QCQP):** An uncertain QCQP constraint is of the form $\|\tilde{\boldsymbol{A}}\boldsymbol{x}\|_2^2 + \tilde{\boldsymbol{b}}'\boldsymbol{x} + \tilde{c} \leq 0$, where $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{b}}$ and $\tilde{c}$ are subject to data uncertainty. The robust counterpart is an SDP if the uncertainty set is a simple ellipsoid, and NP-hard if the set is polyhedral (Ben-Tal and Nemirovski [1, 3]). To the best of our knowledge, there are no available probabilistic bounds.
3. **Second Order Cone Programming (SOCP):** An uncertain SOCP constraint is of the form $\|\tilde{\boldsymbol{A}}\boldsymbol{x} + \tilde{\boldsymbol{b}}\|_2 \leq \tilde{\boldsymbol{c}}'\boldsymbol{x} + \tilde{d}$, where $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{b}}$, $\tilde{\boldsymbol{c}}$ and $\tilde{d}$ are subject to data uncertainty. The robust counterpart is an SDP if $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{b}}$ belong in an ellipsoidal uncertainty set $\mathcal{U}_1$

and $\tilde{\boldsymbol{c}}$, $\tilde{d}$ belong in another ellipsoidal set $\mathcal{U}_2$. The problem has unknown complexity, however, if $\tilde{\boldsymbol{A}}$, $\tilde{\boldsymbol{b}}$, $\tilde{\boldsymbol{c}}$, $\tilde{d}$ vary together in a common ellipsoidal set. Nemirovski [15] proposed tractable approximation in the form of an SDP if $\boldsymbol{c}$ and d are deterministic and showed probabilistic guarantees in this case. However, to the best of our knowledge, there are no available probability bounds to address the problem when $\boldsymbol{c}$ and d are stochastic.

4. **Semidefinite Programming (SDP):** An uncertain SDP constraint of the form $\sum_{j=1}^{n} \tilde{\boldsymbol{A}}_j x_j \succeq \tilde{\boldsymbol{B}}$, where $\tilde{\boldsymbol{A}}_1, ..., \tilde{\boldsymbol{A}}_n$ and $\tilde{\boldsymbol{B}}$ are subject to data uncertainty. The robust counterpart is NP-hard for ellipsoidal uncertainty sets. Nemirovski [15] proposed a tractable approximation in the form of an SDP and showed probabilistic guarantees in this case.
5. **Conic Programming:** An uncertain Conic Programming constraint of the form $\sum_{j=1}^{n} \tilde{\boldsymbol{A}}_j x_j \succeq_{\boldsymbol{K}} \tilde{\boldsymbol{B}}$, where $\tilde{\boldsymbol{A}}_1, ..., \tilde{\boldsymbol{A}}_n$ and $\tilde{\boldsymbol{B}}$ are subject to data uncertainty. The cone $\boldsymbol{K}$ is closed, pointed and with a nonempty interior. To the best of our knowledge, there are no results available regarding tractability and probabilistic guarantees in this case.

Our goal in this paper is to address (a) and (b) above for robust conic optimization problems. Specifically, we propose a new robust counterpart of Problem (1) that has the following properties: (a) It inherits the character of the nominal problem; for example, robust SOCPs remain SOCPs and robust SDPs remain SDPs; (b) under reasonable probabilistic assumptions on data variation we establish probabilistic guarantees for feasibility that lead to explicit ways for selecting parameters that control the robustness; (c) It is applicable for general conic problems.

The structure of the paper is as follows. In Section 2, we describe the proposed robust model and in Section 3, we show that the robust model inherits the character of the nominal problem for LPs, QCQPs, SOCPs and SDPs. In Section 4, we prove probabilistic guarantees for feasibility for these classes of problems. In Section 5, we show tractability and give explicit probabilistic bounds for general conic problems. Section 6 concludes the paper.

2. The Robust model

In this section, we outline the ingredients of the proposed framework for robust conic optimization.

2.1. Model for parameter uncertainty

The model of data uncertainty we consider is

$$\tilde{\boldsymbol{D}} = \boldsymbol{D}^0 + \sum_{j \in N} \Delta \boldsymbol{D}^j \tilde{z}_j, \tag{4}$$

where $\boldsymbol{D}^0$ is the nominal value of the data, $\Delta \boldsymbol{D}^j$, $j \in N$ is a direction of data perturbation, and $\tilde{z}_j$, $j \in N$ are independent and identically distributed random variables with mean equal to zero, so that $E[\tilde{\boldsymbol{D}}] = \boldsymbol{D}^0$. The cardinality of N may be small, modeling situations involving a small collection of primitive independent uncertainties (for example a factor model in a finance context), or large, potentially as large as the number of entries in the data. In the former case, the elements of $\tilde{\boldsymbol{D}}$ are strongly dependent, while in the latter case the elements of $\tilde{\boldsymbol{D}}$ are weakly dependent or even independent (when $|N|$ is equal to the number of entries in the data). The support of $\tilde{z}_j,\ j \in N$ can be unbounded or bounded. Ben-Tal and Nemirovskii [4] and Bertsimas and Sim [9] have considered the case that $|N|$ is equal to the number of entries in the data.

2.2. *Uncertainty sets and related norms*

In the robust optimization framework of (3), we consider the uncertainty set $\mathcal{U}$ as follows:

$$\mathcal{U} = \left\{ \boldsymbol{D} \mid \exists \boldsymbol{u} \in \Re^{|N|} : \boldsymbol{D} = \boldsymbol{D}^0 + \sum_{j \in N} \boldsymbol{\Delta D}^j u_j,\ \|\boldsymbol{u}\| \le \Omega \right\}, \tag{5}$$

where Ω is a parameter, which we will show, is related to the probabilistic guarantee against infeasibility. We restrict the vector norm $\|.\|$ we consider by imposing the condition:

$$\|\boldsymbol{u}\| = \||\boldsymbol{u}|\|, \tag{6}$$

where $|\boldsymbol{u}| = (|u_1|, \ldots, |u_{|N|}|)'$ if $\boldsymbol{u} = (u_1, \ldots, u_{|N|})'$. We call such norm the *absolute norm*. The following norms commonly used in robust optimization are absolute norms :

- The polynomial norm l_k, $k = 1, \ldots, \infty$ (see [1, 4, 18]).
- The $l_2 \cap l_\infty$ norm: $\max\{\|\boldsymbol{u}\|_2, \Omega\|\boldsymbol{u}\|_\infty\}$, $\Omega > 0$ (see [4]). This norm is used in modeling bounded and symmetrically distributed random data.
- The $l_1 \cap l_\infty$ norm: $\max\{\frac{1}{\Gamma}\|\boldsymbol{u}\|_1, \|\boldsymbol{u}\|_\infty\}$, $\Gamma > 0$ (see [9, 8]). Note that this norm is equal to l_∞ if $\Gamma = |N|$, and l_1 if $\Gamma = 1$. This norm is used in modeling bounded and symmetrically distributed random data, and has the additional property that the robust counterpart of an LP is still an LP (Bertsimas et al. [8]).

Note that the norm $\|\boldsymbol{u}\| = \|\boldsymbol{P}\boldsymbol{u}\|_k$, where $\boldsymbol{P}$ is an invertible matrix, is not an absolute norm. However, we can let $\boldsymbol{u} = \boldsymbol{P}^{-1}\boldsymbol{v}$, and modify the uncertainty set of (5) accordingly so that the norm considered remains absolute.

Given a norm $\|.\|$ we consider the dual norm $\|.\|^*$ defined as

$$\|\boldsymbol{s}\|^* = \max_{\|\boldsymbol{x}\| \le 1} \boldsymbol{s}'\boldsymbol{x}.$$

We next show some basic properties of norms satisfying Eq. (6), which we will subsequently use in our development.

Proposition 1. *The absolute norm $\|\cdot\|$ satisfies the following*

(a) $\|\boldsymbol{w}\|^* = \||\boldsymbol{w}|\|^*$.
(b) *For all $\boldsymbol{v}, \boldsymbol{w}$ such that $|\boldsymbol{v}| \le |\boldsymbol{w}|$, $\|\boldsymbol{v}\|^* \le \|\boldsymbol{w}\|^*$.*
(c) *For all $\boldsymbol{v}, \boldsymbol{w}$ such that $|\boldsymbol{v}| \le |\boldsymbol{w}|$, $\|\boldsymbol{v}\| \le \|\boldsymbol{w}\|$.*

The proof is shown in Appendix A.

2.3. The class of functions $f(\boldsymbol{x}, \boldsymbol{D})$

We impose the following restrictions on the class of functions $f(\boldsymbol{x}, \boldsymbol{D})$ in Problem (1) (we drop index i for clarity):

Assumption 1. *The function $f(\boldsymbol{x}, \boldsymbol{D})$ satisfies:*

(a) *The function $f(\boldsymbol{x}, \boldsymbol{D})$ is concave in $\boldsymbol{D}$ for all $\boldsymbol{x} \in \Re^n$.*
(b) $f(\boldsymbol{x}, k\boldsymbol{D}) = kf(\boldsymbol{x}, \boldsymbol{D})$*, for all $k \geq 0$, $\boldsymbol{D}$, $\boldsymbol{x} \in \Re^n$.*

Note that for functions $f(\cdot, \cdot)$ satisfying Assumption 1 we have:

$$f(\boldsymbol{x}, \boldsymbol{A} + \boldsymbol{B}) \geq \frac{1}{2} f(\boldsymbol{x}, 2\boldsymbol{A}) + \frac{1}{2} f(\boldsymbol{x}, 2\boldsymbol{B}) = f(\boldsymbol{x}, \boldsymbol{A}) + f(\boldsymbol{x}, \boldsymbol{B}). \tag{7}$$

The restrictions implied by Assumption 1 still allow us to model LPs, QCQPs, SOCPs and SDPs. Table 1 shows the function $f(\boldsymbol{x}, \boldsymbol{D})$ for these problems. Note that SOCP(1) models situations that only $\boldsymbol{A}$ and $\boldsymbol{b}$ vary, while SOCP(2) models situations that $\boldsymbol{A}$, $\boldsymbol{b}$, $\boldsymbol{c}$ and d vary. Note that for QCQP, the function, $-\|\boldsymbol{A}\boldsymbol{x}\|_2^2 - \boldsymbol{b}'\boldsymbol{x} - c$ does not satisfy the second assumption. However, by extending the dimension of the problem, it is well-known that the QCQP constraint is SOCP constraint representable (see [5]). Finally, the SDP constraint,

$$\sum_{j=1}^{n} \boldsymbol{A}_i x_i \succeq \boldsymbol{B},$$

is equivalent to

$$\lambda_{min}\left(\sum_{j=1}^{n} \boldsymbol{A}_i x_i - \boldsymbol{B}\right) \geq 0,$$

where $\lambda_{min}(\boldsymbol{M})$ is the function that returns the smallest eigenvalue of the symmetric matric $\boldsymbol{M}$.

Table 1. The function $f(\boldsymbol{x}, \boldsymbol{D})$ for different conic optimization problems

Type	Constraint	$\boldsymbol{D}$	$f(\boldsymbol{x}, \boldsymbol{D})$
LP	$\boldsymbol{a}'\boldsymbol{x} \geq b$	$(\boldsymbol{a}, b)$	$\boldsymbol{a}'\boldsymbol{x} - b$
QCQP	$\|\boldsymbol{A}\boldsymbol{x}\|_2^2 + \boldsymbol{b}'\boldsymbol{x}$ $+c \leq 0$	$(\boldsymbol{A}, \boldsymbol{b}, c, d)$ $d^0 = 1$, $\Delta d^j = 0$, $\forall j \in N$	$\frac{d-(\boldsymbol{b}'\boldsymbol{x}+c)}{2}$ $-\sqrt{\|\boldsymbol{A}\boldsymbol{x}\|_2^2 + \left(\frac{d+\boldsymbol{b}'\boldsymbol{x}+c}{2}\right)^2}$
SOCP(1)	$\|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\|_2$ $\leq \boldsymbol{c}'\boldsymbol{x} + d$	$(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}, d)$ $\boldsymbol{\Delta c}^j = \boldsymbol{0}$, $\Delta d^j = 0$, $\forall j \in N$	$\boldsymbol{c}'\boldsymbol{x} + d - \|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\|_2$
SOCP(2)	$\|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\|_2$ $\leq \boldsymbol{c}'\boldsymbol{x} + d$	$(\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}, d)$	$\boldsymbol{c}'\boldsymbol{x} + d - \|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\|_2$
SDP	$\sum_{j=1}^{n} \boldsymbol{A}_i x_i - \boldsymbol{B}$ $\in \boldsymbol{S}_+^m$	$(\boldsymbol{A}_1, ..., \boldsymbol{A}_n, \boldsymbol{B})$	$\lambda_{min}(\sum_{j=1}^{n} \boldsymbol{A}_i x_i - \boldsymbol{B})$

3. The proposed robust framework and its tractability

Specifically, under the model of data uncertainty in Eq. (4) we propose the following constraint for controling the feasibility of stochastic data uncertainty in the constraint $f(\boldsymbol{x}, \tilde{\boldsymbol{D}}) \geq 0$:

$$\min_{(\boldsymbol{v},\boldsymbol{w})\in\mathcal{V}} f(\boldsymbol{x}, \boldsymbol{D}^0) + \sum_{j\in N} \left\{ f(\boldsymbol{x}, \boldsymbol{\Delta D}^j)v_j + f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j)w_j \right\} \geq 0, \tag{8}$$

where

$$\mathcal{V} = \left\{ (\boldsymbol{v}, \boldsymbol{w}) \in \Re_+^{|N|\times|N|} \mid \|\boldsymbol{v} + \boldsymbol{w}\| \leq \Omega \right\}, \tag{9}$$

and the norm $\|.\|$ satisfies Eq. (6). We next show that under Assumption 1, Eq. (8) implies the classical definition of robustness:

$$f(\boldsymbol{x}, \boldsymbol{D}) \geq 0, \qquad \forall \boldsymbol{D} \in \mathcal{U}, \tag{10}$$

where $\mathcal{U}$ is defined in Eq. (5). Moreover, if the function $f(\boldsymbol{x}, \boldsymbol{D})$ is linear in $\boldsymbol{D}$, then Eq. (8) is equivalent to Eq. (10).

Proposition 2. *Suppose the given norm $\|.\|$ satisfies Eq. (6).*

(a) *If $f(\boldsymbol{x}, \boldsymbol{A} + \boldsymbol{B}) = f(\boldsymbol{x}, \boldsymbol{A}) + f(\boldsymbol{x}, \boldsymbol{B})$, then $\boldsymbol{x}$ satisfies (8) if and only if $\boldsymbol{x}$ satisfies (10).*
(b) *Under Assumption 1, if $\boldsymbol{x}$ is feasible in Problem (8), then $\boldsymbol{x}$ is feasible in Problem (10).*

Proof. (a) Under the linearity assumption, Eq. (8) is equivalent to:

$$f\left(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j\in N} \boldsymbol{\Delta D}^j (v_j - w_j)\right) \geq 0, \qquad \forall \|\boldsymbol{v} + \boldsymbol{w}\| \leq \Omega, \quad \boldsymbol{v}, \boldsymbol{w} \geq \boldsymbol{0}, \tag{11}$$

while Eq. (10) can be written as:

$$f\left(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j\in N} \boldsymbol{\Delta D}^j r_j\right) \geq 0, \qquad \forall \|\boldsymbol{r}\| \leq \Omega. \tag{12}$$

Suppose $\boldsymbol{x}$ is infeasible in (12), that is, there exists $\boldsymbol{r}$, $\|\boldsymbol{r}\| \leq \Omega$ such that

$$f\left(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j\in N} \boldsymbol{\Delta D}^j r_j\right) < 0.$$

For all $j \in N$, let $v_j = \max\{r_j, 0\}$ and $w_j = -\min\{r_j, 0\}$. Clearly, $\boldsymbol{r} = \boldsymbol{v} - \boldsymbol{w}$ and since $v_j + w_j = |r_j|$, we have from Eq. (6) that $\|\boldsymbol{v} + \boldsymbol{w}\| = \|\boldsymbol{r}\| \leq \Omega$. Hence, $\boldsymbol{x}$ is infeasible in (11) as well.

Conversely, suppose $\boldsymbol{x}$ is infeasible in (11), then there exist $\boldsymbol{v}, \boldsymbol{w} \geq \boldsymbol{0}$ and $\|\boldsymbol{v}+\boldsymbol{w}\| \leq \Omega$ such that

$$f\left(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j\in N} \boldsymbol{\Delta D}^j (v_j - w_j)\right) < 0.$$

For all $j \in N$, we let $r_j = v_j - w_j$ and we observe that $|r_j| \leq v_j + w_j$. Therefore, for norms satisfying Eq. (6) we have

$$\|\boldsymbol{r}\| = \||\boldsymbol{r}|\| \leq \|\boldsymbol{v} + \boldsymbol{w}\| \leq \Omega,$$

and hence, $\boldsymbol{x}$ is infeasible in (12).

(b) Suppose $\boldsymbol{x}$ is feasible in Problem (8), i.e.,

$$\begin{gathered} f(\boldsymbol{x}, \boldsymbol{D}^0) + \textstyle\sum_{j\in N} \left\{ f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) v_j + f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) w_j \right\} \geq 0, \\ \forall \|\boldsymbol{v} + \boldsymbol{w}\| \leq \Omega, \quad \boldsymbol{v}, \boldsymbol{w} \geq \boldsymbol{0}. \end{gathered}$$

From Eq. (7) and Assumption 1(b)

$$\begin{aligned} 0 &\leq f(\boldsymbol{x}, \boldsymbol{D}^0) + \textstyle\sum_{j\in N} \left\{ f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) v_j + f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) w_j \right\} \\ &\leq f(\boldsymbol{x}, \boldsymbol{D}^0 + \textstyle\sum_{j\in N} \boldsymbol{\Delta D}^j (v_j - w_j)) \end{aligned}$$

for all $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \Omega, \ \ \boldsymbol{v}, \boldsymbol{w} \geq \boldsymbol{0}$. In the proof of part (a) we established that

$$f(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j\in N} \boldsymbol{\Delta D}^j r_j) \geq 0, \qquad \forall \|\boldsymbol{r}\| \leq \Omega$$

is equivalent to

$$f(\boldsymbol{x}, \boldsymbol{D}^0 + \sum_{j\in N} \boldsymbol{\Delta D}^j (v_j - w_j)) \geq 0, \qquad \forall \|\boldsymbol{v} + \boldsymbol{w}\| \leq \Omega, \ \ \boldsymbol{v}, \boldsymbol{w} \geq \boldsymbol{0},$$

and thus $\boldsymbol{x}$ satisfies (10). □

Note that there are other proposals that relax the classical definition of robustness (10) (see for instance Ben-Tal and Nemirovski [2]) and lead to tractable solutions. One natural question is whether the approximation is overly conservative with respect to Problem (10). A way to address this is to show that if $\boldsymbol{x}$ is feasible in Problem (10), it also feasible in Problem (8) in which Ω is reduced to $\sigma\Omega$, $\sigma < 1$. Ideally, σ should not decrease too rapidly with respect to the dimension of the problem. While we do not have theoretical evidence on the closeness of the approximation, Bertsimas and Brown [7] report excellent computational results utilizing Problem (8) for constrained stochastic linear control problems, that is the solutions obtained when solving problem (8) are very close to the solutions obtained when solving Problem (10).

3.1. Tractability of the proposed framework

Unlike the classical definition of robustness (10), which can not be represented in a tractable manner, we next show that Eq. (8) can be represented in a tractable manner.

Theorem 1. *For a norm satisfying Eq. (6) and a function $f(\boldsymbol{x}, \boldsymbol{D})$ satisfying Assumption 1*

(a) *Constraint (8) is equivalent to*

$$f(\boldsymbol{x}, \boldsymbol{D}^0) \geq \Omega \|\boldsymbol{s}\|^*, \tag{13}$$

where

$$s_j = \max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}^j), -f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j)\}, \quad \forall j \in N.$$

(b) *Eq. (13) can be written as:*

$$\begin{aligned}
&f(\boldsymbol{x}, \boldsymbol{D}^0) \geq \Omega y \\
&f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) + t_j \geq 0, \quad \forall j \in N \\
&f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) + t_j \geq 0, \forall j \in N \\
&\|\boldsymbol{t}\|^* \leq y \\
&y \in \Re, \ \boldsymbol{t} \in \Re^{|N|}.
\end{aligned} \tag{14}$$

Proof. (a) We introduce the following problems:

$$\begin{aligned}
z_1 = \max \ & \boldsymbol{a}'\boldsymbol{v} + \boldsymbol{b}'\boldsymbol{w} \\
\text{s.t.} \ & \|\boldsymbol{v} + \boldsymbol{w}\| \leq \Omega \\
& \boldsymbol{v}, \boldsymbol{w} \geq \boldsymbol{0},
\end{aligned} \tag{15}$$

and

$$\begin{aligned}
z_2 = \max \ & \sum_{j \in N} \max\{a_j, b_j, 0\} r_j \\
\text{s.t.} \ & \|\boldsymbol{r}\| \leq \Omega,
\end{aligned} \tag{16}$$

and show that $z_1 = z_2$. Suppose $\boldsymbol{r}^*$ is an optimal solution to (16). For all $j \in N$, let

$$\begin{aligned}
&v_j = w_j = 0 \qquad \text{if } \max\{a_j, b_j\} \leq 0 \\
&v_j = |r_j^*|, w_j = 0 \text{ if } a_j \geq b_j, a_j > 0 \\
&w_j = |r_j^*|, v_j = 0 \text{ if } b_j > a_j, b_j > 0.
\end{aligned}$$

Observe that $a_j v_j + b_j w_j \geq \max\{a_j, b_j, 0\} r_j^*$ and $w_j + v_j \leq |r_j^*|, \ \forall j \in N$. From Proposition 1(c) we have $\|\boldsymbol{v} + \boldsymbol{w}\| \leq \|\boldsymbol{r}^*\| \leq \Omega$, and thus $\boldsymbol{v}, \boldsymbol{w}$ are feasible in Problem (15), leading to

$$z_1 \geq \sum_{j \in N} (a_j v_j + b_j w_j) \geq \sum_{j \in N} \max\{a_j, b_j, 0\} r_j^* = z_2.$$

Conversely, let $\boldsymbol{v}^*$, $\boldsymbol{w}^*$ be an optimal solution to Problem (15). Let $\boldsymbol{r} = \boldsymbol{v}^* + \boldsymbol{w}^*$. Clearly $\|\boldsymbol{r}\| \leq \Omega$ and observe that

$$r_j \max\{a_j, b_j, 0\} \geq a_j v_j^* + b_j w_j^*, \quad \forall j \in N.$$

Therefore, we have

$$z_2 \geq \sum_{j \in N} \max\{a_j, b_j, 0\} r_j \geq \sum_{j \in N} (a_j v_j^* + b_j w_j^*) = z_1,$$

leading to $z_1 = z_2$. We next observe that

$$\begin{aligned}
&\min_{(\boldsymbol{v},\boldsymbol{w}) \in \mathcal{V}} \sum_{j \in N} \left\{ f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) v_j + f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) w_j \right\} \\
&= -\max_{(\boldsymbol{v},\boldsymbol{w}) \in \mathcal{V}} \sum_{j \in N} \left\{ -f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) v_j - f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) w_j \right\} \\
&= -\max_{\{\|\boldsymbol{r}\| \leq \Omega\}} \sum_{j \in N} \left\{ \max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}^j), -f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j), 0\} r_j \right\}
\end{aligned}$$

and using the definition of dual norm, $\|\boldsymbol{s}\|^* = \max_{\|x\| \leq 1} \boldsymbol{s}'\boldsymbol{x}$, we obtain $\Omega\|\boldsymbol{s}\|^* = \max_{\|x\| \leq \Omega} \boldsymbol{s}'\boldsymbol{x}$, i.e., Eq. (13) follows. Note that

$$s_j = \max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}^j), -f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j)\} \geq 0,$$

since otherwise there exists an $\boldsymbol{x}$ such that $s_j < 0$, i.e., $f(\boldsymbol{x}, \boldsymbol{\Delta D}^j) > 0$ and $f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j) > 0$. From Assumption 1(b) $f(\boldsymbol{x}, \boldsymbol{0}) = 0$, contradicting the concavity of $f(\boldsymbol{x}, \boldsymbol{D})$ (Assumption 1(a)).

Suppose that $\boldsymbol{x}$ is feasible in Problem (13). Defining $\boldsymbol{t} = \boldsymbol{s}$ and $y = \|\boldsymbol{s}\|^*$, we can easily check that $(\boldsymbol{x}, \boldsymbol{t}, y)$ are feasible in Problem (14). Conversely, suppose, $\boldsymbol{x}$ is infeasible in (13), that is,

$$f(\boldsymbol{x}, \boldsymbol{D}^0) < \Omega\|\boldsymbol{s}\|^*.$$

Since, $t_j \geq s_j = \max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}^j), -f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j)\} \geq 0$ we apply Proposition 1(b) to obtain $\|\boldsymbol{t}\|^* \geq \|\boldsymbol{s}\|^*$. Thus,

$$f(\boldsymbol{x}, \boldsymbol{D}^0) < \Omega\|\boldsymbol{s}\|^* \leq \Omega\|\boldsymbol{t}\|^* \leq \Omega y,$$

i.e., $\boldsymbol{x}$ is infeasible in (14).

(b) It is immediate that Eq. (13) can be written in the form of Eq. (14). □

In Table 2, we list the common choices of norms, the representation of their dual norms and the corresponding references.

Table 2. Representation of the dual norm for $\boldsymbol{t} \geq \mathbf{0}$

Norms	$\|\boldsymbol{u}\|$	$\|\boldsymbol{t}\|^* \leq y$	Ref.
l_2	$\|\boldsymbol{u}\|_2$	$\|\boldsymbol{t}\|_2 \leq y$	[4]
l_1	$\|\boldsymbol{u}\|_1$	$t_j \leq y,\ \forall j \in N$	[8]
l_∞	$\|\boldsymbol{u}\|_\infty$	$\sum_{j\in N} t_j \leq y$	[8]
$l_p,\ p \geq 1$	$\|\boldsymbol{u}\|_p$	$\left(\sum_{j\in N} t_j^{\frac{q}{q-1}}\right)^{\frac{q-1}{q}} \leq y$	[8]
$l_2 \cap l_\infty$	$\max\{\|\boldsymbol{u}\|_2, \Omega\|\boldsymbol{u}\|_\infty\}$	$\|\boldsymbol{s}-\boldsymbol{t}\|_2 + \frac{1}{\Omega}\sum_{j\in N} s_j \leq y$ $\boldsymbol{s} \in \Re_+^{\|N\|}$	[4]
$l_1 \cap l_\infty$	$\max\{\frac{1}{\Gamma}\|\boldsymbol{u}\|_1, \|\boldsymbol{u}\|_\infty\}$	$\Gamma p + \sum_{j\in N} s_j \leq y$ $s_j + p \geq t_j,\ \forall j \in N$ $p \in \Re_+, \boldsymbol{s} \in \Re_+^{\|N\|}$	[8]

3.2. *Representation of the function* $\max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}), -f(\boldsymbol{x}, -\boldsymbol{\Delta D})\}$

The function

$$g(\boldsymbol{x}, \boldsymbol{\Delta D}^j) = \max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}^j), -f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j)\}$$

naturally arises in Theorem 1. Recall that a norm satisfies $\|\boldsymbol{A}\| \geq 0$, $\|k\boldsymbol{A}\| = |k| \cdot \|\boldsymbol{A}\|$, $\|\boldsymbol{A}+\boldsymbol{B}\| \leq \|\boldsymbol{A}\| + \|\boldsymbol{B}\|$, and $\|\boldsymbol{A}\| = 0$, implies that $\boldsymbol{A} = \mathbf{0}$. We show next that the function $g(\boldsymbol{x}, \boldsymbol{A})$ satisfies all these properties except the last one, i.e., it behaves almost like a norm.

Proposition 3. *Under Assumption 1, the function*

$$g(\boldsymbol{x}, \boldsymbol{A}) = \max\{-f(\boldsymbol{x}, \boldsymbol{A}), -f(\boldsymbol{x}, -\boldsymbol{A})\}$$

satisfies the following properties:

(a) $g(\boldsymbol{x}, \boldsymbol{A}) \geq 0$,
(b) $g(\boldsymbol{x}, k\boldsymbol{A}) = |k| g(\boldsymbol{x}, \boldsymbol{A})$,
(c) $g(\boldsymbol{x}, \boldsymbol{A}+\boldsymbol{B}) \leq g(\boldsymbol{x}, \boldsymbol{A}) + g(\boldsymbol{x}, \boldsymbol{B})$.

Proof. (a) Suppose there exists $\boldsymbol{x}$ such that $g(\boldsymbol{x}, \boldsymbol{A}) < 0$, i.e., $f(\boldsymbol{x}, \boldsymbol{A}) > 0$ and $f(\boldsymbol{x}, -\boldsymbol{A}) > 0$. From Assumption 1(b) $f(\boldsymbol{x}, \mathbf{0}) = 0$, contradicting the concavity of $f(\boldsymbol{x}, \boldsymbol{A})$ (Assumption 1(a)).

(b) For $k \geq 0$, we apply Assumption 1(b) and obtain

$$\begin{aligned} g(\boldsymbol{x}, k\boldsymbol{A}) &= \max\{-f(\boldsymbol{x}, k\boldsymbol{A}), -f(\boldsymbol{x}, -k\boldsymbol{A})\} \\ &= k \max\{-f(\boldsymbol{x}, \boldsymbol{A}), -f(\boldsymbol{x}, -\boldsymbol{A})\} \\ &= kg(\boldsymbol{x}, \boldsymbol{A}). \end{aligned}$$

Similarly, if $k < 0$ we have

$$\begin{aligned} g(\boldsymbol{x}, k\boldsymbol{A}) &= \max\{-f(\boldsymbol{x}, -k(-\boldsymbol{A})), -f(\boldsymbol{x}, -k(\boldsymbol{A}))\} \\ &= -kg(\boldsymbol{x}, \boldsymbol{A}). \end{aligned}$$

(c) Using Eq. (7) we obtain

$$\begin{aligned} g(\boldsymbol{x}, \boldsymbol{A}+\boldsymbol{B}) &= g(\boldsymbol{x}, \tfrac{1}{2}(2\boldsymbol{A}+2\boldsymbol{B})) \\ &\le \tfrac{1}{2} g(\boldsymbol{x}, 2\boldsymbol{A}) + \tfrac{1}{2} g(\boldsymbol{x}, 2\boldsymbol{B}) \\ &= g(\boldsymbol{x}, \boldsymbol{A}) + g(\boldsymbol{x}, \boldsymbol{B}). \end{aligned}$$ □

Note that the function $g(\boldsymbol{x}, \boldsymbol{A})$ does not necessarily define a norm for $\boldsymbol{A}$, since $g(\boldsymbol{x}, \boldsymbol{A}) = 0$ does not necessarily imply $\boldsymbol{A} = \boldsymbol{0}$. However, for LP, QCQP. SOCP(1), SOCP(2) and SDP, and specific direction of data perturbation, $\boldsymbol{\Delta D}^j$, we can map $g(\boldsymbol{x}, \boldsymbol{\Delta D}^j)$ to a function of a norm such that

$$g(\boldsymbol{x}, \boldsymbol{\Delta D}^j) = \|\mathcal{H}(\boldsymbol{x}, \boldsymbol{\Delta D}^j)\|_g,$$

where $\mathcal{H}(\boldsymbol{x}, \boldsymbol{\Delta D}^j)$ is linear in $\boldsymbol{\Delta D}^j$ and defined as follows (see also the summary in Table 3):

(a) LP:
$f(\boldsymbol{x}, \boldsymbol{D}) = \boldsymbol{a}'\boldsymbol{x} - b$, where $\boldsymbol{D} = (\boldsymbol{a}, b)$ and $\boldsymbol{\Delta D}^j = (\boldsymbol{\Delta a}^j, \Delta b^j)$. Hence,

$$\begin{aligned} &g(\boldsymbol{x}, \boldsymbol{\Delta D}^j) \\ &\quad = \max\{-(\boldsymbol{\Delta a}^j)'\boldsymbol{x} + \Delta b^j, (\boldsymbol{\Delta a}^j)'\boldsymbol{x} - \Delta b^j\} \\ &\quad = |(\boldsymbol{\Delta a}^j)'\boldsymbol{x} - \Delta b^j|. \end{aligned}$$

(b) QCQP:
$f(\boldsymbol{x}, \boldsymbol{D}) = (d - (\boldsymbol{b}'\boldsymbol{x} + c))/2 - \sqrt{\|\boldsymbol{A}\boldsymbol{x}\|_2^2 + \big((d + \boldsymbol{b}'\boldsymbol{x} + c)/2\big)^2}$, where $\boldsymbol{D} = (\boldsymbol{A}, \boldsymbol{b}, c, d)$ and $\boldsymbol{\Delta D}^j = (\boldsymbol{\Delta A}^j, \boldsymbol{\Delta b}^j, \Delta c^j, 0)$. Therefore,

Table 3. The function $\mathcal{H}(\boldsymbol{x}, \boldsymbol{\Delta D}^j)$ and the norm $\|\cdot\|_g$ for different conic optimization problems

Type	$\boldsymbol{r} = \mathcal{H}(\boldsymbol{x}, \boldsymbol{\Delta D}^j)$	$g(\boldsymbol{x}, \boldsymbol{\Delta D}^j) = \|\boldsymbol{r}\|_g$
LP	$r = (\boldsymbol{\Delta a}^j)'\boldsymbol{x} - \Delta b^j$	$\|r\|$
QCQP	$\boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_1 \\ r_0 \end{bmatrix}$ $\boldsymbol{r}_1 = \begin{bmatrix} \boldsymbol{\Delta A}^j \boldsymbol{x} \\ ((\boldsymbol{\Delta b}^j)'\boldsymbol{x} + \Delta c^j)/2 \end{bmatrix}$ $r_0 = ((\boldsymbol{\Delta b}^j)'\boldsymbol{x} + \Delta c^j)/2$	$\|\boldsymbol{r}_1\|_2 + \|r_0\|$
SOCP(1)	$\boldsymbol{r} = \boldsymbol{\Delta A}^j \boldsymbol{x} + \boldsymbol{\Delta b}^j$	$\|\boldsymbol{r}\|_2$
SOCP(2)	$\boldsymbol{r} = \begin{bmatrix} \boldsymbol{r}_1 \\ r_0 \end{bmatrix}$ $\boldsymbol{r}_1 = \boldsymbol{\Delta A}^j \boldsymbol{x} + \boldsymbol{\Delta b}^j$ $r_0 = (\boldsymbol{\Delta c}^j)'\boldsymbol{x} + \Delta d^j$	$\|\boldsymbol{r}_1\|_2 + \|r_0\|$
SDP	$\boldsymbol{R} = \sum_{i=1}^n \boldsymbol{\Delta A}_i^j x_i - \boldsymbol{\Delta B}^j$	$\|\boldsymbol{R}\|_2$

$$g(\boldsymbol{x}, \boldsymbol{\Delta D}^j) = \max\left\{\frac{(\boldsymbol{\Delta b}^j)'\boldsymbol{x}+\Delta c^j}{2} + \sqrt{\|\boldsymbol{\Delta A}^j\boldsymbol{x}\|_2^2 + \left(\frac{(\boldsymbol{\Delta b}^j)'\boldsymbol{x}+\Delta c^j}{2}\right)^2},\right.$$
$$\left. -\frac{(\boldsymbol{\Delta b}^j)'\boldsymbol{x}+\Delta c^j}{2} + \sqrt{\|\boldsymbol{\Delta A}^j\boldsymbol{x}\|_2^2 + \left(\frac{(\boldsymbol{\Delta b}^j)'\boldsymbol{x}+\Delta c^j}{2}\right)^2}\right\}$$
$$= \left|\frac{(\boldsymbol{\Delta b}^j)'\boldsymbol{x}+\Delta c^j}{2}\right| + \sqrt{\|\boldsymbol{\Delta A}^j\boldsymbol{x}\|_2^2 + \left(\frac{(\boldsymbol{\Delta b}^j)'\boldsymbol{x}+\Delta c^j}{2}\right)^2}.$$

(c) SOCP(1):
$f(\boldsymbol{x}, \boldsymbol{D}) = \boldsymbol{c}'\boldsymbol{x}+d-\|\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}\|_2^2$, where $\boldsymbol{D} = (\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}, d)$ and $\boldsymbol{\Delta D}^j = (\boldsymbol{\Delta A}^j, \boldsymbol{\Delta b}^j, \boldsymbol{0}, 0)$. Therefore,

$$g(\boldsymbol{x}, \boldsymbol{\Delta D}^j) = \|\boldsymbol{\Delta A}^j\boldsymbol{x} + \boldsymbol{\Delta b}^j\|_2.$$

(d) SOCP(2):
$f(\boldsymbol{x}, \boldsymbol{D}) = \boldsymbol{c}'\boldsymbol{x} + d - \|\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\|_2^2$, where $\boldsymbol{D} = (\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}, d)$ and $\boldsymbol{\Delta D}^j = (\boldsymbol{\Delta A}^j, \boldsymbol{\Delta b}^j, \boldsymbol{\Delta c}^j, d^j)$. Therefore,

$$\begin{aligned} &g(\boldsymbol{x}, \boldsymbol{\Delta D}^j)\\ &= \max\left\{-(\boldsymbol{\Delta c}^j)'\boldsymbol{x} - \Delta d^j + \|\boldsymbol{\Delta A}^j\boldsymbol{x} + \boldsymbol{\Delta b}^j\|_2,\right.\\ &\qquad \left.(\boldsymbol{\Delta c}^j)'\boldsymbol{x} + \Delta d^j + \|\boldsymbol{\Delta A}^j\boldsymbol{x} + \boldsymbol{\Delta b}^j\|_2\right\}\\ &= |(\boldsymbol{\Delta c}^j)'\boldsymbol{x} + \Delta d^j| + \|\boldsymbol{\Delta A}^j\boldsymbol{x} + \boldsymbol{\Delta b}^j\|_2. \end{aligned}$$

(e) SDP:
$f(\boldsymbol{x}, \boldsymbol{D}) = \lambda_{min}(\sum_{j=1}^n \boldsymbol{A}_i x_i - \boldsymbol{B})$, where $\boldsymbol{D} = (\boldsymbol{A}_1, ..., \boldsymbol{A}_n, \boldsymbol{B})$ and $\boldsymbol{\Delta D}^j = (\boldsymbol{\Delta A}_1^j, ..., \boldsymbol{\Delta A}_n^j, \boldsymbol{\Delta B}^j)$. Therefore,

$$\begin{aligned} &g(\boldsymbol{x}, \boldsymbol{\Delta D}^j)\\ &= \max\left\{-\lambda_{min}(\textstyle\sum_{j=1}^n \boldsymbol{\Delta A}_i^j x_i - \boldsymbol{\Delta B}^j),\right.\\ &\qquad \left.-\lambda_{min}\left(-\left(\textstyle\sum_{j=1}^n \boldsymbol{\Delta A}_i^j x_i - \boldsymbol{\Delta B}^j\right)\right)\right\}\\ &= \max\left\{\lambda_{max}\left(-\left(\textstyle\sum_{j=1}^n \boldsymbol{\Delta A}_i^j x_i - \boldsymbol{\Delta B}^j\right)\right),\right.\\ &\quad \left.\lambda_{max}(\textstyle\sum_{j=1}^n \boldsymbol{\Delta A}_i^j x_i - \boldsymbol{\Delta B}^j)\right\}\\ &= \left\|\sum_{j=1}^n \boldsymbol{\Delta A}_i^j x_i - \boldsymbol{\Delta B}^j\right\|_2. \end{aligned}$$

3.3. The nature and size of the robust problem

In this section, we discuss the nature and size of the proposed robust conic problem. Note that in the proposed robust model (14) for every uncertain conic constraint $f(\boldsymbol{x}, \tilde{\boldsymbol{D}})$ we add at most $|N| + 1$ new variables, $2|N|$ conic constraints of the same nature as the nominal problem and an additional constraint involving the dual norm. The nature of this constraint depends on the norm we use to describe the uncertainty set $\mathcal{U}$ defined in Eq. (5).

When all the data entries of the problem have independent random perturbations, by exploiting sparsity of the additional conic constraints, we can further reduce the size of the robust model. Essentially, we can express the model of uncertainty in the form of Eq. (4), for which $\tilde{z}_j$ is the independent random variable associated with the jth data element, and $\boldsymbol{\Delta D}^j$ contains mostly zeros except at the entries corresponding to the data element. As an illustration, consider the following semidefinite constraint,

$$\begin{pmatrix} a_1 & a_2 \\ a_2 & a_3 \end{pmatrix} x_1 + \begin{pmatrix} a_4 & a_5 \\ a_5 & a_6 \end{pmatrix} x_2 \succeq \begin{pmatrix} a_7 & a_8 \\ a_8 & a_9 \end{pmatrix},$$

such that each element in the data $\boldsymbol{d} = (a_1, \ldots, a_9)'$ has an independent random perturbation, that is $\tilde{a}_i = a_i^0 + \Delta a_i \tilde{z}_i$ and $\tilde{z}_i$ are independently distributed. Equivalently, in Eq. (4) we have

$$\tilde{\boldsymbol{d}} = \boldsymbol{d}^0 + \sum_{i=1}^{9} \boldsymbol{\Delta d}^i \tilde{z}_i,$$

where $\boldsymbol{d}^0 = (a_1^0, \ldots, a_9^0)'$ and $\boldsymbol{\Delta d}^i$ is a vector with Δa_i at the ith entry and zero, otherwise. Hence, we can simplify the conic constraint in Eq. (14), $f(\boldsymbol{x}, \boldsymbol{\Delta d}^1) + t_1 \geq 0$ or

$$\lambda_{min}\left(\begin{pmatrix} \Delta a_1 & 0 \\ 0 & 0 \end{pmatrix} x_1 + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} x_2 - \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\right) + t_1 \geq 0,$$

as $t_1 \geq -\min\{\Delta a_1 x_1, 0\}$ or equivalently as linear constraints

$$t_1 \geq -\Delta a_1 x_1, t_1 \geq 0.$$

In Appendix B we derive and in Table 4 we summarize the number of variables and constraints and their nature when the nominal problem is an LP, QCQP, SOCP (1) (only $\boldsymbol{A}, \boldsymbol{b}$ vary), SOCP (2) ($\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}, d$ vary) and SDP for various choices of norms. Note that for the cases of the l_1, l_∞ and l_2 norms, we are able to collate terms so that the number of variables and constraints introduced is minimal. Furthermore, using the l_2 norm results in only one additional variable, one additional SOCP type of constraint, while maintaining the nature of the original conic optimization problem of SOCP and SDP. The use of other norms comes at the expense of more variables and constraints of the order of $|N|$, which is not very appealing for large problems.

Table 4. Size increase and nature of robust formulation when each data entry has independent uncertainty

	l_1	l_∞	$l_1 \cap l_\infty$	l_2	$l_2 \cap l_\infty$
Vars.	$n+1$	1	$2\|N\|+2$	1	$2\|N\|+1$
Linear Const.	$2n+1$	$2n+1$	$4\|N\|+2$	0	$3\|N\|$
SOC Const.	0	0	0	1	1
LP	LP	LP	LP	SOCP	SOCP
QCQP	SOCP	SOCP	SOCP	SOCP	SOCP
SOCP(1)	SOCP	SOCP	SOCP	SOCP	SOCP
SOCP(2)	SOCP	SOCP	SOCP	SOCP	SOCP
SDP	SDP	SDP	SDP	SDP	SDP

4. Probabilistic Guarantees

In this section, we derive a guarantee on the probability that the robust solution is feasible, when the uncertain coefficients obey some natural probability distributions. An important component of our analysis is the relation among different norms. We denote by $\langle\ ,\ \rangle$ the inner product on a vector space, $\Re^m$ or the space of m by m symmetric matrices, $\boldsymbol{S}^{m\times m}$. The inner product induces a norm $\sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$. For a vector space, the natural inner product is the Euclidian inner product, $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \boldsymbol{x}'\boldsymbol{y}$, and the induced norm is the Euclidian norm $\|\boldsymbol{x}\|_2$. For the space of symmetric matrices, the natural inner product is the trace product or $\langle \boldsymbol{X}, \boldsymbol{Y}\rangle = \text{trace}(\boldsymbol{X}\boldsymbol{Y})$ and the corresponding induced norm is the Frobenius norm, $\|\boldsymbol{X}\|_F$ (see [17]).

We analyze the relation of the inner product norm $\sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$ with the norm $\|\boldsymbol{x}\|_g$ defined in Table 3 for the conic optimization problems we consider. Since $\|\boldsymbol{x}\|_g$ and $\sqrt{\langle \boldsymbol{x}, \boldsymbol{x}\rangle}$ are valid norms in a finite dimensional space, there exist finite $\alpha_1, \alpha_2 > 0$ such that

$$\frac{1}{\alpha_1}\|\boldsymbol{r}\|_g \le \sqrt{\langle \boldsymbol{r}, \boldsymbol{r}\rangle} \le \alpha_2\|\boldsymbol{r}\|_g, \tag{17}$$

for all $\boldsymbol{r}$ in the relevant space.

Proposition 4. *For the norm $\|\cdot\|_g$ defined in Table 3 for the conic optimization problems we consider, Eq. (17) holds with the following parameters:*

(a) *LP: $\alpha_1 = \alpha_2 = 1$.*
(b) *QCQP, SOCP(2): $\alpha_1 = \sqrt{2}$ and $\alpha_2 = 1$.*
(c) *SOCP(1): $\alpha_1 = \alpha_2 = 1$.*
(d) *SDP: $\alpha_1 = 1$ and $\alpha_2 = \sqrt{m}$.*

Proof. (a) LP: For $r \in \Re$ and $\|r\|_g = |r|$, leading to Eq. (17) with $\alpha_1 = \alpha_2 = 1$.

(b) QCQP, SOCP(2): For $\boldsymbol{r} = (\boldsymbol{r}_1, r_0)' \in \Re^{l+1}$, let $a = \|\boldsymbol{r}_1\|_2$ and $b = |r_0|$. Since $a, b > 0$, using the inequality $a + b \le \sqrt{2}\sqrt{a^2+b^2}$ and $\sqrt{a^2+b^2} \le a + b$, we have

$$\frac{1}{\sqrt{2}}(\|\boldsymbol{r}_1\|_2 + |r_0|) \le \sqrt{\boldsymbol{r}'\boldsymbol{r}} = \|\boldsymbol{r}\|_2 \le \|\boldsymbol{r}_1\|_2 + |r_0|$$

leading to Eq. (17) with $\alpha_1 = \sqrt{2}$ and $\alpha_2 = 1$.

(c) SOCP(1): For all $\boldsymbol{r}$, Eq. (17) holds with $\alpha_1 = \alpha_2 = 1$.

(d) Let $\lambda_j, j = 1, \ldots, m$ be the eigenvalues of the matrix $\boldsymbol{A}$. Since $\|\boldsymbol{A}\|_F = \sqrt{\text{trace}(\boldsymbol{A}^2)} = \sqrt{\sum_j \lambda_j^2}$ and $\|\boldsymbol{A}\|_2 = \max_j |\lambda_j|$, we have

$$\|\boldsymbol{A}\|_2 \le \|\boldsymbol{A}\|_F \le \sqrt{m}\|\boldsymbol{A}\|_2,$$

leading to Eq. (17) with $\alpha_1 = 1$ and $\alpha_2 = \sqrt{m}$. □

The central result of the section is as follows.

Theorem 2.

(a) *Under the model of uncertainty in Eq. (4), and given a feasible solution* $\boldsymbol{x}$ *in Eq. (8), then*

$$\mathrm{P}(f(\boldsymbol{x}, \tilde{\boldsymbol{D}}) < 0) \le \mathrm{P}\left(\| \sum_{j \in N} \boldsymbol{r}_j \tilde{z}_j \|_g > \Omega \|\boldsymbol{s}\|^* \right),$$

where

$$\boldsymbol{r}_j = \mathcal{H}(\boldsymbol{x}, \boldsymbol{\Delta D}^j), \qquad s_j = \|\boldsymbol{r}_j\|_g, \quad j \in N.$$

(b) *When we use the* l_2*-norm in Eq. (9), i.e.,* $\|\boldsymbol{s}\|^* = \|\boldsymbol{s}\|_2$*, and under the assumption that* z_j *are normally and independently distributed with mean zero and variance one, i.e.,* $\tilde{\boldsymbol{z}} \sim \mathcal{N}(0, \boldsymbol{I})$*, then*

$$\mathrm{P}\left(\left\| \sum_{j \in N} \boldsymbol{r}_j \tilde{z}_j \right\|_g > \Omega \sqrt{\sum_{j \in N} \|\boldsymbol{r}_j\|_g^2}, \right) \le \frac{\sqrt{e}\Omega}{\alpha} \exp\left(-\frac{\Omega^2}{2\alpha^2} \right), \tag{18}$$

where $\alpha = \alpha_1 \alpha_2$*,* α_1, α_2 *derived in Proposition 4 and* $\Omega > \alpha$*.*

Proof. We have

$$\mathrm{P}(f(\boldsymbol{x}, \tilde{\boldsymbol{D}}) < 0)$$

$$\le \mathrm{P}\left(f(\boldsymbol{x}, \boldsymbol{D}^0) + f(\boldsymbol{x}, \sum_{j \in N} \boldsymbol{\Delta D}^j \tilde{z}_j) < 0 \right) \qquad \text{(From (7))}$$

$$\le \mathrm{P}\left(f(\boldsymbol{x}, \textstyle\sum_{j \in N} \boldsymbol{\Delta D}^j \tilde{z}_j) < -\Omega \|\boldsymbol{s}\|^* \right)$$

$$\text{(From (13), } s_j = \|\mathcal{H}(\boldsymbol{x}, \boldsymbol{\Delta D}^j)\|_g\text{)}$$

$$\le \mathrm{P}\left(\min\left(f(\boldsymbol{x}, \sum_{j \in N} \boldsymbol{\Delta D}^j \tilde{z}_j), f(\boldsymbol{x}, -\sum_{j \in N} \boldsymbol{\Delta D}^j \tilde{z}_j) \right) < -\Omega \|\boldsymbol{s}\|^* \right)$$

$$= \mathrm{P}\left(g(\boldsymbol{x}, \sum_{j \in N} \boldsymbol{\Delta D}^j \tilde{z}_j) > \Omega \|\boldsymbol{s}\|^* \right)$$

$$= \mathrm{P}\left(\|\mathcal{H}(\boldsymbol{x}, \sum_{j\in N} \boldsymbol{\Delta D}^j \tilde{z}_j)\|_g > \Omega\|\boldsymbol{s}\|^* \right)$$

$$= \mathrm{P}\left(\|\sum_{j\in N} \mathcal{H}(\boldsymbol{x}, \boldsymbol{\Delta D}^j)\tilde{z}_j\|_g > \Omega\|\boldsymbol{s}\|^* \right) \qquad (\mathcal{H}(\boldsymbol{x}, \boldsymbol{D}) \text{ is linear in } \boldsymbol{D})$$

$$= \mathrm{P}\left(\|\sum_{j\in N} \boldsymbol{r}_j\tilde{z}_j\|_g > \Omega\|\boldsymbol{s}\|^* \right).$$

(b) Using, the relations $\|\boldsymbol{r}\|_g \leq \alpha_1\sqrt{\langle \boldsymbol{r}, \boldsymbol{r}\rangle}$ and $\|\boldsymbol{r}\|_g \geq \frac{1}{\alpha_2}\sqrt{\langle \boldsymbol{r}, \boldsymbol{r}\rangle}$ from Proposition 4, we obtain

$$\mathrm{P}\left(\left\| \sum_{j\in N} \boldsymbol{r}_j\tilde{z}_j \right\|_g > \Omega\sqrt{\sum_{j\in N} \|\boldsymbol{r}_j\|_g^2} \right)$$

$$= \mathrm{P}\left(\left\| \sum_{j\in N} \boldsymbol{r}_j\tilde{z}_j \right\|_g^2 > \Omega^2 \sum_{j\in N} \|\boldsymbol{r}_j\|_g^2 \right)$$

$$\leq \mathrm{P}\left(\alpha_1^2\alpha_2^2 \left\langle \sum_{j\in N} \boldsymbol{r}_j\tilde{z}_j, \sum_{k\in N} \boldsymbol{r}_k\tilde{z}_k \right\rangle > \Omega^2 \sum_{j\in N} \langle \boldsymbol{r}_j, \boldsymbol{r}_j\rangle \right)$$

$$= \mathrm{P}\left(\alpha^2 \sum_{j\in N}\sum_{k\in N} \langle \boldsymbol{r}_j, \boldsymbol{r}_k\rangle \tilde{z}_j\tilde{z}_k > \Omega^2 \sum_{j\in N} \langle \boldsymbol{r}_j, \boldsymbol{r}_j\rangle \right)$$

$$= \mathrm{P}\left(\alpha^2 \tilde{\boldsymbol{z}}'\boldsymbol{R}\tilde{\boldsymbol{z}} > \Omega^2 \sum_{j\in N} \langle \boldsymbol{r}_j, \boldsymbol{r}_j\rangle \right),$$

where $R_{jk} = \langle \boldsymbol{r}_j, \boldsymbol{r}_k\rangle$. Clearly, $\boldsymbol{R}$ is a symmetric positive semidefinite matrix and can be spectrally decomposed such that $\boldsymbol{R} = \boldsymbol{Q}'\boldsymbol{\Lambda}\boldsymbol{Q}$, where $\boldsymbol{\Lambda}$ is the diagonal matrix of the eigenvalues and $\boldsymbol{Q}$ is the corresponding orthonormal matrix. Let $\tilde{\boldsymbol{y}} = \boldsymbol{Q}\tilde{\boldsymbol{z}}$ so that $\tilde{\boldsymbol{z}}'\boldsymbol{R}\tilde{\boldsymbol{z}} = \tilde{\boldsymbol{y}}'\boldsymbol{\Lambda}\tilde{\boldsymbol{y}} = \sum_{j\in N} \lambda_j\tilde{y}_j^2$. Since $\tilde{\boldsymbol{z}} \sim \mathcal{N}(0, \boldsymbol{I})$, we also have $\tilde{\boldsymbol{y}} \sim \mathcal{N}(0, \boldsymbol{I})$, that is, $\tilde{y}_j$, $j \in N$ are independent and normally distributed. Moreover,

$$\sum_{j\in N} \lambda_j = \mathrm{trace}(\boldsymbol{R}) = \sum_{j\in N} \langle \boldsymbol{r}_j, \boldsymbol{r}_j\rangle.$$

Therefore,

$$
\begin{aligned}
& \mathrm{P}\left(\alpha^2 \tilde{z}' \boldsymbol{R} \tilde{z} > \Omega^2 \sum_{j \in N} \langle \boldsymbol{r}_j, \boldsymbol{r}_j \rangle\right) \\
& = \mathrm{P}\left(\alpha^2 \sum_{j \in N} \lambda_j \tilde{y}_j^2 > \Omega^2 \sum_{j \in N} \lambda_j\right) \\
& \leq \frac{E\left(\exp\left(\theta\alpha^2 \sum_{j \in N} \lambda_j \tilde{y}_j^2\right)\right)}{\exp\left(\theta\Omega^2 \sum_{j \in N} \lambda_j\right)} \qquad \text{(From Markov's inequality, } \theta > 0) \\
& = \frac{\prod_{j \in N} E\left(\exp\left(\theta\alpha^2 \lambda_j \tilde{y}_j^2\right)\right)}{\exp\left(\theta\Omega^2 \sum_{j \in N} \lambda_j\right)} \qquad (\tilde{y}_j^2 \text{ are independent}) \\
& = \frac{\prod_{j \in N} E\left(\exp\left(\frac{\tilde{y}_j^2}{\beta}\right)^{\theta\alpha^2\lambda_j\beta}\right)}{\exp\left(\theta\Omega^2 \sum_{j \in N} \lambda_j\right)} \qquad \text{for all } \beta > 2 \text{ and } \theta\alpha^2\lambda_j\beta \leq 1, \forall j \in N \\
& \leq \frac{\prod_{j \in N} \left(E\left(\exp\left(\frac{\tilde{y}_j^2}{\beta}\right)\right)^{\theta\alpha^2\lambda_j\beta}\right)}{\exp\left(\theta\Omega^2 \sum_{j \in N} \lambda_j\right)},
\end{aligned}
$$

where the last inequality follows from Jensen inequality, noting that $x^{\theta\alpha^2\lambda_j\beta}$ is a concave function of x if $\theta\alpha^2\lambda_j\beta \in [0, 1]$. Since $\tilde{y}_j \sim \mathcal{N}(0, 1)$,

$$
E\left(\exp\left(\frac{\tilde{y}_j^2}{\beta}\right)\right) = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} \exp\left(-\frac{y^2}{2}\left(\frac{\beta - 2}{\beta}\right)\right) dy = \sqrt{\frac{\beta}{\beta - 2}}.
$$

Thus, we obtain

$$
\begin{aligned}
\frac{\prod_{j \in N} \left(E\left(\exp\left(\frac{\tilde{y}_j^2}{\beta}\right)\right)^{\theta\alpha^2\lambda_j\beta}\right)}{\exp\left(\theta\Omega^2 \sum_{j \in N} \lambda_j\right)} & = \frac{\prod_{j \in N} \left(\exp\left(\theta\alpha^2\lambda_j\beta\frac{1}{2}\ln\left(\frac{\beta}{\beta-2}\right)\right)\right)}{\exp\left(\theta\Omega^2 \sum_{j \in N} \lambda_j\right)} \\
& = \frac{\exp\left(\theta\alpha^2\beta\frac{1}{2}\ln\left(\frac{\beta}{\beta-2}\right)\sum_{j \in N} \lambda_j\right)}{\exp\left(\theta\Omega^2 \sum_{j \in N} \lambda_j\right)}.
\end{aligned}
$$

We select $\theta = 1/(\alpha^2\beta\lambda^*)$, where $\lambda^* = \max_{j \in N} \lambda_j$, and obtain

$$
\frac{\exp\left(\theta\alpha^2\beta\frac{1}{2}\ln\left(\frac{\beta}{\beta-2}\right)\sum_{j \in N} \lambda_j\right)}{\exp\left(\theta\Omega^2 \sum_{j \in N} \lambda_j\right)} = \exp\left(\rho\left(\frac{1}{2}\ln\left(\frac{\beta}{\beta - 2}\right) - \frac{\Omega^2}{\alpha^2\beta}\right)\right),
$$

where $\rho = (\sum_{j\in N} \lambda_j)/\lambda^*$. Taking derivatives and choosing the best β, we have

$$\beta = \frac{2\Omega^2}{\Omega^2 - \alpha^2},$$

for which $\Omega > \alpha$. Substituting and simplifying, we have

$$\begin{aligned} \exp\left(\rho\left(\frac{1}{2}\ln\left(\frac{\beta}{\beta-2}\right) - \frac{\Omega^2}{\alpha^2\beta}\right)\right) &= \left(\frac{\sqrt{e}\Omega}{\alpha}\exp(-\frac{\Omega^2}{2\alpha^2})\right)^\rho \\ &\le \frac{\sqrt{e}\Omega}{\alpha}\exp(-\frac{\Omega^2}{2\alpha^2}), \end{aligned}$$

where the last inequality follows from $\rho \ge 1$, and from $\frac{\sqrt{e}\Omega}{\alpha}\exp(-\frac{\Omega^2}{2\alpha^2}) < 1$ for $\Omega > \alpha$. □

Remark 1. We note the series of inequalities used in the proof would increase the gap between the actual probability of feasibility with the designed values. In particular, in the last inequality, it is easy to see that ρ can be as large as the rank of the matrix, $\boldsymbol{R}$. Hence, for an uncertain single LP constraint, we have ρ=1, while for an uncertain second order cone constraint, ρ could be as large as the dimension of the cone. Therefore, for such problems, it is conceivable that even if our intended probability bound against infeasibility is ϵ the solution to our proposed robust model may violate the constraints with probability of less than ϵ^n, where n is the dimension of the cone. However, if the errors are small and Ω is not too large, the price to pay for such assurance could be acceptable in practice. A possible tuning approach might be to determine ρ and adapt Ω accordingly. However, such an approach may not be polynomial.

Note that $f(\boldsymbol{x}, \tilde{\boldsymbol{D}}) < 0$, implies that $\|\tilde{z}\| > \Omega$. Thus, when $\tilde{z} \sim \mathcal{N}(0, \boldsymbol{I})$

$$\mathrm{P}(f(\boldsymbol{x}, \tilde{\boldsymbol{D}}) < 0) \le \mathrm{P}(\|\tilde{z}\| > \Omega) = 1 - \chi^2_{|N|}(\Omega^2), \tag{19}$$

where $\chi^2_{|N|}(\cdot)$ is the cdf of a χ-square distribution with $|N|$ degrees of freedom. Note that the bound (19) does not take into account the structure of $f(\boldsymbol{x}, \tilde{\boldsymbol{D}})$ in contrast to bound (18) that depends on $f(\boldsymbol{x}, \tilde{\boldsymbol{D}})$ via the parameter α. To illustrate this, we substitute the value of the parameter α from Proposition 4 in Eq. (18) and report in Table 6 the bound in Eq. (18).

To amplify the previous discussion, we show in Table 6 the value of Ω in order for the bound (18) to be less than or equal to ϵ. The last column shows the value of Ω using

Table 5. Probability bounds of $\mathrm{P}(f(\boldsymbol{x}, \tilde{\boldsymbol{D}}) < 0)$ for $\tilde{z} \sim \mathcal{N}(0, \boldsymbol{I})$

Type	Probability bound of infeasibility
LP	$\sqrt{e}\Omega\exp(-\frac{\Omega^2}{2})$
QCQP	$\sqrt{\frac{e}{2}}\Omega\exp(-\frac{\Omega^2}{4})$
SOCP(1)	$\sqrt{e}\Omega\exp(-\frac{\Omega^2}{2})$
SOCP(2)	$\sqrt{\frac{e}{2}}\Omega\exp(-\frac{\Omega^2}{4})$
SDP	$\sqrt{\frac{e}{m}}\Omega\exp(-\frac{\Omega^2}{2m})$

bound (19) that is independent of the structure of the problem. We choose $|N| = 495000$ which is approximately the maximum number of data entries in a SDP constraint with $n = 100$ and $m = 100$. Although the size $|N|$ is unrealistic for constraints with less data entries such as LP, the derived probability bounds remain valid. Note that bound (19) leads to $\Omega = O(\sqrt{|N|}\ln(1/\epsilon))$.

For LP, SOCP, and QCQP, bound (18) leads to $\Omega = O(\ln(1/\epsilon))$, which is independent of the dimension of the problem. For SDP it leads to we have $\Omega = O(\sqrt{m}\ln(1/\epsilon))$. As a result, ignoring the structure of the problem and using bound (19) leads to very conservative solutions.

Large Deviation Results of Nemirovski

Nemirovski [15, 16] gives bounds on the probability of large deviations in normed spaces, under fairly general distributions for the random variables. He assumes that the random variables $\tilde{z}_j$, $j \in N$ are mutually independent, with zero mean and satisfy the following condition:

$$E(\exp(\tilde{z}_j^2)) \leq \exp(1). \tag{20}$$

Note that $\tilde{z}_j \sim \mathcal{N}(0, \sigma^2)$, where $\sigma^2 = \frac{1}{2\ln(2)}$ satisfies (20). Moreover, bounded random variables such that $|\tilde{z}_j| \leq 1$ and $E(\tilde{z}_j) = 0$ satisfy (20) as well.

Let $(E, \|.\|)$ be a separable Banach space such that there exists a norm $p(\boldsymbol{x})$ satisfying $\|\boldsymbol{x}\| \leq p(\boldsymbol{x}) \leq 2\|\boldsymbol{x}\|$ and that the function $P(\boldsymbol{x}) = \frac{1}{2}p^2(\boldsymbol{x})$ is continuously differentiable and satisfies the relation

$$P(\boldsymbol{x} + \boldsymbol{y}) \leq P(\boldsymbol{x}) + \langle P'(\boldsymbol{x}), \boldsymbol{y}\rangle + \kappa^2 P(\boldsymbol{y}).$$

Theorem 3. *(Nemirovski [15, 16]) Let $\tilde{\boldsymbol{r}}^j$, $j \in N$ be independent random vectors in E with zero mean, such that $E(\exp(\|\tilde{\boldsymbol{r}}^j\|^2/\sigma_j^2)) \leq \exp(1)$.*

(a) *For appropriately chosen absolute constant $c > 0$ and for all $\Omega > 0$,*

$$\Pr\left(\|\sum_{j\in N}\tilde{\boldsymbol{r}}^j\| \geq \Omega\sqrt{\sum_{j\in N}\sigma_j^2}\right) \leq \frac{\exp(-c\Omega^2/\kappa^2)}{c}.$$

(b) *For $E = \Re^n$ and under Euclidian norm, $\|.\|_2$, $\kappa = 1$. For $E = M^{m,n}$ of $m \times n$ matrices, and under the standard matrix norm,*

$$\kappa = O(\sqrt{\ln(\min(n, m) + 1)}).$$

Table 6. Sample calculations of Ω using Probability Bounds of Table 5 for $m = 100$, $n = 100$, $|N| = 495{,}000$

ϵ	LP	QCQP	SOCP(1)	SOCP(2)	SDP	Eq. (19)
10^{-1}	2.76	3.91	2.76	3.91	27.6	704.5
10^{-2}	3.57	5.05	3.57	5.05	35.7	705.2
10^{-3}	4.21	5.95	4.21	5.95	42.1	705.7
10^{-6}	5.68	7.99	5.68	7.99	56.8	706.9

With respect to the bounds of Theorem 2 in which we consider the Euclidian norm, $\|\boldsymbol{s}\|^* = \|\boldsymbol{s}\|_2$, we have exactly the same framework of Theorem 3, for which $\sigma_j = \|\boldsymbol{r}^j\|$. Furthermore, since

$$E(\exp(\|\tilde{\boldsymbol{r}}^j\|^2/\sigma_j^2)) = E(\exp(\|\boldsymbol{r}^j\|^2\tilde{z}_j/\|\boldsymbol{r}^j\|^2)) = E(\exp(\tilde{z}_j)) \leq \exp(1),$$

Theorem 3 directly applies to our proposed framework.

Table 7 shows the desired value of Ω to guarantee that the probability of feasibility is least $1-\epsilon$. We observe that Table 7 provides stronger bounds than our framework.

5. General cones

In this section, we generalize the results in Sections 2-4 to arbitrary conic constraints of the form,

$$\sum_{j=1}^{n} \tilde{\boldsymbol{A}}_j x_j \succeq_{\boldsymbol{K}} \tilde{\boldsymbol{B}}, \tag{21}$$

where $\{\tilde{\boldsymbol{A}}_1, ..., \tilde{\boldsymbol{A}}_n, \tilde{\boldsymbol{B}}\} = \tilde{\boldsymbol{D}}$ constitutes the set of data that is subject to uncertainty, and $\boldsymbol{K}$ is a closed, convex, pointed cone with nonempty interior. For notational simplicity, we define

$$\mathcal{A}(\boldsymbol{x}, \tilde{\boldsymbol{D}}) = \sum_{j=1}^{n} \tilde{\boldsymbol{A}}_j x_j - \tilde{\boldsymbol{B}}$$

so that Eq. (21) is equivalent to

$$\mathcal{A}(\boldsymbol{x}, \tilde{\boldsymbol{D}}) \succeq_{\boldsymbol{K}} \boldsymbol{0}. \tag{22}$$

We assume that the model for data uncertainty is given in Eq. (4) with $\tilde{z} \sim \mathcal{N}(0, \boldsymbol{I})$. The uncertainty set $\mathcal{U}$ satisfies Eq. (5) with the given norm satisfying $\|\boldsymbol{u}\| = \||\boldsymbol{u}|\|$.

Paralleling the earlier development, starting with a cone $\boldsymbol{K}$ and constraint (22), we define the function $f(\cdot,\cdot)$ as follows so that $f(\boldsymbol{x}, \boldsymbol{D}) > 0$ if and only if $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succ_{\boldsymbol{K}} \boldsymbol{0}$.

Proposition 5. *For any $\boldsymbol{V} \succ_{\boldsymbol{K}} 0$, the function*

$$\begin{aligned} f(\boldsymbol{x}, \boldsymbol{D}) = \max\ & \theta \\ \text{s.t. } & \mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succeq_{\boldsymbol{K}} \theta \boldsymbol{V}, \end{aligned} \tag{23}$$

satisfies the properties:

Table 7. The value of Ω to achieve probability of feasibility of at least $1-\epsilon$ obtained by applying Theorem 3

Type	Ω
LP	$O(\sqrt{\ln 1/\epsilon})$
QCQP	$O(\sqrt{\ln 1/\epsilon})$
SOCP(1)	$O(\sqrt{\ln 1/\epsilon})$
SOCP(2)	$O(\sqrt{\ln 1/\epsilon})$
SDP	$O(\sqrt{\ln(m)\ln 1/\epsilon})$

(a) $f(\boldsymbol{x}, \boldsymbol{D})$ *is bounded and concave in* $\boldsymbol{x}$ *and* $\boldsymbol{D}$.
(b) $f(\boldsymbol{x}, k\boldsymbol{D}) = kf(\boldsymbol{x}, \boldsymbol{D}), \ \forall k \geq 0$.
(c) $f(\boldsymbol{x}, \boldsymbol{D}) \geq y$ *if and only if* $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succeq_{\boldsymbol{K}} y\boldsymbol{V}$.
(d) $f(\boldsymbol{x}, \boldsymbol{D}) > y$ *if and only if* $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succ_{\boldsymbol{K}} y\boldsymbol{V}$.

Proof. (a) Consider the dual of Problem (23):

$$\begin{aligned} z^* = \min\ & \langle \boldsymbol{u}, \mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \rangle \\ \text{s.t.}\ & \langle \boldsymbol{u}, \boldsymbol{V} \rangle = 1 \\ & \boldsymbol{u} \succeq_{\boldsymbol{K}^*} \boldsymbol{0}, \end{aligned}$$

where $\boldsymbol{K}^*$ is the dual cone of $\boldsymbol{K}$. Since $\boldsymbol{K}$ is a closed, convex, pointed cone with nonempty interior, so is $\boldsymbol{K}^*$ (see [5]). As $\boldsymbol{V} \succ_{\boldsymbol{K}} \boldsymbol{0}$, for all $\boldsymbol{u} \succeq_{\boldsymbol{K}^*} \boldsymbol{0}$ and $\boldsymbol{u} \neq \boldsymbol{0}$, we have $\langle \boldsymbol{u}, \boldsymbol{V} \rangle > 0$, hence, the dual problem is bounded. Furthermore, since $\boldsymbol{K}^*$ has a nonempty interior, the dual problem is strictly feasible, i.e., there exists $\boldsymbol{u} \succ_{\boldsymbol{K}^*} \boldsymbol{0}, \ \langle \boldsymbol{u}, \boldsymbol{V} \rangle = 1$. Therefore, by conic duality, the dual objective z^* has the same finite objective as the primal objective function $f(\boldsymbol{x}, \boldsymbol{D})$. Since $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D})$ is a linear mapping of $\boldsymbol{D}$ and an affine mapping of $\boldsymbol{x}$, it follows that $f(\boldsymbol{x}, \boldsymbol{D})$ is concave in $\boldsymbol{x}$ and $\boldsymbol{D}$.

(b) Using the dual expression of $f(\boldsymbol{x}, \boldsymbol{D})$, and that $\mathcal{A}(\boldsymbol{x}, k\boldsymbol{D}) = k\mathcal{A}(\boldsymbol{x}, \boldsymbol{D})$, the result follows.

(c) If $\theta = y$ is feasible in Problem (23), we have $f(\boldsymbol{x}, \boldsymbol{D}) \geq \theta = y$. Conversely, if $f(\boldsymbol{x}, \boldsymbol{D}) \geq y$, then $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succeq_{\boldsymbol{K}} f(\boldsymbol{x}, \boldsymbol{D})\boldsymbol{V} \succeq_{\boldsymbol{K}} y\boldsymbol{V}$.

(d) Suppose $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succ_{\boldsymbol{K}} y\boldsymbol{V}$, then there exists $\epsilon > 0$ such that $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) - y\boldsymbol{V} \succeq_{\boldsymbol{K}} \epsilon\boldsymbol{V}$ or $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succeq_{\boldsymbol{K}} (\epsilon + y)\boldsymbol{V}$. Hence, $f(\boldsymbol{x}, \boldsymbol{D}) \geq \epsilon + y > y$. Conversely, since $\boldsymbol{V} \succ_{\boldsymbol{K}} \boldsymbol{0}$, if $f(\boldsymbol{x}, \boldsymbol{D}) > y$ then $(f(\boldsymbol{x}, \boldsymbol{D}) - y)\boldsymbol{V} \succ_{\boldsymbol{K}} \boldsymbol{0}$. Hence, $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succeq_{\boldsymbol{K}} f(\boldsymbol{x}, \boldsymbol{D})\boldsymbol{V} \succ_{\boldsymbol{K}} y\boldsymbol{V}$. □

Remark 2. With $y = 0$, (c) establishes that $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succeq_{\boldsymbol{K}} \boldsymbol{0}$ if and only if $f(\boldsymbol{x}, \boldsymbol{D}) \geq 0$ and (d) establishes that $\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}) \succ_{\boldsymbol{K}} \boldsymbol{0}$ if and only if $f(\boldsymbol{x}, \boldsymbol{D}) > 0$.

The proposed robust model is given in Eqs. (8) and (9). We next derive an expression for $g(\boldsymbol{x}, \Delta\boldsymbol{D}) = \max\{-f(\boldsymbol{x}, \Delta\boldsymbol{D}), -f(\boldsymbol{x}, -\Delta\boldsymbol{D})\}$.

Proposition 6. *Let* $g(\boldsymbol{x}, \Delta\boldsymbol{D}) = \max\{-f(\boldsymbol{x}, \Delta\boldsymbol{D}), -f(\boldsymbol{x}, -\Delta\boldsymbol{D})\}$. *Then*

$$g(\boldsymbol{x}, \Delta\boldsymbol{D}) = \|\mathcal{H}(\boldsymbol{x}, \Delta\boldsymbol{D})\|_g,$$

where $\mathcal{H}(\boldsymbol{x}, \Delta\boldsymbol{D}) = \mathcal{A}(\boldsymbol{x}, \Delta\boldsymbol{D})$ *and*

$$\|\boldsymbol{S}\|_g = \min\{y : y\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{S} \succeq_{\boldsymbol{K}} -y\boldsymbol{V}\}.$$

Proof. We observe that

$$\begin{aligned} g(\boldsymbol{x}, \Delta\boldsymbol{D}) &= \max\{-f(\boldsymbol{x}, \Delta\boldsymbol{D}), -f(\boldsymbol{x}, -\Delta\boldsymbol{D})\} \\ &= \min\{y \mid -f(\boldsymbol{x}, \Delta\boldsymbol{D}) \leq y, -f(\boldsymbol{x}, -\Delta\boldsymbol{D}) \leq y\} \\ &= \min\{y \mid \mathcal{A}(\boldsymbol{x}, \Delta\boldsymbol{D}) \succeq_{\boldsymbol{K}} -y\boldsymbol{V}, -\mathcal{A}(\boldsymbol{x}, \Delta\boldsymbol{D}) \succeq_{\boldsymbol{K}} -y\boldsymbol{V}\} \\ &\qquad \text{(From Proposition 5(c))} \\ &= \|\mathcal{A}(\boldsymbol{x}, \Delta\boldsymbol{D})\|_g. \end{aligned}$$

We also need to show that $\|.\|_g$ is indeed a valid norm. Since $\boldsymbol{V} \succ_{\boldsymbol{K}} \boldsymbol{0}$, then $\|\boldsymbol{S}\|_g \geq 0$. Clearly, $\|\boldsymbol{0}\|_g = 0$ and if $\|\boldsymbol{S}\|_g = 0$, then $\boldsymbol{0} \succeq_{\boldsymbol{K}} \boldsymbol{S} \succeq_{\boldsymbol{K}} \boldsymbol{0}$, which implies that $\boldsymbol{S} = \boldsymbol{0}$. To show that $\|k\boldsymbol{S}\|_g = |k|\|\boldsymbol{S}\|_g$, we observe that for $k > 0$,

$$\begin{aligned}
\|k\boldsymbol{S}\|_g &= \min\{y \mid y\boldsymbol{V} \succeq_{\boldsymbol{K}} k\boldsymbol{S} \succeq_{\boldsymbol{K}} -y\boldsymbol{V}\} \\
&= k \min\left\{\frac{y}{k} \mid \frac{y}{k}\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{S} \succeq_{\boldsymbol{K}} -\frac{y}{k}\boldsymbol{V}\right\} \\
&= k\|\boldsymbol{S}\|_g.
\end{aligned}$$

Likewise, if $k < 0$

$$\begin{aligned}
\|k\boldsymbol{S}\|_g &= \min\{y \mid y\boldsymbol{V} \succeq_{\boldsymbol{K}} k\boldsymbol{S} \succeq_{\boldsymbol{K}} -y\boldsymbol{V}\} \\
&= \min\{y \mid y\boldsymbol{V} \succeq_{\boldsymbol{K}} -k\boldsymbol{S} \succeq_{\boldsymbol{K}} -y\boldsymbol{V}\} \\
&= \|-k\boldsymbol{S}\|_g \\
&= -k\|\boldsymbol{S}\|_g.
\end{aligned}$$

Finally, to verify triangle inequality,

$$\begin{aligned}
&\|\boldsymbol{S}\|_g + \|\boldsymbol{T}\|_g \\
&= \min\{y \mid y\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{S} \succeq_{\boldsymbol{K}} -y\boldsymbol{V}\} + \min\{z \mid z\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{T} \succeq_{\boldsymbol{K}} -z\boldsymbol{V}\} \\
&= \min\{y + z \mid y\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{S} \succeq_{\boldsymbol{K}} -y\boldsymbol{V},\ z\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{T} \succeq_{\boldsymbol{K}} -z\boldsymbol{V}\} \\
&\geq \min\{y + z \mid (y+z)\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{S} + \boldsymbol{T} \succeq_{\boldsymbol{K}} -(y+z)\boldsymbol{V}\} \\
&= \|\boldsymbol{S} + \boldsymbol{T}\|_g. \qquad \square
\end{aligned}$$

For the general conic constraint, the norm, $\|\cdot\|_g$ is dependent on the cone $\boldsymbol{K}$ and a point in the interior of the cone $\boldsymbol{V}$. Hence, we define $\|\cdot\|_{\boldsymbol{K},\boldsymbol{V}} := \|\cdot\|_g$. Using Proposition 5 and Theorem 1 we next show that the robust counterpart for the conic constraint (22) is tractable and provide a bound on the probability that the constraint is feasible.

Theorem 4. *We have*

(a) **(Tractability)** *For a norm satisfying Eq. (6), constraint (8) for general cones is equivalent to*

$$\begin{aligned}
&\mathcal{A}(\boldsymbol{x}, \boldsymbol{D}^0) \succeq_{\boldsymbol{K}} \Omega y \boldsymbol{V}, \\
&t_j \boldsymbol{V} \succeq_{\boldsymbol{K}} \mathcal{A}(\boldsymbol{x}, \boldsymbol{\Delta D}^j) \succeq_{\boldsymbol{K}} -t_j \boldsymbol{V}, \quad j \in N, \\
&\|\boldsymbol{t}\|^* \leq y, \\
&y \in \Re,\ \boldsymbol{t} \in \Re^{|N|}.
\end{aligned} \tag{24}$$

(b) **(Probabilistic guarantee)** *When we use the l_2-norm in Eq. (9), i.e., $\|\boldsymbol{s}\|^* = \|\boldsymbol{s}\|_2$, and under the assumption that $\tilde{\boldsymbol{z}} \sim \mathcal{N}(0, \boldsymbol{I})$, then for all $\boldsymbol{V}$ we have*

$$\mathrm{P}(\mathcal{A}(\boldsymbol{x}, \tilde{\boldsymbol{D}}) \notin \boldsymbol{K}) \leq \frac{\sqrt{e}\Omega}{\alpha_{\boldsymbol{K},\boldsymbol{V}}} \exp\left(-\frac{\Omega^2}{2\alpha^2_{\boldsymbol{K},\boldsymbol{V}}}\right),$$

where

$$\alpha_{\boldsymbol{K},\boldsymbol{V}} = \left(\max_{\sqrt{\langle \boldsymbol{S},\boldsymbol{S}\rangle}=1} \|\boldsymbol{S}\|_{\boldsymbol{K},\boldsymbol{V}} \right) \left(\max_{\|\boldsymbol{S}\|_{\boldsymbol{K},\boldsymbol{V}}=1} \sqrt{\langle \boldsymbol{S}, \boldsymbol{S}\rangle} \right)$$

and

$$\|\boldsymbol{S}\|_{\boldsymbol{K},\boldsymbol{V}} = \min\left\{ y : y\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{S} \succeq_{\boldsymbol{K}} -y\boldsymbol{V} \right\}.$$

Proof. The Theorem follows directly from Propositions 5, 6, Theorems 1, 2. □

From Theorem 4, for any cone $\boldsymbol{K}$, we select $\boldsymbol{V}$ in order to minimize $\alpha_{\boldsymbol{K},\boldsymbol{V}}$, i.e.,

$$\alpha_{\boldsymbol{K}} = \min_{\boldsymbol{V} \succ_{\boldsymbol{K}} \boldsymbol{0}} \alpha_{\boldsymbol{K},\boldsymbol{V}}.$$

We next show that the smallest parameter α is $\sqrt{2}$ and $\sqrt{m}$ for SOCP and SDP respectively. For the second order cone, $\boldsymbol{K} = \boldsymbol{L}^{n+1}$,

$$\boldsymbol{L}^{n+1} = \{\boldsymbol{x} \in \Re^{n+1} : \|\boldsymbol{x}_n\|_2 \leq x_{n+1}\},$$

where $\boldsymbol{x}_n = (x_1, \ldots, x_n)'$. The induced norm is given by

$$\begin{aligned}
&\|\boldsymbol{x}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}} \\
&= \min\left\{ y : y\boldsymbol{v} \succeq_{\boldsymbol{L}^{n+1}} \boldsymbol{x} \succeq_{\boldsymbol{L}^{n+1}} -y\boldsymbol{v} \right\} \\
&= \min\left\{ y : \|\boldsymbol{x}_n + \boldsymbol{v}_n y\|_2 \leq v_{n+1}y + x_{n+1}, \|\boldsymbol{x}_n - \boldsymbol{v}_n y\|_2 \leq v_{n+1}y - x_{n+1}, \right\}
\end{aligned}$$

and

$$\alpha_{\boldsymbol{L}^{n+1},\boldsymbol{v}} = \left(\max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{x}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}} \right) \left(\max_{\|\boldsymbol{x}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}}=1} \|\boldsymbol{x}\|_2 \right).$$

For the symmetric positive semidefinite cone, $\boldsymbol{K} = \boldsymbol{S}^m_+$,

$$\|\boldsymbol{X}\|_{\boldsymbol{S}^m_+,\boldsymbol{V}} = \min\left\{ y : y\boldsymbol{V} \succeq \boldsymbol{X} \succeq -y\boldsymbol{V} \right\},$$

$$\alpha_{\boldsymbol{S}^m_+,\boldsymbol{V}} = \left(\max_{\sqrt{\langle \boldsymbol{X},\boldsymbol{X}\rangle}=1} \|\boldsymbol{x}\|_{\boldsymbol{S}^m_+,\boldsymbol{V}} \right) \left(\max_{\|\boldsymbol{X}\|_{\boldsymbol{S}^m_+,\boldsymbol{V}}=1} \sqrt{\langle \boldsymbol{X}, \boldsymbol{X}\rangle} \right).$$

Case of Homogenous cones

A cone, $\boldsymbol{K} \subseteq \Re^n$ is *homogenous* if for any pair of points $\boldsymbol{A}, \boldsymbol{B} \succ_{\boldsymbol{K}} \boldsymbol{0}$, there exists an invertible linear map $\mathcal{M} : \Re^n \to \Re^n$ such that $\mathcal{M}(\boldsymbol{A}) = \boldsymbol{B}$ and $\mathcal{M}(\boldsymbol{K}) = \boldsymbol{K}$ (see for instance Güler and Tunçel [14]). For general conic optimization, we have shown that the probability bound depends on the the choice of $\boldsymbol{V} \succ \boldsymbol{0}$. However, it turns out that for *homogenous cones*, in which semidefinite and second-order cones are special cases, the probability bound does not depend on $\boldsymbol{V} \succ \boldsymbol{0}$.

Theorem 5. *Suppose the cone* $\boldsymbol{K}$ *is homogenous. For any* $\boldsymbol{V} \succ_{\boldsymbol{K}} \boldsymbol{0}$*, the probability bound of Theorem 4(b) satisfies*

$$\mathrm{P}(\mathcal{A}(\boldsymbol{x}, \tilde{\boldsymbol{D}}) \notin \boldsymbol{K}) \leq \frac{\sqrt{e}\Omega}{\alpha_{\boldsymbol{K}}} \exp\left(-\frac{\Omega^2}{2\alpha_{\boldsymbol{K}}^2}\right).$$

Proof. Let $\boldsymbol{V}^* = \arg\min_{\boldsymbol{V} \succ_{\boldsymbol{K}} \boldsymbol{0}}$. Since the cone is homogenous and $\boldsymbol{V}, \boldsymbol{V}^* \succ_{\boldsymbol{K}} \boldsymbol{0}$, there exists an invertible linear map $\mathcal{M}(\cdot)$ satisfying $\mathcal{M}(\boldsymbol{V}) = \boldsymbol{V}^*$ and $\mathcal{M}(\boldsymbol{K}) = \boldsymbol{K}$. Noting that under the linear mapping, we have

$$\begin{aligned}
&\boldsymbol{X} \succeq_{\boldsymbol{K}} \boldsymbol{Y} \\
&\Rightarrow \boldsymbol{X} - \boldsymbol{Y} \succeq_{\boldsymbol{K}} \boldsymbol{0} \\
&\Rightarrow \mathcal{M}(\boldsymbol{X} - \boldsymbol{Y}) \succeq_{\boldsymbol{K}} \boldsymbol{0} \\
&\Rightarrow \mathcal{M}(\boldsymbol{X}) \succeq_{\boldsymbol{K}} \mathcal{M}(\boldsymbol{Y}).
\end{aligned}$$

Hence, it follows easily that the feasibility of (24) implies

$$\begin{aligned}
&\mathcal{A}_{\mathcal{M}}(\boldsymbol{x}, \boldsymbol{D}^0) \succeq_{\boldsymbol{K}} \Omega y \boldsymbol{V}^*, \\
&t_j \boldsymbol{V}^* \succeq_{\boldsymbol{K}} \mathcal{A}_{\mathcal{M}}(\boldsymbol{x}, \Delta \boldsymbol{D}^j) \succeq_{\boldsymbol{K}} -t_j \boldsymbol{V}^*, \quad j \in N, \\
&\|\boldsymbol{t}\|^* \leq y, \\
&y \in \Re, \ \boldsymbol{t} \in \Re^{|N|},
\end{aligned}$$

where

$$\mathcal{A}_{\mathcal{M}}(\boldsymbol{D}) = \mathcal{M}(\mathcal{A}(\boldsymbol{D})) = \sum_{j=1}^{n} \mathcal{M}(\boldsymbol{A}_j) - \mathcal{M}(\boldsymbol{B}).$$

Hence, the probability bound follows. □

We will next derive $\alpha_{\boldsymbol{K}}$ for semidefinite and second order cones.

Proposition 7. *We have*

(a) *For the second order cone,* $\alpha_{\boldsymbol{L}^{n+1}} = \sqrt{2}$.
(b) *For the symmetric positive semidefinite cone,* $\alpha_{\boldsymbol{S}_+^m} = \sqrt{m}$.

Proof. For any $\boldsymbol{V} \succ_{\boldsymbol{K}} \mathbf{0}$, we observe that

$$\|\boldsymbol{V}\|_{\boldsymbol{K},\boldsymbol{V}} = \min\{y : y\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{V} \succeq_{\boldsymbol{K}} -y\boldsymbol{V}\} = 1.$$

Otherwise, if $\|\boldsymbol{V}\|_{\boldsymbol{K},\boldsymbol{V}} < 1$, there exist $y < 1$ such that $y\boldsymbol{V} \succeq_{\boldsymbol{K}} \boldsymbol{V}$, which implies that $-\boldsymbol{V} \succeq_{\boldsymbol{K}} \mathbf{0}$, contradicting $\boldsymbol{V} \succ_{\boldsymbol{K}} \mathbf{0}$. Hence, $\|\boldsymbol{v}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}} = 1$ and we obtain

$$\left(\max_{\|\boldsymbol{x}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}}=1} \|\boldsymbol{x}\|_2\right) \geq \|\boldsymbol{v}\|_2.$$

Likewise, when $\boldsymbol{x}_n = (\boldsymbol{v}_n)/(\sqrt{2}\|\boldsymbol{v}_n\|_2)$ and $x_{n+1} = -1/(\sqrt{2})$, so that $\|\boldsymbol{x}\|_2 = 1$, we can also verify that the inequalities

$$\|\frac{\boldsymbol{v}_n}{\sqrt{2}\|\boldsymbol{v}_n\|_2} + \boldsymbol{v}_n y\|_2 \leq v_{n+1}y - \frac{1}{\sqrt{2}}$$
$$\|\frac{\boldsymbol{v}_n}{\sqrt{2}\|\boldsymbol{v}_n\|_2} - \boldsymbol{v}_n y\|_2 \leq v_{n+1}y + \frac{1}{\sqrt{2}}$$

hold if and only if $y \geq \sqrt{2}/(v_{n+1} - \|\boldsymbol{v}_n\|_2)$. Hence, $\|\boldsymbol{x}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}} = \sqrt{2}/(v_{n+1} - \|\boldsymbol{v}_n\|_2)$ and we obtain

$$\max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{x}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}} \geq \frac{\sqrt{2}}{v_{n+1} - \|\boldsymbol{v}_n\|_2}.$$

Therefore, since $0 < v_{n+1} - \|\boldsymbol{v}_n\|_2 \leq v_{n+1} \leq \|\boldsymbol{v}\|$, we have

$$\alpha_{\boldsymbol{L}^{n+1},\boldsymbol{v}} = \left(\max_{\|\boldsymbol{x}\|_2=1} \|\boldsymbol{x}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}}\right)\left(\max_{\|\boldsymbol{x}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}}=1} \|\boldsymbol{x}\|_2\right) \geq \frac{\sqrt{2}\|\boldsymbol{v}\|_2}{v_{n+1} - \|\boldsymbol{v}_n\|_2} \geq \sqrt{2}.$$

When $\boldsymbol{v} = (\mathbf{0}, 1)'$, we have

$$\|\boldsymbol{x}\|_{\boldsymbol{L}^{n+1},\boldsymbol{v}} = \|\boldsymbol{x}_n\|_2 + |x_{n+1}|,$$

and from Proposition 4(b), the bound is achieved. Hence, $\alpha_{\boldsymbol{L}^{n+1}} = \sqrt{2}$.
(b) Since $\boldsymbol{V}$ is an invertible matrix, we observe that

$$\begin{aligned}\|\boldsymbol{X}\|_{\boldsymbol{S}^m_+,\boldsymbol{V}} &= \min\{y : y\boldsymbol{V} \succeq \boldsymbol{X} \succeq -y\boldsymbol{V}\}\\ &= \min\left\{y : y\boldsymbol{I} \succeq \boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{X}\boldsymbol{V}^{-\frac{1}{2}} \succeq -y\boldsymbol{I}\right\}\\ &= \|\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{X}\boldsymbol{V}^{-\frac{1}{2}}\|_2.\end{aligned}$$

For any $\boldsymbol{V} \succ \mathbf{0}$, let $\boldsymbol{X} = \boldsymbol{V}$, we have $\|\boldsymbol{X}\|_{\boldsymbol{S}^m_+,\boldsymbol{V}} = 1$ and

$$\langle \boldsymbol{X}, \boldsymbol{X}\rangle = \text{trace}(\boldsymbol{V}\boldsymbol{V}) = \|\boldsymbol{\lambda}\|_2^2,$$

where $\boldsymbol{\lambda} \in \Re^m$ is a vector corresponding to all the eigenvalues of the matrix $\boldsymbol{V}$. Hence, we obtain

$$\left(\max_{\|\boldsymbol{X}\|_{\boldsymbol{S}^m_+,\boldsymbol{V}}=1} \sqrt{\langle \boldsymbol{X}, \boldsymbol{X}\rangle}\right) \geq \|\boldsymbol{\lambda}\|_2.$$

Without loss of generality, let λ_1 be the smallest eigenvalue of $\boldsymbol{V}$ with corresponding normalized eigenvector, $\boldsymbol{q}_1$. Now, let $\boldsymbol{X} = \boldsymbol{q}_1\boldsymbol{q}_1'$. Observe that

$$\begin{aligned}\langle \boldsymbol{X}, \boldsymbol{X}\rangle &= \text{trace}(\boldsymbol{X}\boldsymbol{X}) \\ &= \text{trace}(\boldsymbol{q}_1\boldsymbol{q}_1'\boldsymbol{q}_1\boldsymbol{q}_1') \\ &= \text{trace}(\boldsymbol{q}_1'\boldsymbol{q}_1\boldsymbol{q}_1'\boldsymbol{q}_1) \\ &= 1.\end{aligned}$$

We can express the matrix, $\boldsymbol{V}$ in its spectral decomposition, so that $\boldsymbol{V} = \sum_j \boldsymbol{q}_j\boldsymbol{q}_j'\lambda_j$. Hence,

$$\begin{aligned}\|\boldsymbol{X}\|_{S_+^m,\boldsymbol{V}} &= \|\boldsymbol{V}^{-\frac{1}{2}}\boldsymbol{X}\boldsymbol{V}^{-\frac{1}{2}}\|_2 \\ &= \|\textstyle\sum_j \boldsymbol{q}_j\boldsymbol{q}_j'\lambda_j^{-\frac{1}{2}}\boldsymbol{q}_1\boldsymbol{q}_1'\sum_j \boldsymbol{q}_j\boldsymbol{q}_j'\lambda_j^{-\frac{1}{2}}\|_2 \\ &= \|\lambda_1^{-1}\boldsymbol{q}_1\boldsymbol{q}_1'\|_2 \\ &= \lambda_1^{-1}.\end{aligned}$$

Therefore, we establish that

$$\left(\max_{\sqrt{\langle \boldsymbol{X},\boldsymbol{X}\rangle}=1} \|\boldsymbol{X}\|_{S_+^m,\boldsymbol{V}}\right) \geq \lambda_1^{-1}.$$

Combining the results, we have

$$\alpha_{S^m,\boldsymbol{V}} = \left(\max_{\|\boldsymbol{X}\|_{S_+^m,\boldsymbol{V}}=1} \sqrt{\langle \boldsymbol{X}, \boldsymbol{X}\rangle}\right)\left(\max_{\sqrt{\langle \boldsymbol{X},\boldsymbol{X}\rangle}=1} \|\boldsymbol{X}\|_{S_+^m,\boldsymbol{V}}\right) \geq \frac{\|\boldsymbol{\lambda}\|_2}{\lambda_1} \geq \sqrt{m}.$$

When $\boldsymbol{V} = \boldsymbol{I}$, we have

$$\|\boldsymbol{X}\|_{S^m,\boldsymbol{V}} = \|\boldsymbol{X}\|_2,$$

and from Proposition 4(d), the bound is achieved. Hence, $\alpha_{S^m} = \sqrt{m}$. □

We have shown that for homogeneous cones, while different $\boldsymbol{V}$ lead to the same probability bounds, some choices of $\boldsymbol{V}$ may lead to better objectives. The following theorem suggests an iterative improvement strategy.

Theorem 6. *For any $\boldsymbol{V} \succ_{\boldsymbol{K}} \boldsymbol{0}$, if $\boldsymbol{x}, y, \boldsymbol{t}$ are feasible in (24), then they are also feasible in the same problem in which $\boldsymbol{V}$ is replaced by*

$$\boldsymbol{W} = \mathcal{A}(\boldsymbol{x}, \boldsymbol{D}^0)/(\Omega y).$$

Proof. Observe that $\boldsymbol{W} \succeq_{\boldsymbol{K}} \boldsymbol{V} \succ_{\boldsymbol{K}} \boldsymbol{0}$ and it is trivial to check that the constraints are satisfied. □

Therefore, under this approach, the “best” choice of $\boldsymbol{V}$ satisfies,

$$\Omega y \boldsymbol{V} = \mathcal{A}(\boldsymbol{x}, \boldsymbol{D}^0)$$

Unfortunately, if $\boldsymbol{V}$ is variable, the convexity and possibly tractability of the model would be destroyed. The iterative improvement method of Theorem 6 can be an attractive heuristic.

A similar issue surfaces when we represent quadratic constraints as second order cones. In fact, there are more than one way of representing quadratic constraints as second order conic constraints. In particular, the constraint

$$\|\boldsymbol{A}\boldsymbol{x}\|_2^2 + \boldsymbol{b}'\boldsymbol{x} + c \leq 0$$

is equivalent to

$$\left\| \begin{bmatrix} \boldsymbol{A}\boldsymbol{x} \\ \frac{\lambda + \lambda^{-1}(\boldsymbol{b}'\boldsymbol{x} + c)}{2} \end{bmatrix} \right\|_2 \leq \frac{\lambda - \lambda^{-1}(\boldsymbol{b}'\boldsymbol{x} + c)}{2},$$

for any $\lambda > 0$. Unfortunately, the problem will not be convex if λ is made a variable. We leave it an open problem as to whether this could be done effectively.

6. Conclusions

We proposed a relaxed robust counterpart for general conic optimization problems that we believe achieves the objectives outlined in the introduction, namely:

(a) It preserves the computational tractability of the nominal problem. Specifically the robust conic optimization problem retains its original structure, i.e., robust LPs remain LPs, robust SOCPs remain SOCPs and robust SDPs remain SDPs. Moreover, the size of the proposed robust problem especially under the l_2 norm is practically the same as the nominal problem.

(b) It allows us to provide a guarantee on the probability that the robust solution is feasible, when the uncertain coefficients obey independent and identically distributed normal distributions.

A. Proof of Proposition 1

(a) Let $\boldsymbol{y} \in \arg\max_{\|\boldsymbol{x}\| \leq 1} \boldsymbol{w}'\boldsymbol{x}$, and for every $j \in N$, let $z_j = |y_j|$ if $w_j \geq 0$ and $z_j = -|y_j|$, otherwise. Clearly, $\boldsymbol{w}'\boldsymbol{z} = (|\boldsymbol{w}|)'(|\boldsymbol{y}|) \geq \boldsymbol{w}'\boldsymbol{y}$. Since, $\|\boldsymbol{z}\| = \||\boldsymbol{z}|\| = \||\boldsymbol{y}|\| = \|\boldsymbol{y}\| \leq 1$, and from the optimality of $\boldsymbol{y}$, we have $\boldsymbol{w}'\boldsymbol{z} \leq \boldsymbol{w}'\boldsymbol{y}$, leading to $\boldsymbol{w}'\boldsymbol{z} = (|\boldsymbol{w}|)'(|\boldsymbol{y}|) = \boldsymbol{w}'\boldsymbol{y}$. Since $\|\boldsymbol{w}\| = \||\boldsymbol{w}|\|$, we obtain

$$\|\boldsymbol{w}\|^* = \max_{\|\boldsymbol{x}\| \leq 1} (\boldsymbol{w})'\boldsymbol{x} = \max_{\|\boldsymbol{x}\| \leq 1} (|\boldsymbol{w}|)'(|\boldsymbol{x}|) = \max_{\|\boldsymbol{x}\| \leq 1} (|\boldsymbol{w}|)'\boldsymbol{x} = \||\boldsymbol{w}|\|^*.$$

(b) Note that

$$\|\boldsymbol{w}\|^* = \max_{\|\boldsymbol{x}\|\leq 1} (|\boldsymbol{w}|)'(|\boldsymbol{x}|) = \max_{\substack{\|\boldsymbol{x}\|\leq 1\\ \boldsymbol{x}\geq 0}} (|\boldsymbol{w}|)'\boldsymbol{x}.$$

If $|\boldsymbol{v}| \leq |\boldsymbol{w}|$,

$$\|\boldsymbol{v}\|^* = \max_{\substack{\|\boldsymbol{x}\|\leq 1\\ \boldsymbol{x}\geq 0}} (|\boldsymbol{v}|)'\boldsymbol{x} \leq \max_{\substack{\|\boldsymbol{x}\|\leq 1\\ \boldsymbol{x}\geq 0}} (|\boldsymbol{w}|)'\boldsymbol{x} = \|\boldsymbol{w}\|^*.$$

(c) We apply part (b) to the norm $\|.\|^*$. From the self dual property of norms $\|.\|^{**} = \|.\|$, we obtain part (c). □

B. Simplified formulation under independent uncertainties

In this section, we show that if each data entry of the model has independent uncertainty, we can substantially reduce the size of the robust formulation (14). We focus on the equivalent representation (13),

$$f(\boldsymbol{x}, \boldsymbol{D}^0) \geq \Omega y, \ \|\boldsymbol{s}\|^* \leq y,$$

where, $s_j = \max\{-f(\boldsymbol{x}, \boldsymbol{\Delta D}^j), -f(\boldsymbol{x}, -\boldsymbol{\Delta D}^j)\} = g(\boldsymbol{x}, \boldsymbol{\Delta D}^j)$, for $j \in N$.

Proposition 8. *For LP, QCQP, SOCP(1), SOCP(2) and SDP, we can express $s_j = |\Delta d_j x_{i(j)}|$ for which Δd_j, $j \in N$ are constants and the function, $i : N \to \{0, \ldots, n\}$ maps $j \in N$ to the index of the corresponding variable. We define $x_0 = 1$, to address the case when s_j is not variable dependent.*

Proof. We associate the jth data entry, $j \in N$ with an iid random variable $\tilde{z}_j$. The corresponding expression of $g(\boldsymbol{x}, \boldsymbol{\Delta D}^j)$ is shown in Table 3.

(a) LP:
Uncertain LP data is represented as $\tilde{\boldsymbol{D}} = (\tilde{\boldsymbol{a}}, \tilde{b})$, where

$$\begin{aligned} \tilde{a}_j &= a_j^0 + \Delta a_j \tilde{z}_j, \ \ j = 1, \ldots, n \\ \tilde{b} &= b^0 + \Delta b \tilde{z}_{n+1}. \end{aligned}$$

We have $|N| = n + 1$ and

$$\begin{aligned} s_j &= |\Delta a_j x_j|, \ \ j = 1, \ldots, n \\ s_{n+1} &= |\Delta b|. \end{aligned}$$

(b) QCQP:
Uncertain QCQP data is represented as $\tilde{\boldsymbol{D}} = (\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{b}}, \tilde{c}, 1)$, where

$$\begin{aligned} \tilde{A}_{kj} &= A_{kj}^0 + \Delta A_{kj} \tilde{z}_{n(k-1)+j}, \ \ j 1, \ldots, n, \ \ k = 1, \ldots, l, \\ \tilde{b}_j &= b_j^0 + \Delta b_j \tilde{z}_{nl+j}, \ \ j = 1, \ldots, n, \\ \tilde{c} &= c^0 + \Delta c \tilde{z}_{n(l+1)+1}. \end{aligned}$$

We have $|N| = n(l+1)+1$ and

$$\begin{aligned} s_{n(k-1)+j} &= |\Delta A_{kj} x_j|, \ j = 1, \ldots, n, \ k = 1, \ldots, l, \\ s_{nl+j} &= |\Delta b_j x_j|, \ j = 1, \ldots, n, \\ s_{n(l+1)+1} &= |\Delta c|. \end{aligned}$$

(c) SOCP(1)/SOCP(2):
Uncertain SOCP(2) data is represented as $\tilde{\boldsymbol{D}} = (\tilde{\boldsymbol{A}}, \tilde{\boldsymbol{b}}, \tilde{\boldsymbol{c}}, d)$, where

$$\begin{aligned} \tilde{A}_{kj} &= A_{kj}^0 + \Delta A_{kj} \tilde{z}_{n(k-1)+j}, \ j = 1, \ldots, n, \ k = 1, \ldots, l, \\ \tilde{b}_k &= b_k^0 + \Delta b_k \tilde{z}_{nl+k}, \ k = 1, \ldots, l, \\ \tilde{c}_j &= c_j^0 + \Delta c_j \tilde{z}_{(n+1)l+j}, \ j = 1, \ldots, n, \\ \tilde{d} &= d^0 + \Delta d \tilde{z}_{(n+1)l+n+1}. \end{aligned}$$

We have $|N| = (n+1)l + n + 1$ and

$$\begin{aligned} s_{n(k-1)+j} &= |\Delta A_{kj} x_j|, \ j = 1, \ldots, n, \ k = 1, \ldots, l, \\ s_{nl+k} &= |\Delta b_k|, \ j = 1, \ldots, l, \\ s_{(n+1)l+j} &= |\Delta c_j x_j|, \ j = 1, \ldots, n, \\ s_{(n+1)l+n+1} &= |\Delta d|. \end{aligned}$$

Note that SOCP(1) is a special case of SOCP(2), for which $|N| = (n+1)l$, that is, $s_j = 0$ for all $j > (n+1)l$.

(d) SDP:
Uncertain SDP data is represented as $\tilde{\boldsymbol{D}} = (\tilde{\boldsymbol{A}}_1, \ldots, \tilde{\boldsymbol{A}}_n, \tilde{\boldsymbol{B}})$, where

$$\begin{aligned} \tilde{\boldsymbol{A}}_i &= \boldsymbol{A}_i^0 + \textstyle\sum_{k=1}^m \sum_{j=1}^k [\Delta \boldsymbol{A}_i]_{jk} \boldsymbol{I}_{jk} \tilde{z}_{p(i,j,k)} \ i = 1, \ldots, n, \\ \tilde{\boldsymbol{B}} &= \boldsymbol{B}^0 + \textstyle\sum_{k=1}^n \sum_{j=1}^k [\Delta \boldsymbol{B}]_{jk} \boldsymbol{I}_{jk} \tilde{z}_{p(n+1,j,k)}, \end{aligned}$$

where the index function $p(i, j, k) = (i-1)(m(m+1)/2) + k(k-1)/2 + j$, and the symmetric matrix $\boldsymbol{I}_{jk} \in \Re^{m \times m}$ satisfies,

$$\boldsymbol{I}_{jk} = \begin{cases} (\boldsymbol{e}_j \boldsymbol{e}_k' + \boldsymbol{e}_k \boldsymbol{e}_j') & \text{if } j \neq k \\ \boldsymbol{e}_k \boldsymbol{e}_k' & \text{if } k = j \end{cases}$$

$\boldsymbol{e}_k$ being the kth unit vector. Hence, $|N| = (n+1)(m(m+1))/2$. Note that if $j = k$, $\|\boldsymbol{I}_{jk}\|_2 = 1$. Otherwise, $\boldsymbol{I}_{jk}$ has rank 2 and $(\boldsymbol{e}_j + \boldsymbol{e}_k)/\sqrt{2}$ and $(\boldsymbol{e}_j - \boldsymbol{e}_k)/\sqrt{2}$ are two eigenvectors of $\boldsymbol{I}_{jk}$ with corresponding eigenvalues 1 and -1. Hence, $\|\boldsymbol{I}_{jk}\|_2 = 1$ for all valid indices j and k. Therefore, we have

$$\begin{aligned} s_{p(i,j,k)} &= |[\Delta \boldsymbol{A}_i]_{jk} x_i|, \ \ \forall i \in \{1, \ldots, n\}, j, k \in \{1, \ldots, m\}, j \le k \\ s_{p(n+1,j,k)} &= |[\Delta \boldsymbol{B}]_{jk}|, \ \ \forall j, k \in \{1, \ldots, m\}, j \le k. \end{aligned}$$

□

We define the set $J(l) = \{j : i(j) = l,\ j \in N\}$ for $l \in \{0, \dots, n\}$. From Table 2, we have the following robust formulations under the different norms in the restriction set $\mathcal{V}$ of Eq. (9).

(a) l_∞-norm
The constraint $\|\boldsymbol{s}\|^* \leq y$ for the l_∞-norm is equivalent to

$$\sum_{j\in N} |\Delta d_j x_{i(j)}| \leq y \Leftrightarrow \sum_{l=0}^{n} \left(\sum_{j \in J(l)} |\Delta d_j| \right) |x_l| \leq y$$

or

$$\begin{aligned}
&\textstyle\sum_{j\in J(0)} |\Delta d_j| + \sum_{l=1}^{n} \left(\sum_{j\in J(l)} |\Delta d_j| \right) t_l \leq y \\
&\boldsymbol{t} \geq \boldsymbol{x}, \boldsymbol{t} \geq -\boldsymbol{x} \\
&\boldsymbol{t} \in \Re^n.
\end{aligned}$$

We introduce additional $n + 1$ variables, including the variable y, and $2n + 1$ linear constraints to the nominal problem.

(b) l_1-norm
The constraint $\|\boldsymbol{s}\|^* \leq y$ for the l_1-norm is equivalent to

$$\max_{j\in N} |\Delta d_j x_{i(j)}| \leq y \Leftrightarrow \max_{l\in\{0,\dots,n\}} \left(\max_{j \in J(l)} |\Delta d_j| \right) |x_l| \leq y$$

or

$$\begin{aligned}
\max\nolimits_{j\in J(0)} |\Delta d_j| &\leq y \\
\max\nolimits_{j\in J(l)} |\Delta d_j| x_l &\leq y \quad l = 1, \dots, n \\
-\max\nolimits_{j\in J(l)} |\Delta d_j| x_l &\leq y \quad l = 1, \dots, n.
\end{aligned}$$

We introduce an additional variable and $2n + 1$ linear constraints to the nominal problem.

(c) $l_1 \cap l_\infty$-norm
The constraint $\|\boldsymbol{s}\|^* \leq y$ for the $l_1 \cap l_\infty$-norm is equivalent to

$$\begin{aligned}
&t_j \geq |\Delta d_j| x_{i(j)} && j \in N \\
&t_j \geq -|\Delta d_j| x_{i(j)} && j \in N \\
&\Gamma p + \textstyle\sum_{j\in N} r_j \leq y \\
&r_j + p \geq t_j,\ \forall j \in N \\
&\boldsymbol{r} \in \Re_+^{|N|},\ \boldsymbol{t} \in \Re^{|N|},\ p \in \Re_+,
\end{aligned}$$

leading to an additional of $2|N| + 2$ variables and $4|N| + 2$ linear constraints, including non-negativity constraints, to the nominal problem.

(d) l_2-norm
The constraint $\|\boldsymbol{s}\|^* \leq y$ for the l_2-norm is equivalent to

$$\sqrt{\sum_{j\in N}(\Delta d_j x_{i(j)})^2} \leq y \Leftrightarrow \sqrt{\sum_{j\in J(0)} |\Delta d_j| + \sum_{l=1}^{n}\left(\sum_{j\in J(l)} \Delta d_j^2\right) x_l^2} \leq y.$$

We only introduce an additional variable, y and one SOCP constraint to the nominal problem.

(e) $l_2 \cap l_\infty$-norm
The constraint $\|\boldsymbol{s}\|^* \leq y$ for the $l_2 \cap l_\infty$-norm is equivalent to

$$\begin{aligned} & t_j \geq |\Delta d_j| x_{i(j)} && j \in N \\ & t_j \geq -|\Delta d_j| x_{i(j)} && j \in N \\ & \|\boldsymbol{r} - \boldsymbol{t}\|_2 + \tfrac{1}{\Omega}\textstyle\sum_{j\in N} r_j \leq y \\ & \boldsymbol{t} \in \Re^{|N|},\ \boldsymbol{r} \in \Re_+^{|N|}. \end{aligned}$$

We introduce adds $2|N| + 1$ variables, one SOCP constraint and $3|N|$ linear constraints, including non-negativity constraints, to the nominal problem.

In Table 4, we summarize the size increase and the nature of the robust model for different choices of the given norm.

Acknowledgements. We would like to thank the reviewers of the paper for several very insightful comments.

References

1. Ben-Tal, A., Nemirovski, A.: Robust convex optimization. Math. Oper. Res. **23**, 769–805 (1998)
2. Ben-Tal, A., Nemirovski, A.: On the quality of SDP approximations of uncertain SDP programs, Research Report #4/98 Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology, Israel (1998)
3. Ben-Tal, A., Nemirovski, A.: Robust solutions to uncertain programs. Oper. Res. Let. **25**, 1–13 (1999)
4. Ben-Tal, A., Nemirovski, A.: Robust solutions of Linear Programming problems contaminated with uncertain data, Math. Progr. **88**, 411–424 (2000)
5. Ben-Tal, A., Nemirovski, A.: Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications. MPR-SIAM Series on Optimization, SIAM, Philadelphia 2001
6. Ben-Tal, A., El-Ghaoui, L., Nemirovski, A.: Robust semidefinite programming. In: Saigal, R., Vandenberghe, L., Wolkowicz, H., (eds.), Semidefinite programming and applications, Kluwer Academic Publishers, (2000)
7. Bertsimas, D., Brown, D.: Constrainted stochastic LQC: A tractable approach. submitted for publication, 2004
8. Bertsimas, D., Pachamanova, D., Sim, M.: Robust Linear Optimization under General Norms. Operations Research Letters, **32**, 510–516 (2003)
9. Bertsimas, D., Sim, M.: Price of Robustness. Oper. Res. **52** (1), 35–53 (2004)
10. Bertsimas, D., Sim, M.: Robust Discrete Optimization and Network Flows. Math. Progr. **98**, 49–71 (2003)
11. Birge, J.R., Louveaux, F.: Introduction to Stochastic Programming. Springer, New York, 1997
12. El-Ghaoui, Lebret, H.: Robust solutions to least-square problems to uncertain data matrices. SIAM J. Matrix Anal. Appl. **18**, 1035–1064 (1997)
13. El-Ghaoui, L., Oustry, F., Lebret, H.: Robust solutions to uncertain semidefinite programs. SIAM J. Optim. **9**, 33–52 (1998)
14. Güler, 0, Tunçel, L.: Characterization of the barrier parameter of homogeneous convex cones. Math. Progr. **81**, 55–76 (1998)

15. Nemirovski, A.: On tractable approximations of ramdomly perturbed convex constraints. In: Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, Hawaii, USA, December, pp. 2419–2422, 2003
16. Nemirovski, A.: Regular Banach spaces and large deviations of random sums. http://iew3.technion.ac.il/Home/Users/Nemirovski.html#part4 2004
17. Renegar, J.: A mathematical view of interior-point methods in convex optimization. MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001
18. Soyster, A.L.: Convex programming with set-inclusive constraints and applications to inexact linear programming. Oper. Res. **21**, 1154–1157 (1973)