Robust optimization for models with uncertain SOC and SDP constraints

Jianzhe Zhen*
College of Management of Technology, École Polytechnique Fédérale de Lausanne, Switzerland
Department of Econometrics and Operations Research, Tilburg University, Tilburg, The Netherlands

Frans J.C.T. de Ruiter
CQM, Eindhoven, The Netherlands
Department of Econometrics and Operations Research, Tilburg University, Tilburg, The Netherlands

Ernst Roos Dick den Hertog
Department of Econometrics and Operations Research, Tilburg University, Tilburg, The Netherlands

In this paper we consider uncertain second-order cone (SOC) and semidefinite programming (SDP) constraints with polyhedral uncertainty, which are in general computationally intractable. We propose to reformulate an uncertain SOC or SDP constraint as a set of adjustable robust linear optimization constraints with an ellipsoidal or semidefinite representable uncertainty set, respectively. The resulting adjustable problem can then (approximately) be solved by using adjustable robust linear optimization techniques. For example, we show that if linear decision rules are used, then the final robust counterpart consists of SOC or SDP constraints, respectively, which have the same computational complexity as the nominal version of the original constraints. We propose an efficient method to obtain good lower bounds. Moreover, we extend our approach to other classes of robust optimization problems, such as nonlinear problems that contain wait-and-see variables or linear problems that contain bilinear uncertainty. Numerically, we apply our approach to reformulate the problem on finding the minimum volume circumscribing ellipsoid of a polytope, and solve the resulting reformulation with linear and quadratic decision rules as well as Fourier-Motzkin elimination. We demonstrate the effectiveness and efficiency of the proposed approach by comparing it with the state-of-the-art copositive approach. Moreover, we apply the proposed approach to a robust regression problem and a robust sensor network problem, and use linear decision rules to solve the resulting adjustable robust linear optimization problems, which solves the problem to (near) optimality.

Key words: Robust optimization, second-order cone, semidefinite programming, adjustable robust optimization, linear decision rules.

1. Introduction

Practical optimization problems often contain uncertain parameters. This uncertainty arises, because of, e.g., estimation or prediction errors. One way of dealing with uncertainty is *robust optimization*. The papers El Ghaoui and Lebret (1997), El Ghaoui et al. (1998) and Ben-Tal and Nemirovski (1998) are considered as the birth of this field.

* Corresponding author: jianzhe.zhen@epfl.ch

In robust optimization the uncertainty is not modeled by probability distributions as in *stochastic optimization*, but as uncertainty sets. An uncertainty set contains all scenarios for the uncertain parameters against which the decision maker would like to safeguard herself. The constraints are enforced to hold for all scenarios in this uncertainty set.

The paper Ben-Tal et al. (2004) extends the robust optimization methodology to problems that also contain wait-and-see or adjustable variables. Such variables often occur in multi-stage problems. Adjustable variables model decisions that can be delayed until the values of (a part of) the uncertain parameters have been revealed. Many efficient methods have been proposed in literature to (approximately) solve such *adjustable robust optimization* problems.

The advantages of robust optimization are, among others, the computational tractability and the fact that there is no need to specify a probability distribution. Many classes of robust optimization problems have been shown to be equivalent to tractable formulations. Many of these cases are treated in the book Ben-Tal et al. (2009). A detailed and unified approach to derive computationally tractable reformulations is given in Ben-Tal et al. (2015). In that paper it is shown that, loosely speaking, convex reformulations exist for constraints that are concave in the uncertain parameters, and convex in the optimization variables.

For several problems that contain constraints that are not concave in the uncertain parameters, computationally tractable approximations have also been proposed. For robust **s**econd-**o**rder **c**one (SOC) and robust **s**emi**d**efinite **p**rogramming (SDP) constraints that are convex in the uncertain parameters, exact and approximate convex reformulations for specific (simple) ellipsoidal or norm-bounded uncertainty sets have been proposed by El Ghaoui and Lebret (1997), El Ghaoui et al. (1998), and Ben-Tal et al. (2002), which are summarized in the book Ben-Tal et al. (2009). In all these approaches, both for uncertain SOC and SDP constraints, the final *robust counterpart* contains an SDP constraint. We are not aware of papers that deal with uncertain SOC or SDP constraints with general polyhedral uncertainty.

In this paper we consider uncertain SOC and SDP constraints with polyhedral uncertainty. We propose to reformulate an uncertain SOC or SDP constraint as a set of *adjustable robust linear optimization* constraints with an ellipsoidal or semidefinite representable uncertainty set, respectively. The resulting adjustable problem can then (approximately) be solved by using adjustable robust linear optimization techniques described in the literature. For example, we show that if *linear decision rules* are used, then the final robust counterpart consists of SOC or SDP constraints, respectively, which have the same computational complexity as the nominal version of the original constraints. We also propose an efficient method to obtain good lower bounds. Moreover, we extend our approach to other classes of robust optimization problems, such as nonlinear problems that contain wait-and-see variables and linear problems that contain bilinear uncertainty. Numerically, we apply our approach to reformulate the problem on finding the minimum volume circumscribing ellipsoid of a polytope, and solve the resulting reformulation with linear and quadratic decision rules as well as Fourier-Motzkin elimination. We demonstrate the effectiveness and efficiency of the proposed approach by comparing it with the state-of-the-art copositive approach of Mittal and Hanasusanto (2018). Contrary to existing methods, in our approach we also obtain lower bounds for the minimum volume of the circumscribing ellipsoid. Numerical experiments show that these bounds are very good. Moreover, we apply the proposed approach to a robust regression problem and a robust sensor network problem, and use linear decision rules to solve the resulting adjustable robust linear optimization problems, which solves the problem to (near) optimality.

This paper is organized as follows. In §2 we treat uncertain SOC constraints, and in §3 uncertain SDP constraints with polyhedral uncertainty. §4 describes how to obtain sharp lower bounds in an efficient way. In §5 extensions of our approach to other classes of robust optimization problems are given. §6, §7 and §8 contain the numerical results for finding the minimum volume circumscribing ellipsoid of a polytope, the robust regression and a sensor network problem. §9 contains recommendations for future research.

2. Uncertain second-order cone constraints

Consider the following uncertain second-order cone constraint:

$$\forall \zeta \in \mathcal{U}: \quad a(x)^\top \zeta + \|A(x)\zeta + b(x)\|_2 \leq c(x). \tag{1}$$

Here $||\cdot||_2$ denotes the l_2-norm, $x \in \mathbb{R}^{n_x}$ is the decision (or optimization) variable in a given domain $\mathcal{X} \subseteq \mathbb{R}^{n_x}$, e.g., $\mathcal{X} = \mathbb{R}^{n_x}_+$ or $\mathcal{X} = \mathbb{Z}^{n_x}_+$, $\zeta \in \mathbb{R}^n$ is the uncertain parameter that resides in the uncertainty set $\mathcal{U} \subset \mathbb{R}^n$, and the vectors $a(x) = (a_1(x), \ldots, a_n(x))^\top$ and $b(x) = (b_1(x), \ldots, b_m(x))^\top$, $A(x) \in \mathbb{R}^{m \times n}$, and $c(x) \in \mathbb{R}$ have entries that are general functions of x. We demonstrate the modelling power of (1) via the following examples.

Example 1. **Robust Regression.** In regression models we try to find a vector of coefficients $x \in \mathbb{R}^{n_x}$ such that the norm (or squared norm) of $Ax - b$ is minimized. The standard least-squares solution is the optimal solution to the model:

$$\min_{x \in \mathbb{R}^{n_x}} \|Ax - b\|_2 .$$

Here $A \in \mathbb{R}^{m \times n_x}$ and $b \in \mathbb{R}^m$ are observed data, in which each row of the matrix A is a different observation and the columns refer to the features. The i-th entry of b corresponds to the response, or target value, of the i-th observation. Often, some of the data entries in A and/or b are obtained via measurements, and therefore subject to uncertainty. Suppose there are uncertainties in the entries of the matrix A. For robust regression, we can replace the matrix A in the least-squares model by the term $A+\zeta$, where $\zeta\in\mathbb{R}^{m\times n_x}$ is a matrix with uncertain parameters, and minimize τ subject to:

$$\forall\zeta\in\mathcal{U}\ :\|(A+\zeta)x-b\|_2\leq\tau, \tag{2}$$

where $\tau\in\mathbb{R}$ is an optimization variable. □

Robust regression models and uncertain quadratic constraints with specific norm bounded uncertainty sets were studied by El Ghaoui and Lebret (1997). The method described in this paper can also deal with uncertain quadratic constraints, with polyhedral uncertainty, by reformulating it as an uncertain second-order cone constraint.

Example 2. **Uncertain quadratic constraints.** Consider the following constraint:

$$\forall\zeta\in\mathcal{U}:\quad \zeta^\top H(x)^\top H(x)\zeta+f(x)^\top\zeta\leq g(x),$$

where the entries of $H:\mathbb{R}^{n_x}\to\mathbb{R}^{n\times n}$, $f:\mathbb{R}^{n_x}\to\mathbb{R}^n$ and $g:\mathbb{R}^{n_x}\to\mathbb{R}$ are affine functions. This is equivalent to an uncertain second-order cone constraint in the form of (1):

$$\forall\zeta\in\mathcal{U}:\ \left\|\begin{matrix}\left(1+f(x)^\top\zeta-g(x)\right)/2\\ H(x)\zeta\end{matrix}\right\|_2\leq\left(1-f(x)^\top\zeta+g(x)\right)/2.$$

□

Throughout this paper, we focus on nonempty polyhedral uncertainty sets of the form:

$$\mathcal{U}=\{\zeta\geq 0:\ D\zeta\leq d\}, \tag{3}$$

with $D\in\mathbb{R}^{r\times n}$ and $d\in\mathbb{R}^r$. Constraint (1) is equivalent to:

$$\max_{\zeta\in\mathcal{U}}\ \left\{a(x)^\top\zeta+\|A(x)\zeta+b(x)\|_2\right\}\leq c(x). \tag{4}$$

The left-hand side of inequality (4) is a convex maximization problem, which is in general computationally intractable even if $\mathcal{U}$ is a polyhedron. In the following theorem we propose a novel technique to reformulate the second-order cone constraint (1) into a set of *two-stage robust linear constraints.*

Theorem 1. *Let $\mathcal{U}$ be the nonempty polyhedral uncertainty set given in* (3)*. Then $x\in\mathbb{R}^{n_x}$ satisfies constraint* (1) *if and only if it satisfies the following set of two-stage robust linear constraints:*

$$\forall w\in\mathcal{W}\ \exists\lambda\geq 0:\ \begin{cases}d^\top\lambda+b(x)^\top w\leq c(x)\\ D^\top\lambda\geq a(x)+A(x)^\top w,\end{cases} \tag{5}$$

where $\mathcal{W}=\{w\in\mathbb{R}^m:\|w\|_2\leq 1\}$ and $\lambda\in\mathbb{R}^r$.

Proof. For constraint (1) we can derive the following equivalences:

$$\begin{aligned} & \forall \zeta \in \mathcal{U}: \quad a(x)^\top \zeta + \max_{w:\ \|w\|_2 \le 1} w^\top (A(x)\zeta + b(x)) \le c(x) \\ \iff \quad & \forall w \in \mathcal{W}\ \forall \zeta \in \mathcal{U}: \quad a(x)^\top \zeta + w^\top (A(x)\zeta + b(x)) \le c(x), \end{aligned} \tag{6}$$

with $\mathcal{W} = \{w \in \mathbb{R}^m : \|w\|_2 \le 1\}$. By dualizing over ζ, using strong duality for linear optimization, we can further deduce that (6) is equivalent to:

$$\begin{aligned} & \forall w \in \mathcal{W}: \quad \max_{\zeta \in \mathcal{U}} \left\{ a(x)^\top \zeta + w^\top (A(x)\zeta + b(x)) \right\} \le c(x) \\ \iff \quad & \forall w \in \mathcal{W}: \quad w^\top b(x) + \min_{\lambda \ge 0} \left\{ d^\top \lambda \mid D^\top \lambda \ge a(x) + A(x)^\top w \right\} \le c(x) \\ \iff \quad & \forall w \in \mathcal{W}, \exists \lambda \ge 0: \begin{cases} d^\top \lambda + b(x)^\top w \le c(x) \\ D^\top \lambda \ge a(x) + A(x)^\top w. \end{cases} \end{aligned}$$

□

As the result of the reformulation, the newly introduced variables w and λ appear linearly in constraints (5). The set of constraints (5) can be seen as the constraints of a two-stage robust linear optimization model where w, that resides in an ellipsoidal uncertainty set $\mathcal{W}$, can be considered as the uncertain parameter. The first-stage or *here-and-now* decision x is decided before the realization of the uncertainty parameter w, and the second-stage or *wait-and-see* decision λ is determined after the value of w is revealed. The coefficients of λ (i.e., d and D) are constant, which corresponds to the stochastic optimization format known as *fixed recourse*. Two-stage robust linear optimization models are in general intractable to solve to optimality, because the wait-and-see decision is a decision rule, or infinite dimensional variable, instead of a finite vector of decision variables (see Ben-Tal et al. (2004)).

Theoretically, for several two-stage robust linear models, the structure of the optimal decision rules has been characterized. For instance, one can use Zhen et al. (2018) to establish the optimality of the following decision rules for λ in (5). There exists a piecewise affine function that is optimal for λ in (5). More specifically, if $\mathcal{U}$ is simplicial, there exists a linear decision rule that is optimal for λ in (5); if $\mathcal{U}$ is a box, there exists a two-piecewise affine function that is optimal for λ in (5), and the techniques proposed in Gorissen and den Hertog (2013) to solve (5) approximately. Unfortunately, even when the structure of optimal decision rules is known, it is often hard to find optimal solutions due to the computational intractability of such rules.

Numerically, the main advantage of (5) is that it can be (approximately) solved by any method applicable to two-stage robust linear models such as linear decision rules (see Ben-Tal et al. (2004)), Fourier-Motzkin elimination (see Zhen et al. (2018)), finite adaptability approaches (see Postek and den Hertog (2016), Bertsimas and Dunning (2016), Georghiou et al. (2017)), etc. These solution methods will be discussed in §4.1. Numerical experiments with uncertain second-order cone constraints are conducted in §6, §7 and §8 to evaluate the performance of the proposed methods.

We note that the condition $\zeta \geq 0$ in the uncertainty set $\mathcal{U}$ can be omitted. In that case, the result of Theorem 1 includes equality constraints $D^\top \lambda = a(x) + A(x)^\top w$ instead. These equalities can be used to eliminate some of λ via Gaussian elimination. It is well-known that eliminating the wait-and-see variables in the equalities of a two-stage fixed-recourse robust model is equivalent to imposing linear decision rules (Zhen and den Hertog 2017a, Lemma 2).

3. Uncertain semidefinite programming constraints

Consider the following uncertain semidefinite programming constraint:

$$\forall \zeta \in \mathcal{U}: \quad A(x,\zeta) \succeq 0, \qquad \text{where} \qquad A(x,\zeta) = A^{(0)}(x) + \sum_{i=1}^{n} A^{(i)}(x)\zeta_i, \tag{7}$$

and the components of $A^{(i)} : \mathbb{R}^{n_x} \to \mathbb{R}^{m\times m}$, $i = 0,\ldots,n$, are general functions in $x \in \mathcal{X}$. The following theorem shows that an uncertain semidefinite programming constraint with polyhedral uncertainty can also be reformulated into a set of two-stage robust linear constraints with a semidefinite representable uncertainty set.

Theorem 2. *Let $\mathcal{U}$ be a nonempty polyhedral uncertainty set as in* (3)*. Then $x \in \mathbb{R}^{n_x}$ satisfies constraint* (7) *if and only if it satisfies*

$$\forall W \succeq 0 \ \exists \lambda \geq 0: \quad \begin{cases} Tr\left(A^{(0)}(x)W\right) - d^\top \lambda \geq 0 \\ D_i^\top \lambda \geq -Tr\left(A^{(i)}(x)W\right) \qquad i = 1,\ldots,n, \end{cases} \tag{8}$$

where $Tr(\cdot)$ denotes the trace function, $\lambda \in \mathbb{R}^r$ and D_i is the i-th column of D for $i = 1,\ldots,n$.

Proof. From Lemma 3 (see Appendix A) we know that a matrix $A(x,\zeta)$ is positive semidefinite if and only if the trace of the product with any positive semidefinite matrix is positive.
For constraint (7) we then can derive the following equivalences:

$$\begin{aligned} & \forall \zeta \in \mathcal{U}: \quad A(x,\zeta) \succeq 0 \\ \iff \quad & \forall W \succeq 0 \ \forall \zeta \in \mathcal{U}: \quad \mathrm{Tr}\left(A(x,\zeta)W\right) \geq 0 \\ \iff \quad & \forall W \succeq 0 \ \forall \zeta \in \mathcal{U}: \quad \mathrm{Tr}\left(A^{(0)}(x)W\right) + \sum_{i=1}^{n} \mathrm{Tr}\left(A^{(i)}(x)W\right)\zeta_i \geq 0 \\ \iff \quad & \forall W \succeq 0: \quad \mathrm{Tr}\left(A^{(0)}(x)W\right) + \min_{\zeta \in \mathcal{U}} \left\{ \sum_{i=1}^{n} \mathrm{Tr}\left(A^{(i)}(x)W\right)\zeta_i \right\} \geq 0. \end{aligned}$$

By dualizing over ζ, using strong duality for linear programming, we obtain:

$$\begin{aligned} & \forall W \succeq 0: \quad \mathrm{Tr}\left(A^{(0)}(x)W\right) + \max_{\lambda \geq 0}\left\{-d^\top \lambda \mid D_i^\top \lambda \geq -\mathrm{Tr}\left(A^{(i)}(x)W\right),\ i=1,\dots,n\right\} \geq 0 \\ \iff \quad & \forall W \succeq 0\ \exists \lambda \geq 0: \quad \begin{cases} \mathrm{Tr}\left(A^{(0)}(x)W\right) - d^\top \lambda \geq 0 \\ D_i^\top \lambda \geq -\mathrm{Tr}\left(A^{(i)}(x)W\right) \qquad i=1,\dots,n. \end{cases} \end{aligned}$$

□

Notice that since the system (8) is homogeneous in λ and W, one can in fact replace the unbounded uncertainty set '$\forall W \succeq 0$' by the bounded set '$\forall W : I \succeq W \succeq 0$' without affecting the feasible region of x, where $I \in \mathbb{R}^{m\times m}$ denotes the identity matrix. Any solution method applicable for two-stage robust optimization models can be used to solve problems with constraints (8). These solution methods will be discussed in §4.

4. Convex conservative and progressive approximations

In order to construct conservative and progressive approximations of the constraints (1) and (7) that are convex, we first assume that $-c : \mathbb{R}^{n_x} \to \mathbb{R}$ and $a_i : \mathbb{R}^{n_x} \to \mathbb{R}$, $i = 1,\dots,n$, are convex functions in x, and $b_j : \mathbb{R}^{n_x} \to \mathbb{R}$, $j = 1,\dots,m$, and the components of $A : \mathbb{R}^{n_x} \to \mathbb{R}^{m\times m}$ in constraint (1), and the components of $A^{(i)} : \mathbb{R}^{n_x} \to \mathbb{R}^{m\times m}$, $i = 0,...,n$, in constraint (7) are affine in x.

4.1. Conservative approximation

One popular remedy for the intractability of two-stage robust linear optimization models is to restrict the wait-and-see decisions in (5) and (8) to be simple functions of the uncertain parameters, e.g., linear decision rules (also known as, affine policies, see Ben-Tal et al. (2004)). In the following lemma we present the convex conservative approximation of constraints in (5) via linear decision rules, which is also a conservative approximation of (1).

LEMMA 1. *The vector $x \in \mathcal{X}$ satisfies constraint* (5) *if there exist $v \in \mathbb{R}^r$ and $V \in \mathbb{R}^{r\times m}$ such that x also satisfies:*

$$\begin{cases} d^\top v + \left\|V^\top d + b(x)\right\|_2 \leq c(x) \\ a_i(x) + \left\|A_i(x) - V^\top D_i\right\|_2 \leq D_i^\top v \qquad i=1,\dots,n \\ \left\|(V^\top)_j\right\|_2 \leq v_j, \qquad j=1,\dots,r, \end{cases} \tag{9}$$

where a_i and v_i denote the i-th elements of a and v, respectively, and A_i, D_i and $(V^\top)_j$ denote the i-th column of A, D and $V^\top$, respectively.

Proof. By restricting λ to the linear decision rule in (5):

$$\lambda = v + Vw,$$

we obtain the following conservative approximation of (5):

$$\forall w \in \mathcal{W} \ : \ \begin{cases} d^\top (v + Vw) + b(x)^\top w \leq c(x) \\ D^\top (v + Vw) \geq a(x) + A(x)^\top w \\ v + Vw \geq 0, \end{cases} \tag{10}$$

where the entries of vector $v \in \mathbb{R}^r$ and coefficient matrix $V \in \mathbb{R}^{r\times n}$ are optimization variables. By using well-established reformulation techniques (Ben-Tal et al. (2015)) to get rid of the '$\forall w \in \mathcal{W}$', it can be verified that the convex conservative approximation of constraints (10) are exactly the constraints in (9). □

Since we restrict the decision rule λ to be affine, the set of constraints (9) is indeed a conservative approximation (1). Note that the set of second-order cone constraints (9) has the same computational complexity as the nominal version (that is, with no uncertainty) of (1). The only added computational effort is polynomial, as the single the uncertain constraints is replaced by a set of $n + r + 1$ second-order-cone constraints with mr additional variables.

A simple but powerful enhancement of linear decision rules has been proposed recently by de Ruiter and Ben-Tal (2017), where the authors use a lifted variant of $\mathcal{W}$:

$$\widehat{\mathcal{W}} = \left\{ (w, z) \in \mathbb{R}^m \times \mathbb{R}^m : w_i^2 \leq z_i, i = 1, \ldots, m, \sum_{i=1}^m z_i \leq 1 \right\},$$

and show that the resulting linear decision rule is equivalent to the following nonlinear decision rule:

$$\lambda^\dagger = v + Vw + Uz, \tag{11}$$

where $z_i = w_i^2$ for $i = 1, \ldots, m$ and $U \in \mathbb{R}^{r\times m}$. Notice that the projection of $\widehat{\mathcal{W}}$ onto its w-space is $\mathcal{W}$. The convex reformulation of (5) with $\widehat{\mathcal{W}}$ can be derived by first imposing decision rule (11) on λ, and then applying the standard robust optimization techniques. The resulting robust counterpart is a set of second-order cone constraints (see Appendix B), that is, in the same complexity class as (9), and it is a possibly tighter conservative approximation of (1) than (9).

Similarly, the following lemma gives the convex conservative approximation of the robust semidefinite programming constraints in (8) via linear decision rules.

LEMMA 2. *The vector $x \in \mathcal{X}$ satisfies constraint* (8) *if there exist $v \in \mathbb{R}^r$ and $V^j \in \mathbb{R}^{m\times m}$, $j = 1, \ldots, r$, such that x also satisfies:*

$$\begin{cases} d^\top v \leq 0 \\ A^{(0)}(x) - \sum_{j=1}^r d_j V^{(j)} \succeq 0 \\ D_i^\top v \geq 0 \qquad i = 1, \ldots, n, \\ \sum_{j=1}^r D_{ij} V^{(j)} + A^{(i)}(x) \succeq 0 \qquad i = 1, \ldots, n \\ v \geq 0 \\ V^{(j)} \succeq 0 \qquad j = 1, \ldots, r. \end{cases} \tag{12}$$

Proof. Similarly as for the proof of (9), by restricting λ to a linear decision rule, we obtain the following conservative approximation of (8):

$$\forall W \succeq 0: \begin{cases} \mathrm{Tr}\left(A^{(0)}(x)W\right) - d^\top v - \mathrm{Tr}\left(\sum_{j=1}^r d_j V^{(j)} B\right) \geq 0 \\ D_i^\top v + \mathrm{Tr}\left(\sum_{j=1}^r D_{ij}^\top V^{(j)} B\right) \geq -\mathrm{Tr}\left(A^{(i)}(x)W\right) \qquad i = 1,\dots,n \\ v_j + \mathrm{Tr}\left(V^{(j)} W\right) \geq 0 \qquad j = 1,\dots,r, \end{cases} \tag{13}$$

where the vector $v \in \mathbb{R}^r$ and coefficent matrix $V^{(j)} \in \mathbb{R}^{m\times m}$, $j = 1,\dots,r$, are optimization variables. The convex reformulation of (9) can be easily obtained via well-established reformulation techniques (Ben-Tal et al. (2015)) to get rid of the '$\forall w \in \mathcal{W}$'. □

The set of $n+r+1$ semidefinite constraints (12) again has the same computational complexity as the nominal version of (7), but now with $m^2 r$ additional variables.

As discussed in §2, the inner approximations via linear decision rules are tight if the uncertainty set $\mathcal{U}$ in (1) and (7) is simplicial. In §7 and §8, we show via numerical experiments that linear decision rules give close-to-optimal solutions for problems with uncertain second-order cone constraints. Recently it has been shown by Zhen et al. (2018) that one can use Fourier-Motzkin elimination to eliminate the wait-and-see decisions in (5), and establish guaranteed optimality in a finite number of steps. In this way, problems of small size can be solved to optimality. For larger problems, one could eliminate a subset of the wait-and-see decisions and then apply the existing methods, e.g., linear/quadratic decision rules, to solve the resulting problem. In §6, we use a combination of linear/quadratic decision rules and Fourier-Motzkin elimination to solve the two-stage robust reformulation for the minimum volume circumscribing ellipsoid problem, and compare the efficiency and effectiveness of the proposed approach for the minimum volume circumscribing ellipsoid problem with the state-of-the-art copositive approach of Mittal and Hanasusanto (2018).

Another popular approach for solving two-stage robust optimization problems is finite adaptability, in which the uncertainty set $\mathcal{W}$ is split into a number of smaller subsets, each with its own set of recourse decisions. The number of these subsets can be either fixed a priori or decided by the optimization model (Vayanos et al. (2011), Bertsimas and Caramanis (2010), Hanasusanto et al. (2014), Postek and den Hertog (2016), Bertsimas and Dunning (2016), Georghiou et al. (2017)). In the numerical experiments of paper, we focus on the most effective existing approaches, linear/lifted linear/quadratic decision rule approaches and Fourier-Motzkin elimination.

4.2. Progressive approximation

Since the methods discussed in §4.1 are conservative, the solutions are in general suboptimal. It is therefore important to find effective outer approximations to assess the quality of the conservative approximations in §4.1. In this subsection, we focus on progressive outer approximation methods for uncertain second-order cone constraint (1). We would like to point out that the discussed methods can be applied in an analogous way to uncertain semidefinite constraint (7).

One simple way of obtaining an outer approximation of (1) is to only consider a finite subset of scenarios $\{\zeta^{(1)},\dots,\zeta^{(K)}\}$ from the uncertainty set $\mathcal{U}$. The outer approximation is therefore the "sampled version" of (1):

$$a(x)^\top \zeta^{(k)} + \left\|A(x)\zeta^{(k)} + b(x)\right\|_2 \le c(x) \qquad k = 1,\dots,K. \tag{14}$$

These are standard second-order cone constraints. Clearly the set of constraints (14) is an outer approximation of (1), since a feasible $\hat{x}$ of (14) is only feasible for a finite subset of the uncertainty set. There could be realizations in $\mathcal{U}$ for which $\hat{x}$ is infeasible. For a polyhedral $\mathcal{U} = \{\zeta \ge 0 : \ D\zeta \le d\}$, if the set contains all the extreme points $\zeta^{(1)},\dots,\zeta^{(K)}$ of $\mathcal{U}$, any feasible solution $\hat{x}$ of (14) is also feasible for (1). Of course, the set of extreme points of a polyhedral uncertainty set $\mathcal{U}$ is in practice way too large. As we see in our numerical examples, this is only doable when the uncertainty set has only a few extreme points. We apply the same reasoning to (5) to obtain an outer approximation for the reformulation of the second-order cone constraint:

$$\begin{cases} d^\top \lambda^{(k)} + b(x)^\top w^{(k)} \le c(x) \qquad k = 1,\dots,K \\ D^\top \lambda^{(k)} \ge a(x) + A(x)^\top w^{(k)} \qquad k = 1,\dots,K \\ \lambda^{(k)} \ge 0 \qquad k = 1,\dots,K, \end{cases} \tag{15}$$

is also a valid out approximation of (1). Here $\left\{w^{(1)},\dots,w^{(K)}\right\} \subset \mathcal{W} = \{w \in \mathbb{R}^m : \|w\|_2 \le 1\}$ and $\lambda^{(k)} \in \mathbb{R}^r$ is a here-and-now decision for $k = 1,\dots,K$. In this case there are infinitely many extreme points of the second-order cone $\mathcal{W}$. A complete enumeration of all the extreme points would be impossible. Given two finite scenario sets $\{\zeta^{(1)},\dots,\zeta^{(K)}\}$ and $\left\{w^{(1)},\dots,w^{(K)}\right\}$, one can of course combine the constraints in (14) and (15) to obtain a possibly tighter outer approximation of (1).

Hadjiyiannis et al. (2011) propose a way to obtain a small and effective finite set of scenarios for two-stage fixed-recourse robust linear constraints. For any feasible $(\hat{x},\hat{v},\hat{V})$ of (10), their method takes scenarios that are worst case for the constraints in (10), hoping that the same set of scenarios is also worst case for the optimal (nonlinear) decision rule. For instance, such a scenario of (10) admits the following analytic form:

$$\bar{w} = \arg\max_{w\in\mathcal{W}} \left\{ d^\top \left(\hat{v} + \hat{V}w\right) + b(\hat{x})^\top w \right\} = \frac{\hat{V}^\top d + b(\hat{x})}{\left\|\hat{V}^\top d + b(\hat{x})\right\|_2}, \tag{16}$$

where $(\hat{x},\hat{v},\hat{V})$ satisfies (10). For each constraint one can obtain one such scenario. The obtained scenarios $\left\{\bar{w}^{(1)},\dots,\bar{w}^{(r)}\right\}$ can then be used in (15) to obtain an outer approximation of (1). For more details on the method we refer to the original paper by Hadjiyiannis et al. (2011). One direct extension of the method of Hadjiyiannis et al. (2011) is to use the obtained scenarios $\left\{\bar{w}^{(1)},\ldots,\bar{w}^{(r)}\right\}$ to recover scenarios $\left\{\bar{\zeta}^{(1)},\ldots,\bar{\zeta}^{(r)}\right\}\subseteq\mathcal{U}$, where

$$\bar{\zeta}^{(k)}=\arg\max_{\zeta\in\mathcal{U}}\left\{a(\hat{x})^\top\zeta+(\bar{w}^{(k)})^\top\left(A(\hat{x})\zeta+b(\hat{x})\right)\right\}\qquad k=1,\ldots,r,\tag{17}$$

which can then be used in (14) to obtain an outer approximation of (1). One can again combine constraints (14) with $\{\bar{\zeta}^{(1)},\ldots,\bar{\zeta}^{(r)}\}$ and constraints (15) with $\left\{\bar{w}^{(1)},\ldots,\bar{w}^{(r)}\right\}$ to obtain a possibly tighter outer approximation of (1). However, for a special case of (1) where $a(x)=a$ and $A(x)=A$, the constraints (15) with $\left\{\bar{w}^{(1)},\ldots,\bar{w}^{(r)}\right\}$ are redundant with respect to constraints (14) with $\{\bar{\zeta}^{(1)},\ldots,\bar{\zeta}^{(r)}\}$.

Theorem 3. *Let $a(x)=a$, $A(x)=A$, $\left\{\bar{w}^{(1)},\ldots,\bar{w}^{(r)}\right\}\subseteq\mathcal{W}$ be a finite set of scenarios and $\{\bar{\zeta}^{(1)},\ldots,\bar{\zeta}^{(r)}\}\subseteq\mathcal{U}$ be the corresponding set of scenarios from* (17). *Then $x\in\mathbb{R}^{n_x}$ satisfies the constraints* (14) *with $\{\bar{\zeta}^{(1)},\ldots,\bar{\zeta}^{(r)}\}$ also satisfies the constraints* (15) *with $\left\{\bar{w}^{(1)},\ldots,\bar{w}^{(r)}\right\}$.*

Proof. Let $\bar{x}$ be a vector that satisfies:

$$\begin{aligned}
&a^\top\bar{\zeta}^{(k)}+\left\|A\bar{\zeta}^{(k)}+b(x)\right\|_2\le c(x)\qquad k=1,\ldots,r\\
\iff\qquad &a^\top\bar{\zeta}^{(k)}+\max_{w:\ \|w\|_2\le 1}w^\top\left(A\bar{\zeta}^{(k)}+b(x)\right)\le c(x)\qquad k=1,\ldots,r.
\end{aligned}$$

Since $\left\{\bar{w}^{(1)},\ldots,\bar{w}^{(r)}\right\}\subseteq\mathcal{W}$, then by definition $\bar{x}$ also satisfies:

$$\begin{aligned}
&a^\top\bar{\zeta}^{(k)}+(\bar{w}^{(k)})^\top\left(A\bar{\zeta}^{(k)}+b(x)\right)\le c(x)\qquad k=1,\ldots,r\\
\iff\qquad &\max_{\zeta\in\mathcal{U}}\left\{a^\top\zeta+(\bar{w}^{(k)})^\top\left(A\zeta+b(x)\right)\right\}\le c(x)\qquad k=1,\ldots,r,
\end{aligned}\tag{18}$$

where we have used the definition of $\bar{\zeta}^{(k)}$ from (17). Note that the equivalence here is due to $a(x)=a$ and $A(x)=A$. By dualizing over ζ in (18), using strong duality for linear programming, then $\bar{x}$ also satisfies:

$$\begin{cases}
d^\top\lambda^{(k)}+b(x)^\top\bar{w}^{(k)}\le c(x)\qquad k=1,\ldots,r\\
D^\top\lambda^{(k)}\ge a+A^\top\bar{w}^{(k)}\qquad k=1,\ldots,r\\
\lambda^{(k)}\ge 0\qquad k=1,\ldots,r.
\end{cases}$$

□

Theorem 3 indicates that the scenarios from (17) are effective for the special case of (1), and therefore we use these in the numerical experiments.

5. Extensions

Bilinear uncertainty

Consider the following robust constraint with bilinear uncertainty:

$$\forall w\in\mathcal{W}\ \forall\zeta\in\mathcal{U}:\quad a(x)^\top\zeta+w^\top\left(A(x)\zeta+b(x)\right)\le c(x),\tag{19}$$

where $\mathcal{W}$ is a general convex set. In the following proposition, we reformulate constraint (19) into a set of two-stage robust linear constraints. The proof of this proposition is similar to the proof of Theorem 1, hence, omitted.

Proposition 1 *Let $\mathcal{U}$ be the polyhedral uncertainty set given in* (3)*. Then $x \in \mathbb{R}^{n_x}$ satisfies constraint* (19) *if and only if it satisfies*

$$\forall w \in \mathcal{W}\ \exists \lambda \geq 0: \begin{cases} d^\top \lambda + b(x)^\top w \leq c(x) \\ D^\top \lambda \geq a(x) + A(x)^\top w. \end{cases} \tag{20}$$

Problems with constraints in the form of (20) can then be solved by using the methods described in §4.

Uncertain constraints with wait-and-see decisions

Suppose a set of uncertain second-order cone constraints in the form (1) contains a wait-and-see decision y:

$$\forall \zeta \in \mathcal{U}: \quad a(x)^\top \zeta + h(y) + \|A(x)\zeta + By + b(x)\|_2 \leq c(x), \tag{21}$$

where $h: \mathbb{R}^{n_y} \to \mathbb{R}$ is an affine function. One simple yet crucial observation is that, by imposing linear decision rule $y = u + Y\zeta$, where the vector $u \in \mathbb{R}^{n_y}$ and coefficent matrix $Y \in \mathbb{R}^{n_y \times n}$ are here-and-now decision variables, constraint (21) becomes an instance of (1):

$$\forall \zeta \in \mathcal{U}: \quad a(x)^\top \zeta + h(u + Y\zeta) + \|A(x)\zeta + Bu + BY\zeta + b(x)\|_2 \leq c(x).$$

□

In the following examples, we use the simple algebraic principle (Lobo et al. 1998):

$$u^\top u \leq st, \quad s \geq 0, \quad t \geq 0 \quad \Longleftrightarrow \quad \left\| \begin{matrix} 2u \\ s - t \end{matrix} \right\|_2 \leq s + t, \tag{22}$$

and show that a large class of uncertain convex constraints can be cast as uncertain second-order cone constraints with wait-and-see decisions.

Example 3. **Uncertain quadratic/linear fraction constraints.** Consider uncertain constraints in the form:

$$\forall \zeta \in \mathcal{U}: \quad \begin{cases} \sum_{i=1}^{K} \frac{\left\| A^{(i)}(x)\zeta + b^{(i)}(x) \right\|_2^2}{(a^{(i)}(x))^\top \zeta + c_i(x)} \leq \tau, \\ (a^{(i)}(x))^\top \zeta + c_i(x) > 0 \qquad i = 1, \ldots, K, \end{cases} \tag{23}$$

where $a^{(i)}: \mathbb{R}^{n_x} \to \mathbb{R}^n$, $c_i: \mathbb{R}^{n_x} \to \mathbb{R}$, $A^{(i)}: \mathbb{R}^{n_x} \to \mathbb{R}^{m \times n_x}$, $b^{(i)}: \mathbb{R}^n \to \mathbb{R}^{m \times n_y}$ are (componentwise) affine functions for $i = 1, \ldots, K$, and $\tau \in \mathbb{R}$ is a constant or an optimization variable. By introducing a wait-and-see variable y, we have an equivalent reformulation of (23):

$$\forall \zeta \in \mathcal{U}\ \exists y \geq 0: \begin{cases} \sum_{i=1}^{K} y_i \leq \tau, \\ \left(A^{(i)}(x)\zeta + b^{(i)}(x)\right)^\top \left(A^{(i)}(x)\zeta + b^{(i)}(x)\right) \leq y_i \left((a^{(i)}(x))^\top \zeta + c_i(x)\right) \qquad i = 1, \ldots, K \\ (a^{(i)}(x))^\top \zeta + c_i(x) > 0 \qquad i = 1, \ldots, K, \end{cases}$$

which, by (22), is equivalent to:

$$\forall \zeta \in \mathcal{U}\ \exists y \geq 0: \begin{cases} \sum_{i=1}^{K} y_i \leq \tau, \\ \left\| \begin{matrix} 2A^{(i)}(x)\zeta + 2b^{(i)}(x) \\ (a^{(i)}(x))^\top \zeta + c_i(x) - y_i \end{matrix} \right\|_2 \leq (a^{(i)}(x))^\top \zeta + c_i(x) + y_i & i = 1, \ldots, K \\ (a^{(i)}(x))^\top \zeta + c_i(x) > 0 & i = 1, \ldots, K. \end{cases} \tag{24}$$

Since (24) has constraints of the format in (1) we can use our conservative (and progressive) techniques to find solutions. □

Example 4. **Product of uncertain nonnegative affine functions.** Let us consider uncertain constraints in the form:

$$\forall \zeta \in \mathcal{U}: \begin{cases} \prod_{i=1}^{K} \left((a^{(i)}(x))^\top \zeta + c_i(x) \right)^{\frac{1}{K}} \geq \tau, \\ (a^{(i)}(x))^\top \zeta + c_i(x) \geq 0 & i = 1, \ldots, K, \end{cases} \tag{25}$$

where $a^{(i)} : \mathbb{R}^{n_x} \to \mathbb{R}^n$ and $c_i : \mathbb{R}^{n_x} \to \mathbb{R}$ are affine functions for $i = 1, \ldots, K$. If τ is maximized subject to the constraints (25), then it is equivalent to maximize the geometric mean of uncertain nonnegative affine functions $(a^{(i)}(x))^\top \zeta + c_i(x) \geq 0, i = 1, \ldots, K$. For simplicity, we consider the special case $K = 4$, and first reformulate the problem by introducing new variables y_1 and y_2:

$$\forall \zeta \in \mathcal{U}\ \exists y \geq 0: \begin{cases} y_1 y_2 \geq \tau^2 \\ ((a^{(1)}(x))^\top \zeta + c_1(x))((a^{(2)}(x))^\top \zeta + c_2(x)) \geq y_1^2 \\ ((a^{(3)}(x))^\top \zeta + c_3(x))((a^{(4)}(x))^\top \zeta + c_4(x)) \geq y_2^2 \\ (a^{(i)}(x))^\top \zeta + c_i(x) \geq 0 & i = 1, \ldots, K. \end{cases} \tag{26}$$

The extension to other values of K is straightforward. Applying (22) yields the following set of uncertain SOC constraints:

$$\forall \zeta \in \mathcal{U}\ \exists y \geq 0: \begin{cases} \left\| \begin{matrix} 2\tau \\ y_1 - y_2 \end{matrix} \right\|_2 \leq y_1 + y_2 \\ \left\| \begin{matrix} 2y_1 \\ (a^{(1)}(x))^\top \zeta + c_1(x) - (a^{(2)}(x))^\top \zeta - c_2(x) \end{matrix} \right\|_2 \leq (a^{(1)}(x))^\top \zeta + c_1(x) + (a^{(2)}(x))^\top \zeta + c_2(x) \\ \left\| \begin{matrix} 2y_2 \\ (a^{(3)}(x))^\top \zeta + c_3(x) - (a^{(4)}(x))^\top \zeta - c_4(x) \end{matrix} \right\|_2 \leq (a^{(3)}(x))^\top \zeta + c_3(x) + (a^{(4)}(x))^\top \zeta + c_4(x) \\ (a^{(i)}(x))^\top \zeta + c_i(x) \geq 0 & i = 1, \ldots, K. \end{cases}$$

□

Example 5. **Uncertain logarithmic constraints.** Let us consider uncertain constraints in the form:

$$\forall \zeta \in \mathcal{U}: \ |\log\left((a^{(i)}(x))^\top \zeta\right) - \log(c_i)| \leq \tau \qquad i = 1, \ldots, K, \tag{27}$$

where the entries of $a^{(i)} : \mathbb{R}^{n_x} \to \mathbb{R}^n$ are (componentwise) affine functions for $i = 1, \ldots, K$. We assume $c_i \in \mathbb{R}_+$, and interpret $\log\left((a^{(i)}(x))^\top \zeta\right)$ as $-\infty$ when $(a^{(i)}(x))^\top \zeta \leq 0$. Suppose τ is minimized subject to the constraints (27). Then it can be understood as approximately solving an overdetermined set of uncertain equations $(a^{(i)}(x))^\top \zeta \approx c_i,\ i=1,\dots,K$, measuring the worst case error by the maximum logarithmic deviation between the numbers $(a^{(i)}(x))^\top\zeta$ and c_i. To cast these constraints as a set of uncertain SOC constraints, first note that:

$$|\log\left((a^{(i)}(x))^\top\zeta\right) - \log(c_i)| = \log\max\left(\frac{a^{(i)}(x)^\top\zeta}{c_i}, \frac{c_i}{a^{(i)}(x)^\top\zeta}\right)$$

(assuming $a^{(i)}(x))^\top\zeta > 0$). Then the constraints (27) is therefore equivalent to:

$$\forall\zeta\in\mathcal{U}\ \exists y\ge 0:\quad \begin{cases} y\le\Omega \\ \frac{c_i}{y}\le a^{(i)}(x)^\top\zeta\le yc_i \qquad i=1,\dots,K, \end{cases}$$

where $y\in\mathbb{R}$ and $\Omega=\log\tau$. Applying (22) yields the following set of uncertain SOC constraints:

$$\forall\zeta\in\mathcal{U}\ \exists y\ge 0:\quad \begin{cases} y\le\Omega \\ a^{(i)}(x)^\top\zeta\le yc_i \qquad i=1,\dots,K \\ \left\|\begin{matrix} 2c_i \\ yc_i - a^{(i)}(x)^\top\zeta \end{matrix}\right\|_2 \le yc_i + a^{(i)}(x)^\top\zeta \qquad i=1,\dots,K. \end{cases}$$

□

Uncertain nonlinear constraints with l_p-norms

In this subsection, we extend the result in §2 to uncertain constraints with l_p-norms. Let us consider an uncertain constraint in the form:

$$\forall\zeta\in\mathcal{U}:\quad a(x)^\top\zeta + \|A(x)\zeta + b(x)\|_p \le c(x), \tag{28}$$

where $p\in[1,\infty]$. By definition of the dual norm (see, for example, Lax (2007)), we know:

$$\begin{aligned} &\forall\zeta\in\mathcal{U}:\quad a(x)^\top\zeta + \|A(x)\zeta+b(x)\|_p \le c(x) \\ \iff &\forall\zeta\in\mathcal{U}:\quad a(x)^\top\zeta + \max_{w:\|w\|_q\le 1} w^\top\left(A(x)\zeta+b(x)\right)\le c(x) \\ \iff &\forall w\in\mathcal{W}\ \forall\zeta\in\mathcal{U}:\quad a(x)^\top\zeta + w^\top\left(A(x)\zeta+b(x)\right)\le c(x), \end{aligned}$$

where the set $\mathcal{W}=\left\{w:\|w\|_q\le 1\right\}$ is convex, and $q=\frac{p}{p-1}$ for $p\in[1,\infty]$. Applying Proposition 1 yields a set of two-stage robust linear constraints in the form (20) with a l_q-norm uncertainty set.

6. Minimum Volume Circumscribing Ellipsoid

In this application, we consider the problem of finding the minimum volume circumscribing ellipsoid of a polytope in n-dimensional space. This problem arises in a variety of applications such as outlier detection (Ahipaşaoğlu 2015) and facility location (Elzinga and Hearn 1974) and is known to be NP-hard (Boyd and Vandenberghe 2004). Mathematically, we are interested in solving

$$\begin{aligned} \min_{Q,c}\quad & \log\det Q^{-1} \\ s.t.\quad & \|Qx+c\|_2\le 1 \qquad \forall x\in\mathcal{U}, \end{aligned}$$

for a polyhedron $\mathcal{U}$. This optimization problem can be interpreted as a robust optimization problem with an uncertain SOC constraint, where x is the uncertain parameter residing in a polyhedron $\mathcal{U}$. We can thus apply the theory developed in this paper to reformulate the problem to an adjustable robust optimization problem and approximate it using various decision rules. In particular, in this experiment we have used two different decision rules to illustrate the trade-off between the quality of the resulting approximation and the computation time, and compare their results to the copositive programming approach described in (Mittal and Hanasusanto 2018). First of all, we consider the simple linear decision rule used in Lemma 1. Moreover, we consider a full quadratic decision rule, that is,

$$\lambda = u + Vw + \sum_{i\leq j} r_{ij} w_i w_j. \tag{29}$$

Using such a decision rule yields a static robust optimization problem with conic quadratic inequalities and ball uncertainty, which Ben-Tal et al. (2009) show can be reformulated as a semidefinite program. Therefore, we can thus solve the resulting problem in reasonable time.

6.1. Numerical setting

To test our approach we use the first numerical setup of Mittal and Hanasusanto (2018), who develop a copositive programming approach to approximating the minimum volume circumscribing ellipsoid of a polytope. In this numerical setup, we randomly generate polytopes in L dimensions by intersecting randomly generated halfspaces with the hyperrectangle $\{x \in \mathbb{R}^L : \ \mathbf{0} \leq x \leq \mathbf{1}\}$. The M random halfspaces are generated by generating a vector $s_j \in \mathbb{R}^L$ uniformly distributed on the surface of the hypersphere, and a scalar $r_j \in \mathbb{R}$ uniformly distributed from the interval $\left[-\frac{1}{2}\|s_j\|_1, \frac{1}{2}\|s_j\|_1\right]$. We then consider the constraint $s_j^\top(x - \frac{1}{2}\mathbf{1}) \leq r_j$ if $r_j > 0$ and $s_j^\top(x - \frac{1}{2}\mathbf{1}) \geq r_j$ if $r_j \leq 0$. This construction ensures that the center of the hyperrectangle, $\frac{1}{2}\mathbf{1}$, always remains feasible.

For several values of L and $M = L, 2L, 3L$, we model the problem using the YALMIP toolbox (Löfberg 2004) in MATLAB and solve it with Mosek on a desktop with 8 GB RAM and a 3.4 GHz Intel Core i7 processor. We report the suboptimality of our proposed approximations in low dimensions in Table 1 and the corresponding running times in Table 2. We report results for both the direct reformulations (0 variables eliminated) as well as the results after applying Fourier-Motzkin Elimination (FME) to eliminate the specified number of variables. These latter results are discussed in more detail below. We have also reported the results of the copositive programming approach suggested by Mittal and Hanasusanto (2018), which, to the best of our knowledge, is the current state-of-the-art method for the minimum circumscribing ellipsoid problem.

Clearly, quadratic decision rules offer a significant advantage in terms of approximation quality over linear decision rules. They do, however, perform at least slightly worse than the copositive programming approach, both in approximation quality and computation time. The adjustable robust

Table 1 Average suboptimality over 50 instances of the approximations to the minimum volume circumscribing ellipsoid of a polyhedron in dimension L based on M generated inequalities, relative to the optimal ellipsoid's volume. LB refers to the lower bound from outer approximation (1) based on the linear decision rule approach from Lemma 1 after FME, LDR to the linear decision rule approach from Lemma 1, QDR to using the full quadratic decision rule (29), and CoPos to the copositive programming approach from (Mittal and Hanasusanto 2018). For the LDR and QDR approach we report the results without using FME and after using FME to eliminate the listed number of variables in the adjustable robust optimization formulation.

Method		LB	LDR		QDR		CoPos
Variables Eliminated		2	0	2	0	2	-
$L=2$	$M=L$	-2.3%	37.6%	13.5%	5.3%	4.3%	4.7%
	$M=2L$	-3.2%	39.3%	23.4%	6.6%	5.3%	6.4%
	$M=3L$	-9.1%	33.7%	26.3%	5.8%	4.8%	5.6%
Variables Eliminated		5	0	5	0	5	-
$L=5$	$M=L$	-0.4%	108.6%	10.1%	6.2%	4.7%	4.9%
	$M=2L$	-0.5%	105.6%	19.4%	10.8%	8.3%	8.7%
	$M=3L$	-0.9%	101.2%	29.2%	13.6%	10.6%	11.2%
Variables Eliminated		10	0	10	0	10	-
$L=10$	$M=L$	-0.0%	194.2%	4.0%	2.6%	1.9%	1.8%
	$M=2L$	-0.1%	190.6%	9.4%	6.6%	5.3%	5.2%
	$M=3L$	-0.3%	179.8%	16.5%	10.8%	8.7%	8.6%

Table 2 Average computation time in seconds over 50 instances of the approximations to the minimum volume circumscribing ellipsoid of a polyhedron in dimension L based on M generated inequalities. LB refers to the lower bound from outer approximation (1) based on the linear decision rule approach from Lemma 1 after FME, QDR to using the full quadratic decision rule (29), and CoPos to the copositive programming approach from (Mittal and Hanasusanto 2018). For the LDR and QDR approach we report the results without using FME and after using FME to eliminate the listed number of variables in the adjustable robust optimization formulation.

Method		LB	LDR		QDR		CoPos
Variables Eliminated		2	0	2	0	2	-
$L=2$	$M=L$	0.024	0.115	0.108	0.121	0.112	0.076
	$M=2L$	0.031	0.127	0.114	0.140	0.118	0.080
	$M=3L$	0.038	0.133	0.111	0.152	0.116	0.079
Variables Eliminated		5	0	5	0	5	-
$L=5$	$M=L$	0.020	0.176	0.307	0.229	0.394	0.097
	$M=2L$	0.046	0.203	0.301	0.320	0.466	0.112
	$M=3L$	0.034	0.230	0.355	0.428	0.639	0.144
Variables Eliminated		10	0	10	0	10	-
$L=10$	$M=L$	0.343	0.259	6.572	1.416	71.674	0.183
	$M=2L$	0.480	0.287	7.020	2.948	214.503	0.237
	$M=3L$	0.267	0.362	6.954	4.157	348.168	0.323

optimization approach we propose does have two major advantages that the copositive programming approach lacks: its versatility and the possibility to obtain a lower bound on the optimal objective value. The latter results from the discussion in Section 4.2 on progressive approximation, and is a valuable tool in estimating the suboptimality of a solution in higher dimensions. As can be seen from the results in Table 1, the quality of this lower bound is very good and retains its quality in higher dimensions. Its computation time is also rather low, although an optimal solution to the LDR problem must be known to be able to use the techniques from Section 4.2. Easy access to a good lower bound is a great asset, as it allows for comments on the suboptimality of solutions even in higher dimensions, when we cannot compute the exact optimal solution.

The versatility of our approach stems from the possibility to use different decision rules to make the trade-off between approximation quality and computation time, as well as other techniques from ARO. One of those techniques is Fourier-Motzkin Elimination (FME), which eliminates adjustable variables, resulting in a better approximation with additional constraints that increase the computation time. In fact, when all adjustable variables are eliminated, the optimal minimum volume circumscribing ellipsoid is recovered. Figure 1 illustrates the resulting approximation quality and computation time for both the linear and quadratic decision rule after a variety of variables have been eliminated for $L=5$ and $M=10$, averaged over 50 instances. From the results in Table 1, we know that the copositive programming approach has an average suboptimality of 8.7% for these parameters, also shown in Figure 1 for reference sake. Thus, when at least 5 adjustable variables are eliminated, the quadratic decision rule performs better than the copositive programming approach on average. For this number of eliminated variables, the computation time is still modest, which shows that the approach we propose can compete with the current state-of-the-art techniques available.

The results in Table 1 and 2 in fact indicate that using FME in addition to quadratic decision rules allows us to obtain better approximations than the copositive programming approach for $L=2$ and $L=5$ in reasonable time. This does come at the expense of a higher computation time, especially in higher dimensions. For $L=10$, the copositive programming approach does perform better than the quadratic decision rule, even when FME is used to eliminate 10 of the adjustable variables, which indicates the copositive programming method scales better to higher dimensions. We remark that all results in Table 1 and 2 were obtained using a vanilla implementation of FME, i.e., we did not experiment with cleverly selecting which variables to eliminate, which might improve the results even further.

7. Robust regression

In this section, we first consider the robust regression model in Example 1. We use the Diabetes dataset from Efron et al. (2004), with $m=442$ observations and each with 10 features, i.e., $n_x=11$.

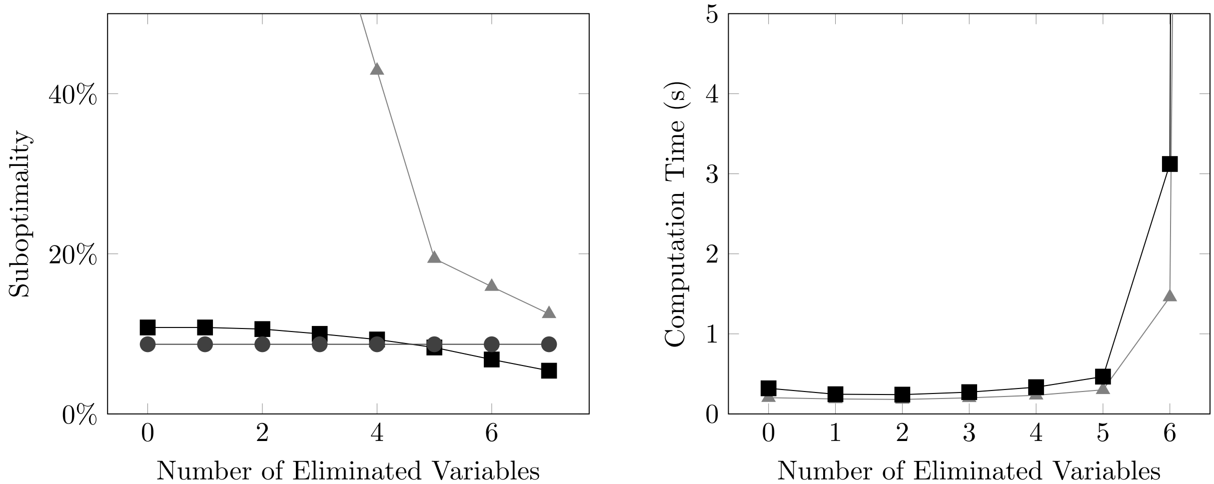

Figure 1 **The average suboptimality and computation time over 50 instances after a number of adjustable variables have been eliminated through Fourier-Motzkin Elimination for the linear decision rule shown in triangles and the quadratic decision rule in squares, for $L = 5$ and $M = 10$. The suboptimality of the copositive programming approach is shown in circles for reference.**

The ten features are age, BMI, average blood pressure, six blood serum measurements, and gender. The response of interest $b \in \mathbb{Z}^m_+$ is a quantitative measure of disease progression one year after baseline. We divide the dataset into a testing set and a training set with the first 132 observations ($\approx 30\%m$) and remaining 310 observations ($\approx 70\%m$), respectively, and solve the considered models only using the training set. The testing set is used to evaluate the obtained solutions.

Let us consider the robust regression model (2) in Example 1. This model admits a two-stage robust linear reformulation (see §2), and can be solved by the methods in §4. For the 10 features in the Diabetes dataset, we take that the gender of the patients is certain, i.e., $\zeta_{i,10} = 0$, $i = 1, \dots, 310$, as well as the last column of A, which is defined by us, the modeler, as the all one vector to represent an intercept, and at most Γ out of 9 remaining features can deviate from the original data A, each feature can deviate up to 1%. We assume the response vector b can be measured exactly. To this end, the following budget uncertainty set is considered:

$$\mathcal{U} = \left\{ \zeta \in \mathbb{R}^{310\times 11} \;:\; \begin{array}{c} \zeta_{i,10} = 0,\ \zeta_{i,11} = 1,\ \zeta_{ij} = 1\% A_{ij}\delta_j \\ \delta \le \xi,\ -\delta \le \xi,\ \xi \le 1,\ \sum_{i=1}^{9} \xi_i \le \Gamma \end{array} \quad i = 1, \dots, 310,\ j = 1, \dots, 9 \right\},$$

where $\delta \in \mathbb{R}^9$, $\xi \in \mathbb{R}^9$, and A_{ij} is the element of A at the i-th row and j-th column. The one-dimensional uncertain parameter δ_j for the j-th feature is the same for all observations. In this way, we can model bias in the measurements of the j-th feature. All computations are carried out with MOSEK 8.0 (MOSEK ApS (2017)) on an Intel Core(TM) i5-4590 Windows computer running at 3.30GHz with 8GB of RAM. All modeling is done using the modeling package XProg (`http://xprog.weebly.com`).

Table 3 The solution quality and computation time for different values of Γ in (2). ROC denotes the (conservatively approximated) objective values obtained from solving (10), and Time(s) is the corresponding computational time in seconds. %Gap denotes the optimality gap, i.e., %Gap = $\frac{\text{LDR-OPT}}{\text{OPT}}$, where OPT denotes the optimal objective values.

Γ	1	2	3	4	5	6	7	8	9
ROC	938.9	942.6	946.2	948.5	949.4	950.4	951.2	951.6	951.7
Time(s)	0.3	0.5	0.3	0.3	0.3	0.4	0.4	0.3	0.2
%Gap	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%

From Table 3 one may observe that the approximated solutions of (2) from solving (10) can be computed efficiently, and they are very close to optimality. Since the extreme points are not too many in $\mathcal{U}$, we enumerate all the extreme points to obtain the optimal solutions.

8. Robust sensor network

In this numerical experiment we consider a problem with N points in $\mathbb{R}^2$ that must be connected by links. Some of the N points are already positioned and the decision maker has to decide where to place the remaining points. The goal is to minimize the total distance of all the links together. An example of this problem can be found in network design, where the points represent wireless sensors and modules on a network that are interconnected and one wants to minimize the total energy needed for the wireless transmission over the links. This example is based on the placement and location problem, for which the nominal case is described in (Boyd and Vandenberghe 2004, Section 8.7). In the nominal case the model can be written as

$$\begin{aligned} \min_{y} \quad & \sum_{(i,j)\in\mathcal{A}} \|y_i - y_j\|_2 \\ s.t. \quad & y_i = \bar{a}_i \qquad \forall i \in L, \end{aligned}$$

where $y_i \in \mathbb{R}^2$ are the locations of the points for all $i = 1,\ldots,N$. The points that are fixed (already positioned) are given by the set L and their locations by $\bar{a}_i \in \mathbb{R}^2$. The remaining points not in L are free to set by the optimizer. The set of prescribed (undirected) links is given by $\mathcal{A}$.

8.1. The robust model

We introduce uncertainty by considering the case where the locations of the fixed sensors $\bar{a}_i$ are not precisely known. This uncertainty in the locations could be due to sea currents for sensors placed at sea, wind drift for sensors that are dropped from planes or other errors due to placement by catapult or missile (see e.g. Akyildiz et al. (2002)). We model the uncertainty in the locations as: $a_i(\zeta) = \bar{a}_i + \hat{a}_i\zeta_i$, where $\hat{a}_i \in \mathbb{R}_+$ is the maximal (absolute) deviation from the nominal value $\bar{a}_i$ for all $i \in L$. The uncertain parameter $\zeta = (\zeta_1,\ldots,\zeta_{|L|})^\top \in \mathbb{R}^{2|L|}$, where $\zeta_i^\top \in \mathbb{R}^2$ for all $i \in L$, resides in a lifted budget uncertainty set $\mathcal{U}$ defined by

$$\mathcal{U} = \left\{ (\zeta,\xi):\ \zeta \le \xi,\ -\zeta \le \xi,\ \xi \le 1, \sum_{i=1}^{2|L|} \xi_i \le \Gamma \right\},$$

where $\xi \in \mathbb{R}^{2|L|}$ and $\Gamma \geq 0$ is called the budget of uncertainty. Projecting $\mathcal{U}$ on the space of ζ, one can recover the classical budget uncertainty set $\left\{\zeta \in \mathbb{R}^{2|L|} : \ -1 \leq \zeta \leq 1, \ \|\zeta\|_1 \leq \Gamma\right\}$ of Bertsimas and Sim (2004). We use the lifted budget uncertainty set here because the improvement in solution quality often worth of the small extra computation effort due to the extra variable (de Ruiter and Ben-Tal (2017)). Some of the modules x_i need to be placed before the exact locations of $a_i(\zeta)$ are known, whereas others can be placed after. We define the set of indices H for those modules that have to be placed before the sensor locations are known. We associate a here-and-now variable $x_i \in \mathbb{R}^2$ for every $i \in H$. The robust model can now be written as:

$$\begin{aligned} \min_{x, y(\zeta,\xi), \tau} \quad & \tau \\ s.t. \quad \forall (\zeta,\xi) \in \mathcal{U} : & \begin{cases} \sum_{(i,j)\in\mathcal{A}} \|y_i(\zeta,\xi) - y_j(\zeta,\xi)\|_2 \leq \tau & \\ y_i(\zeta,\xi) = \bar{a}_i + \hat{a}_i \zeta_i & \forall i \in L \\ y_i(\zeta,\xi) = x_i & \forall i \in H. \end{cases} \end{aligned} \tag{30}$$

For ease of exposition, we use the wait-and-see decision y to represent the location of the sensors and modules. One can eliminate y using the equalities, and model (30) then has constraints of the form (1). We first use Theorem 1 to reformulate the constraints (30) into a set of two-stage robust linear constraints, then we solve the resulting model via the solution methods in §4. The objective value of (30) gives the total energy required for the wireless transmissions in the network.

8.2. Numerical setting

For the experiments we use two sets of data. For illustrative purposes we first consider a small instance with $N = 14$ points, of which 8 nominal sensor locations are uncertain and 6 modules need to be placed. For this, data from (Boyd and Vandenberghe 2004, Section 8.7.3) is used, see Figure 2. Data is obtained from the CVX website `http://web.cvxr.com/cvx/examples`. The maximal deviation from the nominal locations is taken to be $\hat{a}_i \in \{0, 0.2, 0.5, 1\}$ for all $i \in L$. For instance for $\hat{a}_i = 0.2$, $i \in L$, Figure 3 illustrates the robust solution obtained from solving (30) via linear decision rules.

The second set of data shows results for larger instances. We choose $|L| \in \{10, ..., 70\}$ nominal sensor locations $\bar{a}_i$, $i \in L$ uniformly at random from $[-1, 1]^2$. The maximal deviation from the nominal locations is $\hat{a}_i = 0.3$ for all $i \in L$. We have $0.4|L|$ modules that have to collect data from the sensors. Each module is randomly linked with $|L|$ sensors. The modules are randomly linked into a cycle. We link the sensors that are not connected with each of the $0.4|L|$ modules.

8.3. Results

We use the same computer and optimization software as mentioned in §7. Figure 4 depicts the robust solutions for $|L| = 30$. For the small instance with $N = 14$ points, as $\hat{a}_i$, $i \in L = \{1, ..., 8\}$, increases, the obtained objective values ROC from solving (30) via linear decision rules become

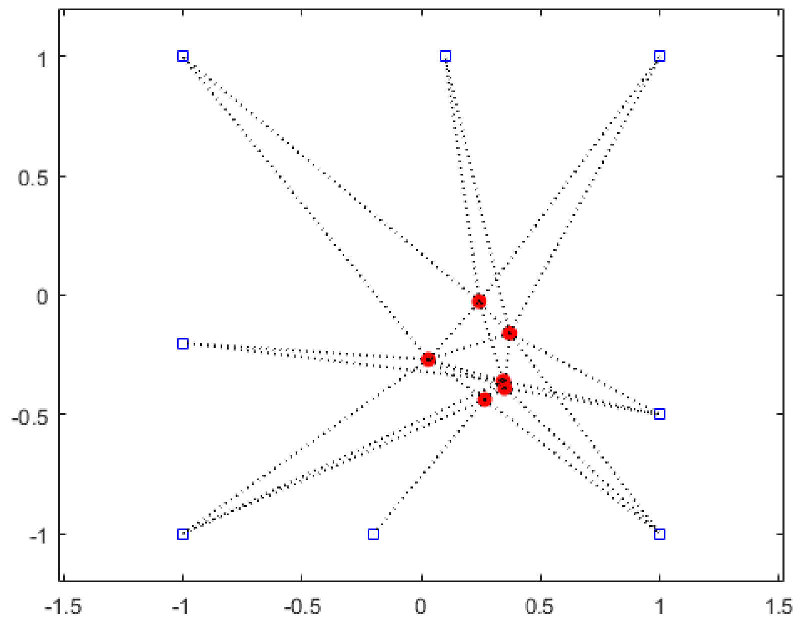

Figure 2 Nominal solution from (Boyd and Vandenberghe 2004, Figure 8.16). The squares represent the sensors. The dots denote the modules. The prescribed (undirected) links between the sensors and modules are plotted using dashed lines.

Table 4 Sensor network model with $N = 14$. ROC is the conservative approximation obtained from solving (9) in §4.1. LB(1) denotes the lower bounds obtained from one scenario. Time(s) reports the computation time (in seconds) for solving ROC.

$\hat{a}_i$	0	0.2	0.5	1
ROC	21.91	23.86	26.88	32.07
LB(1)	21.91	23.79	26.64	31.46
Time(s)	0.1	0.1	0.1	0.1

larger (see Table 4), and the optimal module locations are slightly more spread out (see Figure 3). The **l**ower **b**ounds LB(1) are obtained from solving (14) with only one scenario. This scenario is obtained from (17) using the scenario obtained from (16). The differences between the lower bounds LB(1) and upper bounds ROC are within 1%. This indicates that the optimality gap of the obtained objective values is at most 1% of the optimal value. For medium and large instances considered in Table 5, the approximated solutions of (30) via linear decision rules are near optimal, and can be computed efficiently.

9. Future research

On a theoretical level, one immediate future research direction would be to establish the optimality of the conservative approximations via linear decision rules for uncertain SOC and SDP constraints

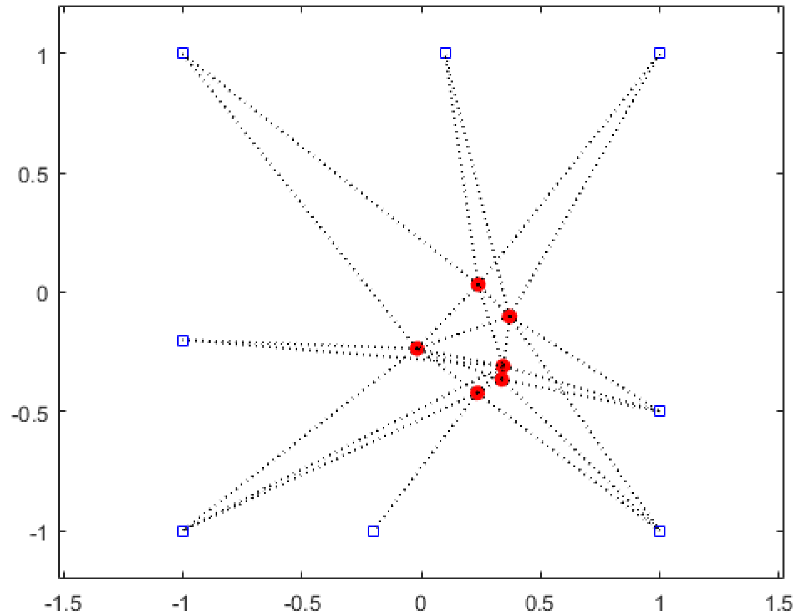

Figure 3 **Robust solution for $\hat{a}_i = 0.2$, $i \in L$, with the nominal sensor locations. The squares represent the sensors. The dots denote the modules. The prescribed (undirected) links between the sensors and modules are plotted using dashed lines.**

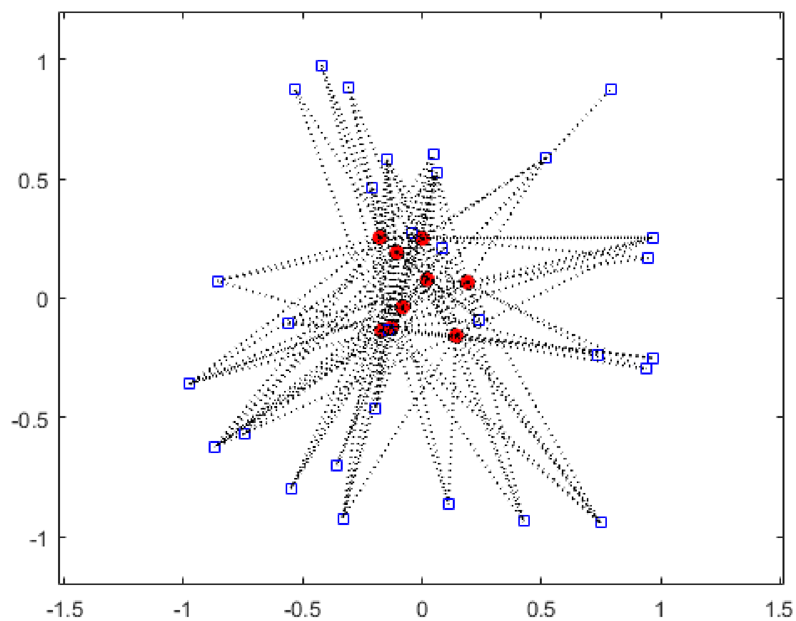

Figure 4 **Robust solution for $\hat{a}_i = 0.3$, $i \in L$, and $|L| = 30$, with the nominal sensor locations. The squares represent the sensors. The dots denote the modules. The prescribed (undirected) links between the sensors and modules are plotted using dashed lines.**

Table 5 Sensor network model with $\hat{a}_i = 0.3, i \in L$, where $|L| = \{10, ..., 70\}$. ROC is the conservative approximation obtained from solving (9) in §4.1. LB(1) denotes the lower bounds obtained from one scenario. Time(s) reports the computation time (in seconds) for obtaining ROC. All the numbers are the average of 10 randomly generated instances.

$\lvert L\rvert$	10	20	30	40	50	60	70
ROC	24.4	70.8	149.5	257.0	392.0	564.1	762.9
LB(1)	24.3	70.8	149.4	256.9	391.9	564.1	762.9
Time(s)	0.35	2.4	10.4	30.6	59.4	102.5	193.3

with uncertainty sets other than a simplex. Another interesting research direction is to extend the proposed approach to uncertain SOC and SDP constraints with non-polyhedral uncertainty sets.

In a follow-up paper (Roos et al. 2019) we show that the approach for uncertain SOC constraints can be extended to general convex uncertain constraints with polyhedral uncertainty.

Acknowledgments

The research of the first author is supported by the Netherlands Organisation for Scientific Research (NWO) Grant 613.001.208. The second author is partially supported by the NWO Research Talent Grant 406.14.067. The research of the third author is supported by the NWO Research Talent Grant 406.17.511.

References

S. Ahipaşaoğlu. Fast algorithms for the minimum volume estimator. *Journal of Global Optimization*, 62(2): 351–370, 2015.

I. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci. Wireless sensor networks: a survey. *Computer Networks*, 38(4):393–422, 2002.

A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4): 769–805, 1998.

A. Ben-Tal, A. Nemirovski, and C. Roos. Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM Journal on Optimization*, 13(2):535–560, 2002.

A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming*, 99(2):351–376, 2004.

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

A. Ben-Tal, D. den Hertog, and J.-Ph. Vial. Deriving robust counterparts of nonlinear uncertain inequalities. *Mathematical Programming*, 149(1):265–299, 2015.

D. Bertsimas and C. Caramanis. Finite adaptability for linear optimization. *IEEE Transactions on Automatic Control*, 55(12):2751–2766, 2010.

D. Bertsimas and I. Dunning. Multistage robust mixed integer optimization with adaptive partitions. *Operations Research*, 64(4):980–998, 2016.

D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

F. de Ruiter and A. Ben-Tal. Tractable nonlinear decision rules for robust optimization. *Working paper*, 2017.

B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani. Least angle regression. *The Annals of Statistics*, 32 (2):407–499, 2004.

L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18(4):1035–1064, 1997.

L. El Ghaoui, F. Oustry, and H. Lebret. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization*, 9:33–53, 1998.

J. Elzinga and D. Hearn. The minimum sphere covering a convex polyhedron. *Naval Research Logistics Quarterly*, 21(4):715–718, 1974.

A. Georghiou, A. Tsoukalas, and W. Wiesemann. A primal-dual lifting scheme for two-stage robust optimization. *Optimization Online*, 2017. URL `http://www.optimization-online.org/DB_FILE/2017/10/6261.pdf`.

B. Gorissen and D. den Hertog. Robust counterparts of inequalities containing sums of maxima of linear functions. *European Journal of Operational Research*, 227(1):30–43, 2013.

M. Hadjiyiannis, P. Goulart, and D. Kuhn. A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization. In *Proceedings IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 7386–7391, 2011.

G. Hanasusanto, D. Kuhn, and W. Wiesemann. K-adaptability in two-stage robust binary programming. *Operations Research*, 63(4):877–891, 2014.

P. Lax. *Linear Algebra and Its Applications*. Wiley, New York, 2 edition, 2007.

M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284(1-3):193–228, 1998.

J. Löfberg. Yalmip: A toolbox for modeling and optimization in matlab. In *Proceedings of the CACSD Conference*, volume 3. Taipei, Taiwan, 2004.

A. Mittal and G.A. Hanasusanto. Finding minimum volume circumscribing ellipsoids using copositive programming. *arXiv preprint arXiv:1807.07507*, 2018.

MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 8*, 2017. URL `http://docs.mosek.com/8.0/toolbox.pdf`.

K. Postek and D. den Hertog. Multi-stage adjustable robust mixed-integer optimization via iterative splitting of the uncertainty set. *INFORMS Journal on Computing*, 28(3):553–574, 2016.

E. Roos, D. den Hertog, A. Ben-Tal, F. de Ruiter, and J. Zhen. Approximation of hard uncertain convex inequalities. *Optimization Online*, 2019. URL `http://www.optimization-online.org/DB_FILE/2018/06/6679.pdf`.

P. Vayanos, D. Kuhn, and B. Rustem. Decision rules for information discovery in multi-stage stochastic programming. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference*, pages 7368–7373, 2011.

J. Zhen and D. den Hertog. Computing the maximum volume inscribed ellipsoid of a polytopic projection. *INFORMS Journal on Computing*, 30(1):31–42, 2017a.

J. Zhen, D. den Hertog, and M. Sim. Adjustable robust optimization via fourier-motzkin elimination. *Operations Research*, 66(4):893–1188, 2018.

A. Proof of Lemma 3

Lemma 3. *A matrix $Q \in \mathbb{R}^{m\times m}$ is positive semidefinite if and only if the trace of the product with any positive semidefinite matrix is positive.*

Proof. "$\Leftarrow$": Suppose for any $c \in \mathbb{R}^m$, $\mathrm{Tr}\,(QW) \geq 0$, where $W = cc^\top \succeq 0$. Since $\mathrm{Tr}\,(QW) = \mathrm{Tr}\,(Qcc^\top) = c^\top Qc \geq 0$ holds for any $c \in \mathbb{R}^m$, by definition, $Q \succeq 0$.

"$\Rightarrow$": Suppose $Q \succeq 0$. For any $W \succeq 0$, there exists a C such that $W = C^\top C$. We then have $\mathrm{Tr}\,(QW) = \mathrm{Tr}\,(QC^\top C) = \mathrm{Tr}\,(CQC^\top) \geq 0$ because $Q \succeq 0$ implies $CQC^\top \succeq 0$. □

B. Robust counterparts of (5) with lifted uncertainty set $\widehat{\mathcal{W}}$

By imposing the linear decision rule (11), we derive the convex reformulation of (5) with the lifted uncertainty set $\widehat{\mathcal{W}}$:

$$\forall (w,z) \in \widehat{\mathcal{W}} \ : \ \begin{cases} d^\top (v + Vw + Uz) + b(x)^\top w \leq c(x) \\ D^\top (v + Vw + Uz) \geq a(x) + A(x)^\top w \\ v + Vw + Uz \geq 0. \end{cases}$$

By using reformulation techniques in Ben-Tal et al. (2015), the following convex reformulation can be derived:

$$\begin{cases} d^\top v + \sum_{i=1}^m \kappa_i^{(1)} + \tau^{(1)} \leq c(x) \\ \left\| \begin{matrix} 2(d^\top V_i + b_i(x)) \\ 4\sigma_i^{(1)} - \kappa_i^{(1)} \end{matrix} \right\|_2 \leq 4\sigma_i^{(1)} + \kappa_i^{(1)} \qquad i = 1,\ldots,m \\ d^\top U_i + \sigma_i^{(1)} - \tau^{(1)} \leq 0 \qquad i = 1,\ldots,m \\ a_j(x) + \sum_{i=1}^m \kappa_{ij}^{(2)} + \tau_j^{(2)} \leq D_j^\top v \qquad j = 1,\ldots,n \\ \left\| \begin{matrix} 2(A_{ij}(x) - V_i^\top D_j) \\ 4\sigma_{ij}^{(2)} - \kappa_{ij}^{(2)} \end{matrix} \right\|_2 \leq 4\sigma_{ij}^{(2)} + \kappa_{ij}^{(2)} \qquad i = 1,\ldots,m \quad j = 1,\ldots,n \\ -D_j^\top U_i + \sigma_{ij}^{(2)} - \tau_j^{(2)} \leq 0 \qquad i = 1,\ldots,m \quad j = 1,\ldots,n \\ \sum_{i=1}^m \kappa_{il}^{(3)} + \tau_l^{(3)} \leq v_l \qquad l = 1,\ldots,r \\ \left\| \begin{matrix} -2V_{li} \\ 4\sigma_{il}^{(3)} - \kappa_{il}^{(3)} \end{matrix} \right\|_2 \leq 4\sigma_{il}^{(3)} + \kappa_{il}^{(3)} \qquad i = 1,\ldots,m \quad l = 1,\ldots,r \\ -U_{li} + \sigma_{il}^{(3)} - \tau_l^{(3)} \leq 0 \qquad i = 1,\ldots,m \quad l = 1,\ldots,r \\ \sigma_i^{(1)}, \sigma_{ij}^{(2)}, \sigma_{il}^{(3)}, \tau^{(1)}, \tau_j^{(2)}, \tau_l^{(3)} \geq 0 \qquad i = 1,\ldots,m \quad j = 1,\ldots,n \quad l = 1,\ldots,r. \end{cases}$$

□