An extension of the Reformulation-Linearization Technique to nonlinear optimization

Jianzhe Zhen
ETH Zürich, Zürich, Switzerland, jizhen@ethz.ch

Danique de Moor
Amsterdam Business School, University of Amsterdam, The Netherlands, d.demoor@uva.nl

Dick den Hertog
Amsterdam Business School, University of Amsterdam, The Netherlands, d.denhertog@uva.nl

We introduce a novel Reformulation-Perspectification Technique (RPT) to obtain convex approximations of nonconvex continuous optimization problems. RPT consists of two steps, those are, a reformulation step and a perspectification step. The reformulation step generates redundant nonconvex constraints from pairwise multiplication of the existing constraints. The perspectification step then convexifies the nonconvex components by using perspective functions. The proposed RPT extends the existing Reformulation-Linearization Technique (RLT) in two ways. First, it can multiply constraints that are not linear or not quadratic, and thereby obtain tighter approximations than RLT. Second, it can also handle more types of nonconvexity than RLT. We demonstrate the applicability of RPT by extensively analyzing all 15 possibilities of pairwise multiplication of the five basic cone constraints (linear cone, second-order cone, power cone, exponential cone, semi-definite cone). We show that many well-known RLT based results can also be obtained and extended by applying RPT. Numerical experiments on dike height optimization and convex maximization problems demonstrate the effectiveness of the proposed approach.

Key words: Reformulation-Linearization Technique, perspective function, conic optimization, conjugate function

---

1. Introduction

Reformulation-Linearization Technique (RLT) is a systematic method to derive hierarchical convex approximations of conic quadratic problems that contain nonconvex components. RLT consists of two steps, those are, a reformulation step and a linearization step. The reformulation step generates redundant nonconvex constraints from pairwise multiplication of the existing linear or quadratic inequalities. The linearization step then substitutes each distinct product of variables by a newly introduced continuous variable. For mixed binary linear programs, a hierarchy of tighter and tighter relaxations can be obtained from RLT, and an explicit algebraic characterization of the convex hull is obtained at the highest level, level-n_x, where n_x is the number of binary variables. The method is also applicable to mixed binary polynomial and to continuous, nonconvex programs (Sherali and Adams, 2013), and has also been extended to mixed binary semi-infinite and convex programs (Sherali and Adams, 1994a). RLT was introduced in Sherali and Adams (1990), and improved by many authors (Sturm and Zhang, 2003; Anstreicher, 2009, 2012, 2017; Bao et al., 2011; Yang and Burer, 2016; Jiang and Li, 2016, 2019). We also refer to Jiang and Li (2020) for an overview of RLT approximations for quadratic optimization problems.

In this paper we propose an extension of RLT, which we call *Reformulation-Perspectification Technique* (RPT). The RPT also consists of two steps, those are, a reformulation step and a perspectification step. Similarly as in RLT, the reformulation step of RPT generates redundant nonconvex constraints from pairwise multiplication of the existing inequalities. In the perspectification step, the nonconvex components are convexified by first reformulating them into their perspective form, and substitutes each distinct product of variables by a newly introduced continuous variable. Overall, the RPT extends RLT in two ways. First, pairwise multiplication of inequalities that are not necessarily linear or quadratic are also considered in the reformulation step of RPT, and thereby obtains tighter approximations than RLT based methods. Second, more types of nonconvexity can be handled in the perspectification step of RPT than in the linearization step of RLT. We show that the RPT approach is able to generate convex approximations if the nonconvex components are *sums of products of convex* functions. This includes nonconvex quadratic functions and difference of convex functions.

We first demonstrate the value of these two extensions to RLT by using the following toy problem:

$$\begin{array}{ll} \min\limits_{x_1,x_2} & 2x_1+3x_2-5x_1x_2-(x_1+2)\ln(x_1+2) \\ \text{s.t.} & x_1+x_2\le 1 \\ & \exp(-x_1)+\exp(-x_2)\le 1+\exp(-1). \end{array} \tag{T}$$

Note that $\ln(x_1+2)$ is well defined if $x_1>-2$, which is ensured by the second inequality. The objective consists of a sum of two nonconvex components, those are, $-5x_1x_2$ and $-(x_1+2)\ln(x_1+2)$. The first nonconvex component can be dealt with by RLT, but the second can not. In the reformulation step of RLT the nonconvex constraint $(1-x_1-x_2)^2\ge 0$ is generated, while in the linearization step, the product of variables x_1x_2 is then substituted by a continuous variable $u_{12}\in\mathbb{R}$, which is unbounded above. The obtained nonconvex program from RLT yields a relaxation of Problem (T) with an optimal objective value $-\infty$. Our RPT approach, however, can deal with both nonconvex components in Problem (T), and can generate additional constraints by multiplying the linear and the nonlinear constraints. Finally, the obtained convex relaxation of Problem (T) has an optimal objective value -1.4830, which coincides with the true optimal objective value of Problem (T).

Most practical problems can be modeled in terms of five basic conic inequalities, those are, linear cone, second-order cone, power cone, exponential cone, semi-definite cone. To demonstrate the applicability of RPT, we extensively analyze all 15 possibilities of multiplying two out of the five basic conic inequalities. We also show that many well-known RLT based techniques for reformulating nonconvex problems into convex ones, can also be obtained by applying RPT. Many existing effective cuts or projection methods can be easily integrated with RPT to further tighten the obtained convex approximations. For instance, a linear matrix inequality from the SDP relaxation of nonconvex quadratic equality can be additionally considered in the obtained convex approximations from RPT.

In Zhen et al. (2018) and Selvi et al. (2020) bilinear and concave maximization problems, respectively, are reformulated into adjustable robust linear ones. In Zhen et al. (2018) it is proven that using linear decision rules for these adjustable problems, is equivalent with applying partial RLT to the original problem. The RPT approach proposed in this paper yields tighter approximations for the bilinear and convex maximization cases studied in Zhen et al. (2018) and Selvi et al. (2020), and can also be applied to a much wider class of optimization problems.

We summarize the main contributions of this paper:

1. We extend the existing RLT approach to nonlinear optimization problems in two significant ways. First, the proposed RPT approach can handle multiplication of constraints that are not linear or not quadratic, and thereby obtain tighter approximations than RLT. Second, it can also handle more types of nonconvexity than RLT.
2. We analyze all 15 possibilities of pairwise multiplication of the five basic cone constraints (linear, second-order, power, exponential, semi-definite). Especially the results for the cases in which a power cone or an exponential cone is involved are new.
3. We show that several existing convex reformulations and relaxations for disjunction optimization, generalized linear optimization, approximate S-lemma for quadratic optimization, and fractional optimization, can also be obtained via the RPT approach.

Moreover, the most important second-level contributions of this paper are:

1. We show that introducing epigraphical variables for the convex components of the nonconvex constraint or objective functions of the optimization problem may yield a tighter approximation.
2. Several ways to deal with the product of two conic quadratic inequalities have been proposed in the literature. In this paper we propose a new one.
3. For quadratic constraints we prove that one obtains the tighest approximation by linearizing all quadratic terms, instead of only the concave part.
4. We prove that if a linear constraint is redundant to existing linear constraints, then the resulting additional constraints from RLT or RPT are also redundant.
5. If we multiply a convex constraint that is conically representable with a linear constraint, then we have two options: we either first multiply the two constraints and then derive a conic reformulation, or we first derive the conic formulation of the convex constraint and then multiply this by the linear one. We prove that both options lead to the same result.

In Section 2 we explain the core idea by applying RPT to the toy problem (T). In Section 3, we introduce the concept of SCC and SLC functions, i.e., functions that are sums of convex times convex functions and linear times convex functions, respectively, to explain the optimization problem class for which our RPT method can be used. In Section 4 we explain the main steps of our RPT method in more detail. In Section 5 we show how to multiply two out of the five basic cones with each other, and how to convexify that. In Section 6, we extend our discussion to difference of convex functions. In Section 7, we show that several convex reformulations for several classes of nonconvex problems derived in the literature can also be obtained via RPT. Section 8 contains numerical results.

Notation. The calligraphic letters $\mathcal{I}$, $\mathcal{J}$, $\mathcal{K}$, $\mathcal{L}$ and the corresponding capital Roman letters I, J, K, L are reserved for finite index sets and their respective cardinalities, *i.e.*, $\mathcal{I} = \{1,\ldots,I\}$ etc. The subscript 0 for an index set indicates that the set additionally includes 0, *i.e.*, $\mathcal{I}_0 = \{0,\ldots,I\}$ etc. Let $\mathbb{R}^{m\times n}$ denote the set of real $m\times n$ matrices, and $\mathbb{S}^n$ the set of real $n\times n$ symmetric matrices. The perspective $h:\mathbb{R}^{n_\nu}\times\mathbb{R}_+\rightarrow[-\infty,+\infty]$ of a proper, closed and convex function $f:\mathbb{R}^{n_\nu}\rightarrow[-\infty,+\infty]$ is defined for all $\boldsymbol{\nu}\in\mathbb{R}^{n_\nu}$ and $t\in\mathbb{R}_+$ as $h(\boldsymbol{\nu},t)=tf(\boldsymbol{\nu}/t)$ if $t>0$, and $h(\boldsymbol{\nu},0)=\delta^*_{\mathrm{dom}(f^*)}(\boldsymbol{\nu})$. For ease of exposition, we use $tf(\boldsymbol{\nu}/t)$ to denote the perspective function $h(\boldsymbol{\nu},t)$ for the rest of this paper.

2. A Simple Example

In this section we demonstrate the core idea of our RPT approach by solving Problem (T). There are two steps in our RPT approach.

In the **reformulation step**, we generate redundant constraints from pairwise multiplication of existing constraints.

(i) Multiply the linear constraint with itself:

$$0 \le (x_1+x_2-1)^2.$$

(ii) Multiply the linear constraint with the nonlinear constraint:

$$(1-x_1-x_2)\exp(-x_1)+(1-x_1-x_2)\exp(-x_2)\le(1+\exp(-1))(1-x_1-x_2).$$

Note that the first constraint is not convex, because the convex quadratic function appears on the right-hand-side of the "$\le$"-inequality, and there are two nonconvex components $(1-x_1-x_2)\exp(-x_1)$ and $(1-x_1-x_2)\exp(-x_2)$ in the second constraint.

In the **perspectification step**, the nonconvex components in (T) and from the reformulation step are "perspectified". The objective function of (T) has two nonconvex components $-5x_1x_2$ and $-(x_1+2)\ln(x_1+2)$. We reformulate these two components into the following equivalent form:

$$\begin{aligned}
&-5x_1x_2 = -5x_1\frac{x_1x_2}{x_1} = -5x_1x_2\\
&-(x_1+2)\ln(x_1+2) = -(x_1+2)\ln\left(\frac{(x_1+2)^2}{x_1+2}\right) = -(x_1+2)\ln\left(\frac{x_1^2+4x_1+4}{x_1+2}\right).
\end{aligned}$$

Analogously, we also reformulate the nonconvex components $(1-x_1-x_2)\exp(-x_1)$ and $(1-x_1-x_2)\exp(-x_2)$ in the second constraint generated from the reformulation step as follows:

$$(1-x_1-x_2)\exp(-x_1) = (1-x_1-x_2)\exp\left(\frac{-x_1+x_1^2+x_1x_2}{1-x_1-x_2}\right)$$
$$(1-x_1-x_2)\exp(-x_2) = (1-x_1-x_2)\exp\left(\frac{-x_2+x_1x_2+x_2^2}{1-x_1-x_2}\right).$$

Finally, all the product of variables x_1^2, x_2^2 and x_1x_2 are substituted with newly introduced variables u_{11}, u_{22} and u_{12}, respectively. The convex relaxation that results from the RPT approach is therefore:

$$\begin{array}{lll} \min\limits_{\substack{x_1,x_2\\ u_{11},u_{12},u_{22}}} & 2x_1+3x_2-5u_{12}-(x_1+2)\ln\left(\frac{u_{11}+4x_1+4}{x_1+2}\right) & \\ \text{s.t.} & x_1+x_2\le 1 & \\ & u_{11}+u_{22}+2u_{12}-2x_1-2x_2+1\ge 0 & \\ & \exp(-x_1)+\exp(-x_2)\le 1+\exp(-1) & \text{(TC)} \\ & (1-x_1-x_2)\exp\left(\dfrac{u_{11}+u_{12}-x_1}{1-x_1-x_2}\right)+(1-x_1-x_2)\exp\left(\dfrac{u_{22}+u_{12}-x_2}{1-x_1-x_2}\right) & \\ & \qquad\qquad \le (1+\exp(-1))(1-x_1-x_2). & \end{array}$$

The solution of (TC) appears to be

$$\boldsymbol{x}' = \begin{bmatrix}1\\0\end{bmatrix} \quad \text{and} \quad \boldsymbol{U}' = \begin{bmatrix}u'_{11} & u'_{12}\\ u'_{21} & u'_{22}\end{bmatrix} = \begin{bmatrix}-7.4 & 8.4\\ 8.4 & -8.4\end{bmatrix},$$

with objective value -35.17, which constitutes a tighter lower bound on the optimal value of (T) than the one obtained from applying RLT, that is, $-\infty$. The obtained $\boldsymbol{x}'$ is a feasible solution to (T), and its corresponding objective value is -1.30, which constitutes an upper bound on the optimal value of (T).

REMARK 1. The nonconvex components $-5x_1x_2$, $-(x_1+2)\ln(x_1+2)$, $(1-x_1-x_2)\exp(-x_1)$ and $(1-x_1-x_2)\exp(-x_2)$ are products of a linear function and a convex function. The class of SLC functions is formally introduced and discussed in the next section.

To further tighten the relaxation (TC), one can consider the convex relaxation for $\boldsymbol{U} = \boldsymbol{x}\boldsymbol{x}^\top$:

$$\begin{pmatrix} u_{11} & u_{12} & x_1\\ u_{12} & u_{22} & x_2\\ x_1 & x_2 & 1\end{pmatrix} \succeq 0.$$

It turns out that if this LMI is added to the constraints of (TC), then the obtained solution is:

$$\boldsymbol{x}^* = \begin{bmatrix}0.80\\0.20\end{bmatrix} \quad \text{and} \quad \boldsymbol{U}^* = \begin{bmatrix}0.64 & 0.16\\ 0.16 & 0.04\end{bmatrix},$$

with objective value -1.4830, which again constitutes a lower bound on the optimal value of (T), while the objective function of (T) evaluated at the feasible solution $\boldsymbol{x}^*$ is also -1.4830, which certifies that $\boldsymbol{x}^*$ is an optimal solution to (T). Extra constraints from nonlinear times nonlinear constraints (see Section 5) can also be considered in (TC), but clearly those constraints are redundant for (TC) with the additional LMI.

3. Problem formulation

3.1. Optimization problems with SLC functions

In the previous section we have discussed a special optimization problem that contains several nonconvex components, which were convexified via RPT. In this section we give a formal description of the type of nonconvex optimization problems that we consider in this paper, and describe how these problems can be convexified by using perspective functions. We focus on optimization problems with nonconvex components that are *sum of linear functions times convex* (SLC) functions.

DEFINITION 1 (SLC FUNCTIONS). A function $f:\mathbb{R}^{n_x}\to[-\infty,+\infty]$ is SLC if f can be written as the sum of linear functions times convex functions, i.e.,

$$f(\boldsymbol{x}) = c_0(\boldsymbol{x}) + \sum_{i\in\mathcal{I}} \left(q_i - \boldsymbol{d}_i^\top \boldsymbol{x}\right) c_i(\boldsymbol{x}),$$

where $q_i\in\mathbb{R}$, $\boldsymbol{d}_i\in\mathbb{R}^{n_x}$, and c_0, $c_i:\mathbb{R}^{n_x}\to[-\infty,+\infty]$, $i\in\mathcal{I}$, are proper, closed and convex. □

Notice that convex functions are also SLC functions.

We consider a generic nonconvex optimization problem of the following form:

$$\begin{array}{lll} \min\limits_{\boldsymbol{x}} & f_0(\boldsymbol{x}) & \\ \text{s.t.} & f_k(\boldsymbol{x}) \le 0 & k\in\mathcal{K} \\ & \boldsymbol{x}\in\mathcal{X}, & \end{array} \tag{P}$$

where $f_k:\mathbb{R}^{n_x}\to[-\infty,+\infty]$ is an SLC function, that is,

$$f_k(\boldsymbol{x}) = c_{0k}(\boldsymbol{x}) + \sum_{i\in\mathcal{I}} \left(q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}\right) c_{ik}(\boldsymbol{x}),$$

and $q_{ik}\in\mathbb{R}$, $\boldsymbol{d}_{ik}\in\mathbb{R}^{n_x}$, and $c_{0k}, c_{ik}:\mathbb{R}^{n_x}\to[-\infty,+\infty]$, $k\in\mathcal{K}_0$, are proper, closed and convex for every $i\in\mathcal{I}$. The nonempty, compact and convex set $\mathcal{X}$ is defined through:

$$\mathcal{X} = \{\boldsymbol{x}\in\mathbb{R}^{n_x} \mid \boldsymbol{A}\boldsymbol{x}\le\boldsymbol{b},\ \boldsymbol{h}(\boldsymbol{x})\le\boldsymbol{0}\},$$

where $\boldsymbol{A}\in\mathbb{R}^{L\times n_x}$, $\boldsymbol{b}\in\mathbb{R}^L$, $\boldsymbol{h}(\boldsymbol{x}) = [h_0(\boldsymbol{x})\ \ h_1(\boldsymbol{x})\ \ \cdots\ \ h_J(\boldsymbol{x})]^\top \subseteq [-\infty,+\infty]^{J+1}$, and $h_j:\mathbb{R}^{n_x}\to[-\infty,+\infty]$ is proper, closed and convex for every $j\in\mathcal{J}$. In order to apply our RPT approach to solve Problem (P) with generally nonconvex SLC constraints, the following assumption is stipulated.

ASSUMPTION 1. *If the function c_{ik} is nonlinear, $i\in\mathcal{I}$ and $k\in\mathcal{K}_0$, then $\boldsymbol{d}_{ik}^\top\boldsymbol{x}\le q_{ik}$ for all $\boldsymbol{x}\in\mathcal{X}$.*

Now we are ready to demonstrate the core idea of RPT. Let f be an SLC function that satisfies Assumption 1, then we can perspectify the generally nonconvex function f by first multiplying and dividing the argument of f by $(q_i - \boldsymbol{d}_i^\top\boldsymbol{x})$ to obtain the following equivalent reformulation of f:

$$f(\boldsymbol{x}) = c_0(\boldsymbol{x}) + \sum_{i\in\mathcal{I}} (q_i - \boldsymbol{d}_i^\top\boldsymbol{x}) c_i\left(\frac{q_i\boldsymbol{x} - \boldsymbol{x}\boldsymbol{x}^\top\boldsymbol{d}_i}{q_i - \boldsymbol{d}_i^\top\boldsymbol{x}}\right).$$

Then, the quadratic terms $\boldsymbol{x}\boldsymbol{x}^\top$ in the argument of the reformulated f can be linearized by substituting $\boldsymbol{x}\boldsymbol{x}^\top$ with $\boldsymbol{U}\in\mathbb{S}^{n_x}$ to obtain the following sum of perspective functions:

$$\sum_{i\in\mathcal{I}}(q_i-\boldsymbol{d}_i^\top\boldsymbol{x})c_i\left(\frac{q_i\boldsymbol{x}-\boldsymbol{U}\boldsymbol{d}_i}{q_i-\boldsymbol{d}_i^\top\boldsymbol{x}}\right),$$

which is jointly convex in $(\boldsymbol{x},\boldsymbol{U})$ because c_i is convex if and only if its perspective is convex (Rockafellar, 1970).

In Table 1, we give some examples of SLC functions that are generally nonconvex and satisfy Assumption 1. Hence, the approach proposed in this paper can deal with Problem (P) containing (sum of) such nonconvex components. In Section 6, we show that under mild regularity conditions an important class of functions, that is, difference of convex functions, can be equivalently represented into SLC functions that satisfy Assumption 1.

Table 1 Examples of SLC representable functions.

f	c	$(q-\boldsymbol{d}^\top\boldsymbol{x})$	Perspectification	Assumptions
$-x\ln x$	$-\ln x$	x	$-x\ln(u/x)$	$x\geq 0$
$\sqrt{x}$	$x^{-1/2}$	x	$x\sqrt{x/u}$	$x\geq 0$
x^θ	$x^{\theta-1}$	x	$x(u/x)^{\theta-1}$	$\theta\in[0,1]$ & $x\geq 0$
$-x^\theta$	$-x^{\theta-1}$	x	$-x(u/x)^{\theta-1}$	$\theta\in[1,2]$ & $x\geq 0$
$-x_1\ln x_2$	$-\ln x_2$	x_1	$-x_1\ln(u_{12}/x_1)$	$x_1,x_2\geq 0$
$\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}$	$(\boldsymbol{Q}\boldsymbol{x})_i$	x_i	$\mathrm{Tr}(\boldsymbol{U}\boldsymbol{Q})$	-
$(q-\boldsymbol{d}^\top\boldsymbol{x})\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}$	$\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}$	$(q-\boldsymbol{d}^\top\boldsymbol{x})$	$\frac{(q\boldsymbol{x}-\boldsymbol{U}\boldsymbol{d})^\top\boldsymbol{Q}(q\boldsymbol{x}-\boldsymbol{U}\boldsymbol{d})}{(q-\boldsymbol{d}^\top\boldsymbol{x})}$	$\boldsymbol{d}^\top\boldsymbol{x}\leq q$ & $\boldsymbol{Q}\succeq\boldsymbol{0}$

3.2. Extension to SCC functions

One can further extend our RPT approach to derive a convex relaxation of Problem (P) with the nonconvex components f_k, $k\in\mathcal{K}_0$, that are *sum of convex times convex functions* (SCC).

Definition 2 (SCC functions). A function $f:\mathbb{R}^{n_x}\to[-\infty,+\infty]$ is SCC if f can be written as the sum of convex functions times convex functions, i.e.,

$$f(\boldsymbol{x})=c_0(\boldsymbol{x})+\sum_{i\in\mathcal{I}}r_i(\boldsymbol{x})c_i(\boldsymbol{x}),$$

where $c_0,r_i,c_i:\mathbb{R}^{n_x}\to[-\infty,+\infty]$ are proper, closed and convex functions for every $i\in\mathcal{I}$. □

In order to apply our RPT approach to solve Problem (P) with SCC functions f_k, $k\in\mathcal{K}_0$, where

$$f_k(\boldsymbol{x})=c_{0k}(\boldsymbol{x})+\sum_{i\in\mathcal{I}}r_{ik}(\boldsymbol{x})c_{ik}(\boldsymbol{x}),$$

and $c_{0k},r_{ik},c_{ik}:\mathbb{R}^{n_x}\to[-\infty,+\infty]$, $k\in\mathcal{K}_0$, are proper, closed and convex functions for every $i\in\mathcal{I}$, the following assumption is stipulated.

Assumption 2. *If r_{ik} and c_{ik} are both nonlinear, then we assume that $r_{ik}(\boldsymbol{x}) \geq 0$ and $c_{ik}(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \mathcal{X}$, for every $i \in \mathcal{I}$ and for every $k \in \mathcal{K}_0$, while if r_{ik} is linear and c_{ik} is nonlinear, then $r_{ik}(\boldsymbol{x}) \geq 0$ for all $\boldsymbol{x} \in \mathcal{X}$, for every $i \in \mathcal{I}$ and for every $k \in \mathcal{K}_0$.*

We now show that Problem (P) with nonconvex but SCC functions f_k, $k \in \mathcal{K}_0$, without loss of generality can be equivalently reformulated as an instance of Problem (P) with only SLC functions. For an SCC function f_k that satisfies Assumption 2, the nonconvex constraint can be reformulated into a set of two constraints, i.e., a nonconvex constraint with SLC functions and a convex constraint:

$$c_{0k}(\boldsymbol{x}) + \sum_{i \in \mathcal{I}} r_{ik}(\boldsymbol{x}) c_{ik}(\boldsymbol{x}) \leq 0 \iff \begin{cases} c_{0k}(\boldsymbol{x}) + \sum_{i \in \mathcal{I}} \tau_{ik} c_{ik}(\boldsymbol{x}) \leq 0 & \\ r_{ik}(\boldsymbol{x}) \leq \tau_{ik} & \text{if } r_{ik} \text{ and } c_{ik} \text{ are nonlinear} \\ r_{ik}(\boldsymbol{x}) = \tau_{ik} & \text{if } r_{ik} \text{ is linear.} \end{cases}$$

An example of SCC functions is fractional optimization (see also in Section 7.4)

$$f(\boldsymbol{x}) = \sum_{i \in \mathcal{I}} \frac{c_i(\boldsymbol{x})}{r_i(\boldsymbol{x})},$$

where $c_i : \mathbb{R}^{n_x} \to \mathbb{R}_+$ is convex and $r_i : \mathbb{R}^{n_x} \to \mathbb{R}_{++}$ is concave for every $i \in \mathcal{I}$. Then f is not necessarily convex or concave. However, the function is SCC, since $1/r_i(\boldsymbol{x})$ is convex and nonnegative.

4. Reformulation-Perspectification Technique and practical issues

4.1. Reformulation-Perspectification Technique

We describe our new RPT approach for Problem (P). It consists of two steps (for an illustration of each of the steps on a toy problem, see Section 2):

Step 1: Reformulation. We first introduce epigraphical variables for the convex components in the nonconvex SLC functions of (P), and we have

$$\begin{array}{lll} \min\limits_{\boldsymbol{x}, \boldsymbol{\tau}} & \tau_0 + \sum_{i \in \mathcal{I}} (q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}) \, c_{i0}(\boldsymbol{x}) & \\ \text{s.t.} & \tau_k + \sum_{i \in \mathcal{I}} (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) \, c_{ik}(\boldsymbol{x}) \leq 0 & k \in \mathcal{K} \\ & (\boldsymbol{x}, \boldsymbol{\tau}) \in \mathcal{T}, & \end{array}$$

where $\mathcal{T} = \{(\boldsymbol{x}, \boldsymbol{\tau}) \in \mathbb{R}^{n_x} \times \mathbb{R}^{K+1} \mid \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{c}_0(\boldsymbol{x}) \leq \boldsymbol{\tau}\}$, and $\boldsymbol{c}_0(\boldsymbol{x}) = [c_{00}(\boldsymbol{x}) \ c_{01}(\boldsymbol{x}) \ \cdots \ c_{0K}(\boldsymbol{x})]^\top \subseteq [-\infty, +\infty]^{K+1}$. We then multiply each pair of constraints in the set $\mathcal{T}$, such that the final result is an SLC constraint function. We consider three possibilities:

(a) *Linear* × *Linear.* This is well-known in RLT: by multiplying the constraints $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$ of (P) with $\boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}$, we obtain $L(L+1)/2$ redundant constraints:

$$\boldsymbol{b}\boldsymbol{x}^\top \boldsymbol{A}^\top + \boldsymbol{A}\boldsymbol{x}\boldsymbol{b}^\top \leq \boldsymbol{A}\boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{A}^\top + \boldsymbol{b}\boldsymbol{b}^\top,$$

since the (i,j)-th constraint is exactly the (j,i)-th constraint. Hence, we only consider the upper triangular of the matrix equations; so $L(L+1)/2$ constraints instead of L^2.

(b) *Linear* $\times$ *Convex.* By multiplying the ℓ-th linear constraint $\boldsymbol{a}_\ell^\top \boldsymbol{x} \le b_\ell$ of (P) with the convex constraints $\boldsymbol{c}_0(\boldsymbol{x}) \le \boldsymbol{\tau}$ and $\boldsymbol{h}(\boldsymbol{x}) \le \boldsymbol{0}$, we obtain $L(J+K+2)$ redundant SLC constraints of the form

$$(b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}) \le \boldsymbol{0} \quad \text{and} \quad (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{c}_0(\boldsymbol{x}) \le (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{\tau} \qquad \ell \in \mathcal{L}.$$

(c) *Convex* $\times$ *Convex.* Just multiplying a nonlinear convex constraint $h_j(\boldsymbol{x}) \le 0$ with another nonlinear convex constraint $h_{j'}(\boldsymbol{x}) \le 0$ results in a constraint $-h_j(\boldsymbol{x})h_{j'}(\boldsymbol{x}) \le 0$ for which the constraint function is not an SLC or SCC function. However, sometimes rewriting the constraints, and then multiplying the left-hand-sides and right-hand sides of the constraints yield convexifiable constraints. This is illustrated in the next section where we multiply pairs of conic constraints.

Step 2: Perspectification. Since the functions in the newly constructed constraints from Step 1 as well as the functions f_k, $k \in \mathcal{K}_0$, are SLC functions, we can perspectify these functions as done in Section 3 analogously. For example, Linear $\times$ Convex constraints can be reformulated into:

$$\begin{aligned}(b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}) \le \boldsymbol{0} \quad &\iff \quad (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{h}\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}}\right) \le \boldsymbol{0} \quad \text{and} \\ (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{c}(\boldsymbol{x}) \le (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{\tau} \quad &\iff \quad (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{c}\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}}\right) \le (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{\tau}.\end{aligned}$$

Finally, the nonlinear quadratic terms $\boldsymbol{x}\boldsymbol{x}^\top$ and the bilinear terms $\boldsymbol{x}\boldsymbol{\tau}^\top$ produced in Steps 1 & 2 are linearized by substituting them with $\boldsymbol{U} \in \mathbb{S}^{n_x}$ and $\boldsymbol{V} \in \mathbb{R}^{n_x \times (K+1)}$.

If the functions f_k, $k \in \mathcal{K}_0$, in (P) satisfy Assumption 1, then one can use RPT to obtain the following convex relaxation of (P):

$$\begin{array}{lll} \min\limits_{\boldsymbol{x},\boldsymbol{\tau},\boldsymbol{U},\boldsymbol{V}} & \tau_0 + \sum\limits_{i\in\mathcal{I}} (q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}) c_{i0}\left(\dfrac{q_{i0}\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}_{i0}}{q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}}\right) & \\ \text{s.t.} & \tau_k + \sum\limits_{i\in\mathcal{I}} (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) c_{ik}\left(\dfrac{q_{ik}\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}}\right) \le 0 & k \in \mathcal{K} \\ & \boldsymbol{b}\boldsymbol{x}^\top \boldsymbol{A}^\top + \boldsymbol{A}\boldsymbol{x}\boldsymbol{b}^\top \le \boldsymbol{A}\boldsymbol{U}\boldsymbol{A}^\top + \boldsymbol{b}\boldsymbol{b}^\top & \\ & (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{h}\left(\dfrac{b_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}}\right) \le \boldsymbol{0} & \ell \in \mathcal{L} \\ & (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{c}_0\left(\dfrac{b_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}}\right) \le b_\ell \boldsymbol{\tau} - \boldsymbol{V}^\top \boldsymbol{a}_\ell & \ell \in \mathcal{L} \\ & (\boldsymbol{x}, \boldsymbol{\tau}) \in \mathcal{T}. & \end{array} \tag{PC}$$

We do not consider *Convex* $\times$ *Convex* functions in (PC) because there are many options for different type of inequalities, and we refer to Section 5.1 for more details. Note that there are $L(J+1+\frac{L+1}{2}) + (L+1)(K+1)$ additional constraints and $n_x^2 + (n_x+1)(K+1)$ additional variables in (PC) compared to (P). The optimal value of (PC) constitutes a lower bound on that of (P), while the optimal solution $\boldsymbol{x}'$ of (PC) constitutes a candidate solution for (P), and the corresponding objective function evaluated at $\boldsymbol{x}'$ yields an upper bound on the optimal value of (P).

REMARK 2 (EPIGRAPH OF NONLINEAR CONVEX COMPONENTS). We remark that introducing epigraphical variables for nonlinear convex components in the reformulation step of RPT indeed tightens the convex relaxation (PC) (see, e.g., Example 3 in Section 6.1), while for linear components, we show via Theorem 2 in Section 4.2.2 that the redundant linear constraints to the existing ones are also redundant for RPT. Therefore, in the reformulation step of RPT, we only introduce epigraphical variables for the nonlinear convex components in the SLC functions of (P).

4.2. Redundant linear constraints

We first formally define *redundant constraints*, and then show in Sections 4.2.1 and 4.2.2 two cases where some of the additional constraints from RPT are redundant.

DEFINITION 3 (REDUNDANT CONSTRAINTS). A constraint $f(\boldsymbol{x}) \leq 0$ or $f(\boldsymbol{x}) = 0$, where $f : \mathbb{R}^{d_{n_x}} \to [-\infty, +\infty]$, is *redundant* to a nonempty set $\mathcal{X} \subseteq \mathbb{R}^{n_x}$ if $\mathcal{X} \subseteq \{\boldsymbol{x} \mid f(\boldsymbol{x}) \leq 0\}$ or $\mathcal{X} \subseteq \{\boldsymbol{x} \mid f(\boldsymbol{x}) = 0\}$, respectively.

4.2.1. Linear equality constraints. When multiplying a convex inequality constraint with a linear equality constraint in RPT, the denominator and coefficient of the resulting perspective function are zero. Fortunately, all additional nonlinear constraints resulting from multiplying a linear equality constraint with a convex inequality constraint are redundant as long as the additional linear constraints resulting from multiplying the linear equality constraint with the variables are considered. This is proven in the following theorem.

THEOREM 1. *Let $\boldsymbol{d}^\top \boldsymbol{x} - q = 0$ be an equality constraint, where $\boldsymbol{d} \in \mathbb{R}^{n_x}$ and $q \in \mathbb{R}$. Let $f : \mathbb{R}^{n_x} \to [-\infty, +\infty]$ be a proper, closed and convex function. Then the constraint $(q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{q\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) = \boldsymbol{0}$ is redundant to*

$$\{(\boldsymbol{x}, \boldsymbol{U}) \mid q\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d} = \boldsymbol{0}\}.$$

Proof. For any closed convex function we have $f^{**} = f$ (Rockafellar, 1970, p. 104), hence by the definition of the conjugate function we have

$$tf\left(\frac{\boldsymbol{x}}{t}\right) = tf^{**}\left(\frac{\boldsymbol{x}}{t}\right) = t \sup_{\boldsymbol{y} \in \mathrm{dom}(f^*)} \left\{\boldsymbol{y}^\top \frac{\boldsymbol{x}}{t} - f^*(\boldsymbol{y})\right\} = \sup_{\boldsymbol{y} \in \mathrm{dom}(f^*)} \left\{\boldsymbol{y}^\top \boldsymbol{x} - tf^*(\boldsymbol{y})\right\}.$$

For $\boldsymbol{x} = \boldsymbol{0}$ and $t = 0$, we therefore have

$$0f\left(\frac{\boldsymbol{0}}{0}\right) = \sup_{\boldsymbol{y} \in \mathrm{dom}(f^*)} \{\boldsymbol{y}^\top \boldsymbol{0} - 0 \cdot f^*(\boldsymbol{y})\} = 0,$$

as long as $\mathrm{dom}(f^*)$ is nonempty. Since $q - \boldsymbol{d}^\top \boldsymbol{x} = 0$ we have $(q - \boldsymbol{d}^\top \boldsymbol{x})\boldsymbol{x} = 0$ implies $q\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d} = \boldsymbol{0}$. f proper and convex implies f^* proper closed and convex (Rockafellar, 1970, Theorem 12.2). The properness of f^* implies $\mathrm{dom}(f^*)$ is nonempty (Rockafellar, 1970, p. 24). Hence the claim follows. □

4.2.2. Linear inequality constraints. We show that if a linear constraint is redundant to existing linear constraints, then the resulting additional constraints from RPT are also redundant.

THEOREM 2. *If the linear constraint $\boldsymbol{d}^\top \boldsymbol{x} \le q$, where $\boldsymbol{d} \in \mathbb{R}^{n_x}$ with $\boldsymbol{d} \ne \boldsymbol{0}$ and $q \in \mathbb{R}$, is redundant to $\{\boldsymbol{x} \mid \boldsymbol{B}\boldsymbol{x} \le \boldsymbol{p}\} \ne \emptyset$, where $\boldsymbol{B} \in \mathbb{R}^{L \times n_x}$ and $\boldsymbol{p} \in \mathbb{R}^L$, then the constraints $\boldsymbol{d}^\top \boldsymbol{x} \le q$, $2q\boldsymbol{d}^\top \boldsymbol{x} \le \boldsymbol{d}^\top \boldsymbol{U}\boldsymbol{d} + q^2$ and $(q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{q\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) \le \boldsymbol{0}$ are redundant to*

$$\left\{ (\boldsymbol{x}, \boldsymbol{U}) \;\middle|\; \begin{array}{l} \boldsymbol{B}\boldsymbol{x} \le \boldsymbol{p} \\ \boldsymbol{p}\boldsymbol{x}^\top \boldsymbol{B}^\top + \boldsymbol{B}\boldsymbol{x}\boldsymbol{p}^\top \le \boldsymbol{B}\boldsymbol{U}\boldsymbol{B}^\top + \boldsymbol{p}\boldsymbol{p}^\top \\ (p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x}) f\left(\frac{p_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{b}_\ell}{p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x}}\right) \le \boldsymbol{0}, \ \ell \in \mathcal{L} \end{array} \right\},$$

where $f : \mathbb{R}^{n_x} \to [-\infty, +\infty]$ is proper, closed and convex, and $\boldsymbol{b}_\ell$ is the ℓ-th column of the matrix $\boldsymbol{B}$.

Proof. Assume that $\boldsymbol{d}^\top \boldsymbol{x} \le q$ is redundant to $\{\boldsymbol{x} \mid \boldsymbol{B}\boldsymbol{x} \le \boldsymbol{p}\}$, then the optimal values of

$$\begin{array}{cl} \min\limits_{\boldsymbol{x}} & q - \boldsymbol{d}^\top \boldsymbol{x} \\ \text{s.t.} & \boldsymbol{B}\boldsymbol{x} \le \boldsymbol{p} \end{array} \qquad \text{and} \qquad \begin{array}{cl} \max\limits_{\boldsymbol{y} \ge \boldsymbol{0}} & q - \boldsymbol{p}^\top \boldsymbol{y} \\ \text{s.t.} & \boldsymbol{B}^\top \boldsymbol{y} = \boldsymbol{d} \end{array}$$

coincide and both are nonnegative thanks to the strong duality of linear programs, which implies that there exists a $\boldsymbol{y} \in \mathbb{R}^L_+$ such that $\boldsymbol{d}^\top \boldsymbol{x} \le q$ is redundant to $\{\boldsymbol{x} \mid \boldsymbol{b}_y^\top \boldsymbol{x} \le p_y\}$, where $\boldsymbol{b}_y = \boldsymbol{B}^\top \boldsymbol{y} = \boldsymbol{d}$ and $p_y = \boldsymbol{p}^\top \boldsymbol{y} \le q$. Clearly, the constraint $\boldsymbol{b}_y^\top \boldsymbol{x} \le p_y$ is redundant to $\{\boldsymbol{x} \mid \boldsymbol{B}\boldsymbol{x} \le \boldsymbol{p}\}$ because linear inequalities are affine invariant. Then, for any $\boldsymbol{x}$ that satisfies $\boldsymbol{b}_y^\top \boldsymbol{x} \le p_y$ and $f(\boldsymbol{x}) \le 0$, we have that

$$\begin{aligned} (q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{(q - \boldsymbol{d}^\top \boldsymbol{x})\boldsymbol{x}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) \le 0 \quad &\Longleftrightarrow \quad (p_y - \boldsymbol{b}_y^\top \boldsymbol{x}) f\left(\frac{(p_y - \boldsymbol{b}_y^\top \boldsymbol{x})\boldsymbol{x}}{p_y - \boldsymbol{b}_y^\top \boldsymbol{x}}\right) \le 0 \\ (q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{q\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) \le 0 \quad &\Longleftrightarrow \quad (p_y - \boldsymbol{b}_y^\top \boldsymbol{x}) f\left(\frac{p_y\boldsymbol{x} - \boldsymbol{U}\boldsymbol{b}_y}{p_y - \boldsymbol{b}_y^\top \boldsymbol{x}}\right) \le 0 \end{aligned}$$

for any $\boldsymbol{U} \in \mathbb{S}^{n_x}$, where the second "$\Longleftrightarrow$" holds because $\boldsymbol{b}_y = \boldsymbol{d}$ so that $\boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{d} = \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{b}_y$ and $\boldsymbol{U}\boldsymbol{d} = \boldsymbol{U}\boldsymbol{b}_y$. Notice that

$$\begin{aligned} (p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x}) f\left(\frac{p_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{b}_\ell}{p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x}}\right) \le 0, \ \ell \in \mathcal{L} \quad &\Longrightarrow \quad \sum_{\ell \in \mathcal{L}} \theta_\ell (p_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}) f\left(\frac{p_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{b}_\ell}{p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x}}\right) \le 0 \\ &\Longrightarrow \quad \left(\sum_{\ell \in \mathcal{L}} \theta_\ell (p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x})\right) f\left(\frac{\sum_{\ell \in \mathcal{L}} \theta_\ell (p_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{b}_\ell)}{\sum_{\ell \in \mathcal{L}} \theta_\ell (p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x})}\right) \le 0 \\ &\Longrightarrow \quad (p_y - \boldsymbol{b}_y^\top \boldsymbol{x}) f\left(\frac{p_y \boldsymbol{x} - \boldsymbol{U}\boldsymbol{b}_y}{p_y - \boldsymbol{b}_y^\top \boldsymbol{x}}\right) \le 0, \end{aligned}$$

where $\theta_\ell = y_\ell / \sum_{\ell \in \mathcal{L}} y_\ell$ for all $\ell \in \mathcal{L}$ (note that $\boldsymbol{\theta} \in \mathbb{R}^L_+$ and $\sum_{\ell \in \mathcal{L}} \theta_\ell = 1$). Here, the second implication follows from the convexity of the perspective functions. Therefore, the constraint $(q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{q\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) \le 0$ is redundant to

$$\left\{ \boldsymbol{x} \;\middle|\; \boldsymbol{b}_y^\top \boldsymbol{x} \le p_y, \ (p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x}) f\left(\frac{p_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{b}_\ell}{p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x}}\right) \le 0, \ \ell \in \mathcal{L} \right\}.$$

Thus, the claim follows. □

Note that Theorem 2 also implies that introducing epigraphical variables for the linear terms of the SLC functions in (P) would only increase the number of optimization variables without tightening the resulting RPT relaxation.

4.3. Choosing the best perspectification

In this section we describe two cases for which there are multiple choices for perspectification. The first one is an indefinite quadratic function, and the second one is a signomial type function.

Quadratic function. We now treat the question: Which part of a quadratic function should be perspectified? Suppose the objective function is quadratic:

$$\min_{\boldsymbol{x}\in\mathcal{X}} \left\{\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x}\right\},$$

where $\boldsymbol{A}$ is not necessarily positive semidefinite. The question then arises whether we should linearize all quadratic terms, or only a part of these terms, such that the remaining part is convex. We will now show that indeed linearizing all quadratic terms yields the tighest approximation. We search for the best value of a semidefinite matrix $\boldsymbol{B}$ such that linearizing $\boldsymbol{x}^\top(\boldsymbol{A}-\boldsymbol{B})\boldsymbol{x}$ and keeping $\boldsymbol{x}^\top\boldsymbol{B}\boldsymbol{x}$ yields the tightest approximation. In other words we consider the following problem

$$\min_{\boldsymbol{x}\in\mathcal{X}}\max_{\boldsymbol{B}\succeq\boldsymbol{0}} \left\{\boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x} + \operatorname{Tr}\left((\boldsymbol{A}-\boldsymbol{B})\boldsymbol{U}\right) + \boldsymbol{b}^\top\boldsymbol{x}\right\} = \min_{\boldsymbol{x}\in\mathcal{X}}\max_{\boldsymbol{B}\succeq\boldsymbol{0}} \left\{\operatorname{Tr}\left(\boldsymbol{B}(\boldsymbol{x}\boldsymbol{x}^\top - \boldsymbol{U})\right) + \operatorname{Tr}\left(\boldsymbol{A}\boldsymbol{U}\right) + \boldsymbol{b}^\top\boldsymbol{x}\right\}.$$

Taking the dual of the inner maximization problem we obtain that this is equivalent to

$$\min_{\boldsymbol{x}\in\mathcal{X}} \left\{\operatorname{Tr}\left(\boldsymbol{A}\boldsymbol{U}\right) + \boldsymbol{b}^\top\boldsymbol{x} \,:\, \begin{pmatrix}\boldsymbol{U} & \boldsymbol{x}\\ \boldsymbol{x}^\top & 1\end{pmatrix}\succeq 0\right\},$$

where the LMI is the same as (3). Therefore, the LMI can be interpreted (from its dual) as obtaining the best $\boldsymbol{B}$, and we do not need to add different LMIs for each RPT-type constraint, as the same LMI implies that already for all RPT-type constraints.

Signomial function. We consider the signomial function

$$c(\boldsymbol{x}) = -\prod_{i=1}^{m} x_i^{\alpha_i},$$

where $1 < \sum_{i=1}^m \alpha_i \leq 2$. As we will see later, such functions arise when power cone constraints are multiplied by linear or other power constraints. Note that $c(\boldsymbol{x})$ is not convex. However, we can perspectify this function in many ways. By substituting $x_ix_j = u_{ij}$ for some of the i's and j's we can lower the total sum of exponents to 1, which makes the function convex again. To this end, we rewrite $c(\boldsymbol{x})$ as follows:

$$c(\boldsymbol{x}) = -\prod_{i=1}^{m}\prod_{j=1}^{i} x_i^{\vartheta_i}(x_ix_j)^{\theta_{ij}} \quad\longrightarrow\quad c(\boldsymbol{x},\boldsymbol{U}) = -\prod_{i=1}^{m}\prod_{j=1}^{i} x_i^{\vartheta_i}u_{ij}^{\theta_{ij}}, \tag{1}$$

where $\boldsymbol{\vartheta}$ and $\boldsymbol{\theta}$ are such that

$$\vartheta_i, \theta_{ij} \geq 0, \theta_{ij} = \theta_{ji}\, \forall i,j \in [m], \quad \vartheta_i + \sum_{j=1}^{m} \theta_{ij} + \theta_{ii} = \alpha_i,\ \forall i \in [m], \quad \sum_{i=1}^{m} \vartheta_i + \sum_{i=1}^{m}\sum_{j=1}^{i} \theta_{ij} = 1, \tag{2}$$

Note that $c(\boldsymbol{x}, \boldsymbol{U})$ is jointly convex in $(\boldsymbol{x}, \boldsymbol{U})$, because of the second equality in (2), and jointly concave in (ϑ, θ). One could also use several different feasible solutions $\boldsymbol{\vartheta}$ and $\boldsymbol{\theta}$, and thereby strengthening the approximation. One could also embed (1) in a robust constraint, where $\boldsymbol{\vartheta}$ and $\boldsymbol{\theta}$ are the uncertain parameters, and one could enforce that the constraint should hold for all $\boldsymbol{\vartheta}$ and $\boldsymbol{\theta}$ that satisfy (2).

4.4. Other practical issues

There are several important practical issues when the RPT is applied:

(a) Valid cuts from SDP relaxations. In order to further tighten the convex relaxation (PC), effective SDP cuts can be considered. In the perspectification step of RPT, the nonconvex quadratic terms $\boldsymbol{x}\boldsymbol{x}^\top$ are linearized by a symmetric matrix $\boldsymbol{U}$. Such a linearization based relaxation for the nonconvex quadratic equality $\boldsymbol{U} = \boldsymbol{x}\boldsymbol{x}^\top$ may be significantly improved by the SDP relaxation $\boldsymbol{U} \succeq \boldsymbol{x}\boldsymbol{x}^\top$, which can be equivalently reformulated as an LMI by using Schur complement (Boyd and Vandenberghe, 2004):

$$\begin{pmatrix} \boldsymbol{U} & \boldsymbol{x} \\ \boldsymbol{x}^\top & 1 \end{pmatrix} \succeq 0. \tag{3}$$

The effectiveness of such SDP relaxations is illustrated by the toy example in Section 2 as well as the numerical experiment in Section 8.2.

(b) First perspectification, then reformulation. If the original constraint function is not convex, then one could first convexify it using the perspectification step of RPT, and then multiply this constraint also with other constraints. However, this will increase the number of extra variables enormously, and hence the question is whether this outweigh the benefit of obtaining tighter approximations.

(c) Retrieving feasible solutions of (P) with a convex feasible region. If there are only nonconvex components in the objective but no nonconvexity in the constraints of (P), which includes but not limited to, difference of convex functions (see Section 6), then we can retrieve feasible solutions from the solution of the RPT relaxation as follows. Firstly, if x_i in (P) has no sign constraint for some $i \in [n_x]$, we replace all those x_i with $(x_i^+ - x_i^-)$, where $x_i^+, x_i^- \geq 0$, everywhere in (P) so that the optimization variables in the lifted (P) are all nonnegative. For instance, the lifted set of $\mathcal{X}$ is

$$\mathcal{X}^+ = \left\{ (\boldsymbol{x}^+, \boldsymbol{x}^-) \in \mathbb{R}^{n_x}_+ \times \mathbb{R}^{n_x}_+ \;\middle|\; \begin{array}{l} \boldsymbol{A}(\boldsymbol{x}^+ - \boldsymbol{x}^-) \leq \boldsymbol{b} \\ h_j(\boldsymbol{x}^+ - \boldsymbol{x}^-) \leq 0,\ j \in \mathcal{J} \end{array} \right\}. \tag{4}$$

Then, we solve the RPT relaxation of the lifted (P). The obtained solution can be used to retrieve feasible solutions of (P). We demonstrate the procedure via Example 1.

EXAMPLE 1 (FEASIBLE SOLUTION CONSTRUCTION). We first lift the toy problem (T) so that the optimization variables are all nonnegative, and then apply RPT to the lifted problem. The solution of the corresponding lifted (TC) appears to be

$$\begin{bmatrix}\boldsymbol{x}^+\\ \boldsymbol{x}^-\end{bmatrix} = \begin{bmatrix}36.18\\37.18\\35.18\\37.18\end{bmatrix} \quad \text{and} \quad \begin{bmatrix}\boldsymbol{U}^+\\ \boldsymbol{U}^-\end{bmatrix} = \begin{bmatrix}41.14 & 38.76 & 27.79 & 17.33\\ 38.76 & 43.43 & 17.22 & 30.59\\ 27.79 & 17.22 & 7.04 & 4.19\\ 17.33 & 30.59 & 4.19 & 7.37\end{bmatrix},$$

and then the recovered feasible solutions and their corresponding objective values are

$$\begin{aligned}\boldsymbol{x}' &= \left[\boldsymbol{x}^+ - \boldsymbol{x}^- \quad (\boldsymbol{U}^+ - \boldsymbol{U}^-)./\begin{bmatrix}\boldsymbol{x}^+\\ \boldsymbol{x}^-\end{bmatrix}^\top\right] = \begin{bmatrix}1 & 0.37 & 0.58 & 0.59 & 0.35\\ 0 & 0.59 & 0.35 & 0.37 & 0.62\end{bmatrix} \quad \text{and}\\ \text{obj} &= [-1.30 \ \ -0.62 \ \ -1.25 \ \ -1.27 \ \ -0.54],\end{aligned}$$

where each column of $\boldsymbol{x}'$ constitutes a feasible solution, and its associated objective value is the corresponding element of the vector obj. □

Two additional second-level practical issues are as follows.

(d) Idempotent identity of binary variables. When the variables x_i are binary for some $i \in [n_x]$, one can employ the fact that $x_i^2 = x_i$, and hence $u_{ii} = x_i$, where u_{ii} is the i-th diagonal element of $\boldsymbol{U}$ (Sherali and Adams, 1990).

(e) Lifting for absolute symmetric convex functions. If some of the constraint functions in $\mathcal{X}$ of (P) are absolute symmetric convex functions $h_{j'}$, $j' \in \mathcal{J}' \subseteq \mathcal{J}$, that is, $h_{j'}(\boldsymbol{x}) = h_{j'}(|\boldsymbol{x}|)$ for all $\boldsymbol{x} \in \mathbb{R}^{d_{n_x}}$, then the lifted set in (4) is equivalent to

$$\mathcal{X}^+ = \left\{ (\boldsymbol{x}^+, \boldsymbol{x}^-) \in \mathbb{R}^{n_x}_+ \times \mathbb{R}^{n_x}_+ \;\middle|\; \begin{array}{l} \boldsymbol{A}(\boldsymbol{x}^+ - \boldsymbol{x}^-) \leq \boldsymbol{b}\\ h_j(\boldsymbol{x}^+ + \boldsymbol{x}^-) \leq 0, \ j \in \mathcal{J}'\\ h_j(\boldsymbol{x}^+ - \boldsymbol{x}^-) \leq 0, \ j \in \mathcal{J} \setminus \mathcal{J}' \end{array} \right\},$$

which can be used to tighten the RPT relaxation for (P). For more details on absolute symmetric convex functions, see Section 14.3.2 of Ben-Tal et al. (2009).

(f) Finding local optimum using mountain climbing. Even if there are nonconvex components in the constraints, one may still retrieve some feasible solutions for (P) by applying the procedure described in Section 4.4 **(c)**, and the obtained feasible solutions can be used as warm starts for existing algorithms, e.g., the mountain climbing algorithm by (Tao and An, 1997), to find a local optimum of (P), e.g., see Section 8.2.

5. Reformulation and linearization for multiplications of fundamental cones

There are five basic cones that play an important role in optimization: the linear, second-order, power, exponential, and semidefinite cone. In this section we show how to apply RPT to these cones. In the first subsection we show how to multiply conic inequalities. In the second subsection we discuss which order is the best: first RPT and then find a conic formulation, or the other way around?

5.1. Products of conic inequalities

It is generally believed that most practical problems can be modeled in terms of the following five cone inequalities:

(**L**) Linear inequality:

$$b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq 0,$$

where $x \in \mathbb{R}^{n_x}$.

(**Q**) Conic quadratic inequality:

$$b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq \left\|\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p}\right\|,$$

where $x \in \mathbb{R}^{n_x}$.

(**P**) Power cone inequality:

$$\prod_{i=1}^{m} x_i^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} x_i^2}, \quad x_1, \cdots, x_m \geq 0,$$

where $n_x > m$, $\alpha_1, \cdots, \alpha_m > 0$ and $\sum_{i=1}^m \alpha_i = 1$.

(**E**) Exponential cone inequality:

$$x_1 \geq x_2 \exp\left(\frac{x_3}{x_2}\right), \quad x_2 \geq 0.$$

(**S**) SDP cone/linear matrix inequality:

$$\boldsymbol{A}(\boldsymbol{x}) \succeq 0,$$

where $\boldsymbol{A}(\boldsymbol{x}) = \boldsymbol{A}_0 + \boldsymbol{A}_1 x_1 + \cdots + \boldsymbol{A}_{n_x} x_{n_x}$. Generalized inequalities are not considered in Sections 3 and 4.1, nevertheless, our RPT techniques can readily be extended to problems with such inequalities.

We demonstrate how to convexify the pairwise products of these inequalities. There are 15 cases of pairwise products in total. Note that all the original inequalities as well as the resulting convexified product of inequalities should be considered in optimization problems.

Multiplying a linear equality with one of the five conic inequalities can be done in the same manner as explained in the previous section. However, pairwise multiplication of the four nonlinear inequalities needs specific investigation. As an example, we consider the following two conic quadratic inequalities:

$$\begin{cases} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq \left\| \boldsymbol{D}_1 \boldsymbol{x} + \boldsymbol{p}_1 \right\| \\ b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq \left\| \boldsymbol{D}_2 \boldsymbol{x} + \boldsymbol{p}_2 \right\|. \end{cases} \tag{5}$$

We now apply RPT to these inequalities. After multiplying the two constraints we obtain:

$$(b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})(b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) + \left\| \boldsymbol{D}_1 \boldsymbol{x} + \boldsymbol{p}_1 \right\| \left\| \boldsymbol{D}_2 \boldsymbol{x} + \boldsymbol{p}_2 \right\| \geq (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x}) \left\| \boldsymbol{D}_2 \boldsymbol{x} + \boldsymbol{p}_2 \right\| + (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) \left\| \boldsymbol{D}_1 \boldsymbol{x} + \boldsymbol{p}_1 \right\|.$$

The term $\left\| \boldsymbol{D}_1 \boldsymbol{x} + \boldsymbol{p}_1 \right\| \left\| \boldsymbol{D}_2 \boldsymbol{x} + \boldsymbol{p}_2 \right\|$ in this inequality is problematic, since also after linearizing this term remains nonconcave. However, if we multiply the left-hand sides with the right-hand sides of the conic quadratic inequalities (5), we can derive useful convex inequalities. The way how to do this for each of the 15 products of conic inequalities, is shown in the Appendix. The final results are summarized in Table 2.

Case	Cone-1	Cone-2	Cone-1 × Cone-2	Reference	Remarks
1	L	L	L	Sherali and Adams (1990)	
2	L	Q	Q	Sturm and Zhang (2003)	
3	L	P	m×L + $(m+1)$×P	This paper	
4	L	E	2×E + L	This paper	
5	L	S	S	Trivial	
6	Q	Q	(i) 5×Q (ii) 4×Q + S (iii) S (iv) S	Jiang and Li (2016) This paper Yang and Burer (2016) Anstreicher (2017)	(ii) is at least as good as (i) (iv) is at least as good as (iii)
7	Q	P	(i) m×Q + $(m+2)$×P (ii) m×Q + $(m+1)$×P + S	This paper This paper	(ii) is at least as good as (i)
8	Q	E	Q + 2×E	This paper	
9	Q	S	S	This paper	
10	P	P	$(2m_1+2m_2+3)$×P	This paper	
11	P	E	2×P + m×E	This paper	
12	P	S	m×S	This paper	
13	E	E	11×E	This paper	
14	E	S	2×S	Trivial	
15	S	S	(i) S (ii) S	Yang and Burer (2016) Anstreicher (2017)	(i) is at least as good and (ii) only if the two cones are of the same size

Table 2 Results of multiplying two cones. L = Linear Cone, Q = Conic Quadratic Cone, P = Power Cone, E = Exponential Cones, S = Semidefinite Cone.

5.2. First RPT and then conic representation?

An important question is which of the following two options yields the tightest approximation:

1. First apply RPT, and then reformulate the result in terms of the five basic cones, or,
2. first reformulate the constraint(s) in terms of the five basic cones, and then apply RPT?

For example, suppose we multiply a linear constraint with a convex quadratic constraint. Then in Option 1 we first apply RPT, and then reformulate the result (the perspective of the quadratic constraint) as conic quadratic, and in Option 2 we first reformulate the quadratic constraint as a conic quadratic constraint and then apply RPT. It appears that both options yield exactly the same approximation. The following lemma states that this result holds for general convex constraints that are conic representable.

We use the definition of conically representable from Serrano (2015), that is, a constraint $f(\boldsymbol{x}) \leq 0$, where $f: \mathbb{R}^n \to \mathbb{R}$, is conically representable if its feasible set can be written as

$$\{\boldsymbol{x} \mid f(\boldsymbol{x}) \leq 0\} = \{\boldsymbol{x} \mid \exists \boldsymbol{u} \in \mathbb{R}^m,\ S(\boldsymbol{x},\boldsymbol{u}) = 0,\ T(\boldsymbol{x},\boldsymbol{u}) \in \mathcal{K}\}, \tag{6}$$

where $S: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{k_1}$ and $T: \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^{k_2}$ are affine mappings and $\mathcal{K}$ is a cone.

Lemma 1. *Suppose the convex constraint $f(\boldsymbol{x}) \leq 0$ is conically representable, and suppose we multiply this constraint with a linear constraint $b - \boldsymbol{a}^\top \boldsymbol{x} \geq 0$. Then Options 1 and 2 yield the same.*

Proof. Let S, T be the affine mappings that define the conic representation of the feasible set $\{\boldsymbol{x} \mid f(\boldsymbol{x}) \leq 0\}$ and let $\mathcal{K}$ be the corresponding cone. Let us denote the linear function $b - \boldsymbol{a}^\top \boldsymbol{x}$ by $\ell(\boldsymbol{x})$ and denote the linear function that results after linearizing $\boldsymbol{x}\ell(\boldsymbol{x})$ by $\tilde{\boldsymbol{\ell}}(\boldsymbol{x},\boldsymbol{U})$, where $\boldsymbol{U} = \boldsymbol{x}\boldsymbol{x}^\top$, and $\tilde{\ell}_i(\boldsymbol{x},\boldsymbol{U}) = bx_i - a^\top \boldsymbol{U}_i$. Then after RPT we obtain the constraint

$$\ell(\boldsymbol{x}) f\left(\frac{\tilde{\boldsymbol{\ell}}(\boldsymbol{x},\boldsymbol{U})}{\ell(\boldsymbol{x})}\right) \leq 0. \tag{7}$$

This is the result of Option 1. On the other hand, suppose we follow Option 2, i.e., we first derive the affine mapping S and T that define the conic representation of the feasible set of the original constraint, and then multiply it by $\ell(\boldsymbol{x})$, and apply RPT. Then we obtain the set

$$\left\{(\boldsymbol{x},\boldsymbol{U}) \;\middle|\; \exists \boldsymbol{u} \in \mathbb{R}^m,\ S\left(\frac{\tilde{\boldsymbol{\ell}}(\boldsymbol{x},\boldsymbol{U})}{\ell(\boldsymbol{x})}, \frac{\boldsymbol{u}}{\ell(\boldsymbol{x})}\right) = 0,\ T\left(\frac{\tilde{\boldsymbol{\ell}}(\boldsymbol{x},\boldsymbol{U})}{\ell(\boldsymbol{x})}, \frac{\boldsymbol{u}}{\ell(\boldsymbol{x})}\right) \in \mathcal{K}\right\}. \tag{8}$$

Moreover we find

$$\begin{aligned}
&\left\{(\boldsymbol{x},\boldsymbol{U}) \;\middle|\; \ell(\boldsymbol{x}) f\left(\frac{\tilde{\boldsymbol{\ell}}(\boldsymbol{x},\boldsymbol{U})}{\ell(\boldsymbol{x})}\right) \leq 0\right\} \\
&\quad = \left\{(\boldsymbol{x},\boldsymbol{U}) \;\middle|\; \exists \boldsymbol{u} \in \mathbb{R}^m,\ S\left(\frac{\tilde{\boldsymbol{\ell}}(\boldsymbol{x},\boldsymbol{U})}{\ell(\boldsymbol{x})}, \boldsymbol{u}\right) = 0,\ T\left(\frac{\tilde{\boldsymbol{\ell}}(\boldsymbol{x},\boldsymbol{U})}{\ell(\boldsymbol{x})}, \boldsymbol{u}\right) \in \mathcal{K}\right\} \\
&\quad = \left\{(\boldsymbol{x},\boldsymbol{U}) \;\middle|\; \exists \boldsymbol{u} \in \mathbb{R}^m,\ S\left(\frac{\tilde{\boldsymbol{\ell}}(\boldsymbol{x},\boldsymbol{U})}{\ell(\boldsymbol{x})}, \frac{\boldsymbol{u}}{\ell(\boldsymbol{x})}\right) = 0,\ T\left(\frac{\tilde{\boldsymbol{\ell}}(\boldsymbol{x},\boldsymbol{U})}{\ell(\boldsymbol{x})}, \frac{\boldsymbol{u}}{\ell(\boldsymbol{x})}\right) \in \mathcal{K}\right\}
\end{aligned}$$

which concludes the proof. □

Lemma 1 only states that Options 1 and 2 for multiplying a linear and convex constraint yield the same. The following example shows that full RPT on Option 2 might be better than on Option 1.

EXAMPLE 2. Consider the following toy problem:

$$\begin{array}{ll} \max_{\boldsymbol{x}} & x_1 + x_2 + x_1 x_2 \\ \text{s.t.} & \max\{|x_1|, |x_2|\} \leq 1. \end{array} \tag{9}$$

Option 1, i.e., first applying RPT does not yield a valuable approximation. For Option 2, we first derive a conic formulation of the constraint:

$$\begin{array}{ll} \max_{\boldsymbol{x}} & x_1 + x_2 + x_1 x_2 \\ \text{s.t.} & x_1 - 1 \leq 0 \\ & -x_1 - 1 \leq 0 \\ & x_2 - 1 \leq 0 \\ & -x_2 - 1 \leq 0. \end{array}$$

Now applying partial RPT (we multiply only the second with the third inequality, and the first with the fourth) yields the following linear optimization problem:

$$\begin{array}{ll} \max_{\boldsymbol{x}, u_{12}} & x_1 + x_2 + u_{12} \\ \text{s.t.} & x_1 - 1 \leq 0 \\ & -x_1 - 1 \leq 0 \\ & x_2 - 1 \leq 0 \\ & -x_2 - 1 \leq 0 \\ & -u_{12} - x_1 + x_2 + 1 \geq 0 \\ & -u_{12} + x_1 - x_2 + 1 \geq 0. \end{array}$$

It can easily be verified that the optimal solution is $u_{12} = x_1 = x_2 = 1$, with optimal objective value 3. This is the optimal solution of the original nonconvex problem (9). □

6. Extension to difference of convex functions

6.1. RPT for difference of convex functions

Another important class of SLC representable functions are difference of convex functions. If the constraint in (P) contains difference of convex functions, that is, $f_k(\boldsymbol{x}) = c_{0k}(\boldsymbol{x}) - c_{1k}(\boldsymbol{x}) \leq 0$, where $c_{0k},\ c_{1k} : \mathbb{R}^{n_x} \to [-\infty, +\infty]$ are proper, closed and convex for some $k \in \mathcal{K}_0$, then we can reformulate the corresponding constraint function into an SLC function. Indeed, for c_{1k} we have

$$c_{1k}(\boldsymbol{x}) = c_{1k}^{**}(\boldsymbol{x}) = \sup_{y \in \mathrm{dom}(c_{1k}^*)} \{\boldsymbol{x}^\top \boldsymbol{y} - c_{1k}^*(\boldsymbol{y})\},$$

where the second equality follows from Rockafellar (1970), hence,

$$f_k(\boldsymbol{x}) \leq 0 \quad \iff \quad \inf_{y \in \mathrm{dom}(c_{1k}^*)} \{c_{0k}(\boldsymbol{x}) - \boldsymbol{x}^\top \boldsymbol{y} + c_{1k}^*(\boldsymbol{y})\} \leq 0. \tag{10}$$

If the infimum in (10) is attained, then one can merge the inf operator in (10) with the min operator in the objective of (P) and consider the SLC representation of $f_k(\boldsymbol{x}) \le 0$, that is,

$$\begin{cases} c_{0k}(\boldsymbol{x}) - \boldsymbol{x}^\top \boldsymbol{y} + c_{1k}^*(\boldsymbol{y}) \le 0 \\ \boldsymbol{y} \in \operatorname{dom}(c_{1k}^*), \end{cases} \tag{11}$$

where $\operatorname{dom}(c_{1k}^*)$ is a convex set (not necessarily closed), c_{1k}^* is a convex function, and $-\boldsymbol{x}^\top \boldsymbol{y}$ is SLC. However, if the infimum of (10) is not attained, we assume that (P) satisfies the following regularity condition.

Assumption 3. *There exists a vector* $\boldsymbol{x}^{\mathrm{S}} \in \operatorname{ri}(\cap_{k\in\mathcal{K}_0} \operatorname{dom}(f_k))$ *such that* $f_k(\boldsymbol{x}^{\mathrm{S}}) < 0$ *for all* $k \in \mathcal{K}$, $\boldsymbol{A}\boldsymbol{x} < \boldsymbol{b}$ *and* $\boldsymbol{h}(\boldsymbol{x}^{\mathrm{S}}) < \boldsymbol{0}$.

Note that Assumption 3 implies that $\boldsymbol{x}^{\mathrm{S}}$ resides in the sets $\cap_{\substack{k\in\mathcal{K}_0\\ i\in\mathcal{I}_0}} \operatorname{ri}(\operatorname{dom}(c_{ik}))$ and $\cap_{j\in\mathcal{J}_0}\operatorname{ri}(\operatorname{dom}(h_j))$ thanks to Proposition 2.42 in Rockafellar and Wets (2009), and thus, $\boldsymbol{x}^{\mathrm{S}}$ is a strict Slater point of (P). Furthermore, there exists a $(\boldsymbol{\tau}^{\mathrm{S}}, \boldsymbol{U}^{\mathrm{S}}, \boldsymbol{V}^{\mathrm{S}})$ such that $(\boldsymbol{x}^{\mathrm{S}}, \boldsymbol{\tau}^{\mathrm{S}}, \boldsymbol{U}^{\mathrm{S}}, \boldsymbol{V}^{\mathrm{S}})$ is a strict Slater point of the corresponding RPT relaxation (PC) of (P) with $f_k(\boldsymbol{x}) \le 0$ is replaced by

$$\begin{cases} c_{0k}(\boldsymbol{x}) - \sup\limits_{y\in\operatorname{dom}(c_{1k}^*)} \{\boldsymbol{x}^\top \boldsymbol{y} - c_{1k}^*(\boldsymbol{y})\} \le 0 \\ \boldsymbol{y} \in \operatorname{dom}(c_{1k}^*). \end{cases} \tag{12}$$

Finally, thanks to Remark 1 and the proof of Theorem 6(iii) of Zhen et al. (2021), the inf operator in the constraint of (PC) can be merged with the inf operator (instead of min operator because the optimal $\boldsymbol{y}$ may not be obtained) in the objective function without affecting the infimum of (PC).

In the case that the constraint function $f(\cdot)$ in (P) is both concave and SLC, e.g., $f(x) = -x^2$ or $\sqrt{x}$, then it may be better to first derive the biconjugate reformulation of $f(\cdot)$ and then apply RPT than directly apply RPT to (P).

Example 3 (RPT for epigraphical biconjugate reformulation). Consider the following convex maximization problem:

$$\begin{array}{ll} \max\limits_{x} & x^2 + 2x \\ \text{s.t.} & 0 \le x \le 1. \end{array} \tag{T$_3$}$$

The optimal solution is obviously $x = 1$. Note that $x^2 + 2x$ is both convex and SLC representable, and its conjugate is $\frac{(y-2)^2}{4}$. In the following, we compare three convex relaxations of (T$_3$) obtained from, (i) directly applying RPT, (ii) applying RPT to the biconjugate reformulation, and (iii) applying RPT to the epigraphical biconjugate reformulation, respectively,

$$\begin{array}{ll} \max\limits_{x,u} & u + 2x \\ \text{s.t.} & 0 \le x \le 1 \\ & 0 \le u \le x \\ & u + 1 \ge 2x, \end{array} \qquad \begin{array}{ll} \max\limits_{x,y,v} & v - \frac{(y-2)^2}{4} \\ \text{s.t.} & 0 \le x \le 1 \end{array} \qquad \text{and} \qquad \begin{array}{ll} \max\limits_{x,y,v,\tau,t} & v - \tau \\ \text{s.t.} & 0 \le x \le 1 \\ & \frac{(y-2)^2}{4} \le \tau \\ & \frac{(v-2x)^2}{4x} \le t \\ & \frac{(y-2-v+2x)^2}{4-4x} \le \tau - t. \end{array}$$

Note that the maximum of the direct RPT relaxation is 3 with $(x^\star, u^\star) = (1,1)$. The maximum for the RPT relaxation of the biconjugate reformulation is ∞ with $x^\star \in [0,1],\ y^\star = 2,\ v^\star = \infty$, and that of the epigraphical biconjugate reformulation is 3 with $(x^\star, y^\star, v^\star, \tau^\star, t^\star) = (1,4,4,1,1)$. Both the direct RPT relaxation and the RPT relaxation of the epigraphical biconjugate reformulation obtain the optimal value and optimal solution of (T_3). The computational complexity of the three relaxations may also be different. The direct RPT relaxation is an LP, while the RPT relaxation of the biconjugate-base reformulations are convex quadratic programs. This example also shows that the epigraphical reformulation for RPT can lead to a tighter approximation. □

6.2. Using first-order conditions

Let us consider the first inequality in (11)

$$c_{0k}(\boldsymbol{x}) - \boldsymbol{x}^\top \boldsymbol{y} + c_{1k}^*(\boldsymbol{y}) \le 0.$$

Now suppose $c_{1k}(\boldsymbol{x})$ is differentiable, then we have

$$\boldsymbol{y} = \nabla_x c_{1k}(\boldsymbol{x}). \tag{13}$$

We may include this extra equation in the RPT approach to get a better approximation for the nonconvex quadratic equality $\boldsymbol{W} = \boldsymbol{x}\boldsymbol{y}^\top$, as illustrated in the following examples.

Example 4. Suppose $c_{1k}(x) = -\ln x$. Then (13) becomes $y = -1/x$, or $xy = w = -1$. This last equality is of course helpful for obtaining tighter approximations. □

Example 5. Suppose $c_{1k}(x) = x\ln x$. Then (13) becomes $y = 1 + \ln x$. Hence, we can add the following convex inequality in the RPT approach: $y \le 1 + \ln x$. □

Example 6. Suppose $c_{1k}(\boldsymbol{x}) = \ln \sum_j \exp(x_j)$. Then (13) becomes

$$y_i = \frac{\exp(x_i)}{\sum_j \exp(x_j)} \iff \ln y_i + \ln \sum_j \exp(x_j) = x_i \iff y_i \ln y_i + y_i \ln \sum_j \exp(x_j) = x_i y_i.$$

Hence, we can add the convex inequality

$$y_i \ln y_i + y_i \ln \sum_j \exp\left(\frac{w_{ij}}{y_i}\right) \le w_{ii}, \quad \forall i$$

to the RPT approximation. Note that $y_i \ln \sum_j \exp\left(\frac{w_{ij}}{y_i}\right)$ is a perspective function of the convex function $\ln \sum_j e^{w_{ij}}$. □

The following example shows that such first-order information is indeed helpful and not included by other constraints resulting from RPT.

Example 7. Consider the following problem:

$$\begin{array}{ll} \min\limits_{x} & -x^2 \\ \text{s.t.} & e^x \leq e \\ & e^{-x} \leq 1, \end{array} \tag{14}$$

that has optimal value -1, and optimal solution $x = 1$. Hence, $c_{1k}(x) = -x^2$. Then (13) becomes $y = -2x$ and $c^*_{1k}(y) = -y^2/4$. Hence, after applying RPT to the biconjugate reformulation and using the first-order information, we obtain the approximation:

$$\begin{array}{ll} \min\limits_{x} & v + y^2/4 \\ \text{s.t.} & e^x \leq e \\ & e^{-x} \leq 1 \\ & y = -2x \\ & v = -2u \\ & x \geq u \\ & \dots . \end{array} \tag{15}$$

The optimal solution for this problem is $x = u = 1$, $v = y = -2$, and optimal value -1, which is the optimal value of Problem (14). However, it can easily be verified that if the first-order information $y = -2x$ is not used, the RPT approach (even the epigraph version) leads to an approximation with value $-\infty$. □

The following example shows that RPT to an indefinite quadratic function yields the same as RPT to the difference of two convex quadratic functions with using first-order information.

Example 8. Suppose the constraint is

$$\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{B}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c \leq 0. \tag{16}$$

where $\boldsymbol{A}$ and $\boldsymbol{B}$ are positive semidefinite matrices. Then Step 2 of RPT yields the linear constraint

$$\mathrm{Tr}\,(\boldsymbol{A}\boldsymbol{U}) - \mathrm{Tr}\,(\boldsymbol{B}\boldsymbol{U}) + \boldsymbol{b}^\top \boldsymbol{x} + c \leq 0. \tag{17}$$

The biconjugate formulation for (16) is:

$$\boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} - \boldsymbol{x}^\top \boldsymbol{y} + \frac{1}{4}\boldsymbol{y}^\top \boldsymbol{B}^{-1}\boldsymbol{y} + \boldsymbol{b}^\top \boldsymbol{x} + c \leq 0. \tag{18}$$

Applying RPT to (18) yields

$$\mathrm{Tr}\,(\boldsymbol{A}\boldsymbol{U}) - \mathrm{Tr}\,(\boldsymbol{W}) + \frac{1}{4}\mathrm{Tr}\left(\boldsymbol{B}^{-1}\boldsymbol{V}\right) + \boldsymbol{b}^\top \boldsymbol{x} + c \leq 0. \tag{19}$$

From the first-order information we get $\boldsymbol{y} = 2\boldsymbol{B}\boldsymbol{x}$. Multiplying this equality on both sides by $x^\top$ and $y^\top$, we obtain

$$\boldsymbol{W} = 2\boldsymbol{B}\boldsymbol{U} \quad \text{and} \quad \boldsymbol{V} = 2\boldsymbol{B}\boldsymbol{W}.$$

Substituting this into (19) yields exactly inequality (17). □

6.3. Applicability of conjugate functions

For many real-life applications, the constraint functions in the optimization problem may not admit a closed form conjugate. However, for many important classes of convex functions, their conjugates and domains are readily available from the literature. Table 3 lists several convex functions and their conjugates and domains. The following example illustrates that the infimal convolution formula (last line in Table 3) is very useful in cases where the conjugate of the function can not be derived.

Table 3 Example of functions $f(\cdot)$ and their corresponding conjugates, defined as $f^*(\boldsymbol{y}) := \sup_{\boldsymbol{x}\in \mathrm{dom}\, f} \{\boldsymbol{x}^\top \boldsymbol{y} - f(\boldsymbol{x})\}$. **CQ** stands for conic quadratic, **PWL** for piecewise linear, **SML** for sum of max of linear functions, **G** for Geometric, and **SC** for sum of concave functions. For SC functions, we assume that there is a common point in $\mathrm{ri}(\mathrm{dom}(f_i))$ for every i.

Type	f	$\mathrm{dom}(f^*)$	f^*
CQ	$f(\boldsymbol{x},\bar{x}) = \lVert\boldsymbol{x}\rVert_2 - \bar{x}$	$\{(\boldsymbol{y},\bar{y}) : \lVert\boldsymbol{y}\rVert_2 \le 1,\ \bar{y}=1\}$	$f^*(\boldsymbol{y},\bar{y}) = 0$
PWL	$f(\boldsymbol{x}) = \max_i x_i$	$\{\boldsymbol{y} : \boldsymbol{y} \ge 0, \sum_i y_i = 1, \forall k\}$	$f^*(\boldsymbol{y}) = 0$
SML	$f(\boldsymbol{x}) = \sum_k \max_{i\in I_k} x_i$	$\{\boldsymbol{y}_k : \boldsymbol{y}_k \ge 0,\ \sum_i y_{ki} = 1, \forall k\}$	$f^*(\boldsymbol{w}) = 0$
G	$f(\boldsymbol{x}) = \ln\left(\sum_i e^{x_i}\right)$	$\{\boldsymbol{y} : \boldsymbol{y} \ge 0, \sum_i y_i = 1\}$	$f^*(\boldsymbol{y}) = \sum_i y_i \ln y_i$
SC	$f(\boldsymbol{x}) = \sum_i f_i(\boldsymbol{x})$	$\{\{\boldsymbol{y}_i\}_i : \sum_i \boldsymbol{y}_i = \boldsymbol{y}, \boldsymbol{y}_i \in \mathrm{dom}(f_i^*), \forall i\}$	$f^*(\boldsymbol{y}) = \min_{\{\boldsymbol{y}_i\}_i} \sum_i f_i^*(\boldsymbol{y}_i)$

Example 9. Consider the following problem:

$$\begin{array}{ll} \min\limits_{x} & \ln x - x^2 \\ \text{s.t.} & e \le x \le e^2, \end{array} \tag{20}$$

that has optimal value $2 - e^4$, and optimal solution $x = e^2$. Hence, $c_{1k}(x) = \ln x - x^2$. Since the conjugate of this function cannot be derived, we use the infimal convolution formula. It can easily be verified that using the biconjugate reformulation, the infimal convolution formula, and first-order information, we obtain

$$\begin{array}{ll} \min\limits_{x,y,y_1,y_2} & -xy - 1 - \ln(-y_1) + y_2^2/4 \\ \text{s.t.} & \exp(1) \le x \le \exp(2) \\ & y_1 + y_2 = y \\ & y_2 = 2x \\ & xy_1 = -1 \\ & y_1 \le 0. \end{array} \tag{21}$$

Now eliminating x, y and xy_1, and applying RPT to (21) we obtain

$$\begin{array}{ll} \min\limits_{y_1,y_2,v_2} & -v_2/2 - \ln(-y_1) \\ \text{s.t.} & 2\exp(1) \le y_2 \le 2\exp(2) \\ & y_2\exp(1) \le v_2 \le y_2\exp(2) \\ & y_1\exp(2) \le -1 \\ & y_1\exp(1) \ge -1 \\ & y_1 \le 0 \\ & \dots. \end{array}$$

The optimal solution for this problem is $y_1 = -\exp(-2)$, $y_2 = 2\exp(2)$, $v_2 = 2\exp(4)$, and optimal value $2 - e^4$, which is the optimal value of Problem (20). □

7. Known convex reformulations and relaxations obtained via RPT

In this section we show that several convex reformulations and relaxations for several classes of nonconvex problems derived in the literature can also be obtained via RPT.

7.1. Disjunctive optimization

A linear description of the convex hull of the union of convex sets can be derived by using RPT. It follows from the definition that the convex hull of the union of nonempty, compact convex sets $\mathcal{X}_k = \{\boldsymbol{x} \mid \boldsymbol{h}_k(\boldsymbol{x}) \le \boldsymbol{0}\}$, $k \in \mathcal{K}$ is:

$$\operatorname{conv}\left(\bigcup_{k\in\mathcal{K}} \mathcal{X}_k\right) = \left\{\boldsymbol{x} \;\middle|\; \exists \boldsymbol{x}_k \in \mathcal{X}_k, \boldsymbol{\lambda} \ge \boldsymbol{0} : \boldsymbol{x} = \sum_{k\in\mathcal{K}} \lambda_k \boldsymbol{x}_k,\ \sum_{k\in\mathcal{K}} \lambda_k = 1 \right\},$$

where $\boldsymbol{h}_k(\boldsymbol{x}) = [h_{1k}(\boldsymbol{x})\ h_{2k}(\boldsymbol{x})\ \cdots\ h_{Jk}(\boldsymbol{x})]^\top$, and $h_{jk} : \mathbb{R}^{n_x} \to [-\infty, +\infty]$ is proper, closed and convex for every $j \in \mathcal{J}$, $k \in \mathcal{K}$. This description is nonlinear and nonconvex, since it contains products of variables $\lambda_k \boldsymbol{x}_k$, $k \in \mathcal{K}$. One can apply RPT to obtain the following convex relaxation

$$\left\{\boldsymbol{x} \;\middle|\; \exists \boldsymbol{u}_k : \boldsymbol{x} = \sum_{k\in\mathcal{K}} \boldsymbol{u}_k,\ \sum_{k\in\mathcal{K}} \lambda_k = 1,\ \boldsymbol{\lambda} \ge \boldsymbol{0},\ \lambda_k \boldsymbol{h}_k(\boldsymbol{u}_k/\lambda_k) \le \boldsymbol{0},\ k \in \mathcal{K} \right\}.$$

This convex relaxation is exact according to Gorissen et al. (2014, Lemma 1), which applies because $\mathcal{X}_k$, $k \in \mathcal{K}$, are nonempty, compact and convex sets. We now use this observation to derive convex relaxation for disjunctive optimization problems with general convex sets. In Sherali and Adams (1994b, Section 4), the authors derive similar result for disjunctive optimization problems with a linear objective function and polyhedral sets $\mathcal{X}_k$, $k \in \mathcal{K}$. Consider a generic disjunctive optimization problem

$$\begin{aligned} \min_{\boldsymbol{x}} \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & \boldsymbol{x} \in \bigcup_{k\in\mathcal{K}} \mathcal{X}_k, \end{aligned} \tag{DP}$$

where $f : \mathbb{R}^{n_x} \to [-\infty, +\infty]$ is proper, closed and convex. Disjunctive optimization problems are in general nonconvex because its feasible region constitutes a union of convex sets $\mathcal{X}_k$. By applying RPT to the feasible region of (DP), we obtain the following convex relaxation:

$$\begin{aligned} \min_{\boldsymbol{y}, \{\boldsymbol{u}_k\}_k} \quad & f\left(\sum_{k\in\mathcal{K}} \boldsymbol{u}_k\right) \\ \text{s.t.} \quad & y_k \boldsymbol{h}_k(\boldsymbol{u}_k/y_k) \le \boldsymbol{0} && k \in \mathcal{K} \\ & \sum_{k\in\mathcal{K}} y_k = 1 \\ & y_k \ge 0 && k \in \mathcal{K}, \end{aligned}$$

which is often refered to as the hull relaxation (Grossmann and Lee, 2003). Note that this hull relaxation is tight if $f(\cdot)$ is a linear function, and $\mathcal{X}_k$, $k \in \mathcal{K}$, are nonempty, compact and convex sets.

7.2. Generalized linear optimization

Consider a generalized linear optimization problem of the following form (Dantzig, 1963, p. 434):

$$\begin{array}{lll} \min\limits_{\boldsymbol{y},\{\boldsymbol{x}_k\}_k} & \boldsymbol{c}^\top \boldsymbol{y} & \\ \text{s.t.} & \sum\limits_{k\in\mathcal{K}_0} \boldsymbol{x}_k y_k \leq \boldsymbol{b} & \\ & \boldsymbol{y}\geq \boldsymbol{0},\ \boldsymbol{x}_k\in\mathcal{X}_k & k\in\mathcal{K}_0, \end{array} \tag{GLP}$$

where $\mathcal{X}_k=\{\boldsymbol{x}\mid \boldsymbol{h}_k(\boldsymbol{x})\leq\boldsymbol{0}\}$, $k\in\mathcal{K}_0$, and $\boldsymbol{h}_k:\mathbb{R}^{n_x}\to[-\infty,+\infty]^J$ is a vector of J proper, closed and convex functions for each $k\in\mathcal{K}_0$. The partial RPT relaxation of (GLP) is:

$$\begin{array}{lll} \min\limits_{\boldsymbol{y},\{\boldsymbol{v}_k\}_k} & \boldsymbol{c}^\top \boldsymbol{y} & \\ \text{s.t.} & \sum\limits_{k\in\mathcal{K}_0} \boldsymbol{v}_k \leq \boldsymbol{b} & \\ & y_k \boldsymbol{h}_k(\boldsymbol{v}_k/y_k)\leq \boldsymbol{0} & k\in\mathcal{K}_0 \\ & \boldsymbol{y}\geq\boldsymbol{0}. & \end{array}$$

The convex program is in general a convex relaxation of (GLP), which has the same optimal value as (GLP) if one of the following regularity conditions is satisfied : (**i**) $\mathcal{X}_k$ is bounded for each $k\in\mathcal{K}_0$ (Gorissen et al., 2014, Lemma 1); (**ii**) there exists a $(\boldsymbol{y},\{\boldsymbol{x}_k\}_k)$ with $\boldsymbol{y}>\boldsymbol{0}$ that is feasible for (GLP) (Zhen et al., 2021, Lemma 6). While for a special case where $\mathcal{X}_k$, $k\in\mathcal{K}$, are (nonempty) boxes, the corresponding linear relaxation of (GLP) is exact due to Dantzig (1963).

7.3. Approximate $\mathcal{S}$-Lemma for quadratically constrained quadratic optimization

Consider a quadratically constrained quadratic optimization problem with only one (quadratic) constraint:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \boldsymbol{x}^\top \boldsymbol{A}_0\boldsymbol{x}+2\boldsymbol{b}_0^\top\boldsymbol{x}+c_0 \\ \text{s.t.} & \boldsymbol{x}^\top \boldsymbol{A}_1\boldsymbol{x}+2\boldsymbol{b}_1^\top\boldsymbol{x}+c_1\leq 0, \end{array} \tag{QCQP}$$

where $\boldsymbol{A}_k\in\mathbb{R}^{n_x\times n_x}$, $\boldsymbol{b}_k\in\mathbb{R}^{n_x}$ and $c_k\in\mathbb{R}$ for each $k\in\{0,1\}$. It is well-known that such a problem admits a convex reformulation via the $\mathcal{S}$-lemma. In the following, we show that the dual of the obtained convex reformulation from the $\mathcal{S}$-lemma can be interpreted as an RPT relaxation. Suppose that there exists an $\boldsymbol{x}\in\mathbb{R}^{n_x}$ with $\boldsymbol{x}^\top\boldsymbol{A}_1\boldsymbol{x}+2\boldsymbol{b}_1^\top\boldsymbol{x}+c_1<0$, then we have

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \boldsymbol{x}^\top \boldsymbol{A}_0\boldsymbol{x}+2\boldsymbol{b}_0^\top\boldsymbol{x}+c_0 \\ \text{s.t.} & \boldsymbol{x}^\top \boldsymbol{A}_1\boldsymbol{x}+2\boldsymbol{b}_1^\top\boldsymbol{x}+c_1\leq 0 \end{array} \iff \begin{array}{ll} \max\limits_{\lambda\geq 0,\gamma} & \gamma \\ \text{s.t.} & \begin{bmatrix}\boldsymbol{A}_0 & \frac{1}{2}\boldsymbol{b}_0\\ \frac{1}{2}\boldsymbol{b}_0^\top & c_0\end{bmatrix}\succeq\gamma\begin{bmatrix}\boldsymbol{O} & \boldsymbol{0}\\ \boldsymbol{0}^\top & 1\end{bmatrix}-\lambda\begin{bmatrix}\boldsymbol{A}_1 & \frac{1}{2}\boldsymbol{b}_1\\ \frac{1}{2}\boldsymbol{b}_1^\top & c_1\end{bmatrix}, \end{array}$$

where $\boldsymbol{O}\in\mathbb{R}^{n_x\times n_x}$ is a matrix of all zeros. Here the "$\iff$" holds due to the $\mathcal{S}$-lemma (Boyd and Vandenberghe, 2004, Appendix B). The dual of the obtained semi-definite program is

$$\begin{array}{ll} \min\limits_{\boldsymbol{X},\boldsymbol{x}} & \mathrm{Tr}\,(\boldsymbol{A}_0\boldsymbol{X})+2\boldsymbol{b}_0^\top\boldsymbol{x}+c_0 \\ \text{s.t.} & \mathrm{Tr}\,(\boldsymbol{A}_1\boldsymbol{X})+2\boldsymbol{b}_1^\top\boldsymbol{x}+c_1\leq 0 \\ & \begin{bmatrix}\boldsymbol{X} & \boldsymbol{x}\\ \boldsymbol{x}^\top & 1\end{bmatrix}\succeq 0, \end{array}$$

which is clearly an RPT relaxation of (QCQP). Consider now a generic quadratically constrained quadratic optimization problem with more than one quadratic inequality constraint:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \boldsymbol{x}^\top \boldsymbol{A}_0 \boldsymbol{x} + 2\boldsymbol{b}_0^\top \boldsymbol{x} + c_0 \\ \text{s.t.} & \boldsymbol{x}^\top \boldsymbol{A}_k \boldsymbol{x} + 2\boldsymbol{b}_k^\top \boldsymbol{x} + c_k \leq 0 \quad k \in \mathcal{K}, \end{array}$$

where $\boldsymbol{A}_k \in \mathbb{R}^{n_x \times n_x}$, $\boldsymbol{b}_k \in \mathbb{R}^{n_x}$ and $c_k \in \mathbb{R}$ for each $k \in \mathcal{K}_0$. Similarly, the dual of the convex relaxation obtained from using the approximate $\mathcal{S}$-lemma coincides with the convex relaxation from RPT:

$$\begin{array}{ll} \min\limits_{\boldsymbol{X},\boldsymbol{x}} & \operatorname{Tr}(\boldsymbol{A}_0 \boldsymbol{X}) + 2\boldsymbol{b}_0^\top \boldsymbol{x} + c_0 \\ \text{s.t.} & \operatorname{Tr}(\boldsymbol{A}_k \boldsymbol{X}) + 2\boldsymbol{b}_k^\top \boldsymbol{x} + c_k \leq 0 \quad k \in \mathcal{K} \\ & \begin{bmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^\top & 1 \end{bmatrix} \succeq 0. \end{array}$$

Note that here the obtained relaxation is not tight in general, and for more details on the approximate $\mathcal{S}$-lemma, we refer to Ben-Tal et al. (2002).

7.4. Fractional optimization

Consider the following generic fractional optimization problem

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & \dfrac{f(\boldsymbol{x})}{g(\boldsymbol{x})} \\ \text{s.t.} & h_k(\boldsymbol{x}) \leq 0 \quad k \in \mathcal{K}, \end{array} \tag{FP}$$

where $f:\mathbb{R}^{n_x} \to \mathbb{R}_+$ is convex and nonnegative, $g:\mathbb{R}^{n_x} \to \mathbb{R}_{++}$ is concave and positive, and $h_k:\mathbb{R}^{n_x} \to [-\infty, +\infty]$ is proper, closed and convex for every $k \in \mathcal{K}$. By first introducing an epigraphical variable τ for the positive convex function $1/g(\boldsymbol{x})$, we obtain the SLC constraint $\tau g(\boldsymbol{x}) \geq 1$, and then apply RPT to obtain:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x},\tau} & \tau f(\boldsymbol{y}/\tau) \\ \text{s.t.} & \tau g(\boldsymbol{y}/\tau) \geq 1 \\ & \tau h_k(\boldsymbol{y}/\tau) \leq 0 \quad k \in \mathcal{K}. \end{array}$$

The obtained convex program is an exact convex reformulation of (FP) (Schaible, 1974).

8. Numerical experiments

In this section we demonstrate the efficiency and effectiveness of RPT on a dike height optimization problem from Eijgenraam et al. (2017), and a sum-of-max-linear-terms maximization problem from Selvi et al. (2020) and Zhen et al. (2018). Numerical experiments are performed on an Intel Core i7-8665U 1.90GHz Windows computer with 32.0GB of RAM. All computations are conducted with MOSEK 9.2.27 (MOSEK ApS, 2020) and Gurobi 9.1.1 and implemented using YALMIP (Löfberg, 2004) in MATLAB (R2020a).

8.1. Dike height optimization

In Eijgenraam et al. (2017) an optimization model is developed to optimize the dike heightening in the Netherlands. The authors show that the optimal solution is periodic, i.e., the dike is heightened with the same amount every t years, and explicit expressions are derived for t and the optimal heightenings. However, in practice there are several reasons to deviate from the periodic solution. For example, it is maybe desired to combine heightenings of several dikes. In this section, we propose to use RPT to solve the dike heightening problem in which the years that the heightening takes place is fixed and may deviate from every t years. Such problems cannot be solved by the approach in Eijgenraam et al. (2017). We consider the following dike height optimization problem, which is the time truncated version of the problem in Eijgenraam et al. (2017):

$$\min_{\boldsymbol{x}\geq\mathbf{0},\boldsymbol{h}} \underbrace{\sum_{k\in\mathcal{K}_0}(c+bx_k)\exp(\lambda h_k-\delta t_k)}_{\text{Investment costs}}+\underbrace{\sum_{k\in\mathcal{K}_0}\frac{S_0}{\beta_\delta}(\exp(\beta_\delta t_{k+1})-\exp(\beta_\delta t_k))\exp(-\theta h_k)}_{\text{Expected damage costs}}+\underbrace{\frac{S_0}{\delta}\exp(\beta_\delta T-\theta h_K)}_{\text{Future damage costs}}, \tag{DHO}$$

where $\boldsymbol{t}$ is the vector of all moments in time at which the dike height is increased, $t_0=0$, $\boldsymbol{x}$ is the vector of all increases in dike height, where x_k is the increment of the dike height at time t_k, h_k is the increase in dike height after t_k years, i.e., $h_k=\sum_{i=0}^k x_i$, $h_K=\sum_{k\in\mathcal{K}_0}x_k$ and $\beta_\delta,\delta,\theta,\lambda,b,c,T$ and S_0 are constants, which are explained in more detail in Appendix E.1.

The objective of (DHO) is to find an optimal balance between investment costs and the total expected costs of flooding, both as a result of heightening dikes. The first term in the objective represents the investment costs, the second term represents the expected damage costs and the third term represents the future damage costs. We refer the reader to Eijgenraam et al. (2017) for a full description.

Since $\boldsymbol{t}$ is fixed, the objective of (DHO) is SLC, as it consists of two convex terms (expected damage costs and future damage costs) and a sum of linear times convex functions. Hence we can apply RPT.

Moreover we add the SDP relaxation as described in Section 4.4 (a) to obtain tighter bounds. The RPT-SDP approximation gives a lower bound of the optimal objective value of problem (DHO), see Appendix C.1. Since the optimal solution for the RPT-SDP relaxation is feasible for (DHO), the objective value of (DHO) evaluated at this solution constitutes an upper bound of the optimal objective value. The results for the homogeneous dike rings 10, 15 and 16 in the Netherlands are shown in Table 4.

Our method yields a global optimum for each dike segment and each considered $\boldsymbol{t}$, within a computation time of 0.5 seconds. Observe that in Table 4 we only state the optimal value, since for each case the upper and lower bound are the same. In the first two columns, i.e., for periodic $\boldsymbol{t}$, we see that except at the beginning and the end of the planning period, the heightenings are

	Every 25 years			**Every 50 years**			**Irregular $\boldsymbol{t}$**		
No.	10	15	16	10	15	16	10	15	16
x_k	2.69	41.72	34.05	11.30	55.41	47.53	0.73	38.27	30.64
	23.55	27.11	24.91				22.83	28.92	26.83
	24.85	25.95	24.37	48.08	51.95	48.61	33.16	36.19	34.39
	24.93	25.86	24.34						
	24.94	25.85	24.33	49.78	51.72	48.66	39.50	40.81	38.55
	24.94	25.85	24.33				34.07	32.73	30.40
	24.93	25.85	24.33	49.83	51.71	48.67	25.54	26.54	24.79
	24.93	25.85	24.33				27.11	29.24	27.61
	25.15	25.85	24.33	49.85	51.83	48.75	35.09	38.29	36.61
	24.76	25.87	24.34						
	25.68	26.24	24.64	55.95	59.52	56.46	32.16	27.93	25.39
	30.26	35.13	33.58				29.32	35.78	34.45
h_T	281.62	337.14	311.88	264.79	322.13	298.68	279.50	334.69	309.65
Exp. Inv.	38.02	473.48	1029.59	32.22	415.73	879.36	38.72	473.51	1029.84
Damage	23.29	136.44	240.03	23.28	129.50	220.71	23.27	135.24	238.27
Sum	61.31	609.92	1269.63	55.50	545.23	1100.07	61.98	608.74	1268.11
Time	0.39	0.50	0.47	0.45	0.42	0.42	0.41	0.42	0.50

Table 4 Results for dike rings 10, 15 and 16 in the Netherlands. In the first two columns $\boldsymbol{t}$ is such that $t_{k+1} - t_k$ equals 25 and 50 respectively for every $k \in \mathcal{K}_0$, and in the third column we consider the non-periodic $\boldsymbol{t} = (0, 20, 50, 90, 130, 155, 180, 210, 255, 279)^\top$. For the planning horizon we choose $T = 300$. The total increase in dike heights at the end of the planning horizon is given by h_T. Exp. Inv. represents the exponential investment costs, Damage reflects the total damage costs, i.e., expected plus future damage costs, and Time reflects the computation time.

approximately the same. We do not observe this for irregular $\boldsymbol{t}$, as we can see in the third column. From this result we conjecture that if the moments of heightening are chosen to be periodic, then the amount of heightening is always the same, except for the beginning and the end of the planning period. This has not been observed by the authors of Eijgenraam et al. (2017).

8.2. Sum-of-max-linear-terms maximization over convex constraints

We consider the following generic sum-of-max-linear-terms maximization problem from Zhen et al. (2018) and Selvi et al. (2020):

$$\max_{\boldsymbol{x} \in \mathcal{X}} \quad \sum_{k \in \mathcal{K}} \max_{j \in \mathcal{J}_k} \{\boldsymbol{A}_j^\top \boldsymbol{x} + b_j\}, \tag{CM}$$

where $\boldsymbol{A} \in \mathbb{R}^{n_x \times n_y}$, $\boldsymbol{b} \in \mathbb{R}^{n_y}$ and $\mathcal{J}$ is the union of the mutually disjoint sets $\mathcal{J}_k, k \in \mathcal{K}$. We consider three cases of $\mathcal{X}$, those are, a set defined by linear constraints, a set defined by an additional geometric constraint, and a set defined by an additional non-convex constraint, i.e., $\mathcal{X} = \mathcal{X}_1$, $\mathcal{X} = \mathcal{X}_2$, and $\mathcal{X} = \mathcal{X}_3$, where

$$\begin{aligned}
\mathcal{X}_1 &= \{\boldsymbol{x} \in \mathbb{R}^{n_x}_+ \mid \boldsymbol{D}^\top \boldsymbol{x} \leq \boldsymbol{d}\} \\
\mathcal{X}_2 &= \left\{ \boldsymbol{x} \in \mathcal{X}_1 \;\middle|\; \ln\left(\sum_{i=1}^{n_x} \exp(x_i)\right) \leq a \right\} \\
\mathcal{X}_3 &= \left\{ \boldsymbol{x} \in \mathcal{X}_1 \middle| \|\boldsymbol{x}\|_2 + \sum_{i=1}^{n_x} \sqrt{x_i} \leq c \right\}.
\end{aligned}$$

Here $\boldsymbol{D} \in \mathbb{R}^{n_x \times L}$ and $\boldsymbol{d} \in \mathbb{R}^L$. Since the objective of (CM) is a closed convex function, we can replace the objective by its biconjugate function (see Section 6.1) and obtain the following equivalent maximization problem:

$$\max_{\substack{\boldsymbol{x}\in\mathcal{X} \\ \boldsymbol{y}\in\mathcal{Y}}} \quad (\boldsymbol{A}^\top \boldsymbol{x} + \boldsymbol{b})^\top \boldsymbol{y}, \tag{CM$_\mathrm{B}$}$$

where $\mathcal{Y}$ equals the domain of the conjugate function of the objective of (CM), i.e.,

$$\mathcal{Y} = \left\{ \boldsymbol{y} \in \mathbb{R}^{n_y}_+ \middle| \sum_{j\in\mathcal{J}_k} y_j = 1, k \in \mathcal{K} \right\}.$$

Observe that $n_y = |\mathcal{J}|$. Because $(\boldsymbol{A}^\top \boldsymbol{x})^\top \boldsymbol{y}$ is SLC, we first apply RPT to derive a convex relaxation of (CM$_\mathrm{B}$) (see Appendix C.2), and compute an upper bound for (CM). Subsequently, a lower bound is obtained using the mountain climbing procedure as discussed in Section 4.4 (**f**) (see Appendix D).

We compare our RPT approach with our RPT-SDP approach, which is the RPT approach in which we consider the additional SDP relaxation as discussed in 4.4 (a). Furthermore, for $\mathcal{X} = \mathcal{X}_1$

#	RPT			RPT-SDP			R4B			Gurobi	
	Lb	Gap	Time	Lb	Gap	Time	Lb	Gap	Time	Opt	Time
1	23.29	0.00	0.12	23.29	0.00	0.48	23.29	0.00	0.08	23.29	0.12
1a	22.72	0.00	0.10	22.72	0.00	0.30	22.72	0.00	0.09	22.72	0.09
2	233.94	0.00	0.07	233.94	0.00	1.75	233.94	0.00	0.07	233.94	0.25
2a	211.87	0.11	0.10	211.87	0.00	1.31	211.87	0.11	0.10	211.87	0.22
3	1081.62	8.12	1.62	1081.62	0.00	111.06	1081.62	0.00	6.99	1081.62	5.64
3a	1159.90	5.81	1.27	1159.90	0.02	84.69	1159.90	0.00	15.19	1159.90	3.72
7	113.71	0.00	0.15	113.71	0.00	0.36	113.71	0.00	0.09	113.71	0.23
7a	83.78	0.31	0.28	83.78	0.00	0.58	83.78	0.00	0.40	83.78	0.26
11	3002.43	0.98	2.25	3002.43	0.04	168.06	3002.52	0.98	600*	3002.43	100.09
11a	2878.62	4.33	1.73	2878.62	1.06	224.97	2878.70	4.32	600*	2878.62	131.19

Table 5 Comparison of RPT for $\mathcal{X} = \mathcal{X}_1$ with RPT-SDP, R4B, and Gurobi for problem instances 1, 2, 3, 7, and 11. The results for problem instances 1a, 2a, 3a, 7a, and 11a reflect the average of 10 randomly generated instances corresponding to problems 1, 2, 3, 7, and 11 respectively. The gap (Gap) in % is obtained by $\frac{100(Ub-Lb)}{Ub}$, where Lb and Ub stand for the lower and upper bound respectively. Time represents the computation time, which equals the sum of the computation time of the upper and lower bound for RPT and RPT-SDP. We set the maximum time limit equal to 600 seconds, hence if the computation time equals 600*, the optimum cannot be found within 600 seconds and Gurobi returns the best value it can obtain within 600 seconds.

and $\mathcal{X} = \mathcal{X}_2$, we compare our RPT and RPT-SDP approaches to the R4B method of Zhen et al. (2018, Algorithm 2), and the exact mixed integer optimization reformulation, given by

$$\begin{array}{lll} \max\limits_{\boldsymbol{\lambda},\boldsymbol{z}} & \lambda_k & \\ \text{s.t.} & \lambda_k \leq \boldsymbol{A}_j^\top \boldsymbol{x} + b_j + M(1 - z_j) & j \in \mathcal{J}_k, k \in \mathcal{K} \\ & \sum\limits_{j \in \mathcal{J}_k} z_j = 1 & k \in \mathcal{K} \\ & \boldsymbol{z} \in \{0,1\}^{n_y}. & \end{array}$$

We use Gurobi to solve all linear optimization problems, while we use MOSEK for the nonlinear optimization problems. The problem settings and test instances are adopted from Selvi et al. (2020) and Zhen et al. (2018); see Appendix E.2. for more detail.

The results for $\mathcal{X} = \mathcal{X}_1$ are given in Table 5. Our RPT approach finds the optimum for instances 1, 1a, 2, and 7. By considering the additional SDP relaxation, we also find an optimum for instances 2a, 3, and 7a, and we improve the gap of RPT for instances 3a, 11 and 11a. However, the computation time is much longer, e.g., for instance 3, the computation time of RPT-SDP is 111.06 against 1.62 for RPT. R4B improves the gap of RPT for instances 3, 3a, and 7, for which it finds the optimum. For instance 11 and 11a, R4B is not able to find the optimum within the time limit of 600 seconds. Moreover, for those instances, the lower bound computed by R4B is higher than the optimum value, as a result of numerical errors. Gurobi finds the optimum for every instance. For instances 3, 3a and 7a, Gurobi is faster than both R4B and RPT-SDP, while for instances 1, 1a, 2, 2a, and 7, R4B finds the optimum within the lowest computation time.

#	RPT			RPT-SDP			R4B			MOSEK		
	Lb	Gap	Time	Lb	Gap	Time	Lb	Gap	Time	Lb	Gap	Time
1	14.58	0.00	0.33	14.58	0.00	0.36	14.58	0.00	0.33	14.58	0.00	0.30
1a	14.54	0.00	0.33	14.54	0.00	0.34	14.54	0.00	0.33	14.54	0.00	0.29
2	136.22	0.09	1.35	136.22	0.00	4.34	136.22	0.77	600*	136.22	125.13	600*
2a	122.21	4.40	3.85	122.21	0.34	6.96	122.21	5.83	600*	122.21	126.83	600*
3	837.94	11.06	4.24	837.94	0.00	340.11	837.94	11.06	600*	837.94	526.55	600*
3a	890.07	10.49	5.37	890.07	0.07	303.28	890.07	10.49	600*	887.27	515.78	600*
7	33.73	7.61	1.75	33.73	1.67	2.50	33.73	0.00	2.11	33.73	0.00	0.33
7a	31.81	4.93	1.58	31.81	0.73	1.68	31.81	0.00	1.93	31.81	0.00	0.27
11	1610.69	11.66	9.25	1610.69	5.78	285.64	1610.69	13.10	600*	1610.69	300.09	600*
11a	1660.75	8.69	7.16	1660.80	4.01	297.68	1660.75	10.48	600*	1658.33	284.13	600*

Table 6 Comparison of our RPT approximation for $\mathcal{X} = \mathcal{X}_2$ with the RPT-SDP approximation, R4B, and MOSEK for problem instances 1, 2, 3, 7, and 11. The results for problem instances 1a, 2a, 3a, 7a, and 11a reflect the average of 10 randomly generated instances corresponding to problems 1, 2, 3, 7, and 11 respectively. The gap (Gap) in % is obtained by $\frac{100(Ub-Lb)}{Ub}$, where Lb and Ub stand for lower and upper bound respectively. Time represents the computation time, which equals the sum of the computation time of the upper and lower bound for RPT and RPT-SDP. We set the maximum time limit equal to 600 seconds, hence if the computation time equals 600^*, the global optimum cannot be found within 600 seconds and MOSEK returns the best value it can obtain within 600 seconds.

The results for $\mathcal{X} = \mathcal{X}_2$ are given in Table 6. All approaches find the optimum for instances 1 and 1a. RPT-SDP finds the best solution for instances 2, 2a, 3, 3a, 11, and 11a, where for instances 2 and 3, RPT-SDP even finds the optimum. However, for bigger problem instances, like 3, 3a, 11, and 11a, RPT is much faster without much of compromise on the quality of the obtained solution. R4B and MOSEK only improve the gap of RPT and RPT-SDP for instances 7 and 7a, for which they find an optimum. For instances 2, 2a, 3, 3a, 11, and 11a, both R4B and MOSEK cannot find the optimum within 600 seconds.

The results for $\mathcal{X} = \mathcal{X}_3$ are given in Table 7. Note that this problem is extremely hard, since not only the objective function is concave, but it also contains a highly nonconvex constraint function. RPT and RPT-SDP both find an upper bound for every instance. It is possible that the convex relaxation from the RPT or RPT-SDP approach returns a solution that is infeasible for the original problem. Therefore, not always a feasible solution is found, in which case we cannot derive a lower bound. For instance 7a, RPT could not find a lower bound for one of the 10 randomly generated instances, while RPT-SDP could not find a lower bound for instances 11 and 11a. For the remaining instances, RPT-SDP improves the gap of RPT for all instances except for 1a, for which the lower bound found by the RPT approach is better than the lower bound found by the RPT-SDP approach. Especially for instances 3, 3a, and 7, the gap for RPT-SDP is much smaller than the gap for RPT. However, for instances 3 and 3a, the computation time for RPT-SDP is much longer.

#	RPT				RPT-SDP			
	Lb	Ub	Gap	Time	Lb	Ub	Gap	Time
1	13.44	16.09	19.69	11.66	13.44	14.65	8.96	8.64
1a	14.46	16.46	13.82	12.01	13.49	15.89	17.75	13.03
2	140.89	146.04	3.66	12.93	140.89	145.59	3.34	24.97
2a	128.61	139.92	8.79	16.74	128.61	135.90	5.67	24.37
3	768.97	880.61	14.52	24.84	768.95	791.35	2.91	398.46
3a	804.69	906.39	12.69	11.54	805.95	833.80	3.46	387.37
7	45.34	69.17	52.57	13.36	45.34	62.60	38.06	14.65
7a	43.50*	61.66	41.13*	17.02*	39.22	56.44	43.91	16.44
11	2175.60	3008.27	38.27	25.25	-	2907.81	-	503.43
11a	2315.50	2909.02	29.55	15.08	-	2890.47	-	426.55

Table 7 Comparison of our RPT approximation for $\mathcal{X} = \mathcal{X}_3$ with our RPT-SDP approximation for problem instances 1, 2, 3, 7, and 11. The results for problem instances 1a, 2a, 3a, 7a, and 11a reflect the average of 10 randomly generated instances corresponding to problems 1, 2, 3, 7, and 11 respectively. The gap (Gap) in % is obtained by $\frac{100(Ub-Lb)}{Ub}$, where Lb and Ub stand for lower and upper bound respectively. Time represents the computation time, which equals the sum of the computation time of the upper and lower bound. * indicates that for 9 out of the 10 randomly generated instances, a feasible solution is found. Hence for the corresponding Lb, Gap and Time, the average of those 9 instances is reported. - indicates no feasible solution is found. Hence, for those problem instances we only report the upper bound and the computation time of the upper bound.

For instances 4-6, 8-10 and 12-13 in Selvi et al. (2020), RPT-SDP took too much time because of the LMI. In such cases one can use RPT without the LMI.

Acknowledgment

We would like to thank Bram Gorissen for the suggestion to consider fractional optimization (Section 7.4) from a RPT lens. We thank Jean Pauphilet for providing us with the new idea of case 6 (ii) in Table 2. Finally, we thank Kurt Anstreicher and Daniel Kuhn for their valuable comments on a draft version of this paper.

References

K.M. Anstreicher. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. *Journal of Global Optimization*, 43(2–3):471–484, 2009.

K.M. Anstreicher. On convex relaxations for quadratically constrained quadratic programming. *Mathematical Programming*, 136(2):233–251, 2012.

K.M. Anstreicher. Kronecker product constraints with an application to the two-trust-region subproblem. *SIAM Journal on Optimization*, 27(1):368–378, 2017.

X. Bao, N.V. Sahinidis, and M. Tawarmalani. Semidefinite relaxations for quadratically constrained quadratic programming: A review and comparisons. *Mathematical Programming*, 129:129–157, 2011.

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

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

G. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ, 1963.

C. Eijgenraam, R. Brekelmans, D. den Hertog, and C. Roos. Optimal strategies for flood prevention. *Management Science*, 63(5):1644–1656, 2017.

B. Gorissen and D. Den Hertog. Robust nonlinear optimization via the dual. *Optimization Online*, 2015. URL http://www.optimization-online.org/DB_FILE/2015/04/4886.pdf.

B. Gorissen, A. Ben-Tal, H. Blanc, and D. den Hertog. Deriving robust and globalized robust solutions of uncertain linear programs with general convex uncertainty sets. *Operations Research*, 62(3):672–679, 2014.

I.E. Grossmann and S Lee. Generalized convex disjunctive programming: Nonlinear convex hull relaxation. *Computational Optimization and Applications*, 26:83–100, 2003.

R.A. Horn and C.R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

R. Jiang and D. Li. Convex relaxations with second order cone constraints for nonconvex quadratically constrained quadratic programming. *arXiv*, 2016. URL https://arxiv.org/pdf/1608.02096.pdf.

R. Jiang and D. Li. Second order cone constrained convex relaxations for nonconvex quadratically constrained quadratic programming. *Journal of Global Optimization*, 75:461–494, 2019.

R. Jiang and D. Li. Semidefinite programming based convex relaxation for nonconvex quadratically constrained quadratic programming. In Hoai An Le Thi, Hoai Minh Le, and Tao Pham Dinh, editors, *Optimization of Complex Systems: Theory, Models, Algorithms and Applications*, pages 213–220, Cham, 2020. Springer International Publishing. ISBN 978-3-030-21803-4.

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

MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 9.2*, 2020. URL http://docs.mosek.com/9.2/toolbox.pdf.

R. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

R. Rockafellar and R. Wets. *Variational Analysis*. Springer, 2009.

S. Schaible. Parameter-free convex equivalent and dual programs of fractional programming problems. *Zeitschrift f ur Operations Research*, 18(5):187–196, 1974.

I. Schur. Bemerkungen zur theorie der beschränkten bilinearformen mit unendlich vielen veränderlichen. *Journal für die Reine und Angewandte Mathematik*, 140:1–28, 1911.

A. Selvi, A. Ben-Tal, R. Brekelmans, and D. den Hertog. Convex maximization via adjustable robust optimization. *Optimization Online*, 2020. URL http://www.optimization-online.org/DB_HTML/2020/07/7881.html.

S.A. Serrano. *Algorithms for unsymmetric cone optimization and an implementation for problems with the exponential cone*. PhD thesis, Stanford University, 2015.

H. Sherali and W. Adams. A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. *SIAM Journal on Discrete Mathematics*, 3(3):411–430, 1990.

H. Sherali and W. Adams. A reformulation-linearization technique for semi-infinite and convex programs under mixed 0-1 and general discrete restrictions. *Discrete Applied Mathematics*, 157(6):1319–1333, 1994a.

H. Sherali and W. Adams. A hierarchy of relaxations leading to the convex hull representation for general discrete optimization problems. *Discrete Applied Mathematics*, 52:83–106, 1994b.

H. Sherali and W. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, volume 31. Springer Science & Business Media, 2013.

H. Sherali and A. Alameddine. A new reformulation-linearization technique for bilinear programming problems. *Journal of Global Optimization*, 2(4):379–410, 1992.

J.F. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research*, 28(2):246–267, 2003.

P. D. Tao and L.T.H. An. Convex analysis approach to d. c. programming: Theory, algorithms and applications. *Acta Mathematica Vietnamica*, 22(1):289–355, 1997.

B. Yang and S. Burer. A two-variable approach to the two-trust-region subproblem. *SIAM Journal on Optimization*, 26(1):661–680, 2016.

J. Zhen, A. Marandi, D. de Moor, D. den Hertog, and L. Vandenberghe. Disjoint bilinear programming: a two-stage robust optimization perspective. *Optimization Online*, 2018. URL http://www.optimization-online.org/DB_FILE/2018/06/6685.pdf.

J. Zhen, D. Kuhn, and W. Wiesemann. Mathematical foundations of robust and distributionally robust optimization. *Optimization Online*, 2021. URL http://www.optimization-online.org/DB_HTML/2021/04/8360.html.

Appendix A. Technical lemma

LEMMA 2. *Suppose* $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are positive semidefinite matrices. Then also the Kronecker product is positive semidefinite, i.e.,*

$$\boldsymbol{A}(\boldsymbol{x}) \otimes \boldsymbol{B}(\boldsymbol{x}) \succeq 0. \tag{22}$$

Moreover, if $\boldsymbol{A}$ *and* $\boldsymbol{B}$ *are positive semidefinite matrices of the same size, then the Hadamard product is also positive semidefinite, i.e.,*

$$\boldsymbol{A}(\boldsymbol{x}) \circ \boldsymbol{B}(\boldsymbol{x}) \succeq 0. \tag{23}$$

Proof The first statement follows from Theorem 4.2.12 of Horn and Johnson (1991). The last statement follows from the Schur Product Theorem (Schur, 1911). □

Appendix B. Examples for products of conic inequalities

1. **(L)** × **(L)**

   Consider two linear inequalities

$$\begin{cases} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq 0 \\ b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq 0. \end{cases}$$

   We now apply RPT to these inequalities, and obtain (Sherali and Alameddine, 1992):

$$\begin{aligned} (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})(b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) \geq 0 &\iff b_1 b_2 - b_1 \boldsymbol{a}_2^\top \boldsymbol{x} - b_2 \boldsymbol{a}_1^\top \boldsymbol{x} + \boldsymbol{a}_1^\top \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{a}_2 \geq 0 \\ &\implies b_1 b_2 - b_1 \boldsymbol{a}_2^\top \boldsymbol{x} - b_2 \boldsymbol{a}_1^\top \boldsymbol{x} + \boldsymbol{a}_1^\top \boldsymbol{U} \boldsymbol{a}_2 \geq 0. \end{aligned}$$

2. **(L)** × **(Q)**

   Consider one linear inequality and one conic quadratic inequality

$$\begin{cases} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq 0 \\ b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} - \left\|\boldsymbol{D}\boldsymbol{x} + \boldsymbol{d}\right\| \geq 0. \end{cases}$$

   We now apply RPT to these inequalities, and obtain (Sturm and Zhang (2003)) :

$$\begin{aligned} &(b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})\left\|\boldsymbol{D}\boldsymbol{x} + \boldsymbol{d}\right\| \leq (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})(b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) \\ \iff &\left\|(b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})(\boldsymbol{D}\boldsymbol{x} + \boldsymbol{d})\right\| \leq (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})(b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) \\ \implies &\left\|b_1 \boldsymbol{D}\boldsymbol{x} + b_1 \boldsymbol{d} - \boldsymbol{D}\boldsymbol{U}\boldsymbol{a}_1 - \boldsymbol{a}_1^\top \boldsymbol{x}\boldsymbol{d}\right\| \leq b_1 b_2 - b_1 \boldsymbol{a}_2^\top \boldsymbol{x} - b_2 \boldsymbol{a}_1^\top \boldsymbol{x} + \boldsymbol{a}_1^\top \boldsymbol{U} \boldsymbol{a}_2. \end{aligned}$$

3. **(L)** × **(P)**

   Consider one linear inequality and one power cone inequality

$$\begin{cases} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq 0 \\ \displaystyle\prod_{i=1}^{m} x_i^{\alpha_i} - \sqrt{\sum_{i=m+1}^{n_x} x_i^2} \geq 0 \\ x_i \geq 0, \quad i = 1, \ldots, m, \end{cases}$$

where $\sum_{i=1}^m \alpha_i = 1$. Multiplying the linear constraints with the nonnegativity constraints and the nonnegativity constraints with each other yields m additional linear constraints, and a nonnegativity constraint for $\boldsymbol{U}$. We omit the derivations, since this can be done in the same way as explained under case 1, i.e., (L)×(L). We now apply RPT to these inequalities, and obtain (only linear constraints times the power cone inequality):

$$\begin{cases} (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x}) \prod_{i=1}^m x_i^{\alpha_i} \geq (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x}) \sqrt{\sum_{i=m+1}^{n_x} x_i^2} \\ x_j \prod_{i=1}^m x_i^{\alpha_i} \geq x_j \sqrt{\sum_{i=m+1}^{n_x} x_i^2}, \quad j = 1, \ldots, m \end{cases}$$

$$\Longleftrightarrow \begin{cases} \prod_{i=1}^m (b_1 x_i - \boldsymbol{a}_1^\top \boldsymbol{x} x_i)^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} (b_1 x_i - \boldsymbol{a}_1^\top \boldsymbol{x} x_i)^2} \\ \prod_{i=1}^m (x_j x_i)^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} (x_j x_i)^2}, \ j = 1, \ldots, m \end{cases}$$

$$\Longrightarrow \begin{cases} \prod_{i=1}^m (b_1 x_i - \boldsymbol{a}_1^\top \boldsymbol{u}_i)^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} (b_1 x_i - \boldsymbol{a}_1^\top \boldsymbol{u}_i)^2} \\ \prod_{i=1}^m u_{ij}^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} u_{ij}^2}, \ j = 1, \ldots, m. \end{cases}$$

Hence, finally we obtain $m+1$ additional power cone inequalities. Note that there are many more possibilities to perspectify the left-hand sides of the above inequalities. Section 4.3 describes how to find the best choice.

4. (**L**) × (**E**)

Consider one linear inequality and one exponential cone inequality

$$\begin{cases} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq 0 \\ x_1 - x_2 \exp(x_3/x_2) \geq 0 \\ x_2 \geq 0. \end{cases}$$

Multiplying the linear constraint with the nonnegativity constraint and the nonnegativity constraint with itself yields 1 additional linear constraint, and a nonnegativity constraint for u_{22}, see case 1, i.e., (L)×(L). We now apply RPT to these inequalities, and obtain (only the linear constraints times the exponential cone inequality):

$$\begin{cases} (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x}) x_1 \geq (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x}) x_2 \exp\left(\frac{(b_1 - \boldsymbol{a}_1^\top \boldsymbol{x}) x_3}{(b_1 - \boldsymbol{a}_1^\top \boldsymbol{x}) x_2}\right) \\ x_1 x_2 \geq x_2^2 \exp\left(\frac{x_2 x_3}{x_2^2}\right) \end{cases}$$

$$\Longrightarrow \begin{cases} b_1 x_1 - \boldsymbol{a}_1^\top \boldsymbol{u}_1 \geq (b_1 x_2 - \boldsymbol{a}_1^\top \boldsymbol{u}_2) \exp\left(\frac{b_1 x_3 - \boldsymbol{a}_1^\top \boldsymbol{u}_3}{b_1 x_2 - \boldsymbol{a}_1^\top \boldsymbol{u}_2}\right) \\ u_{12} \geq u_{22} \exp\left(\frac{u_{23}}{u_{22}}\right). \end{cases}$$

Hence, we finally get two additional exponential conic inequalities.

5. (**L**) × (**S**)

   Consider one linear inequality and one LMI

$$\begin{cases} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq 0 \\ \boldsymbol{A}(\boldsymbol{x}) \succeq 0. \end{cases}$$

We now apply RPT to these inequalities (only linear constraint times the LMI), and obtain:

$$\begin{aligned} & (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})\boldsymbol{A}(\boldsymbol{x}) \succeq 0 \\ \Longleftrightarrow\ & (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})\boldsymbol{A}_0 + (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})\boldsymbol{A}_1 x_1 + \cdots + (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})\boldsymbol{A}_{n_x} x_{n_x} \succeq 0 \\ \Longrightarrow\ & (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})\boldsymbol{A}_0 + (b_1 x_1 - \boldsymbol{a}_1^\top \boldsymbol{u}_1)\boldsymbol{A}_1 + \cdots + (b_1 x_{n_x} - \boldsymbol{a}_1^\top \boldsymbol{u}_{n_x})\boldsymbol{A}_{n_x} \succeq 0. \end{aligned}$$

6. (**Q**) × (**Q**)

   Consider two conic quadratic inequalities

$$\begin{cases} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq \left\| \boldsymbol{D}_1 \boldsymbol{x} + \boldsymbol{p}_1 \right\| \\ b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq \left\| \boldsymbol{D}_2 \boldsymbol{x} + \boldsymbol{p}_2 \right\|. \end{cases} \tag{24}$$

We now apply RPT to these inequalities. Multiplying the valid linear constraints $b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq 0$ and $b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq 0$ with the conic quadratic constraints (24) yields four additional conic quadratic inequalities. We multiply the left-hand sides and right-hand sides of the inequalities, and obtain:

$$\begin{aligned} & (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})(b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) \geq \left\| \boldsymbol{D}_1 \boldsymbol{x} + \boldsymbol{p}_1 \right\| \left\| \boldsymbol{D}_2 \boldsymbol{x} + \boldsymbol{p}_2 \right\| \\ \Longleftrightarrow\quad & b_1 b_2 - b_1 \boldsymbol{a}_2^\top \boldsymbol{x} - b_2 \boldsymbol{a}_1^\top \boldsymbol{x} + \boldsymbol{a}_1^\top \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{a}_2 \geq \left\| (\boldsymbol{D}_1 \boldsymbol{x} + \boldsymbol{p}_1)(\boldsymbol{D}_2 \boldsymbol{x} + \boldsymbol{p}_2)^\top \right\|_2 \\ \Longrightarrow\quad & b_1 b_2 - b_1 \boldsymbol{a}_2^\top \boldsymbol{x} - b_2 \boldsymbol{a}_1^\top \boldsymbol{x} + \boldsymbol{a}_1^\top \boldsymbol{U} \boldsymbol{a}_2 \geq \left\| \boldsymbol{D}_1 \boldsymbol{U} \boldsymbol{D}_2^\top + \boldsymbol{p}_1 \boldsymbol{x}^\top \boldsymbol{D}^\top + \boldsymbol{D}\boldsymbol{x}\boldsymbol{p}_2^\top + \boldsymbol{p}_1 \boldsymbol{p}_2^\top \right\|_2, \end{aligned}$$

in which $\|\cdot\|_2$ denotes the 2-norm of a matrix defined as its largest singular value, i.e., $\|\boldsymbol{M}\|_2 := \sigma_{\max}(\boldsymbol{M}) = \sqrt{\lambda_{\max}(\boldsymbol{M}\boldsymbol{M}^\top)}$. In particular, we used the fact that for rank-1 matrices $\boldsymbol{M} = \boldsymbol{u}\boldsymbol{v}^\top$, $\|\boldsymbol{M}\|_2 = \|\boldsymbol{u}\|\|\boldsymbol{v}\|$. In practice, constraints of the form $\|\boldsymbol{M}\|_2 \leq \lambda$ can be modeled as semidefinite constraints using Schur complement:

$$\begin{pmatrix} \lambda^2 \boldsymbol{I} & \boldsymbol{M} \\ \boldsymbol{M}^\top & I \end{pmatrix} \succeq 0.$$

This results in case 6(ii) of Table 2, which is a new result to the best of our knowledge. To improve scalability, however, one can replace the 2-norm in the first constraint by the Frobenius norm of the matrix. Indeed, $\|\boldsymbol{M}\|_2 \leq \|\boldsymbol{M}\|_F$, and the equality holds for rank-1 and zero matrices, so

$$b_1 b_2 - b_1 \boldsymbol{a}_2^\top \boldsymbol{x} - b_2 \boldsymbol{a}_1^\top \boldsymbol{x} + \boldsymbol{a}_1^\top \boldsymbol{U} \boldsymbol{a}_2 \geq \left\| \boldsymbol{D}_1 \boldsymbol{U} \boldsymbol{D}_2^\top + \boldsymbol{p}_1 \boldsymbol{x}^\top \boldsymbol{D}^\top + \boldsymbol{D}\boldsymbol{x}\boldsymbol{p}_2^\top + \boldsymbol{p}_1 \boldsymbol{p}_2^\top \right\|_F \tag{25}$$

is a valid, yet looser, constraint. This is case 6(i) in Table 2.

In the literature also two LMIs are proposed. First observe that the two conic quadratic inequalities (24) can be written as

$$b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} \geq \left\|\boldsymbol{D}_1\boldsymbol{x} + \boldsymbol{p}_1\right\| \iff \begin{bmatrix} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} & (\boldsymbol{D}_1\boldsymbol{x} + \boldsymbol{p}_1)^\top \\ \boldsymbol{D}_1\boldsymbol{x} + \boldsymbol{p}_1 & (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})\boldsymbol{I} \end{bmatrix} \succeq 0$$

and

$$b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq \left\|\boldsymbol{D}_2\boldsymbol{x} + \boldsymbol{p}_2\right\| \iff \begin{bmatrix} b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} & (\boldsymbol{D}_2\boldsymbol{x} + \boldsymbol{p}_2)^\top \\ \boldsymbol{D}_2\boldsymbol{x} + \boldsymbol{p}_2 & (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})\boldsymbol{I} \end{bmatrix} \succeq 0.$$

We now assume that, without loss of generality, the matrices $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$ are of the same size. Indeed, suppose that $\boldsymbol{D}_1$ has less rows than $\boldsymbol{D}_2$, then we can extend matrix $\boldsymbol{D}_1$ by zero rows or by copying scaled versions of some of the original rows. Using Lemma 2, it follows for the Hadamard product that

$$\begin{bmatrix} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} & (\boldsymbol{D}_1\boldsymbol{x} + \boldsymbol{p}_1)^\top \\ \boldsymbol{D}_1\boldsymbol{x} + \boldsymbol{p}_1 & (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})\boldsymbol{I} \end{bmatrix} \circ \begin{bmatrix} b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} & (\boldsymbol{D}_2\boldsymbol{x} + \boldsymbol{p}_2)^\top \\ \boldsymbol{D}_2\boldsymbol{x} + \boldsymbol{p}_2 & (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})\boldsymbol{I} \end{bmatrix} \succeq 0,$$

which implies

$$\begin{bmatrix} \alpha & \boldsymbol{\beta}^\top \\ \boldsymbol{\beta} & \alpha\boldsymbol{I} \end{bmatrix} \succeq 0, \tag{26}$$

where

$$\alpha = b_1 b_2 - b_2 \boldsymbol{a}_1^\top \boldsymbol{x} - b_1 \boldsymbol{a}_2^\top \boldsymbol{x} + a_1^\top \boldsymbol{U} a_2 \tag{27}$$

$$\beta_i = \boldsymbol{d}_{1i} U \boldsymbol{d}_{1i} + p_{1i} p_{2i} + p_{2i} \boldsymbol{d}_{1i} \boldsymbol{x} + p_{1i} \boldsymbol{d}_{2i} \boldsymbol{x}, \quad i = 1, \ldots, r, \tag{28}$$

and $\boldsymbol{d}_{1i}$ and $\boldsymbol{d}_{2i}$ is the ith row of $\boldsymbol{D}_1$ and $\boldsymbol{D}_2$, respectively. Notice that the matrix in the left-hand side of (26) has an arrow structure, and hence LMI (26) is equivalent with the following conic quadratic inequality:

$$\|\beta\|_2 \leq \alpha. \tag{29}$$

It can easily be verified that (29) is a weaker inequality than (25). Since inequality (29) neither has computational advantages, it is therefore not useful.

Using Lemma 2 for the Kronecker product, and linearizing each element of the product, we obtain

$$\begin{bmatrix} b_1 - \boldsymbol{a}_1^\top \boldsymbol{x} & (\boldsymbol{D}_1\boldsymbol{x} + \boldsymbol{p}_1)^\top \\ \boldsymbol{D}_1\boldsymbol{x} + \boldsymbol{p}_1 & (b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})\boldsymbol{I} \end{bmatrix} \otimes \begin{bmatrix} b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} & (\boldsymbol{D}_2\boldsymbol{x} + \boldsymbol{p}_2)^\top \\ \boldsymbol{D}_2\boldsymbol{x} + \boldsymbol{p}_2 & (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})\boldsymbol{I} \end{bmatrix} \succeq 0$$

$$\implies \left[\begin{array}{cc|cc|c|cc} \alpha & \boldsymbol{\gamma}^\top & \delta_1 & \boldsymbol{\eta}_1^\top & \cdots & \delta_r & \boldsymbol{\eta}_r^\top \\ \boldsymbol{\gamma} & \alpha I & \boldsymbol{\eta}_1 & \delta_1 I & & \boldsymbol{\eta}_r & \delta_r I \\ \hline \delta_1 & \boldsymbol{\eta}_1^\top & \alpha & \boldsymbol{\gamma}^\top & & & \\ \boldsymbol{\eta}_1 & \delta_1 I & \boldsymbol{\gamma} & \alpha I & & & \\ \hline \vdots & & & & \ddots & & \\ \hline \delta_r & \boldsymbol{\eta}_r^\top & & & & \alpha & \boldsymbol{\gamma}^\top \\ \boldsymbol{\eta}_r & \delta_r I & & & & \boldsymbol{\gamma} & \alpha I \end{array}\right] \succeq 0,$$

where

$$\begin{aligned}
\alpha &= b_1 b_2 - b_2 \boldsymbol{a}_1^\top \boldsymbol{x} - b_1 \boldsymbol{a}_2^\top \boldsymbol{x} + a_1^\top U a_2 \\
\boldsymbol{\gamma} &= b_1(\boldsymbol{D}_2 \boldsymbol{x} + \boldsymbol{p}_2) - (a_1^\top \boldsymbol{x})\boldsymbol{p}_2 - \boldsymbol{D}_2 \boldsymbol{U} \boldsymbol{a}_1 \\
\delta_i &= b_2(\boldsymbol{d}_{1i}^\top \boldsymbol{x}) + p_{1i}(b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) - \boldsymbol{d}_{1i}^\top \boldsymbol{U} \boldsymbol{a}_1, \quad i = 1, \dots, r \\
\boldsymbol{\eta}_i &= (\boldsymbol{d}_{1i}^\top \boldsymbol{x} + p_{1i})\boldsymbol{p}_2 + p_{1i} \boldsymbol{D}_2 \boldsymbol{x} + \boldsymbol{D}_2 \boldsymbol{U} \boldsymbol{d}_{1i}, \quad i = 1, \dots, r.
\end{aligned}$$

7. (**Q**) $\times$ (**P**)

   Consider one conic quadratic inequality and one power cone inequality

$$\begin{cases} b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq \|\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p}\| \\ \prod_{i=1}^{m} x_i^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} x_i^2} \\ x_1, \cdots, x_m \geq 0. \end{cases}$$

We now apply RPT to these inequalities. Multiplying the linear constraints with the conic quadratic and the power cone inequalities yields m conic quadratic and m power cone inequalities, respectively. Multiplying the power cone inequality with the left-hand side of the conic quadratic inequality and multiplying the left-hand-sides and the right-hand sides of the two inequalities with each other, yields:

$$\begin{aligned}
&\begin{cases} (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) \prod_{i=1}^{m} x_i^{\alpha_i} \geq (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) \sqrt{\sum_{i=m+1}^{n_x} x_i^2} \\ (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x}) \prod_{i=1}^{m} x_i^{\alpha_i} \geq \|\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p}\| \sqrt{\sum_{i=m+1}^{n_x} x_i^2} \end{cases} \\
\iff &\begin{cases} \prod_{i=1}^{m} (b_2 x_i - \boldsymbol{a}_2^\top \boldsymbol{x} x_i)^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} (b_2 x_i - \boldsymbol{a}_2^\top \boldsymbol{x} x_i)^2} \\ \prod_{i=1}^{m} (b_2 x_i - \boldsymbol{a}_2^\top \boldsymbol{x} x_i)^{\alpha_i} \geq \left\|(\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p})\boldsymbol{x}_{[m+1]}^\top\right\|_F \end{cases} \\
\implies &\begin{cases} \prod_{i=1}^{m} (b_2 x_i - \boldsymbol{a}_2^\top \boldsymbol{u}_i)^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} (b_2 x_i - \boldsymbol{a}_2^\top \boldsymbol{u}_i)^2} \\ \prod_{i=1}^{m} (b_2 x_i - \boldsymbol{a}_2^\top \boldsymbol{u}_i)^{\alpha_i} \geq \left\|\boldsymbol{D}\boldsymbol{U}_{[m+1]} + \boldsymbol{p}\boldsymbol{x}_{[m+1]}^\top\right\|_F, \end{cases}
\end{aligned}$$

where $\boldsymbol{x}_{[m+1]} = [x_{m+1} \ \cdots \ x_{n_x}]$ and $\boldsymbol{U}_{[m+1]} = [\boldsymbol{u}_{m+1} \ \cdots \ \boldsymbol{u}_{n_x}]$. Hence, we have obtained two extra power cone inequalities. This results in case 7(i) in Table 2. The Frobenius-norm in the last power cone inequality can also be replaced by the 2-norm, which yields the tighter case 7(ii) in Table 2.

8. $(\mathbf{Q}) \times (\mathbf{E})$ Consider one conic quadratic inequality and one exponential cone inequality

$$\begin{cases} b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq \|\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p}\| \\ x_1 \geq x_2 \exp^{x_3/x_2} \\ x_2 \geq 0. \end{cases}$$

We now apply RPT to these inequalities. Multiplying the linear constraint with the conic quadratic and the exponential cone inequality yields another conic quadratic and another exponential cone inequality, respectively. Multiplying the left-hand side of the conic quadratic inequality with the exponential cone inequality yields:

$$(b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})x_1 \geq (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})x_2 \exp^{(b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})x_3/(b_1 - \boldsymbol{a}_1^\top \boldsymbol{x})x_2}$$

$$\implies b_2 x_1 - \boldsymbol{a}_2^\top \boldsymbol{u}_1 \geq (b_2 x_2 - \boldsymbol{a}_2^\top \boldsymbol{u}_2) \exp^{(b_2 x_3 - \boldsymbol{a}_2^\top \boldsymbol{u}_3)/(b_2 x_2 - \boldsymbol{a}_2^\top \boldsymbol{u}_2)},$$

which is again an exponential cone inequality.

9. $(\mathbf{Q}) \times (\mathbf{S})$

Consider one conic quadratic inequality and one LMI

$$\begin{cases} b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq \|\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p}\| \\ \boldsymbol{A}(\boldsymbol{x}) \succeq 0. \end{cases} \tag{30}$$

First observe that

$$b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} \geq \|\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p}\| \iff \begin{bmatrix} b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} & (\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p})^\top \\ \boldsymbol{D}\boldsymbol{x} + \boldsymbol{p} & (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})\boldsymbol{I} \end{bmatrix} \succeq 0.$$

We now apply RPT to these inequalities, use the fact that the Kronecker product of two positive semidefinite matrices is also positive semidefinite (Lemma 2), and obtain:

$$\begin{aligned}
&\left(b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} - \|\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p}\|\right) \boldsymbol{A}(\boldsymbol{x}) \succeq 0 \implies \begin{bmatrix} b_2 - \boldsymbol{a}_2^\top \boldsymbol{x} & (\boldsymbol{D}\boldsymbol{x} + \boldsymbol{p})^\top \\ \boldsymbol{D}\boldsymbol{x} + \boldsymbol{p} & (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})\boldsymbol{I} \end{bmatrix} \otimes \boldsymbol{A}(\boldsymbol{x}) \succeq 0 \\
&\iff \begin{bmatrix} (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})\boldsymbol{A}(\boldsymbol{x}) & (\boldsymbol{d}_1^\top \boldsymbol{x} + \boldsymbol{p}_1)\boldsymbol{A}(\boldsymbol{x}) & \cdots & (\boldsymbol{d}_r^\top \boldsymbol{x} + \boldsymbol{p}_r)\boldsymbol{A}(\boldsymbol{x}) \\ (d_1^\top \boldsymbol{x} + \boldsymbol{p}_1)\boldsymbol{A}(\boldsymbol{x}) & (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})\boldsymbol{A}(\boldsymbol{x}) & & \\ \vdots & & \ddots & \\ (d_r^\top \boldsymbol{x} + \boldsymbol{p}_r)\boldsymbol{A}(\boldsymbol{x}) & & & (b_2 - \boldsymbol{a}_2^\top \boldsymbol{x})\boldsymbol{A}(\boldsymbol{x}) \end{bmatrix} \succeq 0 \\
&\implies \begin{bmatrix} \boldsymbol{A}(b_2 \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_2) & \boldsymbol{A}(p_1 \boldsymbol{x} + \boldsymbol{U}\boldsymbol{d}_1) & \cdots & \boldsymbol{A}(p_r \boldsymbol{x} + \boldsymbol{U}\boldsymbol{d}_r) \\ \boldsymbol{A}(p_1 \boldsymbol{x} + \boldsymbol{U}\boldsymbol{d}_1) & \boldsymbol{A}(b_2 \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_2) & & \\ \vdots & & \ddots & \\ \boldsymbol{A}(p_r \boldsymbol{x} + \boldsymbol{U}\boldsymbol{d}_r) & & & \boldsymbol{A}(b_2 \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_2) \end{bmatrix} \succeq 0,
\end{aligned} \tag{31}$$

where $\boldsymbol{d}_i$ is the ith row of $\boldsymbol{D}$. Notice that (31) is an LMI (linear in $\boldsymbol{x}$ and $\boldsymbol{U}$). One could also multiply the left-hand-side of the conic quadratic inequality with the LMI in (30), and then one obtains

$$A(b_2 \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_2) \succeq 0,$$

which is also implied by (31).

10. $(\mathbf{P}) \times (\mathbf{P})$

Consider two power cone inequalities

$$\begin{cases} \prod_{i=1}^{m_1} x_i^{\alpha_{1i}} \geq \sqrt{\sum_{i=m_1+1}^{n_x} x_i^2} \\ x_i \geq 0,\ i=1,\ldots,m_1 \\ \prod_{j=1}^{m_2} x_{\sigma(j)}^{\alpha_{2j}} \geq \sqrt{\sum_{j=m_2+1}^{n_x} x_{\sigma(j)}^2} \\ x_{\sigma(j)} \geq 0,\ j=1,\ldots,m_2, \end{cases}$$

where σ be an arbitrary permutation, $n_x > m_1, m_2,\ \alpha_{11}, \cdots, \alpha_{1m}, \alpha_{21}, \cdots, \alpha_{2m} > 0$ and $\sum_{i=1}^{m_1} \alpha_{1i} = \sum_{j=1}^{m_2} \alpha_{2j} = 1$. We now apply RPT to these inequalities. Multiplying the linear constraints with the two power cone inequalities yields $2m_1 + 2m_2$ power cone inequalities. Multiplying the right-hand sides and the left-hand sides of the two power cone inequalities yields:

$$\prod_{i=1}^{m_1} x_i^{\alpha_{1i}} \prod_{j=1}^{m_2} x_{\sigma(j)}^{\alpha_{2j}} \geq \sqrt{\sum_{i=m_1+1}^{n_x} x_i^2} \sqrt{\sum_{j=m_2+1}^{n_x} x_{\sigma(j)}^2} \iff \prod_{i=1}^{m_1} x_i^{\alpha_{1i}} \prod_{j=1}^{m_2} x_{\sigma(j)}^{\alpha_{2j}} \geq \sqrt{\sum_{i=m_1+1}^{n_x} x_i^2 \sum_{j=m_2+1}^{n_x} x_{\sigma(j)}^2} \tag{32}$$

$$\implies \prod_{i=1}^{m_1} \prod_{j=1}^{m_2} u_{i,\sigma(j)}^{\theta_{ij}} \geq \sqrt{\sum_{i=m_1+1}^{n_x} \sum_{j=m_2+1}^{n_x} u_{i,\sigma(j)}^2}, \tag{33}$$

where θ is such that

$$\sum_j \theta_{ij} = \alpha_{1i},\ \forall i, \quad \sum_i \theta_{ij} = \alpha_{2j},\ \forall j, \quad \theta_{ij} \geq 0,\ \forall i,j. \tag{34}$$

As discussed in Section 4.3, there are many choices for θ, and (34) is a special version of (2). Note that the constraints in (34) are the constraints for a transportation problem. In principle one can find many feasible $\boldsymbol{\theta}$. Two heuristics to find a feasible solution is the Vogel's Approximation method and the North-West Corner rule. Which $\boldsymbol{\theta}$ yields the tightest approximation is subject for further research. One could also use several different feasible solutions $\boldsymbol{\theta}$. One could also consider (33) as a robust constraint, where $\boldsymbol{\theta}$ is the uncertain parameter, and one could enforce that the constraint should hold for all $\boldsymbol{\theta}$ that satisfy (34). Hence, the inequality becomes

$$\prod_{i=1}^{m_1} \prod_{j=1}^{m_2} u_{i,\sigma(j)}^{\theta_{ij}} \geq \sqrt{\sum_{i=m_1+1}^{n_x} \sum_{j=m_2+1}^{n_x} u_{i,\sigma(j)}^2}, \quad \forall \theta \in \mathcal{U}, \tag{35}$$

where $\mathcal{U}$ is defined by (34). This robust inequality is concave in $\boldsymbol{\theta}$, hence one could apply the adversarial approach, or try to find a tractable robust counterpart. With respect to the iterative adversarial approach, in each iteration one has to find the worst-case value for $\boldsymbol{\theta}$.

Minimizing the left-hand side of (35) is equivalent with minimizing the ln of the left-hand side. Hence, the final problem to find the worst-case for $\boldsymbol{\theta}$ is the following transportation problem:

$$\min_{\boldsymbol{\theta}} \left\{ \sum_{i,j} \theta_{ij} \ln u_{i,\sigma(j)} \mid \sum_{j} \theta_{ij} = \alpha_{1i},\ \forall i, \quad \sum_{i} \theta_{ij} = \alpha_{2j},\ \forall j, \quad \theta_{ij} \geq 0,\ \forall i,j \right\}. \tag{36}$$

In each iteration the worst-case value for $\boldsymbol{\theta}$ is found, and the corresponding power cone inequality is added. One could also derive the tractable robust counterpart of (35); see Gorissen and Den Hertog (2015) for deriving the conjugate of a power function.

Multiplying the right-hand side and the left-hand side of the first cone inequality with itself yields another power cone inequality in a similar way. The same holds for the second cone inequality.

Finally, it follows from the nonnegativity of $x_1, \cdots, x_{m_1}$ and $x_{\sigma(1)}, \cdots, x_{\sigma(m_2)}$ that $u_{1,\sigma(1)}, \cdots, u_{m_1,\sigma(m_2)} \geq 0$.

11. (**P**) $\times$ (**E**)

Consider one power cone inequality and one exponential cone inequality

$$\begin{cases} \prod_{i=1}^{m} x_i^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} x_i^2} \\ x_1, \cdots, x_m \geq 0 \\ x_1 \geq x_2 \exp^{x_3/x_2}. \end{cases}$$

We do not know how to multiply the right-hand side of the exponential cone constraint with the power cone constraint. However, we can combine cases 3 and 4, i.e., (L)$\times$(P) and (L)$\times$(E) respectively, such that applying RPT we obtain:

$$\begin{cases} \prod_{i=1}^{m} (x_1 x_i)^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} (x_1 x_i)^2} \\ \prod_{i=1}^{m} (x_2 x_i)^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} (x_2 x_i)^2} \\ x_1 x_1 \geq x_2 x_1 \exp^{x_3 x_1 / x_2 x_1} \\ \qquad \vdots \\ x_1 x_m \geq x_2 x_m \exp^{x_3 x_m / x_2 x_m} \end{cases} \Longrightarrow \begin{cases} \prod_{i=1}^{m} u_{1i}^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} u_{1i}^2} \\ \prod_{i=1}^{m} u_{2i}^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} u_{2i}^2} \\ x_1, \cdots, x_m \geq 0 \\ u_{11} \geq u_{21} \exp^{u_{31}/u_{21}} \\ \qquad \vdots \\ u_{1m} \geq u_{2m} \exp^{u_{3m}/u_{2m}} \\ u_{21}, \cdots, u_{2m}, u_{31}, \cdots, u_{3m} \geq 0. \end{cases}$$

Note that there are many more possibilities to perspectify the left-hand sides of the above power cone inequalities. Section 4.3 describes how to find the best choice.

12. (**P**) × (**S**)

Consider one power cone inequality and one LMI

$$\begin{cases} \prod_{i=1}^{m} x_i^{\alpha_i} \geq \sqrt{\sum_{i=m+1}^{n_x} x_i^2} \\ x_1, \cdots, x_m \geq 0 \\ \boldsymbol{A}(\boldsymbol{x}) \succeq 0. \end{cases}$$

We do not know how to multiply the power cone constraint with the semidefinite cone constraint. However, we can additionally multiply the nonnegativity constraints with the semidefinite cone constraint, such that applying RPT to these inequalities we obtain:

$$x_i \boldsymbol{A}(\boldsymbol{x}) \succeq 0 \quad \Longrightarrow \quad \boldsymbol{A}(\boldsymbol{u}_i) \succeq 0, \qquad\qquad i = 1, \cdots, m.$$

13. (**E**) × (**E**)

Consider two exponential cone inequalities

$$\begin{cases} x_1 \geq x_2 \exp(x_3/x_2) \\ x_4 \geq x_5 \exp(x_6/x_5) \\ x_2, x_5 \geq 0. \end{cases}$$

We now apply RPT to these inequalities. We have four nonnegative variables x_1, x_2, x_4, x_5, and multiplying this with the two exponential cone inequalities yields eight other exponential cone inequalities. Moreover, we can multiply the left-hand sides and right-hand sides of the two inequalities with each other and with itself, and obtain three new exponential cone inequalities:

$$\begin{cases} x_1 x_4 \geq x_2 x_5 \exp(x_3 x_5/x_2 x_5 + x_6 x_2/x_2 x_5) \\ x_1^2 \geq x_2^2 \exp(2 x_3 x_2/x_2^2) \\ x_4^2 \geq x_5^2 \exp(2 x_6 x_5/x_5^2) \end{cases} \quad \Longrightarrow \quad \begin{cases} u_{14} \geq u_{25} \exp((u_{35} + u_{26})/u_{25}) \\ u_{11} \geq u_{22} \exp(2 u_{23}/u_{22}) \\ u_{44} \geq u_{55} \exp(2 u_{56}/u_{55}). \end{cases}$$

Hence, in total we obtain $8 + 3 = 11$ new exponential cone inequalities.

14. (**E**) × (**S**)

Consider one exponential cone inequality and one LMI

$$\begin{cases} x_1 \geq x_2 \exp^{x_3/x_2} \\ x_2 \geq 0 \\ \boldsymbol{A}(\boldsymbol{x}) \succeq 0. \end{cases}$$

We now apply RPT to these inequalities. Both x_1 and x_2 are nonnegative, and we multiply these inequalities with the LMI and obtain two LMIs:

$$\begin{cases} x_1 \boldsymbol{A}(\boldsymbol{x}) \succeq 0 \\ x_2 \boldsymbol{A}(\boldsymbol{x}) \succeq 0 \end{cases} \quad \Longrightarrow \quad \begin{cases} \boldsymbol{A}(\boldsymbol{u}_1) \succeq 0 \\ \boldsymbol{A}(\boldsymbol{u}_2) \succeq 0. \end{cases}$$

15. $(\mathbf{S}) \times (\mathbf{S})$

Consider two LMIs

$$\begin{cases} \boldsymbol{A}(\boldsymbol{x}) \succeq 0 \\ \boldsymbol{B}(\boldsymbol{x}) \succeq 0. \end{cases}$$

We now apply RPT to these inequalities. It follows from Lemma 2 that also the Kronecker product of $\boldsymbol{A}(\boldsymbol{x})$ and $\boldsymbol{B}(\boldsymbol{x})$ is positive semidefinite:

$$\boldsymbol{A}(\boldsymbol{x}) \otimes \boldsymbol{B}(\boldsymbol{x}) \succeq 0. \tag{37}$$

Notice that each element in the Kronecker product is the multiplication of two affine functions in $\boldsymbol{x}$. Let us denote the matrix obtained after linearization of the quadratic terms in $\boldsymbol{A}(\boldsymbol{x}) \otimes \boldsymbol{B}(\boldsymbol{x})$ by $\boldsymbol{C}(\boldsymbol{U}, \boldsymbol{x})$, which is linear in $\boldsymbol{x}$ and $\boldsymbol{U}$. This results in case 15(i) of Table 2.

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are of the same size, then it follows from Lemma 2 that

$$\boldsymbol{A}(\boldsymbol{x}) \circ \boldsymbol{B}(\boldsymbol{x}) \succeq 0. \tag{38}$$

Notice that each element in the Hadamard product is the multiplication of two affine functions in $\boldsymbol{x}$. Let us denote the matrix obtained after linearization of the quadratic terms in $\boldsymbol{A}(\boldsymbol{x}) \circ \boldsymbol{B}(\boldsymbol{x})$ by $\boldsymbol{D}(\boldsymbol{U}, \boldsymbol{x})$, which is linear in $\boldsymbol{x}$ and $\boldsymbol{U}$. This results in case 15(ii) of Table 2.

The following lemma shows that the Kronecker product relaxation is at least as tight as the Hadamard product relaxation.

LEMMA 3. *Suppose $\boldsymbol{A}$ and $\boldsymbol{B}$ are of the same size. The relaxation obtained after linearization of the quadratic terms in the Kronecker product (37) is at least as tight as the relaxation obtained after linearization of the quadratic terms in the Hadamard product (38).*

Proof. Let us denote the matrix obtained after linearization of the quadratic terms in $\boldsymbol{A}(\boldsymbol{x}) \otimes \boldsymbol{B}(\boldsymbol{x})$ by $\boldsymbol{C}(\boldsymbol{U}, \boldsymbol{x})$. Let us denote the matrix obtained after linearization of the quadratic terms in $\boldsymbol{A}(\boldsymbol{x}) \circ \boldsymbol{B}(\boldsymbol{x})$ by $\boldsymbol{D}(\boldsymbol{U}, \boldsymbol{x})$. It can easily be checked that $\boldsymbol{D}(\boldsymbol{U}, \boldsymbol{x})$ is a minor of $\boldsymbol{C}(\boldsymbol{U}, \boldsymbol{x})$ obtained by the elements in rows and columns $1, n+2, 2n+3, \ldots, n^2$. Since $C(\boldsymbol{U}, \boldsymbol{x})$ is positive semidefinite, and each minor of a positive semidefinite matrix is positive semidefinite, we have that $\boldsymbol{D}(\boldsymbol{U}, \boldsymbol{x})$ is positive semidefinite. □

Appendix C. RPT-SDP formulations of the numerical experiments

C.1. RPT-SDP formulation of Problem (DHO)

We perspectify the nonconvex part of the objective, multiply the nonnegativity constraint with itself, linearize the product terms $\boldsymbol{x}\boldsymbol{x}^\top$ by $\boldsymbol{U}$, and relax $\boldsymbol{U} = \boldsymbol{x}\boldsymbol{x}^\top$ by the LMI as described in Section 4.4 (a), such that we obtain the following RPT-SDP formulation:

$$
\begin{aligned}
\min_{\boldsymbol{x},\tau,\boldsymbol{U},\boldsymbol{h},\boldsymbol{v}} \quad & \sum_{k\in\mathcal{K}}(C+bx_k)\exp\left(\frac{\lambda Ch_k+\lambda b\sum_{i=0}^{k}U_{ik}-\delta t_k(C+bx_k)}{C+bx_k}\right)+\tau \\
\text{s.t.} \quad & \sum_{k\in\mathcal{K}}\frac{S0}{\beta_\delta}\left(\exp\left(\beta_\delta t_{k+1}\right)-\exp\left(\beta_\delta t_k\right)\right)\exp\left(-\theta h_k\right)+\frac{S_0}{\delta}\exp\left(\beta_\delta T-\theta h_K\right)\le\tau \\
& \begin{pmatrix}\boldsymbol{U} & \boldsymbol{x}\\ \boldsymbol{x}^\top & 1\end{pmatrix}\succeq\boldsymbol{0} \\
& \boldsymbol{x}\ge\boldsymbol{0} \\
& \boldsymbol{U}\ge\boldsymbol{0}.
\end{aligned}
$$

C.2. RPT-SDP formulation of Problem (CM)

Replacing the objective function with the biconjugate function in (CM) we obtain the following equivalent maximization problem

$$
\max_{\substack{\boldsymbol{x}\in\mathcal{X}\\ \boldsymbol{y}\in\mathcal{Y}}} \quad (\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b})^\top\boldsymbol{y}, \tag{CM$_B$}
$$

where we consider three cases of $\mathcal{X}$, those are $\mathcal{X}=\mathcal{X}_1$, $\mathcal{X}=\mathcal{X}_2$, and $\mathcal{X}_3$, where

$$
\begin{aligned}
\mathcal{X}_1 &= \{\boldsymbol{x}\in\mathbb{R}_+^{n_x} \mid \boldsymbol{D}^\top\boldsymbol{x}\le\boldsymbol{d}\} \\
\mathcal{X}_2 &= \left\{\boldsymbol{x}\in\mathcal{X}_1 \;\middle|\; \log\left(\sum_{i=1}^{n_y}\exp(x_i)\right)\le a\right\} \\
\mathcal{X}_3 &= \left\{\boldsymbol{x}\in\mathcal{X}_1 \;\middle|\; \|\boldsymbol{x}\|_2+\sum_{i=1}^{n_x}\sqrt{x_i}\le c\right\},
\end{aligned}
$$

and $\mathcal{Y}$ is given by

$$
\mathcal{Y}=\left\{\boldsymbol{y}\in\mathbb{R}_+^{n_y} \mid \sum_{j\in\mathcal{J}_k}y_j=1, k\in\mathcal{K}\right\}.
$$

$\boldsymbol{\mathcal{X}=\mathcal{X}_1}$. By Theorem 1 we only need to multiply the equality constraint in $\mathcal{Y}$ with $\boldsymbol{x}$ and $\boldsymbol{y}$. Next, multiplying each constraint in $\mathcal{X}_1$ with the nonnegativity constraint in $\mathcal{Y}$, pairwise multiplying the constraints in $\mathcal{X}_1$ with each other, multiplying the nonnegativity constraint in $\mathcal{Y}$ with itself, linearizing the terms $\boldsymbol{x}\boldsymbol{y}^\top$ by $\boldsymbol{U}$, $\boldsymbol{x}\boldsymbol{x}^\top$ by $\boldsymbol{X}$ and $\boldsymbol{y}\boldsymbol{y}^\top$ by $\boldsymbol{Y}$ and adding the SDP relaxation as described in Section 4.4 (a), we obtain

$$
\max_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y}} \quad \mathrm{Tr}(\boldsymbol{U}\boldsymbol{A})+\boldsymbol{b}^\top\boldsymbol{y}
$$

$$
\text{s.t.} \quad \sum_{j\in\mathcal{J}_k}\boldsymbol{U}_j-\boldsymbol{x}=\boldsymbol{0}, \qquad k\in\mathcal{K} \tag{39a}
$$

$$
\sum_{j\in\mathcal{J}_k}\boldsymbol{Y}_j-\boldsymbol{y}=\boldsymbol{0} \qquad k\in\mathcal{K} \tag{39b}
$$

$$
\boldsymbol{D}^\top\boldsymbol{U}_j-\boldsymbol{d}y_j\le\boldsymbol{0}, \qquad j\in[n_y] \tag{39c}
$$

$$
\boldsymbol{D}^\top\boldsymbol{X}_i-\boldsymbol{d}x_i\le\boldsymbol{0} \qquad i\in[n_x] \tag{39d}
$$

$$
\boldsymbol{d}\boldsymbol{x}^\top\boldsymbol{D}+\boldsymbol{D}^\top\boldsymbol{x}\boldsymbol{d}^T\le\boldsymbol{D}^\top\boldsymbol{X}\boldsymbol{D}+\boldsymbol{d}\boldsymbol{d}^\top \tag{39e}
$$

$$
\sum_{j\in\mathcal{J}_k}y_j=1 \qquad k\in\mathcal{K} \tag{39f}
$$

$$
\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y}\ge\boldsymbol{0} \tag{39g}
$$

$$
\begin{pmatrix}\boldsymbol{X} & \boldsymbol{U} & \boldsymbol{x}\\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{y}\\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & 1\end{pmatrix}\succeq\boldsymbol{0}. \tag{39h}
$$

Observe that $\boldsymbol{x}\in\mathcal{X}_1$ is redundant. The constraint $\boldsymbol{D}^\top\boldsymbol{x}\le\boldsymbol{d}$ is redundant by (39a), (39f) and (39c):

$$\boldsymbol{D}^\top\boldsymbol{x}\le\boldsymbol{d}\iff\boldsymbol{D}^\top\sum_{j\in\mathcal{J}_k}\boldsymbol{U}_j-\boldsymbol{d}\le\boldsymbol{0}\iff\boldsymbol{D}^\top\sum_{j\in\mathcal{J}_k}\boldsymbol{U}_j-\boldsymbol{d}\sum_{j\in\mathcal{J}_k}y_j\le\boldsymbol{0}\iff\sum_{j\in\mathcal{J}_k}\left(\boldsymbol{D}^\top\boldsymbol{U}_j-\boldsymbol{d}y_j\right)\le\boldsymbol{0}.$$

The nonnegativity constraint $\boldsymbol{x}\ge\boldsymbol{0}$ is redundant by (39a) and (39g). Moreover, the nonnegativity constraint $\boldsymbol{y}\ge\boldsymbol{0}$ is redundant by (39b) and (39g). Hence, these constraints are not included in the above formulation.

$\mathcal{X}=\mathcal{X}_2$. From Section 5.2 it follows that first reformulating the constraints in terms of the five basic cones, and then apply RPT might be a better option than first applying RPT and then reformulate the results in terms of the five basic cones. Hence we consider the following equivalent form of $\mathcal{X}_2$:

$$\mathcal{X}_2=\left\{\boldsymbol{x}\in\mathcal{X}_1,\boldsymbol{z}\in\mathbb{R}^{n_x}\,\middle|\,z_i\ge\exp\left(x_i-a\right),\sum_{i=1}^{n_x}z_i\le 1\right\}.$$

We can expand the formulation for $\mathcal{X}=\mathcal{X}_1$. Next to multiplying the equality constraint in $\mathcal{Y}$ with $\boldsymbol{x}$ and $\boldsymbol{y}$ we now also need to multiply the equality constraint with $\boldsymbol{z}$. Next, multiplying each constraint in $\mathcal{X}_2\setminus\mathcal{X}_1$ with the nonnegativity constraint in $\mathcal{Y}$, pairwise multiplying the constraints in $\mathcal{X}_1$ with the constraints in $\mathcal{X}_2\setminus\mathcal{X}_1$, pairwise multiplying the constraints in $\mathcal{X}_2\setminus\mathcal{X}_1$ with each other, where for the multiplication of the exponential cone constraint with itself we use the result in Appendix B, case 13, linearizing the extra terms $\boldsymbol{x}\boldsymbol{y}^\top$ by $\boldsymbol{U}$, $\boldsymbol{x}\boldsymbol{z}^\top$ by $\boldsymbol{V}$, $\boldsymbol{y}\boldsymbol{z}^\top$ by $\boldsymbol{W}$, $\boldsymbol{x}\boldsymbol{x}^\top$ by $\boldsymbol{X}$, $\boldsymbol{y}\boldsymbol{y}^\top$ by $\boldsymbol{Y}$ and $\boldsymbol{z}\boldsymbol{z}^\top$ by $\boldsymbol{Z}$, where $\boldsymbol{X},\boldsymbol{Y}$ and $\boldsymbol{Z}$ are symmetric matrices, and replacing the LMI in (39h) by the LMI including also the newly introduced variables, we obtain

$$
\begin{aligned}
\max_{\substack{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\\ \boldsymbol{U},\boldsymbol{V},\boldsymbol{W}\\ \boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}}}\quad & \mathrm{Tr}(\boldsymbol{U}\boldsymbol{A})+\boldsymbol{b}^\top\boldsymbol{y} \\
\text{s.t.}\quad & (39a)-(39g) \\
& \sum_{i=1}^{n_x} z_i\le 1 && && (40a)\\
& z_i\ge\exp(x_i-a) && i\in[n_x] && (40b)\\
& \sum_{j\in\mathcal{J}_k} W_{ji}-z_i=0 && i\in[n_x],k\in\mathcal{K} && (40c)\\
& \boldsymbol{W}\ge\boldsymbol{0} && && (40d)\\
& \sum_{i=1}^{n_x}(\boldsymbol{W}^\top)_i-\boldsymbol{y}\le 0 && && (40e)\\
& W_{ji}\ge y_j\exp\left(\frac{U_{ij}-y_ja}{y_j}\right) && i\in[n_x],j\in[n_y] && (40f)\\
& \sum_{i=1}^{n_x}\boldsymbol{V}_i-\boldsymbol{x}\le\boldsymbol{0} && && (40g)\\
& x_{i'}\exp\left(\frac{X_{ii'}-ax_{i'}}{x_{i'}}\right)-V_{i'i}\le 0 && i,i'\in[n_x] && (40h)\\
& \boldsymbol{D}^\top\sum_{i=1}^{n_x}\boldsymbol{V}_i-\boldsymbol{D}^\top\boldsymbol{x}-\boldsymbol{d}\sum_{i=1}^{n_x}z_i+\boldsymbol{d}\ge\boldsymbol{0} && && (40i)\\
& (d_\ell-\boldsymbol{D}_\ell^\top\boldsymbol{x})\exp\left(\frac{d_\ell x_i-\boldsymbol{D}_\ell^\top\boldsymbol{X}_i-a(d_\ell-\boldsymbol{D}_\ell^\top\boldsymbol{x})}{d_\ell-\boldsymbol{D}_\ell^\top\boldsymbol{x}}\right)-d_\ell z_i+\boldsymbol{D}_\ell^\top\boldsymbol{V}_i\le 0 && i\in[n_x],\ell\in\mathcal{L} && (40j)
\end{aligned}
$$

$$\sum_{i=1}^{n_x}\sum_{i'=1}^{n_x} Z_{ii'} - 2\sum_{i=1}^{n_x} z_i + 1 \geq 0 \tag{40k}$$

$$\sum_{i'=1}^{n_x} Z_{ii'} - z_i + \left(1 - \sum_{i'=1}^{n_x} z_{i'}\right) \exp\left(\frac{x_i - \sum_{i'=1}^{n_x} V_{ii'} - a + a\sum_{i'=1}^{n_x} z_{i'}}{1 - \sum_{i'=1}^{n_x} z_{i'}}\right) \leq 0 \qquad i \in [n_x] \tag{40l}$$

$$Z_{ii'} \geq \exp\left(x_i + x_{i'} - 2a\right) \qquad i \leq i' \in [n_x] \tag{40m}$$

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} & \boldsymbol{V} & \boldsymbol{x} \\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{W} & \boldsymbol{y} \\ \boldsymbol{V}^\top & \boldsymbol{W}^\top & \boldsymbol{Z} & \boldsymbol{z} \\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & \boldsymbol{z}^\top & 1 \end{pmatrix} \succeq \boldsymbol{0}. \tag{40n}$$

$\boldsymbol{\mathcal{X} = \mathcal{X}_3}$**.** We reformulate the non-convex constraint via the biconjugate and subsequently linearize the product terms. Next, we use an epigraph variable for the convex part of the new relaxed constraint, such that we obtain the following relaxed set of constraints

$$\mathcal{X}_3^* = \left\{ \boldsymbol{x} \in \mathcal{X}_1, \boldsymbol{V} \in \mathbb{R}^{n_x \times n_x}, \boldsymbol{z} \in \mathbb{R}^{n_x}_{++}, s \in \mathbb{R} \;\middle|\; \begin{array}{l} s + \sum_{i=1}^{n_x} V_{ii} \leq c \\ \left\|\boldsymbol{x}\right\|_2 + \sum_{i=1}^{n_x} \frac{1}{4z_i} \leq s \end{array} \right\}.$$

From Example 2 in Section 5.2, it follows that first reformulating the constraints in terms of the five basic cones, and then apply RPT might be a better option than first applying RPT and then reformulate the results in terms of the five basic cones. Hence we consider the following equivalent form of $\mathcal{X}_3^*$:

$$\mathcal{X}_3^* = \left\{ \boldsymbol{x} \in \mathcal{X}_1, \boldsymbol{V} \in \mathbb{R}^{n_x \times n_x}, s \in \mathbb{R}, \boldsymbol{t} \in \mathbb{R}^{n_x}_{+}, \boldsymbol{z} \in \mathbb{R}^{n_x}_{+} \;\middle|\; \begin{array}{l} s + \sum_{i=1}^{n_x} V_{ii} \leq c \\ \left\|\boldsymbol{x}\right\|_2 + \sum_{i=1}^{n_x} t_i \leq s \\ \left\|(z_i - t_i, 1)^\top\right\|_2 \leq z_i + t_i, \quad i \in [n_x] \end{array} \right\}.$$

We choose c to be large enough such that (CM_B) with $\mathcal{X} = \mathcal{X}_3$ satisfies Assumption 3. We can expand the formulation for $\mathcal{X} = \mathcal{X}_1$. We do not multiply the linear constraint in $\mathcal{X}_3$ with the other constraints, since this will increase the number of variables enormously. However, we do pairwise multiply the other constraints in $\mathcal{X}_3^*$ with all other existing constraints to obtain valuable bounds on the newly introduced variables resulting from the perspectification step.

Next to multiplying the equality constraint in $\mathcal{Y}$ with $\boldsymbol{x}$ and $\boldsymbol{y}$ we now also need to multiply the equality constraint with s, $\boldsymbol{t}$ and $\boldsymbol{z}$. Next, except for the linear constraint in $\mathcal{X}_3^*$, multiplying the constraints in $\mathcal{X}_3^* \setminus \mathcal{X}_1$ with the nonnegativity constraint in $\mathcal{Y}$, pairwise multiplying the constraints in $\mathcal{X}_3^* \setminus \mathcal{X}_1$ with the constraints in $\mathcal{X}_1$ and $\mathcal{X}_3^* \setminus \mathcal{X}_1$, where for the pairwise multiplication of the second order cone constraints with each other and itself we use the result in Appendix B, case 6, linearizing the extra terms $\boldsymbol{x}\boldsymbol{x}^\top$ by $\boldsymbol{X}$, $\boldsymbol{x}\boldsymbol{y}^\top$ by $\boldsymbol{U}$, $\boldsymbol{x}\boldsymbol{z}^\top$ by $\boldsymbol{V}$, $\boldsymbol{y}\boldsymbol{y}^\top$ by $\boldsymbol{Y}$, $\boldsymbol{y}\boldsymbol{z}^\top$ by $\boldsymbol{W}$, $\boldsymbol{z}\boldsymbol{z}^\top$ by $\boldsymbol{Z}$,

$\boldsymbol{x}\boldsymbol{t}^\top$ by $\boldsymbol{Q}$, $\boldsymbol{y}\boldsymbol{t}^\top$ by $\boldsymbol{R}$, $\boldsymbol{z}\boldsymbol{t}^\top$ by $\boldsymbol{P}$, $\boldsymbol{t}\boldsymbol{t}^\top$ by $\boldsymbol{T}$ – where the matrices $\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}$ and $\boldsymbol{T}$ are symmetric – $s\boldsymbol{x}$ by $\boldsymbol{\alpha}$, $s\boldsymbol{y}$ by $\boldsymbol{\gamma}$, $s\boldsymbol{z}$ by $\boldsymbol{\theta}$, $s\boldsymbol{t}$ by $\boldsymbol{\lambda}$, ss by β and replacing the LMI in (39h) by the LMI including also the newly introduced variables, we obtain

$$
\begin{aligned}
\max_{\substack{\boldsymbol{\alpha},\beta,\boldsymbol{\gamma},\boldsymbol{\theta},\boldsymbol{\lambda},\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\\ \boldsymbol{Q},\boldsymbol{R},\boldsymbol{T},\boldsymbol{U},\boldsymbol{V}\\ \boldsymbol{W},\boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}}} \quad & \mathrm{Tr}(\boldsymbol{U}\boldsymbol{A}) + \boldsymbol{b}^\top \boldsymbol{y} \\
\text{s.t.} \quad & (39a) - (39g) \\
& s + \sum_{i=1}^{n_x} V_{ii} \le c && && (41a) \\
& \|\boldsymbol{x}\|_2 + \sum_{i=1}^{n_x} t_i \le s && && (41b) \\
& \left\|(z_i - t_i, 1)^\top\right\|_2 \le z_i + t_i, && i \in [n_x] && (41c) \\
& \sum_{j\in\mathcal{J}_k} \gamma_j - s = 0 && i \in [n_x], k \in \mathcal{K} && (41d) \\
& \sum_{j\in\mathcal{J}_k} R_{ji} - t_i = 0 && i \in [n_x], k \in \mathcal{K} && (41e) \\
& \sum_{j\in\mathcal{J}_k} W_{ji} - z_i = 0 && i \in [n_x], k \in \mathcal{K} && (41f) \\
& \|\boldsymbol{U}_j\|_2 + \sum_{i=1}^{n_x} R_{ji} \le \gamma_j && j \in [n_y] && (41g) \\
& \left\|(W_{ji} - R_{ji}, y_j)^\top\right\|_2 \le W_{ji} + R_{ji} && i \in [n_x], j \in [n_y] && (41h) \\
& \|\boldsymbol{X}_i\|_2 + \sum_{i'=1}^{n_x} Q_{ii'} \le \alpha_i && i \in [n_x] && (41i) \\
& \left\|(V_{ii'} - Q_{ii'}, x_{i'})^\top\right\|_2 \le V_{ii'} + Q_{ii'} && i, i' \in [n_x] && (41j) \\
& \boldsymbol{D}^\top \boldsymbol{V}_i - \boldsymbol{d} z_i \le 0 && i \in [n_x] && (41k) \\
& \boldsymbol{D}^\top \boldsymbol{Q}_i - \boldsymbol{d} t_i \le 0 && i \in [n_x] && (41l) \\
& \left\|d_\ell \boldsymbol{x} - \boldsymbol{D}_\ell^\top \boldsymbol{X}\right\|_2 + d_\ell \sum_{i=1}^{n_x} t_i - \boldsymbol{D}_\ell^\top \sum_{i=1}^{n_x} \boldsymbol{Q}_i \le d_\ell s - \boldsymbol{D}_\ell^\top \boldsymbol{\alpha} && \ell \in \mathcal{L} && (41m) \\
& \left\|\left(d_\ell (z_i - t_i) + \boldsymbol{D}_\ell^\top (Q_i - V_i), d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}\right)\right\|_2 \le d_\ell (z_i + t_i) - \boldsymbol{D}_\ell^\top (Q_i + V_i) && i \in [n_x], \ell \in \mathcal{L} && (41n) \\
& \|\boldsymbol{V}_i\|_2 + \sum_{i'=1}^{n_x} P_{ii'} \le \theta_i && i \in [n_x] && (41o) \\
& \|\boldsymbol{Q}_i\|_2 + \sum_{i'=1}^{n_x} T_{ii'} \le \lambda_i && i \in [n_x] && (41p) \\
& \|\boldsymbol{\alpha}\|_2 + \sum_{i=1}^{n_x} \lambda_i \le \beta && && (41q) \\
& \|\boldsymbol{X}\|_2 + 2\sum_{i=1}^{n_x} \|\boldsymbol{Q}_i\|_2 + \sum_{i=1}^{n_x} \sum_{i=1'}^{n_x} T_{ii'} \le \beta && && (41r) \\
& \left\|\left(\boldsymbol{V}_i - \boldsymbol{Q}_i,\ \boldsymbol{x}\right)\right\|_2 + \sum_{i'=1}^{n_x} \left\|(P_{ii'} - T_{ii'}, t_{i'})^\top\right\|_2 \le \theta_i + \lambda_i && i \in [n_x] && (41s)
\end{aligned}
$$

$$\|\boldsymbol{V}_i+\boldsymbol{Q}_i\|_2+\sum_{i'=1}^{n_x}(P_{ii'}+T_{ii'})\le\theta_i+\lambda_i \qquad i\in[n_x] \tag{41t}$$

$$\|(\theta_i-\lambda_i,s)\|_2\le\theta_i+\lambda_i \qquad i\in[n_x] \tag{41u}$$

$$\|(Z_{ii'}-P_{ii'},z_{i'})\|_2\le Z_{ii'}+P_{ii'} \qquad i,i'\in[n_x] \tag{41v}$$

$$\|(P_{ii'}-T_{ii'},t_{i'})\|_2\le P_{ii'}+T_{ii'} \qquad i,i'\in[n_x] \tag{41w}$$

$$\|(Z_{ii'}-T_{ii'},z_{i'}+t_{i'})\|\le Z_{ii'}+P_{ii'}+P_{i'i}+T_{ii'} \qquad i,i'\in[n_x] \tag{41x}$$

$$\left\|\begin{pmatrix}Z_{ii'}-P_{ii'}-P_{i'i}+T_{ii'} & z_i-t_i\\ z_{i'}-t_{i'} & 1\end{pmatrix}\right\|_2\le Z_{ii'}+P_{ii'}+P_{i'i}+T_{ii'} \qquad i,i'\in[n_x] \tag{41y}$$

$$\boldsymbol{P},\boldsymbol{Q},\boldsymbol{R},\boldsymbol{T},\boldsymbol{V},\boldsymbol{W}\ge 0 \tag{41z}$$

$$\boldsymbol{z}>0 \tag{43aa}$$

$$\begin{pmatrix}\boldsymbol{X} & \boldsymbol{U} & \boldsymbol{V} & \boldsymbol{Q} & \boldsymbol{\alpha} & \boldsymbol{x}\\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{W} & \boldsymbol{R} & \boldsymbol{\gamma} & \boldsymbol{y}\\ \boldsymbol{V}^\top & \boldsymbol{W}^\top & \boldsymbol{Z} & \boldsymbol{P} & \boldsymbol{\theta} & \boldsymbol{z}\\ \boldsymbol{Q}^\top & \boldsymbol{R}^\top & \boldsymbol{P}^\top & \boldsymbol{T} & \boldsymbol{\lambda} & \boldsymbol{t}\\ \boldsymbol{\alpha}^\top & \boldsymbol{\gamma}^\top & \boldsymbol{\theta}^\top & \boldsymbol{\lambda}^\top & \beta & s\\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & \boldsymbol{z}^\top & \boldsymbol{t}^\top & s & 1\end{pmatrix}\succeq 0 \tag{43bb}$$

Appendix D: Mountain climbing procedure

We use a mountain climbing procedure based on the algorithm from Tao and An (1997), to find a lower bound for (CM).

The input of the MC procedure for $\mathcal{X}=\mathcal{X}_1$ and $\mathcal{X}=\mathcal{X}_2$ is a finite set $\mathcal{X}'=\{\boldsymbol{x}_{ub},\boldsymbol{x}_1,\dots,\boldsymbol{x}_{n_y}\}$, where

$$\boldsymbol{x}_j=\begin{cases}\frac{\boldsymbol{U}_j}{(y_{ub})_j} & \text{if } (y_{ub})_j\ge 0\\ \boldsymbol{x}_{ub} & \text{if } (y_{ub})_j=0\end{cases},\quad \text{for all } j\in[n_y],$$

and $\boldsymbol{x}_{ub}$, $\boldsymbol{y}_{ub}$, and $\boldsymbol{U}$ are part of the solution obtained from the RPT relaxation. Observe that for $\mathcal{X}=\mathcal{X}_1$ and $\mathcal{X}=\mathcal{X}_2$ we have $\mathcal{X}'\subseteq\mathcal{X}$. The MC procedure for $\mathcal{X}=\mathcal{X}_1$ and $\mathcal{X}=\mathcal{X}_2$ is given by Algorithm 1.

For the mountain climbing procedure for $\mathcal{X}=\mathcal{X}_3$ we consider the equivalent biconjugate reformulation of $\mathcal{X}_3$, i.e.,

$$\widehat{\mathcal{X}}_3=\left\{\boldsymbol{x}\in\mathcal{X}_1,\boldsymbol{z}\in\mathbb{R}^{n_x}_{++},s\in\mathbb{R}\;\middle|\;\begin{array}{l}s+\boldsymbol{x}^\top\boldsymbol{z}\le c\\ \|\boldsymbol{x}\|_2+\sum_{i=1}^{n_x}\frac{1}{4z_i}\le s\end{array}\right\}.$$

The mountain climbing procedure for $\mathcal{X}=\mathcal{X}_3$ follows the same steps as in Algorithm 1, with a few differences, due to the nonconvexity in the original problem formulation.

The input is a finite set $\mathcal{X}_3'$ constructed in the following way. We first construct the set $\mathcal{X}_3''=\{\boldsymbol{x}_{ub},\boldsymbol{x}_1^z,\dots,\boldsymbol{x}_{n_x}^z,\boldsymbol{x}_1^y,\dots,\boldsymbol{x}_{n_y}^y\}$ of candidate solutions for (CM), where

$$\boldsymbol{x}_i^z=\begin{cases}\frac{\boldsymbol{V}_i}{(z_{ub})_i} & \text{if } (z_{ub})_i\ge 0\\ \boldsymbol{x}_{ub} & \text{if } (z_{ub})_i=0,\end{cases}\quad\text{and}\quad \boldsymbol{x}_j^y=\begin{cases}\frac{\boldsymbol{U}_j}{(y_{ub})_j} & \text{if } (y_{ub})_j\ge 0\\ \boldsymbol{x}_{ub} & \text{if } (y_{ub})_j=0,\end{cases}\quad\text{for all } i\in[n_z], j\in[n_y],$$

Algorithm 1 Mountain climbing procedure

Input: $\mathcal{X}'$, $\mathcal{L} = \emptyset$

for $\boldsymbol{x} \in \mathcal{X}'$ **do**
$\quad \boldsymbol{y} \leftarrow \arg\max_{\boldsymbol{y} \in \mathcal{Y}} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}^\top \boldsymbol{y}$
$\quad \varepsilon \leftarrow 1$
$\quad$ **while** $\varepsilon > 0.001$ **do**
$\quad\quad \text{Lb} \leftarrow \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}^\top \boldsymbol{y}$
$\quad\quad \boldsymbol{x} \leftarrow \arg\max_{\boldsymbol{x} \in \mathcal{X}} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}^\top \boldsymbol{y}$
$\quad\quad \boldsymbol{y} \leftarrow \arg\max_{\boldsymbol{y} \in \mathcal{Y}} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}^\top \boldsymbol{y}$
$\quad\quad \text{Lb}_{\boldsymbol{x}} \leftarrow \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}^\top \boldsymbol{y}$
$\quad\quad \varepsilon \leftarrow \text{Lb}_{\boldsymbol{x}} - \text{Lb}$
$\quad$ **end while**
$\quad \mathcal{L} \leftarrow \mathcal{L} \cup \{(\boldsymbol{x}, \boldsymbol{y})\}$
end for
$(\boldsymbol{x}^\star, \boldsymbol{y}^\star) \leftarrow \arg\max_{(\boldsymbol{x},\boldsymbol{y}) \in \mathcal{L}} \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}^\top \boldsymbol{y}$
$\text{Lb}^\star = (\boldsymbol{x}^\star)^\top \boldsymbol{A}\boldsymbol{y}^\star + \boldsymbol{b}^\top \boldsymbol{y}^\star$
return $(\text{Lb}^\star, \boldsymbol{x}^\star, \boldsymbol{y}^\star)$

and $\boldsymbol{x}_{ub}$, $\boldsymbol{y}_{ub}$, $\boldsymbol{U}$, and $\boldsymbol{V}$ are part of the solution from the RPT relaxation. Note that $\mathcal{X}_3'' \not\subseteq \mathcal{X}_3$, since the solution from the RPT relaxation might not be feasible for (CM). Therefore, we select the set of candidate solutions in $\mathcal{X}_3''$ that satisfy the nonconvex constraint in $\mathcal{X}_3$, such that we obtain the finite set

$$\mathcal{X}_3' = \left\{ \boldsymbol{x} \in \mathcal{X}_3'' \;\middle|\; \|\boldsymbol{x}\|_2 + \sum_{i=1}^{n_x} \sqrt{x_i} \leq c \right\}.$$

Observe that it is possible that $\mathcal{X}_3' = \emptyset$. In this case, we cannot find a feasible solution.

Moreover, instead of fixing $\boldsymbol{x}$ and maximizing for $\boldsymbol{y}$ and vice versa, we now first fix $\boldsymbol{x}$ and maximize for $\boldsymbol{y}$ and $\boldsymbol{z}$, and then maximize for $\boldsymbol{x}$, in which $\boldsymbol{y}$ and $\boldsymbol{z}$ are fixed.

Appendix E: Data generation of numerical experiments

E.1. Data generation of numerical experiment of Section 8.1

We take the data from Eijgenraam et al. (2017). The values for the parameters for each distinct dike ring are given in Table 8. Here, α is the exponential distribution parameter for extreme water levels (1/cm), ζ is the increase of loss per centimeter dike heightening (1/cm), η is the structural increase of the water level in cm per year, S_0 is the expected loss at time $t = 0$, γ is the rate of growth per year of wealth in dike ring, δ is the discount rate, T is the time horizon and b, c and λ

are positive constants. Moreover, $\theta = \alpha - \zeta$, β denotes the growth rate of the damage costs and is given by $\beta = \alpha\eta + \gamma$, and $\beta_\delta = \beta - \delta$.

No.	α	b	c	λ	ζ	η	S_0	γ	δ	T
10	0.033027	0.6258	16.6939	0.0014	0.003774	0.32	$\frac{1564.9}{2270}$	0.02	0.04	300
15	0.0502	1.1268	125.6422	0.0098	0.003764	0.76	$\frac{11810.4}{729}$	0.02	0.04	300
16	0.0574	2.1304	324.6287	0.01	0.002032	0.76	$\frac{22656.5}{906}$	0.02	0.04	300

Table 8 Values for the parameters for each distinct dike ring considered.

E.2. Data generation of numerical experiment of Section 8.2

We use the data generated by Selvi et al. (2020, Appendix F.5). In every problem, every max-term has the same number of elements, i.e., $|\mathcal{J}_k| = |\mathcal{J}_{k'}|$ for every $k, k' \in \mathcal{K}$. We denote the problems in which $\mathcal{X} = \mathcal{X}_2$. For $\mathcal{X} = \mathcal{X}_1$ and $\mathcal{X} = \mathcal{X}_3$ we take the same problem but without the geometric constraint and with the geometric constraint replaced by the constraint $\|\boldsymbol{x}\|_2 + \sum_{i=1}^{n_x} \sqrt{x_i} \leq c$, respectively.

Problem instances 1, 2 and 3 are defined by:

$$\begin{array}{lll} \max\limits_{\boldsymbol{x}} & \sum\limits_{k\in\mathcal{K}} \max\limits_{j\in\mathcal{J}_k}\{\boldsymbol{A}_j \boldsymbol{x}\} & \\ \text{s.t.} & x_i \leq \frac{n}{i} & i \in [n_x] \\ & \ln\left(\sum\limits_{i=1}^{n_x} \exp(x_i)\right) \leq a, & \end{array}$$

where $A_{ij} \sim [-5, 5]$.

Problem instances 7 and 11 are defined by:

$$\begin{array}{ll} \max\limits_{\boldsymbol{x}} & \sum\limits_{k\in\mathcal{K}} \max\limits_{j\in\mathcal{J}_k}\{\boldsymbol{A}_j \boldsymbol{x} + b_j\} \\ \text{s.t.} & \boldsymbol{D}^\top \boldsymbol{x} \leq \boldsymbol{d} \\ & \ln\left(\sum\limits_{i=1}^{n_x} \exp(x_i)\right) \leq a, \end{array}$$

where:

Problem instance 7: $A_{ij} \sim [-5,5], b_j \sim [-10,10], D_{ij} \sim [0,1], d_i \sim [5,15]$

Problem instance 11: $A_{ij} \sim [-5,10], b_j \sim [-10,10]$, and $\boldsymbol{D}$ and $\boldsymbol{d}$ are given by:

$$\boldsymbol{D} = \begin{bmatrix} -3 & 7 & 0 & -5 & 1 & 1 & 0 & 2 & -1 & 1 \\ 7 & 0 & -5 & 1 & 1 & 0 & 2 & -1 & -1 & 1 \\ 0 & -5 & 1 & 1 & 0 & 2 & -1 & -1 & -9 & 1 \\ -5 & 1 & 1 & 0 & 2 & -1 & -1 & -9 & 3 & 1 \\ 1 & 1 & 0 & 2 & -1 & -1 & -9 & 3 & 5 & 1 \\ 1 & 0 & 2 & -1 & -1 & -9 & 3 & 5 & 0 & 1 \\ 0 & 2 & -1 & -1 & -9 & 3 & 5 & 0 & 0 & 1 \\ 2 & -1 & -1 & -9 & 3 & 5 & 0 & 0 & 1 & 1 \\ -1 & -1 & -9 & 3 & 5 & 0 & 0 & 1 & 7 & 1 \\ -1 & -9 & 3 & 5 & 0 & 0 & 1 & 7 & -7 & 1 \\ -9 & 3 & 5 & 0 & 0 & 1 & 7 & -7 & -4 & 1 \\ 3 & 5 & 0 & 0 & 1 & 7 & -7 & -4 & -6 & 1 \\ 5 & 0 & 0 & 1 & 7 & -7 & -4 & -6 & -3 & 1 \\ 0 & 0 & 1 & 7 & -7 & -4 & -6 & -3 & 7 & 1 \\ 0 & 1 & 7 & -7 & -4 & -6 & -3 & 7 & 0 & 1 \\ 1 & 7 & -7 & -4 & -6 & -3 & 7 & 0 & -5 & 1 \\ 7 & -7 & -4 & -6 & -3 & 7 & 0 & -5 & 1 & 1 \\ -7 & -4 & -6 & -3 & 7 & 0 & -5 & 1 & 1 & 1 \\ -4 & -6 & -3 & 7 & 0 & -5 & 1 & 1 & 0 & 1 \\ -6 & -3 & 7 & 0 & -5 & 1 & 1 & 0 & 2 & 1 \end{bmatrix}, \quad \text{and} \quad \boldsymbol{d} = \begin{bmatrix} -5 \\ 2 \\ -1 \\ -3 \\ 5 \\ 4 \\ -1 \\ 0 \\ 9 \\ 40 \end{bmatrix},$$

respectively. Moreover, the values for the parameters of each distinct problem are given in Table 9.

	#1	#2	#3	#7	#11
n_x	5	5	20	10	20
$\lvert\mathcal{K}\rvert$	1	10	10	2	10
$\lvert\mathcal{J}_k\rvert$	5	5	10	5	10
a	3	3	11	3	5
c	6	6	25	7	30
M	100	100	1000	1000	1000

Table 9 Values for the parameters for each distinct problem instance considered.