A novel algorithm for a broad class of nonconvex optimization problems

Dimitris Bertsimas
Operations Research Center, Massachusetts Institute of Technology, United States, d.bertsim@mit.edu

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

Thodoris Koukouvinos
Operations Research Center, Massachusetts Institute of Technology, United States, tkoukouv@mit.edu

Jianzhe Zhen*
ETH Zürich, Zürich, Switzerland, trevorzhen@gmail.com

In this paper, we propose a new global optimization approach for solving nonconvex optimization problems in which the nonconvex components are sums of products of convex functions. A broad class of nonconvex problems can be formulated in this way, such as concave minimization problems, problems with difference of convex functions in the objective and constraints, and fractional optimization problems. Our approach leverages two techniques: first, we introduce a new technique, called the Reformulation-Perspectification Technique (RPT), to obtain a convex approximation of the considered nonconvex continuous optimization problem. Next, we employ a spatial Branch and Bound scheme, utilizing RPT, to obtain a global optimal solution. Numerical experiments on three different convex maximization problems, as well as a quadratic constrained quadratic optimization problem, and a dike height optimization problem demonstrate the effectiveness of the proposed approach. In particular, our approach solves more instances to global optimality for the considered problems than BARON and SCIP. Moreover, for large-dimensional problem instances, our approach outperforms both BARON and SCIP in computation time for most cases, while for smaller dimensions, BARON overall performs better in terms of computation time.

Key words: Reformulation-Linearization Technique, perspective function, nonconvex optimization, conjugate function, branch and bound.

1. Introduction

We introduce a novel global optimization approach for nonconvex optimization problems where the nonconvex elements are sums of products of convex functions. This formulation covers a wide

* Corresponding author.

range of optimization problems, including nonconvex quadratic optimization, mixed binary linear optimization, concave minimization, difference of convex programming, and fractional optimization.

For nonconvex quadratic optimization problems and mixed binary linear optimization problems, hierarchical convex approximations can be obtained from the Reformulation-Linearization Technique (RLT) (Sherali and Adams, 1999). 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, 2019). RLT is also applicable to mixed binary polynomial and to continuous nonconvex polynomial optimization problems (Sherali and Adams, 1999), and has been extended to mixed binary semi-infinite and convex optimization problems (Sherali and Adams, 2009). 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 continuous variable. We also refer to Jiang and Li (2020) for an overview of RLT approximations for quadratic optimization problems.

We propose an extension of RLT, which we call Reformulation-Perspectification Technique (RPT), to obtain a convex relaxation of the original nonconvex optimization problem. RPT consists of a reformulation and a perspectification step. Similarly to RLT, the reformulation step of RPT generates redundant nonconvex constraints from pairwise multiplication of the existing inequalities. Where in RLT only multiplications of linear or quadratic inequalities are considered, RPT also considers pairwise multiplications of not necessarily linear or quadratic convex inequalities, thereby obtaining tighter approximations than RLT based methods. 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. Hence, RPT can handle more types of nonconvexity than RLT based methods.

Moreover, in this paper we use a spatial branch and bound scheme, leveraging RPT and the eigenvector branching strategy proposed in Anstreicher (2022), to obtain a global optimal solution of the original nonconvex problem. A branch and bound algorithm was first introduced by Falk and Soland (1969), addressing optimization problems with continuous nonconvex separable objectives, and extended by Horst (1976) to non-separable functions, leveraging a different partitioning rule. In the context of nonconvex quadratically constrained quadratic problems (QCQPs), Al-Khayyal et al. (1995) develop a method for solving nonconvex QCQPs based on branch and bound, leveraging a linearization technique. Chen and Burer (2012) develop a branch and bound method, utilizing co-positive programming, addressing nonconvex quadratic problems over linear constraints. Anstreicher (2022) develops a spatial branching method termed eigenvector branching in order to strengthen semidefinite relaxations of problems with a nonconvex quadratic objective and/or constraints, leveraging the multiplication of linear with conic quadratic inequalities. In the context of nonlinear problems (NLPs) and mixed integer nonlinear problems (MINLPs), Ryoo and Sahinidis (1996) propose the branch-and-reduce algorithm, which is implemented in BARON (Sahinidis, 1996). The latter uses a branch and bound algorithm, that iteratively solves convex relaxations of the initial problem and finds tighter variable bounds. BARON has been very successful so far and is in fact considered a state-of-the-art method for nonconvex optimization problems. Another state-of-the-art method that utilizes branch and bound, addressing nonconvex problems with an emphasis on integer problems, is the global optimization algorithm SCIP, developed by Achterberg (2009).

Although the idea of using branch and bound to obtain the global optimal solution of nonconvex optimization problems has thus been already present, in this paper we deviate from previous works in the implementation of it. Namely, we propose a spatial branch and bound scheme that leverages RPT and the eigenvector branching strategy proposed in Anstreicher (2022), which we refer to as RPT-BB, standing for Reformulation-Perspectification Technique - Branch and Bound.

For the problem of maximizing a twice continuously differentiable convex function over a convex compact feasible region, Selvi et al. (2022) develop an algorithm based on adjustable robust optimization, which can be shown to be a special case of our RPT approach. Additionally, Ben-Tal and Roos (2022) develop an algorithm called CoMax, which is based on gradient ascent. The latter is also applicable to integer optimization problems where the feasible set is a polytope. While both methods can find high-quality bounds on the optimal solution, neither can guarantee global optimality. Specifically, Selvi et al. (2022) is limited to problems with only one norm constraint or multiple linear constraints, while CoMax only obtains a lower bound for the optimal solution and imposes strict assumptions on the constraints.

Our main contributions can be summarized as follows:

1. We extend the existing RLT approach to a broader class of nonconvex optimization problems, namely optimization problems in which the nonconvex components are sums of products of convex functions. The proposed RPT approach can handle multiplication of constraints that are neither linear nor quadratic, and thereby obtains tighter approximations than RLT. Moreover, it can also handle more types of nonconvexity than RLT.
2. We introduce a new global optimization approach, by incorporating the RPT framework within branch and bound, using the eigenvector branching strategy of Anstreicher (2022). We show that the proposed RPT-BB approach converges to the optimal solution of the original nonconvex optimization problems in which the nonconvex components are sums of products of convex functions.
3. We provide several theoretical insights for our approach. We show that using epigraphical variables for the nonlinear convex components in the nonconvex objective and/or constraints,

yields an RPT relaxation that is at least as tight as without the introduction of epigraphical variables. Further, we show that adding linear constraints that are redundant to existing linear constraints does not tighten the RPT relaxation, while adding linear constraints that are redundant to existing nonlinear constraints can be useful.

4. We demonstrate the effectiveness of the proposed RPT-BB approach, by conducting numerical experiments on three different convex maximization problems, a quadratic constraint quadratic optimization problem, and a dike height optimization problem. We show that our approach solves more instances to global optimality for the considered problems than BARON and SCIP. Moreover, for the larger problem instances our approach overall performs better on computation time than both BARON and SCIP, while for smaller dimension BARON overall outperforms our approach and SCIP on computation time. In addition, for a convex maximization problem that allows for a mixed integer reformulation, we show that RPT-BB outperforms MOSEK for most problem instances when considering a nonlinear convex feasible region. Finally, for the quadratic constraint quadratic optimization problem with convex feasible region, we show that both RPT-BB and CPLEX are comparable on computation time.

This paper is structured as follows: In Section 2, we describe the generic nonconvex optimization problem we consider. In Section 3, we describe the RPT-BB approach to obtain a global optimal solution of the considered nonconvex optimization problem. In Section 4, we demonstrate the RPT-BB approach on the basis of a simple example. In Section 5, we present several additional ways to strenghten the RPT-BB approach. In Section 6, we present the convergence analysis. In Section 7, we asses the numerical performance of the approach. We end the paper by a short discussion and conclude our findings in Section 8.

Notation. We generally use bold faced characters such as $\boldsymbol{a} \in \mathbb{R}^n$ and $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ to represent vectors and matrices, respectively, a_i to denote the i-th element of the vector $\boldsymbol{a}$, $\boldsymbol{A}_i \in \mathbb{R}^m$ to denote the i-th column of matrix $\boldsymbol{A}$, and A_{ij} to denote the entry of $\boldsymbol{A}$ in the i-th row and j-th column, unless specified otherwise. 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. We use $\mathrm{ri}(\mathcal{V})$ to denote the relative interior of a set $\mathcal{V} \subseteq \mathbb{R}^{n_{\boldsymbol{\nu}}}$. The *domain* of a function $f : \mathbb{R}^{n_{\boldsymbol{\nu}}} \to [-\infty, +\infty]$ is defined as $\mathrm{dom}(f) = \{\boldsymbol{\nu} \in \mathbb{R}^{n_{\boldsymbol{\nu}}} \mid f(\boldsymbol{\nu}) < +\infty\}$. The function f is *proper* if $f(\boldsymbol{\nu}) > -\infty$ for all $\boldsymbol{\nu} \in \mathbb{R}^{n_{\boldsymbol{\nu}}}$ and $f(\boldsymbol{\nu}) < +\infty$ for at least one $\boldsymbol{\nu} \in \mathbb{R}^{n_{\boldsymbol{\nu}}}$, implying that $\mathrm{dom}(f) \neq \emptyset$. In addition, f is *closed* if f is lower semicontinuous and either $f(\boldsymbol{\nu}) > -\infty$ for all $\boldsymbol{\nu} \in \mathbb{R}^{n_{\boldsymbol{\nu}}}$ or $f(\boldsymbol{\nu}) = -\infty$ for all $\boldsymbol{\nu} \in \mathbb{R}^{n_{\boldsymbol{\nu}}}$. The *conjugate* of a function $f : \mathbb{R}^{n_{\boldsymbol{\nu}}} \to [-\infty, +\infty]$ is the function

$f^*:\mathbb{R}^{n_{\boldsymbol{\nu}}}\to[-\infty,+\infty]$ defined through $f^*(\boldsymbol{w})=\sup_{\boldsymbol{\nu}}\left\{\boldsymbol{\nu}^\top\boldsymbol{w}-f(\boldsymbol{\nu})\right\}$. The conjugate $(f^*)^*$ of f^* is called the *biconjugate* of f and is abbreviated as f^{**}. The *indicator function* $\delta_{\mathcal{V}}:\mathbb{R}^{n_{\boldsymbol{\nu}}}\to[-\infty,+\infty]$ of a set $\mathcal{V}\subseteq\mathbb{R}^{n_{\boldsymbol{\nu}}}$ is defined through $\delta_{\mathcal{V}}(\boldsymbol{\nu})=0$ if $\boldsymbol{\nu}\in\mathcal{V}$ and $\delta_{\mathcal{V}}(\boldsymbol{\nu})=+\infty$ if $\boldsymbol{\nu}\notin\mathcal{V}$. The *support function* $\delta^*_{\mathcal{V}}:\mathbb{R}^{n_{\boldsymbol{\nu}}}\to[-\infty,+\infty]$ of a set $\mathcal{V}\subseteq\mathbb{R}^{n_{\boldsymbol{\nu}}}$ is defined through $\delta^*_{\mathcal{V}}(\boldsymbol{w})=\sup_{\boldsymbol{\nu}\in\mathcal{V}}\ \{\boldsymbol{\nu}^\top\boldsymbol{w}\}$. The *perspective function* of a proper, closed and convex function $f:\mathbb{R}^{n_\nu}\to(-\infty,+\infty]$ is defined as $h(\boldsymbol{\nu},t)=tf(\boldsymbol{\nu}/t)$ if $t>0$, and $h(\boldsymbol{\nu},0)=\delta^*_{\mathrm{dom}(f^*)}(\boldsymbol{\nu})$ for all $\boldsymbol{\nu}\in\mathbb{R}^{n_\nu}$ and $t\in\mathbb{R}_+$. 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. Generic problem formulation

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

$$\begin{array}{lll}\min\limits_{\boldsymbol{x}} & f_0(\boldsymbol{x}) & \\ \text{s.t.} & f_k(\boldsymbol{x})\le 0, & k\in\mathcal{K}, \\ & \boldsymbol{x}\in\mathcal{X}, & \end{array}\tag{1}$$

where $f_k:\mathbb{R}^{n_x}\to[-\infty,\infty]$ is a sum of convex times convex (SCC) function for all $k\in\mathcal{K}_0$, that is,

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

and $c_{0k},r_{ik},c_{ik}:\mathbb{R}^{n_x}\to(-\infty,+\infty]$, are proper, closed and convex functions for every $i\in\mathcal{I}$, $k\in\mathcal{K}_0$. The set $\mathcal{X}\subseteq\mathbb{R}^{n_x}$ is defined by:

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

where $\boldsymbol{A}\in\mathbb{R}^{n_x\times m_1}$, $\boldsymbol{P}\in\mathbb{R}^{n_x\times m_2}$, $\boldsymbol{b}\in\mathbb{R}^{m_1}$, $\boldsymbol{s}\in\mathbb{R}^{m_2}$, $\boldsymbol{h}(\boldsymbol{x})=[h_0(\boldsymbol{x})\ h_1(\boldsymbol{x})\ \cdots\ h_J(\boldsymbol{x})]^\top$, and $h_j:\mathbb{R}^{n_x}\to(-\infty,+\infty]$ are proper, closed and convex for every $j\in\mathcal{J}_0$. We make the following assumptions.

Assumption 1. *The set $\mathcal{X}$ is nonempty and compact.*

Assumption 2. *If r_{ik} and c_{ik} are both nonlinear, then $r_{ik}(\boldsymbol{x})\ge 0$ and $c_{ik}(\boldsymbol{x})\ge 0$ for all $\boldsymbol{x}\in\mathcal{X}$, for every $i\in\mathcal{I}$ and $k\in\mathcal{K}_0$. If r_{ik} is linear and c_{ik} is nonlinear, then $r_{ik}(\boldsymbol{x})\ge 0$ for all $\boldsymbol{x}\in\mathcal{X}$, for every $i\in\mathcal{I}$ and $k\in\mathcal{K}_0$. If both r_{ik} and c_{ik} are linear, then we do not impose any assumption on these functions.*

Observe that we can reformulate an SCC function in the following way:

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

Hence in the remainder we can assume, without loss of generality, that the functions f_k in (1) are sum of linear times convex (SLC) functions for all $k\in\mathcal{K}_0$, that is,

$$f_k(\boldsymbol{x})=c_{0k}(\boldsymbol{x})+\sum_{i\in\mathcal{I}_k}\left(q_{ik}-\boldsymbol{d}_{ik}^\top\boldsymbol{x}\right)c_{ik}(\boldsymbol{x}),\tag{2}$$

and $q_{ik} \in \mathbb{R}$, $\boldsymbol{d}_{ik} \in \mathbb{R}^{n_x}$, and $c_{0k}, c_{ik} : \mathbb{R}^{n_x} \to (-\infty, +\infty]$ are proper, closed and convex for every $i \in \mathcal{I}$, and $k \in \mathcal{K}_0$.

We now present some examples of functions that are SLC or can be equivalently written as an SLC function.

Example 1 (Difference of convex functions). Consider the Difference of Convex (DC) function $f_k(\boldsymbol{x}) = c_{0k}(\boldsymbol{x}) - c_{1k}(\boldsymbol{x}) \le 0$, where c_{0k}, $c_{1k} : \mathbb{R}^{n_x} \to (-\infty, +\infty]$ are proper, closed and convex for some $k \in \mathcal{K}_0$. We can then reformulate the corresponding constraint function into an SLC function using the biconjugate reformulation (Rockafellar, 1970) and obtain

$$\begin{aligned} f_k(\boldsymbol{x}) \le 0 \quad &\iff \quad \inf_{\boldsymbol{y} \in \mathrm{dom}(c_{1k}^*)} \{c_{0k}(\boldsymbol{x}) - \boldsymbol{x}^\top \boldsymbol{y} + c_{1k}^*(\boldsymbol{y})\} \le 0 \\ &\iff \quad \exists \boldsymbol{y} \in \mathrm{dom}(c_{1k}^*) : \ c_{0k}(\boldsymbol{x}) - \boldsymbol{x}^\top \boldsymbol{y} + c_{1k}^*(\boldsymbol{y}) \le 0, \end{aligned} \tag{3}$$

as long as the infimum is attained, since we can then remove the inf operator. In case the infimum is not attained we refer to Appendix A. We note that for many important classes of convex functions, their conjugates and domains are readily available from the literature. We summarize several of them in Table 1. We note that DC functions constitute an important class of SLC representable functions. For example, every twicely differentiable continuous function has a DC decomposition (Hartman, 1959) and can therefore be written as an SLC function. Moreover, every concave function is a DC function, as we can take $c_{0k}(\boldsymbol{x}) = 0$.

Table 1 Example of functions $f(\cdot)$ and their corresponding conjugates. For the functions in #8, we assume that $\cap_i \mathrm{ri}(\mathrm{dom}(g_i)) \neq \emptyset$.

#	f	$\mathrm{dom}(f^*)$	f^*
1	$f(\boldsymbol{x}, \bar{x}) = \\|\boldsymbol{x}\\|_2 - \bar{x}$	$\{(\boldsymbol{y}, \bar{y}) : \\|\boldsymbol{y}\\|_2 \le 1,\ \bar{y} = 1\}$	$f^*(\boldsymbol{y}, \bar{y}) = 0$
2	$f(x) = x \log(x)$	$\{y : y \in \mathbb{R}\}$	$f^*(y) = \exp(y-1)$
3	$f(x) = -\log(x)$	$\{y : y < 0\}$	$f^*(y) = -\log(-y) - 1$
4	$f(x) = \sqrt{x}$	$\{y : y < 0\}$	$f^*(y) = -\frac{1}{4y}$
5	$f(\boldsymbol{x}) = \max_i x_i$	$\{\boldsymbol{y} : \boldsymbol{y} \ge 0, \sum_i y_i = 1, \forall k\}$	$f^*(\boldsymbol{y}) = 0$
6	$f(\boldsymbol{x}) = \sum_i \max_{k \in \mathcal{K}_i} x_i$	$\{\{\boldsymbol{y}_i\}_i : \boldsymbol{y}_i \ge 0, \sum_{k \in \mathcal{K}_i} y_{ik} = 1, \forall i\}$	$f^*(\boldsymbol{y}) = 0$
7	$f(\boldsymbol{x}) = \log(\sum_i \exp(x_i))$	$\{\boldsymbol{y} : \boldsymbol{y} \ge 0, \sum_i y_i = 1\}$	$f^*(\boldsymbol{y}) = \sum_i y_i \log(y_i)$
8	$f(\boldsymbol{x}) = \sum_i g_i(\boldsymbol{x})$	$\{\{\boldsymbol{y}_i\}_i : \sum_i \boldsymbol{y}_i = \boldsymbol{y}, \boldsymbol{y}_i \in \mathrm{dom}(g_i^*), \forall i\}$	$f^*(\boldsymbol{y}) = \min_{\{\boldsymbol{y}_i\}_i} \sum_i g_i^*(\boldsymbol{y}_i)$

Example 2 (Fractional optimization). Consider the following fractional function

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

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. □

EXAMPLE 3 (SOME EXAMPLES OF SLC FUNCTIONS). In Table 2, we give some more examples of SLC functions that are generally nonconvex and satisfy Assumption 2. Hence, our proposed approach can deal with Problem (1) containing (sum of) such nonconvex components. □

#	f	c	$(q-\boldsymbol{d}^\top\boldsymbol{x})$	Perspectification	Assumptions
1	$\sqrt{q-\boldsymbol{d}^\top\boldsymbol{x}}$	$(q-\boldsymbol{d}^\top\boldsymbol{x})^{-1/2}$	$q-\boldsymbol{d}^\top\boldsymbol{x}$	$(q-\boldsymbol{d}^\top\boldsymbol{x})\sqrt{\dfrac{q-\boldsymbol{d}^\top\boldsymbol{x}}{q^2-2q\boldsymbol{d}^\top\boldsymbol{x}+\boldsymbol{d}^\top\boldsymbol{X}\boldsymbol{d}}}$	$\boldsymbol{d}^\top\boldsymbol{x}\leq q$
2	$(q-\boldsymbol{d}^\top\boldsymbol{x})^\theta$	$(q-\boldsymbol{d}^\top\boldsymbol{x})^{\theta-1}$	$q-\boldsymbol{d}^\top\boldsymbol{x}$	$(q-\boldsymbol{d}^\top\boldsymbol{x})\left(\dfrac{q^2-2q\boldsymbol{d}^\top\boldsymbol{x}+\boldsymbol{d}^\top\boldsymbol{X}\boldsymbol{d}}{q-\boldsymbol{d}^\top\boldsymbol{x}}\right)^{\theta-1}$	$\theta\in[0,1]$ & $\boldsymbol{d}^\top\boldsymbol{x}\leq q$
3	$-(q-\boldsymbol{d}^\top\boldsymbol{x})^\theta$	$-(q-\boldsymbol{d}^\top\boldsymbol{x})^{\theta-1}$	$q-\boldsymbol{d}^\top\boldsymbol{x}$	$-(q-\boldsymbol{d}^\top\boldsymbol{x})\left(\dfrac{q^2-2q\boldsymbol{d}^\top\boldsymbol{x}+\boldsymbol{d}^\top\boldsymbol{X}\boldsymbol{d}}{q-\boldsymbol{d}^\top\boldsymbol{x}}\right)^{\theta-1}$	$\theta\in[1,2]$ & $\boldsymbol{d}^\top\boldsymbol{x}\leq q$
4	$-(q_1-\boldsymbol{d}_1^\top\boldsymbol{x})\ln(q_2-\boldsymbol{d}_2^\top\boldsymbol{x})$	$-\ln(q_2-\boldsymbol{d}_2^\top\boldsymbol{x})$	$(q_1-\boldsymbol{d}_1^\top\boldsymbol{x})$	$-(q_1-\boldsymbol{d}_1^\top\boldsymbol{x})\ln\left(\dfrac{q_1q_2-q_1\boldsymbol{d}_2^\top\boldsymbol{x}-q_2\boldsymbol{d}_1^\top\boldsymbol{x}+\boldsymbol{d}_1^\top\boldsymbol{X}\boldsymbol{d}_2}{q_1-\boldsymbol{d}_1^\top\boldsymbol{x}}\right)$	$\boldsymbol{d}_i^\top\boldsymbol{x}\leq q_i,\quad i\in\{1,2\}$
5	$\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}$	$(\boldsymbol{Q}\boldsymbol{x})_i$	x_i	$\mathrm{Tr}(\boldsymbol{X}\boldsymbol{Q})$	-
6	$(q-\boldsymbol{d}^\top\boldsymbol{x})\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}$	$\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}$	$(q-\boldsymbol{d}^\top\boldsymbol{x})$	$\frac{(q\boldsymbol{x}-\boldsymbol{X}\boldsymbol{d})^\top\boldsymbol{Q}(q\boldsymbol{x}-\boldsymbol{X}\boldsymbol{d})}{(q-\boldsymbol{d}^\top\boldsymbol{x})}$	$\boldsymbol{d}^\top\boldsymbol{x}\leq q$ & $\boldsymbol{Q}\succeq\boldsymbol{0}$

Table 2 Examples of SLC representable functions.

3. Reformulation-Perspectification Technique and Branch and Bound

In this section, we describe our new approach, called RPT-BB, to obtain a global optimal solution of (1). Our approach comprises five steps:

Step 1: Preprocessing. Introduce epigraphical variables for every nonlinear convex component c_{0k}, $k\in\mathcal{K}_0$, in the nonconvex SLC functions.

Step 2: Reformulation and perspectification. Generate additional redundant nonconvex constraints from pairwise multiplication of the existing convex inequalities in (1). Next, convexify all nonconvex components in (1) and all nonconvex components in the additional generated constraints by reformulating them in their perspective form and subsequently linearizing all product terms.

Step 3 (Optional): SDP relaxation. Add an additional LMI inequality from the SDP relaxation of the linearization of all product terms.

Step 4: Obtaining upper bounds. Solve the convex RPT relaxation. From the solution of the RPT relaxation, construct a set of candidate solutions for (1), substitute these candidate solutions in Problem (1) and choose the best upper bound obtained.

Step 5: Branch and bound. Solve Problem (1) to optimality by means of a spatial branch and bound method. In the next sections, we describe each of these steps in more detail.

3.1. Preprocessing step

We introduce epigraphical variables for the convex component in the nonconvex SLC functions of (1), and from (2) we have

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\tau}} \quad & \tau_0 + \sum_{i\in\mathcal{I}_k} \left(q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}\right) c_{i0}(\boldsymbol{x}) \\ \text{s.t.} \quad & \tau_k + \sum_{i\in\mathcal{I}_k} \left(q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}\right) c_{ik}(\boldsymbol{x}) \le 0, \quad k\in\mathcal{K}, \\ & (\boldsymbol{x},\boldsymbol{\tau}) \in \mathcal{T}, \end{aligned} \tag{4}$$

where $\mathcal{T} = \{(\boldsymbol{x},\boldsymbol{\tau}) \in \mathbb{R}^{n_x} \times \mathbb{R}^{K+1} \mid \boldsymbol{x}\in\mathcal{X},\ \boldsymbol{c}_0(\boldsymbol{x}) \le \boldsymbol{\tau}\}$, and $\boldsymbol{c}_0(\boldsymbol{x}) = [c_{00}(\boldsymbol{x})\ c_{01}(\boldsymbol{x})\ \cdots\ c_{0K}(\boldsymbol{x})]^\top \subseteq (-\infty,+\infty]^{K+1}$. As we will see later in Theorem 1, we can multiply these extra epigraphical constraints with the existing convex constraints to obtain a tighter convex relaxation.

3.2. Reformulation and perspectification

Now we are ready to explain the core idea of RPT. We will first explain the core idea of RPT on the univariate case. The intuition is as follows: Consider the convex function $c:\mathbb{R}\to\mathbb{R}$ and the nonconvex constraint set $\mathcal{X} = \{x\in\mathbb{R}_+ :\ x\,c(x)\le 0\}$. The constraint set can be written as $\mathcal{X} = \left\{(x,x')\in\mathbb{R}_+\times\mathbb{R}_+ :\ x\,c\left(\frac{x'}{x}\right)\le 0,\ x' = x^2\right\}$. Observe that the new set $\mathcal{X}$ is also nonconvex due to the constraint $x' = x^2$. However, observe that for $x\ge 0$, the function $h(x,x') = x\,c\left(\frac{x'}{x}\right)$ is the perspective function of $c(\cdot)$ and therefore it is jointly convex in x and x'. Thus, by either completely relaxing $x' = x^2$ and adding the hitherto redundant $x'\ge 0$, or by relaxing it as $x'\ge x^2$ to preserve convexity (and nonnegativity), one obtains a convex outer approximation of $\mathcal{X}$.

Remark 1. Observe that the RLT approach is a sub-case of this, assuming the function $c(\cdot)$ is linear. Since the perspective of a linear function is the function itself the perspectification step is not needed in RLT. However, for nonlinear functions this is not the case and therefore the perspectification step is necessary.

In the general case, let f be an SLC function as given by (2), that satisfies Assumption 2. Then we can perspectify the generally nonconvex function f by first multiplying and dividing the argument of c_i by $(q_i - \boldsymbol{d}_i^\top\boldsymbol{x})$ for every $i\in\mathcal{I}$ 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 $\boldsymbol{X}\in\mathbb{S}^{n_x}$ to obtain the following sum of perspective functions:

$$\sum_{i\in\mathcal{I}} (q_i - \boldsymbol{d}_i^\top\boldsymbol{x}) c_i\left(\frac{q_i\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}_i}{q_i - \boldsymbol{d}_i^\top\boldsymbol{x}}\right), \tag{5}$$

which is jointly convex in $(\boldsymbol{x}, \boldsymbol{X})$ because c_i is convex if and only if its perspective is convex (Rockafellar, 1970). Observe that if Assumption 2 is not satisfied, i.e., $q_i - \boldsymbol{d}_i^\top \boldsymbol{x} \le 0$ for some $\boldsymbol{x} \in \mathcal{X}$ and $i \in \mathcal{I}$, then the above sum of perspective functions might not be convex. We obtain the following convex relaxation:

$$\begin{array}{lll} \min\limits_{\boldsymbol{x},\boldsymbol{\tau},\boldsymbol{X}} & \tau_0 + \sum\limits_{i\in\mathcal{I}_k} (q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}) c_{i0}\left(\dfrac{q_{i0}\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}_{i0}}{q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}}\right) & \\ \text{s.t.} & \tau_k + \sum\limits_{i\in\mathcal{I}_k} (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) c_{ik}\left(\dfrac{q_{ik}\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}}\right) \le 0, & k \in \mathcal{K}, \\ & (\boldsymbol{x}, \boldsymbol{\tau}) \in \mathcal{T}. & \end{array} \tag{6}$$

By pairwise multiplying inequalities in the set $\mathcal{T}$, we can obtain additional redundant SLC constraints, which can then be convexified in a manner similar to what was described above. Once convexified, these SLC constraints are no longer redundant and actually serve as bounds on the newly introduced variables corresponding to the product terms. We can pairwise multiply the linear inequality constraints in the set $\mathcal{T}$, similar to the approach in RLT, to derive bounds on the newly introduced variables $\boldsymbol{X} \in \mathbb{S}^{n_x}$. However, with RPT, we can incrementally improve this approximation by also considering the pairwise multiplication of the linear and convex constraints in the set $\mathcal{T}$, followed by the pairwise multiplication of the convex inequalities in the set $\mathcal{T}$.

To be more precise, by considering the following cases of pairwise multiplication of the constraints in the set $\mathcal{T}$, we incrementally improve the convex relaxation of (1) derived from RPT:

Linear inequality $\times$ Linear inequality. This is well-known in RLT: we multiply the constraints $\boldsymbol{A}^\top \boldsymbol{x} \le \boldsymbol{b}$ of (1) with $\boldsymbol{A}^\top \boldsymbol{x} \le \boldsymbol{b}$, and obtain the redundant constraints:

$$\boldsymbol{b}\boldsymbol{x}^\top \boldsymbol{A} + \boldsymbol{A}^\top \boldsymbol{x}\boldsymbol{b}^\top \le \boldsymbol{A}^\top \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{A} + \boldsymbol{b}\boldsymbol{b}^\top.$$

Since the (i,j)-th constraint is exactly the (j,i)-th constraint, we only consider the upper triangular of the matrix equations; so $m_1(m_1+1)/2$ constraints instead of m_1^2. Next, the nonlinear quadratic terms $\boldsymbol{x}\boldsymbol{x}^\top$ are linearized by substituting them with $\boldsymbol{X} \in \mathbb{S}^{n_x}$. We then obtain the following additional convex constraints:

$$\boldsymbol{b}\boldsymbol{x}^\top \boldsymbol{A} + \boldsymbol{A}^\top \boldsymbol{x}\boldsymbol{b}^\top \le \boldsymbol{A}^\top \boldsymbol{X}\boldsymbol{A} + \boldsymbol{b}\boldsymbol{b}^\top. \tag{7}$$

Moreover, we include the additional constraints

$$X_{ii} \ge 0, \quad i \in \{1, \dots, n_x\},$$

since $x_i^2 \ge 0$ for all $i \in \{1, \dots, n_x\}$.

Linear inequality × Convex inequality. By multiplying each ℓ-th linear inequality $\boldsymbol{A}_\ell^\top \boldsymbol{x} \le b_\ell$ of (1) with the convex constraints $\boldsymbol{c}_0(\boldsymbol{x}) \le \boldsymbol{\tau}$ and $\boldsymbol{h}(\boldsymbol{x}) \le \boldsymbol{0}$, we obtain $m_1(J+K+2)$ redundant SLC constraints of the form

$$(b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{c}_0(\boldsymbol{x}) \le (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{\tau} \quad \text{and} \quad (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}) \le \boldsymbol{0} \qquad \ell \in \{1, \ldots, m_1\}.$$

Next, the redundant SLC constraints can be reformulated into:

$$(b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}) \le \boldsymbol{0} \iff (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{h}\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{A}_\ell}{b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x}}\right) \le \boldsymbol{0} \quad \text{and}$$

$$(b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{c}_0(\boldsymbol{x}) \le (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{\tau} \iff (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{c}_0\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{A}_\ell}{b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x}}\right) \le (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{\tau}.$$

Finally, the nonlinear quadratic terms $\boldsymbol{x}\boldsymbol{x}^\top$ and the bilinear terms $\boldsymbol{\tau}\boldsymbol{x}^\top$ are linearized by substituting them with $\boldsymbol{X} \in \mathbb{S}^{n_x}$ and $\boldsymbol{V} \in \mathbb{R}^{(K+1)\times n_x}$, to obtain the following additional convex constraints:

$$(b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{h}\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{X}\boldsymbol{A}_\ell}{b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x}}\right) \le \boldsymbol{0} \quad \text{and} \quad (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{c}_0\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{X}\boldsymbol{A}_\ell}{b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x}}\right) \le b_\ell \boldsymbol{\tau} - \boldsymbol{V}\boldsymbol{A}_\ell.$$

Linear equality × Convex inequality. When multiplying a linear equality constraint with a convex inequality constraint, 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 we consider the pairwise multiplication of the linear equality constraints with all variables (see Lemma 1). For quadratic problems, a similar observation was first mentioned by Sherali and Adams (1999, Remark 8.1). Before we formally prove Lemma 1, we first define redundant constraints.

Definition 1 (Redundant equality constraints). An equality constraint $f(\boldsymbol{x}) = 0$, where $f : \mathbb{R}^{n_x} \to (-\infty, +\infty]$, is *redundant* to the nonempty set $\mathcal{X} \subseteq \mathbb{R}^{n_x}$, if $\mathcal{X} \subseteq \{\boldsymbol{x} \mid f(\boldsymbol{x}) = 0\}$.

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

Proof. Since $\boldsymbol{d}^\top \boldsymbol{x} = q$, $\boldsymbol{X}\boldsymbol{d} = q\boldsymbol{x}$, and $f : \mathbb{R}^{n_x} \to (-\infty, +\infty]$ is proper, closed and convex, it then follows from the definition of the perspective function that

$$(q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{q\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) = \delta^*_{\mathrm{dom}(f^*)}(\boldsymbol{0}).$$

If $\mathrm{dom}(f^*)$ is nonempty, then $\delta^*_{\mathrm{dom}(f^*)}(\boldsymbol{0}) = 0$. The set $\mathrm{dom}(f^*)$ is indeed nonempty because of the properness of f^* (Rockafellar, 1970, p. 24), which is implied by Theorem 12.2 of Rockafellar (1970) thanks to the properness and convexity of f. □

Thanks to Lemma 1, it suffices to multiply each ℓ-th linear equality constraint $\boldsymbol{P}_\ell^\top \boldsymbol{x} = s_\ell$ with $\boldsymbol{x}$ and $\boldsymbol{\tau}$ respectively. We then obtain $m_2 n_x + m_2 n_\tau$ redundant SLC constraints of the form

$$(s_\ell - \boldsymbol{P}_\ell^\top \boldsymbol{x})\boldsymbol{x} = \boldsymbol{0} \quad \text{and} \quad (s_\ell - \boldsymbol{P}_\ell^\top \boldsymbol{x})\boldsymbol{\tau} = \boldsymbol{0} \qquad \ell \in \{1, \ldots, m_2\}.$$

Finally, the nonlinear quadratic terms $\boldsymbol{x}\boldsymbol{x}^\top$ and the bilinear terms $\boldsymbol{\tau}\boldsymbol{x}^\top$ are linearized by substituting them with $\boldsymbol{X} \in \mathbb{S}^{n_x}$ and $\boldsymbol{V} \in \mathbb{R}^{(K+1)\times n_x}$. Including all additional constraints from pairwise multiplying the linear constraints with the linear and nonlinear convex constraints in Problem (6) we obtain the following convex relaxation:

$$\begin{array}{lll}
\min\limits_{\boldsymbol{x},\boldsymbol{\tau},\boldsymbol{X},\boldsymbol{V}} & \tau_0 + \sum\limits_{i\in\mathcal{I}_k} (q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}) c_{i0}\left(\dfrac{q_{i0}\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}_{i0}}{q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}}\right) & \\
\text{s.t.} & \tau_k + \sum\limits_{i\in\mathcal{I}_k} (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) c_{ik}\left(\dfrac{q_{ik}\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}}\right) \le 0, & k \in \mathcal{K}, \\
& \boldsymbol{b}\boldsymbol{x}^\top \boldsymbol{A} + \boldsymbol{A}^\top \boldsymbol{x}\boldsymbol{b}^\top \le \boldsymbol{A}^\top \boldsymbol{X}\boldsymbol{A} + \boldsymbol{b}\boldsymbol{b}^\top, & \\
& X_{ii} \ge 0, & i \in \{1, \ldots, n_x\}, \\
& (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{h}\left(\dfrac{b_\ell \boldsymbol{x} - \boldsymbol{X}\boldsymbol{A}_\ell}{b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x}}\right) \le \boldsymbol{0}, & \ell \in \{1, \ldots, m_1\}, \\
& (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{c}_0\left(\dfrac{b_\ell \boldsymbol{x} - \boldsymbol{X}\boldsymbol{A}_\ell}{b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x}}\right) \le b_\ell \boldsymbol{\tau} - \boldsymbol{V}\boldsymbol{A}_\ell, & \ell \in \{1, \ldots, m_1\}, \\
& s_\ell \boldsymbol{x} - \boldsymbol{X}\boldsymbol{P}_\ell = \boldsymbol{0}, & \ell \in \{1, \ldots, m_2\}, \\
& s_\ell \boldsymbol{\tau} - \boldsymbol{V}\boldsymbol{P}_\ell = \boldsymbol{0}, & \ell \in \{1, \ldots, m_2\}, \\
& (\boldsymbol{x}, \boldsymbol{\tau}) \in \mathcal{T}. &
\end{array} \tag{8}$$

Note that there are $m_1(J + 1 + \frac{m_1+1}{2}) + (m_1 + 1)(K + 1) + (m_2 + 1)n_x + m_2 n_\tau$ additional constraints and $n_x^2 + (n_x + 1)(K + 1)$ additional variables in (8) compared to (1).

The following theorem demonstrates that introducing epigraphical variables for the convex components of the SLC functions in (1) (as is done in the preprocessing step of RPT; see (4)), can result in a potentially tighter RPT relaxation.

Theorem 1. *The convex relaxation (8) obtained from RPT with the preprocessing step is at least as tight as the relaxation obtained from RPT without the preprocessing step.*

Proof. By applying RPT without the preprocessing step, one can obtain the following relaxation of (1):

$$\begin{array}{lll}
\min\limits_{\boldsymbol{x},\boldsymbol{X}} & c_{00}(\boldsymbol{x}) + \sum\limits_{i\in\mathcal{I}_k} (q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}) c_{i0}\left(\dfrac{q_{i0}\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}_{i0}}{q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}}\right) & \\
\text{s.t.} & c_{0k}(\boldsymbol{x}) + \sum\limits_{i\in\mathcal{I}_k} (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) c_{ik}\left(\dfrac{q_{ik}\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}}\right) \le 0, & k \in \mathcal{K}, \\
& \boldsymbol{b}\boldsymbol{x}^\top \boldsymbol{A} + \boldsymbol{A}^\top \boldsymbol{x}\boldsymbol{b}^\top \le \boldsymbol{A}^\top \boldsymbol{X}\boldsymbol{A} + \boldsymbol{b}\boldsymbol{b}^\top, & \\
& X_{ii} \ge 0, & i \in \{1, \ldots, n_x\}, \\
& (b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x})\boldsymbol{h}\left(\dfrac{b_\ell \boldsymbol{x} - \boldsymbol{X}\boldsymbol{A}_\ell}{b_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x}}\right) \le \boldsymbol{0}, & \ell \in \{1, \ldots, m_1\}, \\
& s_\ell \boldsymbol{x} - \boldsymbol{X}\boldsymbol{P}_\ell = \boldsymbol{0}, & \ell \in \{1, \ldots, m_2\}, \\
& \boldsymbol{x} \in \mathcal{X}. &
\end{array} \tag{9}$$

It can be directly verified that from any feasible solution $(\boldsymbol{x}', \boldsymbol{X}')$ to (9), there exists a $(\boldsymbol{\tau}', \boldsymbol{V}')$, where $\boldsymbol{\tau}' = \boldsymbol{c}_0(\boldsymbol{x}')$ and $\boldsymbol{V}' = \boldsymbol{\tau}'\boldsymbol{x}'^\top$, such that $(\boldsymbol{x}', \boldsymbol{\tau}', \boldsymbol{X}', \boldsymbol{V}')$ is feasible to (8), and attains the same objective value. □

The following example shows that introducing epigraphical variables for the nonlinear convex components in the preprocessing step tightens the RPT relaxation.

Example 4 (Effectiveness of preprocessing). Consider the following convex maximization problem:

$$\begin{array}{ll} \max\limits_{x} & x_1 x_2 - \frac{(x_2-2)^2}{4} \\ \text{s.t.} & 0 \le x_1 \le 1. \end{array} \tag{10}$$

In the following we compare two convex relaxations of (10) obtained from (i) applying RPT without the epigraphical reformulation, and (ii) applying RPT with the epigraphical reformulation, respectively,

$$\begin{array}{ll} \max\limits_{\boldsymbol{x},\boldsymbol{X}} & X_{12} - \frac{(x_2-2)^2}{4} \\ \text{s.t.} & 0 \le x_1 \le 1, \\ & X_{11} - 2x_1 + 1 \ge 0, \\ & X_{11} \ge 0, \end{array} \quad \text{and} \quad \begin{array}{ll} \max\limits_{\boldsymbol{x},\boldsymbol{X},\tau,t} & X_{12} - \tau \\ \text{s.t.} & 0 \le x_1 \le 1, \\ & X_{11} - 2x_1 + 1 \ge 0, \\ & X_{11} \ge 0, \\ & \dfrac{(x_2-2)^2}{4} \le \tau, \\ & \dfrac{(X_{12}-2x_1)^2}{4x_1} \le t, \\ & \dfrac{(x_2 - 2 - X_{12} + 2x_1)^2}{4 - 4x_1} \le \tau - t. \end{array}$$

Note that the maximum objective value of the convex relaxation without the epigraphical reformulation is ∞ with $x_1^\star \in [0,1]$, $x_2^\star = 2$, $X_{11}^\star = \infty$, and that of the epigraphical reformulation is 3 with $(x_1^\star, x_2^\star, X_{12}^\star, \tau^\star, t^\star) = (1,4,4,1,1)$. □

However, as demonstrated in the following lemma, introducing epigraphical variables for the linear components does not enhance the convex relaxation obtained from RPT.

Lemma 2. *Introducing epigraphical variables for the linear components in the preprocessing step of RPT does not result in a tighter convex relaxation.*

Proof. Suppose $c_{0k} : \mathbb{R}^{n_x} \to \mathbb{R}$ is a linear function in (1) for some $k \in \mathcal{K}_0$. In the preprocessing step of RPT, one can introduce an epigraphical variable τ_k for the linear component c_{0k} and consider the linear equality $\tau_k = c_{0k}(\boldsymbol{x})$ in the constraints. From Lemma 1, we know that it suffices to consider the multiplication of this constraint by $\boldsymbol{x}$ and $\boldsymbol{\tau}$ to obtain the convex relaxation from RPT. Specifically, from the reformulation and perspectification step of RPT, we obtain:

$$V_{ki} = \boldsymbol{x}_i c_{0k}\left(\frac{\boldsymbol{X}_i}{x_i}\right), \quad i \in \{1, \ldots, n_x\}, \quad \text{and} \quad T_{kk'} = \boldsymbol{\tau}_{k'} c_{0k}\left(\frac{(\boldsymbol{V}^\top)_{k'}}{\tau_{k'}}\right), \quad k' \in \mathcal{K}_0,$$

where $\boldsymbol{V}_k$ and $\boldsymbol{T}_k$ are variables introduced to linearize the product terms $\tau_k \boldsymbol{x}$ and $\tau_k \boldsymbol{\tau}$, respectively. Since $\boldsymbol{V}_k$ and $\boldsymbol{T}_k$ do not appear anywhere else but in these linear equalities in the convex relaxation from RPT, these equalities exactly determine $\boldsymbol{V}_k$ and $\boldsymbol{T}_k$ and are therefore redundant. □

Convex inequality × Convex inequality. Just multiplying a nonlinear convex constraint $h_j(\boldsymbol{x}) \leq 0$ with another nonlinear convex constraint $h_{j'}(\boldsymbol{x}) \leq 0$ results in a constraint $-h_j(\boldsymbol{x})h_{j'}(\boldsymbol{x}) \leq 0$ for which the constraint function is not an SCC function, since in this case $-h_j(\boldsymbol{x})$ is concave instead of convex. However, sometimes rewriting the constraints, and then multiplying the left-hand sides and right-hand sides of the constraints yields convexifiable constraints. Consider for example the following two exponential constraints:

$$\exp(x_1) \leq x_2 \quad \text{and} \quad \exp(x_3) \leq x_4.$$

We can then multiply the left-hand sides and the right-hand sides, and multiply the right-hand side of each constraint with the other exponential constraint to obtain the following convexified constraints:

$$\begin{cases} \exp(x_1 + x_3) \leq X_{24}, \\ x_4 \exp\left(\dfrac{X_{14}}{x_4}\right) \leq X_{24}, \\ x_2 \exp\left(\dfrac{X_{23}}{x_2}\right) \leq X_{24}. \end{cases}$$

Also, several ways of obtaining a convexifiable constraint from pairwise multiplication of conic quadratic constraints are readily available in the literature (Yang and Burer, 2016; Anstreicher, 2017; Jiang and Li, 2019).

3.3. Additional SDP relaxation

In order to further tighten the convex relaxation, 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 $\boldsymbol{X}$. Such a linearization based relaxation for the nonconvex quadratic equality $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^\top$ may be significantly improved by the SDP relaxation $\boldsymbol{X} \succeq \boldsymbol{x}\boldsymbol{x}^\top$, which can be equivalently reformulated as an LMI by using Schur complement (Boyd and Vandenberghe, 2004):

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^\top & 1 \end{pmatrix} \succeq 0. \tag{11}$$

Because we also have epigraphical constraints, we can consider including the following LMI:

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{V}^\top & \boldsymbol{x} \\ \boldsymbol{V} & \boldsymbol{T} & \boldsymbol{\tau} \\ \boldsymbol{x}^\top & \boldsymbol{\tau}^\top & 1 \end{pmatrix} \succeq 0,$$

where $\boldsymbol{T} \in \mathbb{S}^{K+1}$ denotes the matrix that substitutes the quadratic terms $\boldsymbol{\tau}\boldsymbol{\tau}^\top$. Although including the LMI might tighten the convex RPT relaxation, it can significantly increase computation time. Hence, this step is optional. Observe that if the above LMI is included, the additional constraints $X_{ii} \geq 0,\ i \in \{1, \dots, n_x\}$, are redundant to the LMI.

3.4. Obtaining upper bounds

Let $(\boldsymbol{x}^*, \boldsymbol{\tau}^*, \boldsymbol{X}^*, \boldsymbol{V}^*)$ be the solution of the convex RPT relaxation (8). We propose to construct the set $\mathcal{X}' = \left\{\boldsymbol{x}^*, \boldsymbol{x}_1^X, \dots, \boldsymbol{x}_{n_x}^X, \boldsymbol{x}_1^V, \dots, \boldsymbol{x}_{n_\tau}^V\right\}$ of candidate solutions for (1), where

$$\boldsymbol{x}_i^X = \begin{cases} \boldsymbol{x}^* & \text{if } x_i^* = 0, \\ \frac{\boldsymbol{X}_i^*}{x_i^*} & \text{otherwise,} \end{cases} \quad \text{and} \quad \boldsymbol{x}_j^V = \begin{cases} \boldsymbol{x}^* & \text{if } \tau_j^* = 0, \\ \frac{(\boldsymbol{V}^{*\top})_j}{\tau_j^*} & \text{otherwise,} \end{cases} \quad \text{for all } i \in \{1, \dots, n_x\}, j \in \{1, \dots, n_\tau\}.$$

Note that by definition $\boldsymbol{x}^* \in \mathcal{X}$ is always satisfied.

LEMMA 3. *If $\mathcal{X}$ includes the non-negativity constraints $\boldsymbol{x} \geq \boldsymbol{0}$, then $\mathcal{X}' \subset \mathcal{X}$.*

Proof. By construction, if $x_i^* \neq 0$, it follows from $x_i^* \geq 0$ and $\boldsymbol{x}_i^X = \frac{\boldsymbol{X}_i^*}{x_i^*}$ that

$$\begin{cases} \boldsymbol{A}^\top \boldsymbol{X}_i^* \leq x_i^* \boldsymbol{b} \\ \boldsymbol{P}^\top \boldsymbol{X}_i^* = x_i^* \boldsymbol{s} \\ x_i^* \boldsymbol{h}\left(\frac{\boldsymbol{X}_i^*}{x_i^*}\right) \leq 0 \end{cases} \implies \begin{cases} \boldsymbol{A}^\top \frac{\boldsymbol{X}_i^*}{x_i^*} \leq \boldsymbol{b} \\ \boldsymbol{P}^\top \frac{\boldsymbol{X}_i^*}{x_i^*} = \boldsymbol{s} \\ \boldsymbol{h}\left(\frac{\boldsymbol{X}_i^*}{x_i^*}\right) \leq 0. \end{cases}$$

Otherwise, if $x_i^* = 0$, we have $\boldsymbol{x}_i^X = \boldsymbol{x}^*$, hence $\boldsymbol{x}_i^X \in \mathcal{X}$. Therefore, $\boldsymbol{x}_i^X \in \mathcal{X}$ for every $i \in \{1, \dots, n_x\}$. Analogously, one can show that $\boldsymbol{x}_j^V \in \mathcal{X}$ for every $j \in \{1, \dots, n_\tau\}$. This concludes the proof. □

In cases where all constraints are convex (i.e., $\mathcal{K} = \emptyset$), substituting the candidate solutions into the original Problem (1) yields upper bounds corresponding to each candidate solution, allowing us to choose the best upper bound obtained. However, if $\boldsymbol{x}$ is not assumed to be nonnegative, or if we also have nonconvexity in the constraints, then the candidate solutions in $\mathcal{X}'$ obtained from the RPT relaxation may not be feasible for Problem (1). In such scenarios, to determine the best upper bound, we propose considering only those candidate solutions that are feasible for Problem (1).

Observe that if we have nonconvexities in the constraints, it is possible that none of the candidate solutions is feasible. Furthermore, note that if only some of the x_i, $i \in \{1, \cdots, n_x\}$, are assumed to be nonnegative, it is possible to obtain an upper bound by considering only those candidate solutions $\frac{\boldsymbol{X}_i^*}{x_i^*}$ where the indices i correspond to positive x_i^*. This can be proved similarly to the proof of Lemma 3.

The obtained feasible solutions could also be used as warm starts for existing algorithms, to improve the upper bound. Namely, we can use a local optimization algorithm, such as the Ipopt solver (Waechter et al., 2009), initialized at the candidate feasible solution, to obtain a local optimum. We then replace the candidate solution in $\mathcal{X}'$ by the obtained local optimum. Note that we can also initialize it from an infeasible solution, and if the solver finds a feasible solution, we also add the solution to $\mathcal{X}''$. Moreover, for the problem of minimizing a concave or a difference of convex function, using the biconjugate reformulation, the problem can be formulated as a disjoint bilinear optimization problem, where the bilinear function is in fact an SLC function (see Example 1). Hence, we can leverage the mountain climbing algorithm by (Tao and An, 1997), to find a local optimum of (1), see Appendix B.

3.5. Spatial branch and bound method

We can solve the original Problem (1) to optimality using the following branch and bound scheme: At the root node, denoted by N_0, we solve an RPT relaxation of Problem (1) and obtain upper and lower bounds to Problem (1). If the lower bound does not equal the upper bound, we search for a hyperplane that cuts the feasible region into two parts. We use the branching strategy proposed by Anstreicher (2022). The intuition behind the branching strategy stems from Saxena et al. (2010). Let $\{\boldsymbol{c}_1, .., \boldsymbol{c}_n\}$ be any collection of mutually orthogonal unit vectors in $\mathbb{R}^n$. Then, Saxena et al. (2010) showed that

$$\left\{(\boldsymbol{x}, \boldsymbol{X}) \mid \boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^\top\right\} = \left\{(\boldsymbol{x}, \boldsymbol{X}) \mid \boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top \succeq \boldsymbol{0}\right\} \cap \left\{(\boldsymbol{x}, \boldsymbol{X}) \mid \boldsymbol{c}_j^\top \boldsymbol{X} \boldsymbol{c}_j \le \left(\boldsymbol{c}_j^\top \boldsymbol{x}\right)^2, \forall j = 1, .., n\right\}.$$

We take the eigenvectors of the matrix $\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top$ as the collection $\{\boldsymbol{c}_1, .., \boldsymbol{c}_n\}$. Observe that if $\boldsymbol{a}$ is a unit eigenvector of $\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top$, with corresponding eigenvalue λ, then it satisfies

$$\left(\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top\right) \boldsymbol{a} = \lambda \boldsymbol{a} \implies \boldsymbol{a}^\top \left(\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top\right) \boldsymbol{a} = \lambda.$$

Now observe that we want to have

$$\boldsymbol{c}_j^\top \left(\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top\right) \boldsymbol{c}_j \le 0, \quad \forall j = 1, .., n.$$

Therefore, the maximum infeasibility corresponds to the eigenvector associated with the maximum eigenvalue of the matrix $\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top$. The branching strategy is as follows: Suppose we solve the RPT relaxation at a node and obtain an optimal solution $(\boldsymbol{x}^*, \boldsymbol{X}^*)$. Then, unless we are at the optimal solution we have $\boldsymbol{X}^* \neq \boldsymbol{x}^*(\boldsymbol{x}^*)^\top$. We then compute the eigenvector $\boldsymbol{f}$ that corresponds to the largest eigenvalue of the matrix $\boldsymbol{X}^* - \boldsymbol{x}^*(\boldsymbol{x}^*)^\top$. Let $l = \boldsymbol{f}^\top \boldsymbol{x}^*$, and $H = \left\{\boldsymbol{x} \mid \boldsymbol{f}^\top \boldsymbol{x} = l\right\}$ be the hyperplane.

After solving the RPT relaxation at the root node and computing the hyperplane H, we create two new "child" nodes N_1 and N_2 from the root-node N_0, where at each child node we solve Problem (1) with its feasible region $\mathcal{X}$ intersected with one of the closed half spaces of the hyperplane H, i.e.,

$$\begin{aligned}\mathcal{X}_l^0 &= \left\{\boldsymbol{x} \in \mathcal{X} \mid \boldsymbol{f}^\top \boldsymbol{x} \le l\right\},\\ \mathcal{X}_r^0 &= \left\{\boldsymbol{x} \in \mathcal{X} \mid \boldsymbol{f}^\top \boldsymbol{x} \ge l\right\}.\end{aligned}$$

The proposed branching strategy is summarized in pseudocode in Algorithm 1, where $\mathcal{X}_N$ denotes the feasible region of node N.

Subsequently, we apply RPT to each child node and obtain a lower and an upper bound for each child node. If for the child node with the lowest lower bound of the two child nodes it holds that it equals the upper bound, we have found the optimal solution. If not, we can repeat this procedure for each child node, i.e., for each constructed child node N_k, we solve the RPT relaxation and then find a hyperplane H_k as described earlier to create again two new child nodes, and repeat the process.

A key element of the branch and bound algorithm is pruning parts of the tree in order to speed up the method. The condition that we use for pruning is when a lower bound is greater than the current best upper bound. Another important aspect is which node to select from the unexplored ones. We propose picking the one with the smallest lower bound. In Algorithm 2, we summarize the described spatial branch and bound procedure via RPT to obtain a global optimal solution to Problem (1).

Finally, note that if we introduce additional epigraphical variables $\boldsymbol{\tau}$, then branching can be performed based on the eigenvectors of the matrix $\begin{pmatrix} \boldsymbol{X} & \boldsymbol{V} \\ \boldsymbol{V}^\top & \boldsymbol{T} \end{pmatrix} - (\boldsymbol{x},\boldsymbol{\tau})(\boldsymbol{x},\boldsymbol{\tau})^\top$, where $(\boldsymbol{x},\boldsymbol{\tau})$ denotes the vector formed by stacking $\boldsymbol{x}$ and $\boldsymbol{\tau}$.

Algorithm 1 Branching strategy

Input: $(N, \mathcal{X}_N)$.

Output: (N_1, N_2).

1: Solve RPT relaxation at N and obtain optimal solutions $(\boldsymbol{x}^*, \boldsymbol{X}^*)$

2: Take f as the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{X}^* - \boldsymbol{x}^*(\boldsymbol{x}^*)^\top$

3: Take $l = \boldsymbol{f}^\top \boldsymbol{x}^*$

4: Take N_1 as $\left\{\boldsymbol{x} \in \mathcal{X}_N \mid \boldsymbol{f}^\top \boldsymbol{x} \leq l\right\}$

5: Take N_2 as $\left\{\boldsymbol{x} \in \mathcal{X}_N \mid \boldsymbol{f}^\top \boldsymbol{x} \geq l\right\}$

6: **return** N_1, N_2

4. A simple example

In this section we demonstrate the approach by solving the following toy problem:

$$\begin{array}{lll} \min\limits_{x_1,x_2,x_3} & 3x_1 - 3x_2 + 3x_3 + (x_1 + x_2 + 1)\exp(x_1) + (x_1 + x_2 + 1)\exp(x_3) & \\ \text{s.t.} & x_1 + x_2 \geq -1, & \\ & x_i \leq 10, & i \in \{1,2,3\}, \\ & \exp(x_2 - x_3) \leq x_1, & \\ & 2\exp\left(\frac{-x_1}{2}\right) + 2\exp\left(\frac{-x_2}{2}\right) \leq 2 + \exp(-1). & \end{array} \tag{12}$$

Algorithm 2 Branch and bound via RPT

Input: $(N_0, \mathrm{Lb}^0, \mathrm{Ub}^0, \delta)$.

Output: $(\boldsymbol{x}^*, \mathrm{Lb}, \mathrm{Ub})$.

1: $\mathrm{Lb} \leftarrow \mathrm{Lb}^0$
2: $\mathrm{Ub} \leftarrow \mathrm{Ub}^0$
3: $\mathrm{ACTIVE} \leftarrow \{N_0\}$
4: **while** $\mathrm{Ub} - \mathrm{Lb} > \delta$ **do**
5: $\quad j \leftarrow \arg\min_{i \in \mathrm{ACTIVE}} \mathrm{Lb}^i$
6: $\quad$ Partition node N_j into two child nodes N_{j_1}, N_{j_2} by applying Algorithm 1
7: $\quad$ **for** $i = 1, 2$ **do**
8: $\quad\quad$ Solve N_{j_i} by applying steps 1-4 and obtain Lb^{j_i} and Ub^{j_i}.
9: $\quad$ **end for**
10: $\quad \mathrm{Ub} \leftarrow \min\{\mathrm{Ub}^j, \mathrm{Ub}^{j_1}, \mathrm{Ub}^{j_2}\}$
11: $\quad$ **for** $i = 1, 2$ **do**
12: $\quad\quad$ **if** $\mathrm{Lb}^{j_i} < \mathrm{Ub}$ **then**
13: $\quad\quad\quad \mathrm{ACTIVE} \leftarrow \mathrm{ACTIVE} \cup \{j_i\}$
14: $\quad\quad$ **end if**
15: $\quad$ **end for**
16: $\quad \mathrm{Lb} \leftarrow \min\{\mathrm{Lb}^{j_1}, \mathrm{Lb}^{j_2}\}$
17: $\quad \mathrm{ACTIVE} \leftarrow \mathrm{ACTIVE} \setminus \{j\}$
18: **end while**

Let $\mathcal{X}_\mathrm{T}$ denote the feasible set of toy problem (12), consisting of a linear constraint and two convex exponential constraints. The objective is nonconvex, however it is SLC, hence we can apply the proposed framework to find the global optimum.

Linear × Linear. First, we perspectify the SLC objective. Next, the following constraints are generated:

$$
\begin{aligned}
&(x_1+x_2+1)^2 = x_1^2+x_2^2+2x_1x_2+2x_1+2x_2+1 \geq 0, && \\
&(x_i-10)(x_{i'}-10) = x_ix_{i'} - 10x_i - 10x_{i'} + 100 \geq 0, && i \leq i' \in \{1,2,3\}, \\
&(10-x_i)(x_1+x_2+1) = 10x_1+10x_2+10-x_ix_1-x_ix_2-x_i \geq 0, && i \in \{1,2,3\}, \\
&x_i^2 \geq 0, && i \in \{1,2,3\}.
\end{aligned}
$$

Finally, the product of variables x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3 and x_2x_3 in both the perspectified objective as well as the additional generated constraint are substituted by continuous variables X_{11}, X_{22},

X_{33}, X_{12}, X_{13}, and $X_{23} \in \mathbb{R}$ respectively to obtain the following convex relaxation:

$$\begin{array}{lll} \min\limits_{\boldsymbol{x},\boldsymbol{X}} & 3x_1 - 3x_2 + 3x_3 + (x_1+x_2+1)\exp\left(\frac{X_{11}+X_{12}+x_1}{x_1+x_2+1}\right) + (x_1+x_2+1)\exp\left(\frac{X_{13}+X_{23}+x_3}{x_1+x_2+1}\right) & \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X}_{\mathrm{T}}, & \\ & X_{11} + 2X_{12} + X_{22} + 2x_1 + 2x_2 + 1 \geq 0, & \\ & X_{ii'} - 10x_i - 10x_{i'} + 100 \geq 0, & i \leq i' \in \{1,2,3\}, \\ & 10x_1 + 10x_2 + 10 - X_{i1} - X_{i2} - x_i \geq 0, & i \in \{1,2,3\}, \\ & X_{ii} \geq 0, & i \in \{1,2,3\}. \end{array} \tag{13}$$

The solution of (13) appears to be

$$\boldsymbol{x}' = \begin{bmatrix} 1 \\ 1.10 \\ 1.10 \end{bmatrix} \quad \text{and} \quad \boldsymbol{X}' = \begin{bmatrix} X'_{11} & X'_{12} & X'_{13} \\ X'_{21} & X'_{22} & X'_{23} \\ X'_{31} & X'_{32} & X'_{33} \end{bmatrix} = \begin{bmatrix} 12.94 & -48.08 & -41.10 \\ -48.08 & 78.01 & -40.07 \\ -41.10 & -40.07 & 0 \end{bmatrix},$$

with objective value 3, which constitutes a lower bound on the optimal value of (12). Since $\mathcal{X}_{\mathrm{T}}$ consists of only convex constraints, the obtained $\boldsymbol{x}'$ is contained in the set of feasible candidate solutions to (12), and its corresponding objective value is 20.796, which constitutes an upper bound on the optimal value of (12).

Linear × Convex. Let $\mathcal{X}_{\mathrm{TLL}}$ denote the feasible set of (13). We pairwise multiply the linear with the nonlinear constraints and obtain the SLC constraints

$$\begin{array}{ll} (x_1+x_2+1)\exp(x_2-x_3) \leq (x_1+x_2+1)x_1, & \\ (x_1+x_2+1)2\exp\left(\dfrac{-x_1}{2}\right) + (x_1+x_2+1)2\exp\left(\dfrac{-x_2}{2}\right) \leq (x_1+x_2+1)(2+\exp(-1)), & \\ (10-x_i)\exp(x_2-x_3) \leq (10-x_i)x_1, & i \in \{1,2,3\}, \\ (10-x_i)2\exp\left(\dfrac{-x_1}{2}\right) + (10-x_i)2\exp\left(\dfrac{-x_2}{2}\right) \leq (10-x_i)(2+\exp(-1)), & i \in \{1,2,3\}. \end{array}$$

Next, the nonconvex components in the LHS of the above SLC constraints can be reformulated as:

$$\begin{array}{ll} (x_1+x_2+1)\exp(x_2-x_3) = (x_1+x_2+1)\exp\left(\dfrac{x_1x_2 - x_1x_3 + x_2^2 - x_2x_3 + x_2 - x_3}{x_1+x_2+1}\right), & \\ (x_1+x_2+1)2\exp\left(\dfrac{-x_1}{2}\right) = 2(x_1+x_2+1)\exp\left(\dfrac{-x_1^2 - x_1x_2 - x_1}{2(x_1+x_2+1)}\right), & \\ (x_1+x_2+1)2\exp\left(\dfrac{-x_2}{2}\right) = 2(x_1+x_2+1)\exp\left(\dfrac{-x_1x_2 - x_2^2 - x_2}{2(x_1+x_2+1)}\right), & \\ (10-x_i)\exp(x_2-x_3) = (10-x_i)\exp\left(\dfrac{10x_2 - 10x_3 - x_ix_2 + x_ix_3}{10-x_i}\right), & i \in \{1,2,3\}, \\ (10-x_i)2\exp\left(\dfrac{-x_1}{2}\right) = 2(10-x_i)\exp\left(\dfrac{-10x_1 + x_ix_1}{2(10-x_i)}\right), & i \in \{1,2,3\}, \\ (10-x_i)2\exp\left(\dfrac{-x_2}{2}\right) = 2(10-x_i)\exp\left(\dfrac{-10x_2 + x_ix_2}{2(10-x_i)}\right), & i \in \{1,2,3\}. \end{array}$$

Finally, all the product of variables x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3 and x_2x_3 are substituted with newly introduced variables X_{11}, X_{22}, X_{33}, X_{12}, X_{13}, and X_{23} respectively. The convex relaxation that results from the RPT approach is therefore:

$$\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{X}} \quad & 3x_1 - 3x_2 + 3x_3 + (x_1+x_2+1)\exp\left(\frac{X_{11}+X_{12}+x_1}{x_1+x_2+1}\right) \\
& + (x_1+x_2+1)\exp\left(\frac{X_{13}+X_{23}+x_3}{x_1+x_2+1}\right) \\
\text{s.t.} \quad & \boldsymbol{x} \in \mathcal{X}_{\text{TLL}}, \\
& (x_1+x_2+1)\exp\left(\frac{X_{12}-X_{13}+X_{22}-X_{23}+x_2-x_3}{x_1+x_2+1}\right) \le X_{11}+X_{12}+x_1, \\
& 2(x_1+x_2+1)\exp\left(\frac{-X_{11}-X_{12}-x_1}{2(x_1+x_2+1)}\right) + 2(x_1+x_2+1)\exp\left(\frac{-X_{12}-X_{22}-x_2}{2(x_1+x_2+1)}\right) \\
& \qquad \le (2+\exp(-1))(x_1+x_2+1), \\
& (10-x_i)\exp\left(\frac{10x_2-10x_3-X_{2i}+X_{3i}}{10-x_i}\right) \le 10x_1 - X_{1i}, \qquad i\in\{1,2,3\}, \\
& 2(10-x_i)\exp\left(\frac{-10x_1+X_{1i}}{2(10-x_i)}\right) + 2(10-x_i)\exp\left(\frac{-10x_2+X_{2i}}{2(10-x_i)}\right) \\
& \qquad \le (2+\exp(-1))(10-x_i), \qquad i\in\{1,2,3\}.
\end{aligned} \tag{14}$$

The solution of (14) is

$$\boldsymbol{x}' = \begin{bmatrix}1.17\\0.93\\0.77\end{bmatrix} \quad \text{and} \quad \boldsymbol{X}' = \begin{bmatrix}X'_{11} & X'_{12} & X'_{13}\\ X'_{21} & X'_{22} & X'_{23}\\ X'_{31} & X'_{32} & X'_{33}\end{bmatrix} = \begin{bmatrix}1.77 & 0.75 & 0.23\\ 0.75 & 1.16 & 1.31\\ 0.23 & 1.31 & 0.57\end{bmatrix},$$

with objective value 19.778, which constitutes a tighter lower bound on the optimal value of (12) than (13). The obtained $\boldsymbol{x}'$ is contained in the set of feasible candidate solutions to (12), and its corresponding objective value is 19.809, which constitutes a tighter upper bound on the optimal value of (12) than (13).

Set of candidate solutions. We have the following candidate solutions:

$$\boldsymbol{x}' = \begin{pmatrix}1.17\\0.93\\0.77\end{pmatrix}, \; \boldsymbol{x}_1^X = \begin{pmatrix}1.51\\0.64\\0.20\end{pmatrix}, \; \boldsymbol{x}_2^X = \begin{pmatrix}0.81\\1.25\\1.22\end{pmatrix}, \; \boldsymbol{x}_3^X = \begin{pmatrix}0.30\\1.47\\0.61\end{pmatrix}.$$

Observe that only $\boldsymbol{x}'$ is feasible, hence the set of candidate feasible solutions is given by $\mathcal{X}'' = \{\boldsymbol{x}'\}$.

Branch and Bound. We have $\text{Ub}^0 = 19.809$ and $\text{Lb}^0 = 19.778$ and $\boldsymbol{x}^0 = \boldsymbol{x}'$, $\boldsymbol{X}^0 = \boldsymbol{X}'$. The eigenvector corresponding to the largest eigenvalue of the matrix $\boldsymbol{X}^0 - \boldsymbol{x}^0(\boldsymbol{x}^0)^\top$ is $\boldsymbol{f} = (0.62, -0.54, -0.56)^\top$. We have $l = \boldsymbol{f}^\top \boldsymbol{x}^0 = -0.21$ and we obtain the hyperplane $H = \{\boldsymbol{x} \in \mathbb{R}^3 \mid 0.62x_1 - 0.54x_2 - 0.56x_3 = -0.21\}$. Hence we create two child nodes N_1 and N_2 from the root-node N_0 such that N_1 represents Problem (12) in which the feasible region $\mathcal{X}_{\text{T}}$ is intersected with $\mathcal{X}_l^0 = \{\boldsymbol{x} \in \mathcal{X}_{\text{T}} \mid 0.62x_1 - 0.54x_2 -$

$0.56x_3 \leq -0.21\}$ and N_2 represents Problem (12) in which the feasible region $\mathcal{X}_{\mathrm{T}}$ is intersected with $\mathcal{X}_r^0 = \{\boldsymbol{x} \in \mathcal{X}_{\mathrm{T}} \mid 0.62x_1 - 0.54x_2 - 0.56x_3 \geq -0.21\}$. We apply steps 1-4 on N_1 and N_2 and obtain:

$$\mathrm{Lb}_1 = 19.787,\ \mathrm{Ub}_1 = 19.788\ ,\ \mathrm{Lb}_2 = 19.790,\ \mathrm{Ub}_2 = 19.790.$$

We set $\mathrm{UB} = \min\{\mathrm{Ub}_0, \mathrm{Ub}_1, \mathrm{Ub}_2\} = 19.788$. Moreover, node N_1 becomes active, node N_2 remains inactive, since $\mathrm{Lb}_1 > \mathrm{Ub}$, and we delete node N_0 from the list of active nodes, i.e., $\mathrm{ACTIVE} = \{N_1\}$. We set $\mathrm{LB} = \min\{\mathrm{Lb}_1, \mathrm{Lb}_2\} = 19.787$.

Since UB - LB $= 0.001 > \delta$, we select node N_2 from the list of active nodes. Let $(\boldsymbol{x}^1, \boldsymbol{X}^1)$ denote the optimal solution at node N_1. The eigenvector corresponding to the largest eigenvalue of the matrix $\boldsymbol{X}^1 - \boldsymbol{x}^1(\boldsymbol{x}^1)^\top$ is $\boldsymbol{f} = (0.78, -0.58, -0.22)^\top$. We have $l = \boldsymbol{f}^\top \boldsymbol{x}^1 = 0.23$ and we obtain the hyperplane $H = \{\boldsymbol{x} \in \mathbb{R}^3 \mid 0.78x_1 - 0.58x_2 - 0.22x_3 = 0.23\}$. Hence we create two child nodes N_3 and N_4 from N_1 such that N_3 represents Problem (12) in which the feasible region $\mathcal{X}$ is intersected with $\mathcal{X}_l^1 = \{\boldsymbol{x} \in \mathcal{X}_l^0 \mid 0.78x_1 - 0.58x_2 - 0.22x_3 \leq 0.23\}$ and N_4 represents Problem (12) in which the feasible set $\mathcal{X}$ is intersected with $\mathcal{X}_r^1 = \{\boldsymbol{x} \in \mathcal{X}_l^0 \mid 0.78x_1 - 0.58x_2 - 0.22x_3 \geq 0.23\}$. We apply steps 1-4 on N_3 and N_4 and obtain:

$$\mathrm{Lb}_3 = 19.787,\ \mathrm{Ub}_3 = 19.787\ ,\ \mathrm{Lb}_4 = 19.787,\ \mathrm{Ub}_4 = 19.787.$$

We set $\mathrm{UB} = 19.787$ and $\mathrm{LB} = 19.787$ and therefore obtain the optimal solution

$$\boldsymbol{x}' = \begin{bmatrix} 1.18 \\ 0.92 \\ 0.75 \end{bmatrix} \quad \text{and} \quad \boldsymbol{X}' = \begin{bmatrix} X'_{11} & X'_{12} & X'_{13} \\ X'_{21} & X'_{22} & X'_{23} \\ X'_{31} & X'_{32} & X'_{33} \end{bmatrix} = \begin{bmatrix} 1.41 & 1.09 & 0.89 \\ 1.09 & 0.85 & 0.69 \\ 0.89 & 0.69 & 0.56 \end{bmatrix},$$

with optimal objective value 19.787.

5. Improving the RPT-BB approach

In this section, we describe several ways to strengthen the RPT-BB approach.

5.1. Multiplying with redundant linear constraints

We show that adding linear constraints that are redundant to existing linear constraints does not tighten the RPT relaxation, while adding linear constraints that are redundant to existing nonlinear constraints might be useful.

Definition 2 (Redundant inequality constraints). An inequality constraint $f(\boldsymbol{x}) \leq 0$, where $f : \mathbb{R}^{n_x} \to (-\infty, +\infty]$, is *redundant* to a nonempty set $\mathcal{X} \subseteq \mathbb{R}^{n_x}$, if $\mathcal{X} \subseteq \{\boldsymbol{x} \mid f(\boldsymbol{x}) \leq 0\}$.

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

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

Here, $f : \mathbb{R}^{n_x} \to (-\infty, +\infty]$ is proper, closed and convex; $\boldsymbol{b}_\ell$ is the ℓ-th column of the matrix $\boldsymbol{B}$; and the last two inequalities result from pairwise multiplication of the linear constraints $p_\ell - \boldsymbol{b}_\ell^\top \boldsymbol{x}$ with itself and $f(\boldsymbol{x}) \le 0$, respectively.

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

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

coincide and both are nonnegative thanks to the strong duality of linear optimization and the redundancy of $\boldsymbol{d}^\top \boldsymbol{x} \le q$ to $\{\boldsymbol{x} \mid \boldsymbol{B}^\top \boldsymbol{x} \le \boldsymbol{p}\}$, 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 $\boldsymbol{b}_y = \boldsymbol{B}\boldsymbol{y} = \boldsymbol{d}$ and $p_y = \boldsymbol{p}^\top \boldsymbol{y} \le q$. 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

$$(q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{(q - \boldsymbol{d}^\top \boldsymbol{x})\boldsymbol{x}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) \le 0 \quad \text{and} \quad (p_y - \boldsymbol{b}_y^\top \boldsymbol{x}) f\left(\frac{(p_y - \boldsymbol{b}_y^\top \boldsymbol{x})\boldsymbol{x}}{p_y - \boldsymbol{b}_y^\top \boldsymbol{x}}\right) \le 0.$$

Moreover, for any $\boldsymbol{X} \in \mathbb{S}^{n_x}$ we have

$$(q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{q\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) \le 0 \iff (p_y - \boldsymbol{b}_y^\top \boldsymbol{x}) f\left(\frac{p_y \boldsymbol{x} - \boldsymbol{X}\boldsymbol{b}_y}{p_y - \boldsymbol{b}_y^\top \boldsymbol{x}}\right) \le 0,$$

because $\boldsymbol{b}_y = \boldsymbol{d}$ so that $\boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{d} = \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{b}_y$ and $\boldsymbol{X}\boldsymbol{d} = \boldsymbol{X}\boldsymbol{b}_y$. Notice that

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

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 $(q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{q\boldsymbol{x} - \boldsymbol{X}\boldsymbol{d}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) \le 0$ is redundant to

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

Thus, the claim follows. □

Note that adding linear constraints that are redundant to existing nonlinear constraints might be useful, as is demonstrated in Example 5.

EXAMPLE 5. Consider the following nonconvex problem:

$$\begin{array}{ll}\min\limits_{x_1,x_2} & 2x_1+3x_2-5x_1x_2-(x_1+2)\ln(x_1+2)\\ \text{s.t.} & x_1+x_2\le 1,\\ & \exp(-x_1)+\exp(-x_2)\le 1+\exp(-1).\end{array} \tag{15}$$

Note that $\ln(x_1+2)$ is well defined if $x_1>-2$, which is ensured by the second inequality. The objective contains a sum of two SLC functions, those are, $-5x_1x_2$ and $-(x_1+2)\ln(x_1+2)$. The obtained convex relaxation of Problem (15) from RPT without the optional SDP relaxation has an objective value of -35.17. The obtained optimal solution $\boldsymbol{x}'=(1,0)^\top$ is a feasible solution to (15), and its corresponding objective value is 1.30, which constitutes an upper bound on the optimal value of (15).

The linear constraints $x_1\ge -1$ and $x_2\ge -1$ are redundant to the second inequality. However, adding those constraints to (15) and subsequently applying RPT results in a convex relaxation of Problem (15) with objective value of -4.47. Again, the obtained optimal solution $\boldsymbol{x}'=(0.5,0.5)^\top$ is a feasible solution to (15), and its corresponding objective value is -1.04, which constitutes an upper bound on the optimal value of (15). Hence, by adding the redundant linear constraints $x_1\ge -1$ and $x_2\ge -1$, we obtain a tighter lower- and upper bound on the optimal objective value of (15). □

5.2. Handling biconvex problems with convex constraints

We consider the following generic instance of (1):

$$\begin{array}{ll}\min\limits_{\boldsymbol{x},\boldsymbol{y}} & c_0(\boldsymbol{x})+\sum_k r_k(\boldsymbol{y})c_k(\boldsymbol{x})\\ \text{s.t.} & \boldsymbol{x}\in\mathcal{X},\boldsymbol{y}\in\mathcal{Y},\end{array} \tag{16}$$

where the functions $r_k(\cdot)$ and $c_k(\cdot)$ are assumed to be convex, and $r_k(\boldsymbol{y}),c_k(\boldsymbol{x})\ge 0$. We show that the constraints in the RPT relaxation of (16) resulting from pairwise multiplication in the same set, i.e., either in $\mathcal{X}$ or in $\mathcal{Y}$, are redundant as long as the LMI as given in Section 3.3 is not included.

LEMMA 4. *The additional constraints in the RPT relaxation of (16) resulting from pairwise multiplication in $\mathcal{X}\times\mathcal{X}$ and $\mathcal{Y}\times\mathcal{Y}$ are redundant, if the additional constraints resulting from pairwise multiplication in $\mathcal{X}\times\mathcal{Y}$ are included, and the following LMI*

$$\begin{pmatrix}\boldsymbol{X} & \boldsymbol{U} & \boldsymbol{x}\\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{y}\\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & 1\end{pmatrix}\succeq \boldsymbol{0},$$

where $\boldsymbol{x}\boldsymbol{x}^\top$, $\boldsymbol{x}\boldsymbol{y}^\top$ and $\boldsymbol{y}\boldsymbol{y}^\top$ are linearized by $\boldsymbol{X}$, $\boldsymbol{U}\in\mathbb{R}^{n_x\times n_y}$ and $\boldsymbol{Y}\in\mathbb{R}^{n_y\times n_y}$, respectively, is not included in the RPT relaxation.

Proof. Notice that any feasible solution for the problem involving all constraint multiplications is also feasible for the one involving only those in $\mathcal{X} \times \mathcal{Y}$. On the other hand, if a solution is feasible for the problem involving only the multiplications in $\mathcal{X} \times \mathcal{Y}$, since we are not using the LMI, we can take $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^\top$, $\boldsymbol{Y} = \boldsymbol{y}\boldsymbol{y}^\top$ and therefore have a feasible solution for the problem involving all multiplications. Therefore, we conclude that the two formulations are equivalent, which shows that the constraint multiplications in $\mathcal{X} \times \mathcal{X}$ and $\mathcal{Y} \times \mathcal{Y}$ are redundant. $\square$

Observe that the problem of minimizing a DC function is a special case of (16) where $\mathcal{Y} = \mathrm{dom}(c_1^*)$ and both r_k and c_k are linear, as a result of the biconjugate reformulation as explained in Example 1. Finally, we have the following remark in terms of branching, when we solve Problem (16) with Algorithm 2.

REMARK 2. When we apply Algorithm 2 to minimization problems of the form (16) in which we include the LMI, we typically generate hyperplanes by computing the eigenvector corresponding to the largest eigenvalue of $\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top$, i.e., we only generate hyperplanes in the $\boldsymbol{x}$-space. Alternatively, hyperplanes can be generated in the $\boldsymbol{x}\boldsymbol{y}$-space by considering the eigenvectors of the matrix

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} \\ \boldsymbol{U}^\top & \boldsymbol{Y} \end{pmatrix} - (\boldsymbol{x}, \boldsymbol{y})\,(\boldsymbol{x}, \boldsymbol{y})^\top,$$

where $(\boldsymbol{x}, \boldsymbol{y})$ denotes the vector formed by stacking $\boldsymbol{x}$ and $\boldsymbol{y}$. We observe that also generating hyperplanes in the $\boldsymbol{y}$-space adds many more constraints in each branch-and-bound iteration, which can increase the computation time at each successive child node. On the other hand, generating hyperplanes in the $\boldsymbol{y}$-space might reduce the total number of hyperplanes that need to be generated, potentially decreasing the overall computation time. The question of whether generating hyperplanes in the $\boldsymbol{y}$-space could be beneficial is left for future research. Finally, we note that when we apply Algorithm 2 to minimization problems of the form (16) in which we do not include the LMI, one can branch based on the left and right eigenvalues of $\boldsymbol{U}$. $\square$

5.3. Strengthening upper bounds with eigenvectors of the optimal solution

At optimality we will always have $\boldsymbol{X} = \boldsymbol{x}\boldsymbol{x}^\top$. If we multiply both sides with $\boldsymbol{x}$ we obtain $\boldsymbol{X}\boldsymbol{x} = (\boldsymbol{x}^\top\boldsymbol{x})\boldsymbol{x}$. Therefore, we notice that $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{X}$ with corresponding eigenvalue $\boldsymbol{x}^\top\boldsymbol{x}$. Hence, we can add the eigenvectors of $\boldsymbol{X}$ to the set of candidate feasible solutions $\mathcal{X}'$, as described in Section 3.4. Example 6 illustrates a case where the tightest upper bound can be obtained from an eigenvector of $\boldsymbol{X}$.

EXAMPLE 6. Consider the following toy problem:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & (x_1 + x_2 + 1)\exp(x_2) \\ \text{s.t.} & x_1 + x_2 + 1 \geq 0, \\ & x_1 x_2 \geq -1. \end{array} \tag{17}$$

The optimal solution of (17) is $(-1,0)$ with optimal value 0. After applying RPT we obtain the following relaxation

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & (x_1+x_2+1)\exp\left(\frac{X_{12}+X_{22}+x_2}{x_1+x_2+1}\right) \\ \text{s.t.} & x_1+x_2+1\geq 0, \\ & X_{12}\geq -1, \\ & X_{11}+X_{22}+2X_{12}+2x_1+2x_2+1\geq 0. \end{array} \tag{18}$$

The optimal solution of (18) is

$$\boldsymbol{x}^* = \begin{bmatrix} 0.71 \\ -1.57 \end{bmatrix} \quad \text{and} \quad \boldsymbol{X}^* = \begin{bmatrix} 2.73 & -1.00 \\ -1.00 & 0.00 \end{bmatrix},$$

with optimal value 0, which gives us a lower bound for the optimal value of (17). We have the following candidate vectors

$$\boldsymbol{x}_1^X = \frac{\boldsymbol{X}_1^*}{x_1^*} = \begin{bmatrix} 3.85 \\ -1.41 \end{bmatrix}, \quad \boldsymbol{x}_2^X = \frac{\boldsymbol{X}_2^*}{x_2^*} = \begin{bmatrix} 0.64 \\ 0.00 \end{bmatrix}.$$

The eigenvectors of $\boldsymbol{X}^*$ are

$$\boldsymbol{x}_3^{EV} = \begin{bmatrix} -0.31 \\ -0.95 \end{bmatrix}, \quad \boldsymbol{x}_4^{EV} = \begin{bmatrix} -0.95 \\ 0.31 \end{bmatrix}.$$

We observe that $\boldsymbol{x}^*$ is infeasible as $x_1^* x_2^* = -1.11 < -1$. Moreover, $\boldsymbol{x}_1^X$ is infeasible as $(x_1^X)_1(x_1^X)_2 = -5.44 < -1$ and $\boldsymbol{x}_3^X$ is infeasible as $(x_3^{EV})_1 + (x_3^{EV})_2 + 1 = -0.26 < 0$. Finally, we notice that $\boldsymbol{x}_2^X$ is feasible and gives an upper bound of 1.64, while $\boldsymbol{x}_4^{EV}$ is also feasible and gives an upper bound of 0.49. Therefore, in this example the tightest upper bound is obtained from the second eigenvector of $\boldsymbol{X}^*$. □

6. Convergence analysis of the RPT-BB approach

In the spatial B&B approach the feasible region in each leaf is becoming smaller and smaller. However, for convergence we need that the feasible region of that leaf is becoming smaller and smaller in each coordinate direction. Indeed, adding cuts via the eigenvector branching strategy does not necessarily decreases the feasible region in each coordinate direction. Suppose, for example that we constantly add hyperplanes that are more or less parallel to one of the constraints. Hence, for convergence we need additional cutting planes, and that is stated in the following adaptation to Algorithm 2.

Adaptation A: For a leaf in depth $j \in \bar{\mathcal{J}}$ of the B&B tree, where $\bar{\mathcal{J}} = \{1d, 2d, \cdots, Jd\}$ and $d \in \mathbb{Z}_{++}$, we calculate the corresponding range of $\boldsymbol{x}$ by solving $x_i^{max} = \max x_i$ and $x_i^{min} = \min x_i$, for all $i \in \{1,\ldots,n_x\}$, subject to the feasible region of this leaf. We then separate the feasible region of this leaf by adding the hyperplane $x_{i'} = \frac{1}{2}(x_{i'}^{\max} + x_{i'}^{\min})$, where i' is the entry of $\boldsymbol{x}$ for which the value is the largest, i.e., $i' = \max_{i\in\{1,\ldots,n_x\}} x_i^{\max} + x_i^{\min}$, instead of the hyperplane proposed in Section 3.5. We apply this adaptation to all the leaves in depth j of the B&B tree for every $j \in \bar{\mathcal{J}}$.

Note that the index set $\bar{\mathcal{J}}$ contains integers that are multiples of $d \in \mathbb{Z}_{++}$. Thanks to Assumption 1, there exists a n_x-dimensional box that contains the feasible region of each leaf in depth j of the B&B tree for every $j \in \bar{\mathcal{J}}$.

THEOREM 3. *If for each $i \in \mathcal{I}$ and $k \in \mathcal{K}_0$, the function c_{ik} and its corresponding recession function $\delta^*_{\mathrm{dom}(c^*_{ik})}$ are Lipschitz continuous, then the spatial B&B Algorithm 2 with Adaptation A converges to a global optimal solution of Problem (1).*

Proof. The proof consists of three steps:
Step 1) We first show that as $j \to \infty$, the feasible region of each leaf in depth j of the B&B tree becomes smaller and smaller. Indeed, it follows from Adaptation A that the feasible region of a leaf in depth j of the B&B tree is contained in a n_x-dimensional box:

$$\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{\alpha}\|_\infty \leq \epsilon\}, \tag{19}$$

for $\boldsymbol{\alpha} = \frac{1}{2}(\boldsymbol{x}^{max} + \boldsymbol{x}^{min})$ and $\epsilon = \max_i\{x_i^{max} - x_i^{min}\}$. Note that $\epsilon \to 0$ as $j \to \infty$, as we branch along a given coordinate. Initially, we may continue branching along this coordinate for several iterations. However, after a finite number of iterations, we switch to branching along a different coordinate.

Step 2) In this step we show that

$$|X_{ij} - x_i x_j| \leq 4\epsilon(|\alpha_j| + \epsilon). \tag{20}$$

To prove this, first observe that for $i, j \in \{1, \ldots, n_x\}$, we have

$$x_i - \alpha_i \leq \epsilon \tag{21}$$

$$x_i - \alpha_i \geq -\epsilon \tag{22}$$

$$x_j - \alpha_j \geq -\epsilon \tag{23}$$

$$x_j - \alpha_j \leq \epsilon. \tag{24}$$

Clearly, Constraints (21) – (24) are redundant with respect to the corresponding feasible region of the leaf of the B&B tree. It follows from Theorem 2 that the inequalities that are obtained after multiplying these redundant constraints and perspectification are also redundant. We multiply the inequalities (22) and (23) and apply perspectification to obtain:

$$0 \leq (x_i - \alpha_i + \epsilon)(x_j - \alpha_j + \epsilon) = X_{ij} - \alpha_i x_j + \epsilon x_j + (x_i - \alpha_i + \epsilon)(-\alpha_j + \epsilon).$$

From this inequality we obtain

$$X_{ij} - x_i x_j \geq (\alpha_i - x_i)x_j - \epsilon x_j - (x_i - \alpha_i + \epsilon)(-\alpha_j + \epsilon) \geq -2\epsilon(|\alpha_j| + \epsilon) - (\epsilon + \epsilon)(|\alpha_j| + \epsilon) = -4\epsilon(|\alpha_j| + \epsilon).$$

By multiplying (22) and (24) we can prove in a similar way:

$$X_{ij} - x_i x_j \le 4\epsilon(|\alpha_j| + \epsilon).$$

Hence, we obtain (20).

Step 3) In this step we prove that also the perspective approximation of the linear × convex function converges to the right value. Since each function c_{ik} is Lipschitz continuous with Lipschitz constant L_{ik}, we have when $q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x} > 0$,

$$\begin{aligned}
\tilde{f}_k(\boldsymbol{x}, \boldsymbol{X}') - \tilde{f}_k(\boldsymbol{x}, \boldsymbol{X}'') &= \sum_i (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) \left(c_{ik}\left(\frac{q_{ik}\boldsymbol{x} - \boldsymbol{X}'\boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}} \right) - c_{ik}\left(\frac{q_{ik}\boldsymbol{x} - \boldsymbol{X}''\boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}} \right) \right) \\
&\le \sum_i (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) L_{ik} \left\| \frac{q_{ik}\boldsymbol{x} - \boldsymbol{X}'\boldsymbol{d}_{ik} - q_{ik}\boldsymbol{x} + \boldsymbol{X}''\boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}} \right\| \\
&= \sum_i L_{ik} \|(\boldsymbol{X}' - \boldsymbol{X}')\boldsymbol{d}_{ik}\| \\
&\le \sum_i L_{ik} \|\boldsymbol{X} - \boldsymbol{X}''\| \|\boldsymbol{d}_{ik}\| \\
&\le \tilde{L_k} \|\boldsymbol{X}' - \boldsymbol{X}''\|,
\end{aligned}$$

where $\tilde{L_k} = \sum_i L_{ik}\|\boldsymbol{d}_{ik}\|$. In particular if we take $\boldsymbol{X}' = \boldsymbol{X}$ and $\boldsymbol{X}'' = \boldsymbol{x}\boldsymbol{x}^\top$, and using (20) we get

$$\tilde{f}_k(\boldsymbol{x}, \boldsymbol{X}) - \tilde{f}_k(\boldsymbol{x}, \boldsymbol{x}\boldsymbol{x}^\top) \le \tilde{L_k}\|\boldsymbol{X} - \boldsymbol{x}\boldsymbol{x}^\top\| \le 4\epsilon n(\tilde{\alpha} + \epsilon)\tilde{L_k},$$

where $\tilde{\alpha} = \max_j |\alpha_j|$. This means that $\tilde{f}_k(\boldsymbol{x}, \boldsymbol{X}) \to \tilde{f}_k(\boldsymbol{x}, \boldsymbol{x}\boldsymbol{x}^\top)$ when $\epsilon \to 0$. Analogously, when $q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x} = 0$, a similar result can be obtained by using the fact the recession function is assumed to be Lipschitz continuous. □

For the theorem we need Lipschitz continuity for the recession functions of c_{ik}. The following example shows that convergence may not hold without this assumption:

$$c(x) = xe^x,\ x \ge 0.$$

The perspective approximation is:

$$\tilde{c}(x, x') = xe^{x'/x}.$$

Now take $x = 0$. Even if $x' \ge 0$ is very close to $x = 0$, we have $\tilde{c}(0, x') = \infty$.

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 $\mathcal{X}_k = \{\boldsymbol{x} \mid \boldsymbol{h}_k(\boldsymbol{x}) \leq \boldsymbol{0}\}$, $k \in \mathcal{K}$ is:

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

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

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

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 (2005, 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

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & f(\boldsymbol{x}) \\ \text{s.t.} & \boldsymbol{x} \in \bigcup\limits_{k\in\mathcal{K}} \mathcal{X}_k, \end{array} \tag{DP}$$

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:

$$\begin{array}{lll} \min\limits_{\boldsymbol{y},\{\boldsymbol{u}_k\}_k} & f\left(\sum\limits_{k\in\mathcal{K}} \boldsymbol{u}_k\right) & \\ \text{s.t.} & y_k \boldsymbol{h}_k(\boldsymbol{u}_k/y_k) \leq \boldsymbol{0}, & k \in \mathcal{K}, \\ & \sum\limits_{k\in\mathcal{K}} y_k = 1, & \\ & y_k \geq 0, & k \in \mathcal{K}, \end{array}$$

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):

$$\begin{aligned} \min_{\boldsymbol{y},\{\boldsymbol{x}_k\}_k} \quad & \boldsymbol{c}^\top \boldsymbol{y} \\ \text{s.t.} \quad & \sum_{k\in\mathcal{K}_0} \boldsymbol{x}_k y_k \leq \boldsymbol{b}, \\ & \boldsymbol{y} \geq \boldsymbol{0}, \\ & \boldsymbol{x}_k \in \mathcal{X}_k, \qquad\qquad k \in \mathcal{K}_0, \end{aligned} \tag{GLP}$$

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:

$$\begin{aligned} \min_{\boldsymbol{y},\{\boldsymbol{u}_k\}_k} \quad & \boldsymbol{c}^\top \boldsymbol{y} \\ \text{s.t.} \quad & \sum_{k\in\mathcal{K}_0} \boldsymbol{u}_k \leq \boldsymbol{b}, \\ & y_k \boldsymbol{h}_k(\boldsymbol{u}_k/y_k) \leq \boldsymbol{0}, \qquad k \in \mathcal{K}_0, \\ & \boldsymbol{y} \geq \boldsymbol{0}. \end{aligned}$$

The convex problem 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., 2023, 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. Convex hull representation for 0-1 mixed-integer convex programs

Consider the following mixed-integer 0-1 constrained set (Sherali and Adams, 2009):

$$\mathcal{X} = \left\{ \boldsymbol{x} \equiv (\boldsymbol{x}_B, \boldsymbol{x}_C) \;\middle|\; \begin{array}{ll} h_k(\boldsymbol{x}) \leq 0, & k \in \mathcal{K} \\ x_j \in \{0,1\}, & j \in \mathcal{B} \\ 0 \leq x_j \leq 1, & j \in \mathcal{C} \end{array} \right\},$$

where the vector $\boldsymbol{x} \in \mathbb{R}^{n_x}$ is expressed as the concatenation of the binary vector $\boldsymbol{x}_B \in \{0,1\}^{|\mathcal{B}|}$ and the continuous vector $\boldsymbol{x}_C \in [0,1]^{|\mathcal{C}|}$, with $n_x = |\mathcal{B}| + |\mathcal{C}|$ and $\mathcal{B} \cap \mathcal{C} = \emptyset$. Moreover, $h_k : \mathbb{R}^{n_x} \to \mathbb{R}$ are bounded-valued, convex, and differentiable functions and the set $\{\boldsymbol{x} \mid h_k(\boldsymbol{x}) \leq 0\}$ is nonempty and compact for every $k \in \mathcal{K}$. Similarly as in Sherali and Adams (2009), we can define the polynomial factors

$$F_{\mathcal{J}}(\boldsymbol{x}_B) = \prod_{j\in\mathcal{J}} x_j \prod_{j\in\mathcal{B}\setminus\mathcal{J}} (1 - x_j),$$

where $\mathcal{J} \subseteq \mathcal{B}$, and linearize them by a newly introduced variable $\lambda_{\mathcal{J}}$. We can apply partial RPT in which we generate additional redundant nonconvex constraints by multiplying each of the defining inequalities in $\mathcal{X}$ by each $\lambda_{\mathcal{J}}$, for $\mathcal{J} \subseteq \mathcal{B}$. We then obtain the following representation where the binary restrictions on x_j for $j \in \mathcal{B}$ are now relaxed

$$\mathcal{X}_{\mathrm{R}} = \left\{ \boldsymbol{x} \in \mathbb{R}^{n_x}, \boldsymbol{\lambda} \in \mathbb{R}^{2^{|\mathcal{B}|}} \;\middle|\; \begin{array}{lr} \lambda_{\mathcal{J}} h_k(\boldsymbol{x}) \leq 0, & k \in \mathcal{K}, \mathcal{J} \subseteq \mathcal{B} \\ \lambda_{\mathcal{J}} \geq 0, & \mathcal{J} \subseteq \mathcal{B} \\ 0 \leq x_j \lambda_{\mathcal{J}} \leq \lambda_{\mathcal{J}}, & j \in \mathcal{C}, \mathcal{J} \subseteq B \end{array} \right\}. \tag{25}$$

Subsequently convexifying all nonconvex components in $\mathcal{X}_{\mathrm{R}}$ by reformulating them in their perspective form and linearizing all product terms we obtain

$$\mathcal{X}_{\mathrm{RPT}} = \left\{ \boldsymbol{\lambda} \in \mathbb{R}^{2^{|\mathcal{B}|}}, \boldsymbol{y} \in \mathbb{R}^{n_x \cdot 2^{|\mathcal{B}|}} \;\middle|\; \begin{array}{lr} \lambda_{\mathcal{J}} h_k\left(\frac{\boldsymbol{y}_{\mathcal{J}}}{\lambda_{\mathcal{J}}}\right) \leq 0, & k \in \mathcal{K}, \mathcal{J} \subseteq \mathcal{B} \\ \lambda_{\mathcal{J}} \geq 0, & \mathcal{J} \subseteq \mathcal{B} \\ 0 \leq (y_{\mathcal{J}})_j \leq \lambda_{\mathcal{J}}, & j \in \mathcal{C}, \mathcal{J} \subseteq \mathcal{B} \end{array} \right\}, \tag{26}$$

where $x_j \lambda_{\mathcal{J}}$ is linearized by $(y_{\mathcal{J}})_j$ for $j \in \mathcal{C}$, $\mathcal{J} \subseteq \mathcal{B}$, and for all $j \in \mathcal{B}$, $\mathcal{J} \subseteq B$, we have

$$(y_{\mathcal{J}})_j = \begin{cases} \lambda_{\mathcal{J}} & \text{if } j \in \mathcal{J}, \;\; \mathcal{J} \subseteq \mathcal{B} \\ 0 & \text{otherwise,} \end{cases}$$

since in this case $x_j^2 = x_j$. Observe that $\sum_{\mathcal{J} \subseteq \mathcal{B}} F_{\mathcal{J}}(\boldsymbol{x}_B) = 1$. To see this, define $\mathcal{J}' = \{j \mid (x_B)_j = 1\}$. As $\boldsymbol{x}_B$ are binary variables, $F_{\mathcal{J}}(\boldsymbol{x}_B)$ takes the value 1 only when $\mathcal{J} = \mathcal{J}'$ and 0 otherwise. Hence, $\sum_{\mathcal{J} \subseteq \mathcal{B}} \lambda_{\mathcal{J}} = 1$. Adding this constraint to $\mathcal{X}_{\mathrm{RPT}}$, together with $\boldsymbol{x} = \sum_{\mathcal{J} \subseteq \mathcal{B}} \boldsymbol{y}_{\mathcal{J}}$, which follows from $\sum_{\mathcal{J} \subseteq \mathcal{B}} \lambda_{\mathcal{J}} = 1$ and $\boldsymbol{y}_{\mathcal{J}} = \boldsymbol{x} \lambda_{\mathcal{J}}$, we obtain

$$\mathcal{X}_{\mathrm{conv}} = \left\{ \boldsymbol{x} \in \mathbb{R}^{n_x} \;\middle|\; \begin{array}{lr} \lambda_{\mathcal{J}} h_k\left(\frac{\boldsymbol{y}_{\mathcal{J}}}{\lambda_{\mathcal{J}}}\right) \leq 0, & k \in \mathcal{K}, \mathcal{J} \subseteq \mathcal{B} \\ \lambda_{\mathcal{J}} \geq 0, & \mathcal{J} \subseteq \mathcal{B} \\ 0 \leq y_{\mathcal{J}j} \leq \lambda_{\mathcal{J}}, & j \in \mathcal{C}, \mathcal{J} \subseteq \mathcal{B} \\ \sum_{\mathcal{J} \subseteq \mathcal{B}} \lambda_{\mathcal{J}} = 1 & \\ \boldsymbol{x} = \sum_{\mathcal{J} \subseteq \mathcal{B}} \boldsymbol{y}_{\mathcal{J}} & \end{array} \right\}, \tag{27}$$

which is precisely the convex hull of $\mathcal{X}$ as given in Sherali and Adams (2009). This demonstrates that by iteratively applying the proposed RPT approach (without Branch and Bound), where we consider all $|\mathcal{B}|$ combinations of the constraints $0 \leq x_j \leq 1$ for $j \in \mathcal{B}$, we construct the convex hull of the 0-1 mixed-integer convex set $\mathcal{X}$.

7.4. 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

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

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

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

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

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

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:

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

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).

7.5. Fractional optimization

Consider the following generic fractional optimization problem

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

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:

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

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

8. Numerical experiments

In this section, we demonstrate the efficiency and effectiveness of our RPT-BB approach on several nonconvex optimization problems, including a sum-of-max-of-linear-terms maximization problem, a Euclidean norm maximization problem, a log-sum-exp maximization problem, a linear multiplicative optimization problem, a quadratically constrained quadratic optimization problem, and a dike height optimization problem. Using the biconjugate reformulation, we show that the first four nonconvex optimization problems can be formulated as bilinear optimization problems subject to convex and nonconvex, though SLC, constraints. The latter two nonconvex optimization problems are already in generic form (1). In the implementation of RPT-BB, all problems are assumed to be minimization problems, by switching to the minus of the objective if necessary. Moreover, in all problems that we address, except for the linear multiplicative optimization problem, the conditions for convergence of RPT-BB are satisfied.

Numerical experiments are performed on one Intel i9 2.3GHz CPU core with 16 GB RAM. All computations for RPT-BB and SCIP are conducted with MOSEK version 9.2.45 (MOSEK ApS, 2020), Gurobi version 9.0.2 (Gurobi Optimization, 2019), SCIP version 8.0.2 (Achterberg, 2009), and implemented using Julia 1.5.3 and the Julia package `JuMP.jl` version 0.21.6. All computations for BARON are conducted with BARON version 20.10.16 (Sahinidis, 1996) implemented using the Python package `pyomo` version 6.4.1. Finally, all computations for CPLEX are conducted with CPLEX version 22.1.0 (ILOG,Inc., 2017) and implemented using the Python package `docplex` version 2.23.222. Inside RPT-BB, we use Gurobi for the linear optimization problems and MOSEK for the nonlinear optimization problems. In all branch and bound implementations the optimality gap is set to 10^{-4}. We compare our approach, when applied with (RPT-SDP-BB) and without the LMI (RPT-BB), with BARON, and with SCIP, applied either on the direct formulation (BARON-Dir, SCIP-Dir) or the biconjugate reformulation (BARON-Bic, SCIP-Bic), where applicable.

In the remainder of this section, in the tables depicting the results, we use the abbreviations "Opt", "Gen Hyp", and "Time" to represent the optimal value, the total number of hyperplanes generated during branch and bound, and the computation time in seconds, respectively. Moreover, we set the maximum time limit equal to 3600 seconds, hence if the computation time equals 3600^*, the optimum cannot be found within 3600 seconds and all approaches return the best value they can obtain within 3600 seconds. $**$ denotes that some of the instances were solved within one hour, while others were not and returned the best value obtained. $***$ denotes that no feasible solution was found and just the lower bound was returned. Finally, a "-" indicates that no solution was returned after one hour. To bridge the gap between theory and practice, we have made our code freely available on *GitHub*[a].

[a] `https://github.com/ThKoukouv/RPT_BB`

8.1. Sum-of-max-linear-terms maximization over convex and nonconvex constraints

We consider the following generic sum-of-max-linear-terms maximization problem from Zhen et al. (2022) and Selvi et al. (2022):

$$\max_{\boldsymbol{x}\in\mathcal{X}} \quad \sum_{\ell\in\mathcal{L}} \max_{j\in\mathcal{J}_\ell}\{\boldsymbol{A}_j^\top \boldsymbol{x} + b_j\}, \tag{28}$$

where $\boldsymbol{A}\in\mathbb{R}^{n_x\times n_y}$, $\boldsymbol{b}\in\mathbb{R}^{n_y}$. 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 nonconvex constraint, i.e., $\mathcal{X}=\mathcal{X}_1$, $\mathcal{X}=\mathcal{X}_2$, and $\mathcal{X}=\mathcal{X}_3$, where

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

Here $\boldsymbol{D}\in\mathbb{R}^{n_x\times m}$ and $\boldsymbol{d}\in\mathbb{R}^m$. Since the objective of (28) is a closed convex function, we can replace it by its biconjugate function and obtain the following equivalent maximization problem:

$$\max_{\substack{\boldsymbol{x}\in\mathcal{X}\\ \boldsymbol{y}\in\mathcal{Y}}} \quad (\boldsymbol{A}^\top\boldsymbol{x}+\boldsymbol{b})^\top\boldsymbol{y}, \tag{29}$$

where $\mathcal{Y}$ equals the domain of the conjugate function of the objective of (28), i.e.,

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

Note that $n_y=\sum_{\ell\in\mathcal{L}}|\mathcal{J}_\ell|$. We compare RPT-BB, RPT-SDP-BB, and BARON. Furthermore, for $\mathcal{X}=\mathcal{X}_1$ and $\mathcal{X}=\mathcal{X}_2$, we also compare them with the exact mixed integer optimization reformulation (MIR), given by

$$\begin{array}{lll}
\max\limits_{\boldsymbol{\lambda},\boldsymbol{z}} & \sum\limits_{\ell\in\mathcal{L}} \lambda_\ell & \\
\text{s.t.} & \lambda_\ell \le \boldsymbol{A}_j^\top\boldsymbol{x} + b_j + M(1-z_j), & j\in\mathcal{J}_\ell, \ell\in\mathcal{L},\\
& \sum\limits_{j\in\mathcal{J}_\ell} z_j = 1, & \ell\in\mathcal{L},\\
& \boldsymbol{z}\in\{0,1\}^{n_y}. &
\end{array} \tag{30}$$

We solve Problem (30) with Gurobi for $\mathcal{X}_1$ and Mosek for $\mathcal{X}_2$. We refer to Appendices C.1 and D.1 for the convex RPT relaxation and problem instances, respectively. The results are illustrated in Table 3.

From Table 3, we observe that for $\mathcal{X}=\mathcal{X}_1$, MIR is able to solve all instances the fastest, except for instances 5 and 5a, for which RPT-BB has the lowest computation time. We notice that BARON finds the optimum for instances 3, 3a, 5, and 5a, but cannot prove optimality within the time limit.

$\mathcal{X}$	#	RPT-BB			RPT-SDP-BB			BARON		GUROBI/MOSEK	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Opt	Time
$\mathcal{X}_1$	1	23.29	0.03	0	23.29	0.06	0	23.29	0.05	23.29	0.01
	1a	22.72	0.03	0	22.72	0.06	0	22.72	0.05	22.72	0.01
	2	233.94	0.32	0	233.94	1.49	0	233.94	0.28	233.94	0.04
	2a	211.87	0.65	0.2	211.87	1.47	0	211.87	1.08	211.87	0.06
	3	1081.62	562.13	104	1081.62	42.32	0	1081.62	3600*	1081.62	1.63
	3a	1159.90	296.77	47.7	1159.90	88.29	0.3	1159.90	3600*	1159.90	1.53
	4	113.71	0.09	0	113.71	0.14	0	113.71	4.01	113.71	0.01
	4a	83.78	0.52	0.5	83.78	0.19	0	83.78	4.09	83.78	0.02
	5	3002.44	6.57	1	3002.44	213.93	1	3002.44	3600*	3002.44	647.18
	5a	2898.05	44.36	7.8	2898.05	238.14	1.3	2898.05	3600*	2898.05	734.65
$\mathcal{X}_2$	1	14.58	0.06	0	14.58	0.08	0	14.58	0.09	14.58	0.04
	1a	14.54	0.09	0	14.54	0.11	0	14.54	0.07	14.54	0.04
	2	136.22	5.43	1	136.22	3.43	0	136.22	3600*	136.22	778.65
	2a	122.21	8.84	2.6	122.21	9.21	0.8	122.21	3600*	122.21	977.05
	3	837.94	3600*	283	837.94	201.34	0	837.94	3600*	837.94	3600*
	3a	890.07	3600*	278.3	890.07	616.93	0.6	890.07	3600*	886.41	3600*
	4	33.73	4.95	7	33.73	3.15	2	33.73	75.44	33.73	0.09
	4a	31.81	2.35	2.7	31.81	1.69	0.8	31.81	39.33	31.81	0.11
	5	1610.69	3600*	286	1610.69	3600*	9	1610.69	3600*	1610.69	3600*
	5a	1670.92	3600*	291.3	1670.92	3600*	9.6	1670.92	3600*	1670.92	3600*
$\mathcal{X}_3$	1	13.44	98.51	441	13.44	13.19	6	13.44	0.31		
	1a	15.02	862.32	1165.7	15.02	13.43	8.7	15.02	0.13		
	2	140.89	146.74	241	140.89	101.04	4	140.89	8.48		
	2a	129.31	442.59	739.1	129.31	103.44	3.9	129.31	417.41		
	3	768.96	855.87	43	768.96	3600*	3	768.96	3600*		
	3a	805.95	1369.82	66.4	805.95	3600*	4.1	805.95	3600*		
	4	45.34	218.51	123	45.34	133.29	36	45.34	114.87		
	4a	43.79	2336.64**	677.1	44.97	352.84	67.9	44.97	97.11		
	5	2256.91	3600*	136	-	3600*	4	1700.68	3600*		
	5a	2353.19	3600*	228.7	-	3600*	4.3	2000.09	3600*		

Table 3 Results for Problem (29) over the feasible regions $\mathcal{X}_1, \mathcal{X}_2$, and $\mathcal{X}_3$. The exact mixed integer reformulation is solved with Gurobi for $\mathcal{X} = \mathcal{X}_1$ and MOSEK for $\mathcal{X} = \mathcal{X}_2$.

For $\mathcal{X} = \mathcal{X}_2$, MIR still finds the optimum within the lowest computation time for instances 1, 1a, and 4. Further, for instances 2, 2a, 3, 3a, and 4a our approach achieves the lowest computation time. Finally, for instances 5 and 5a all methods find the optimum, however they are not able to prove optimality within the time limit.

For $\mathcal{X} = \mathcal{X}_3$, MOSEK cannot solve the instances because of the nonconvex constraint. For instances 1, 1a, 2, 2a, and 4, all approaches find the optimum. For instances 1, 1a, 2, and 4, BARON has the lowest computation time, while for instance 2a RPT-SDP-BB has the lowest computation time. Moreover, RPT-BB solves the problem to optimality for instances 3 and 3a, while BARON could only find the optimum but not prove optimality within the time limit. BARON has the lowest computation time for instance 4a. Finally, for instances 5 and 5a, RPT-BB is able to compute a better bound than BARON within one hour.

8.2. Log-sum-exp maximization over linear constraints

We consider the log-sum-exp maximization problem subject to linear constraints:

$$\max_{\boldsymbol{x}\in\mathcal{X}} \quad \log\left(\sum_{i=1}^{n_x}\exp(x_i)\right), \tag{31}$$

where $\mathcal{X} = \{\boldsymbol{x}\in\mathbb{R}_+^{n_x} \mid \boldsymbol{D}^\top\boldsymbol{x}\le\boldsymbol{d}\}$, $\boldsymbol{D}\in\mathbb{R}^{n_x\times L'}$ and $\boldsymbol{d}\in\mathbb{R}^{L'}$. Since the objective of (31) is a closed convex function, we can replace it by its biconjugate function and obtain the following equivalent maximization problem:

$$\max_{\substack{\boldsymbol{x}\in\mathcal{X}\\ (\boldsymbol{y},\boldsymbol{w})\in\mathcal{Y}}} \quad \boldsymbol{x}^\top\boldsymbol{y}+\sum_{i=1}^{n_x} w_i, \tag{32}$$

where $\mathcal{Y}$ equals the domain of the conjugate function of the objective of (31), i.e.,

$$\mathcal{Y}=\left\{\boldsymbol{y}\in\mathbb{R}_+^{n_x},\,\boldsymbol{w}\in\mathbb{R}^{n_x}\;\middle|\; y_i\exp\left(\frac{w_i}{y_i}\right)\le 1,\, i\in\{1,\dots,n_x\},\, \sum_{i=1}^{n_x} y_i=1\right\}.$$

Observe that here we make use of case 7 in Table 1 and introduce epigraphical variables w_i for every $i\in\{1,\dots,n_x\}$. We refer to Appendices C.2 and D.2 for the convex RPT relaxation and problem instances, respectively. The results are illustrated in Table 4.

$\mathcal{X}$	#	RPT-BB			RPT-SDP-BB			BARON-Dir		SCIP-Dir		BARON-Bic		SCIP-Bic	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Opt	Time	Opt	Time	Opt	Time
	1	10.01	0.14	0	10.01	0.48	0	10.01	0.06	10.01	0.07	6.48	3600*	10.01	3600*
	2	40.00	1.51	0	40.00	158.44	0	40.00	360.65	3.69	3600*	30.27	3600*	-	3600*
	3	6.09	28.22	7.4	6.09	11.07	2.1	6.09	0.12	6.09	0.17	5.46	3600*	6.08	3600*
	4	21.96	0.96	0	21.96	3.37	0	21.78	1082.38**	19.31	2520.37**	5.88	3600*	20.09	3600*
	5	34.76	13.46	0	34.76	368.69	0	34.76	1450.84**	15.43	3600*	-	3600*	31.91	3600*

Table 4 Results for Problem (31).

From Table 4 we observe that both RPT-BB and RPT-SDP-BB find the optimum for all instances, whereas BARON could not find the optimum for some of the generated instances in 4 and 5, and SCIP could not find the optimum for some of the generated instances in 4 and for all of the generated instances in 2 and 5 within one hour. Observe that while BARON-Dir is not able to prove optimality within the time limit for instance 5, it does find the optimum, whereas for instance 4 it does not. For instances 1 and 3, BARON-Dir performs best on computation time, while for instances 2, 4, and 5, RPT-BB has the lowest computation time. Moreover, we notice that in all instances BARON-Bic and SCIP-Bic cannot find the optimum within one hour. Hence, if we had started with Problem (32), our approach would have significantly outperformed both BARON and SCIP for all instances. We further observe that RPT-BB has lower computation time than RPT-SDP-BB in all instances except for instance 3, in which case more hyperplanes are needed during branch and bound when the LMI is not included.

8.3. Linear multiplicative optimization

We consider the following linear multiplicative optimization problem from Ryoo and Sahinidis (1996):

$$\min_{\boldsymbol{x}\in\mathcal{X}} \quad \prod_{i=1}^{n_y} \boldsymbol{A}_i^\top \boldsymbol{x} + b_i, \tag{33}$$

where $\boldsymbol{A}\in\mathbb{R}^{n_x\times n_y}$, $\boldsymbol{b}\in\mathbb{R}^{n_y}$, $\mathcal{X}=\{\boldsymbol{x}\in\mathbb{R}^{n_x}_+ \mid \boldsymbol{D}^\top\boldsymbol{x}\le\boldsymbol{d},\ \boldsymbol{A}_i^\top\boldsymbol{x}+b_i\ge 0\}$, $\boldsymbol{D}\in\mathbb{R}^{n_x\times L}$, and $\boldsymbol{b}\in\mathbb{R}^L$. Without loss of generality we assume $\boldsymbol{A}_i^\top\boldsymbol{x}+b_i>0$ for all $i\in\mathcal{I}$. Utilizing a log transformation, as in Ryoo and Sahinidis (1996), Problem (33) can be equivalently reformulated as

$$\min_{\boldsymbol{x}\in\mathcal{X}} \quad \sum_{i=1}^{n_y} \log\left(\boldsymbol{A}_i^\top\boldsymbol{x}+b_i\right). \tag{34}$$

Since the objective of (34) is a closed concave function, we can replace it by its biconjugate function and obtain the following equivalent maximization problem:

$$\max_{\substack{\boldsymbol{x}\in\mathcal{X}\\(\boldsymbol{y},\boldsymbol{w})\in\mathcal{Y}}} \quad -(\boldsymbol{A}^\top\boldsymbol{x}+\boldsymbol{b})^\top\boldsymbol{y}-\sum_{i=1}^{n_y} w_i, \tag{35}$$

where $\mathcal{Y}$ equals the domain of the conjugate function of the objective of (34), i.e.,

$$\mathcal{Y}=\left\{\boldsymbol{y}\in\mathbb{R}^{n_y}_+,\boldsymbol{w}\in\mathbb{R}^{n_y}\;\middle|\;\exp(-w_i-1)\le y_i,\, i\in\{1,\dots,n_y\}\right\}.$$

Observe that here we make use of case 3 in Table 1 and introduce epigraphical variables w_i for every $i\in\{1,\dots,n_y\}$. Moreover, observe that $\mathcal{Y}$ is not bounded. Hence, the conditions for convergence of RPT-BB are not satisfied. We refer to Appendices C.3 and D.3 for the convex RPT relaxation and problem instances, respectively. The results are illustrated in Table 5.

$\mathcal{X}$	#	RPT-BB			RPT-SDP-BB			BARON-Dir		SCIP-Dir		BARON-Bic		SCIP-Bic	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Opt	Time	Opt	Time	Opt	Time
	1	12.14	8.90	35.0	12.14	9.18	20.9	12.14	0.23	12.14	0.35	12.14	42.16	12.14	2889.96**
	2	23.47	75.99	103.1	23.47	83.49	41.6	23.47	3600*	33.65	3600*	23.47	2766.69	33.17	2284.53**
	3	20.96	1141.31	516.7	20.96	160.08	55.6	20.96	2526.99	32.93	3240.03**	20.96	2288.78	20.96	3600*
	4	18.99	3600*	2338.3	18.99	648.59	78.2	18.99	3600*	28.20	3240.49**	18.99	3442.49**	18.99	3600*
	5	8.89	3600*	453.3	8.89	723.38	43.1	8.89	466.13	8.89	0.76	8.89	67.94	8.90	3600*

Table 5 Results for Problem (33).

From Table 5, we observe that our approach often outperforms BARON and SCIP on linear multiplicative optimization problems. Specifically, in instances 2, 3, and 4, which correspond to the multiplication of 10, 9, and 8 linear terms, respectively, RPT-BB or RPT-SDP-BB achieve the best computation times. However, in instances 1 and 5, which involve 5 and 40 variables, and 5 and 4 linear terms in the objective, respectively, both BARON and SCIP solve the problems faster than RPT-BB and RPT-SDP-BB. We notice that BARON achieves the best computation time in instance 1 and SCIP in instance 5.

8.4. Quadratic constraint quadratic optimization

We consider the following quadratic constraint quadratic optimization problem:

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathcal{X}_1} \quad & \boldsymbol{x}^\top \boldsymbol{P}_0 \boldsymbol{x} + \boldsymbol{q}_0^\top \boldsymbol{x} + r_0 \\ \text{s.t.} \quad & \boldsymbol{x}^\top \boldsymbol{P}_k \boldsymbol{x} + \boldsymbol{q}_k^\top \boldsymbol{x} + r_k \leq 0, \qquad k \in \mathcal{K}, \end{aligned} \tag{36}$$

where $\boldsymbol{P}_k \in \mathbb{R}^{n_x \times n_x}$, $k \in \mathcal{K}_0$, are not necessarily positive semi-definite and as a result Problem (36) is not necessarily convex. However, the nonconvex quadratic functions are SLC, hence we can apply RPT. In the first five instances involving nonconvex QPs over linear constraints, we also compare with CPLEX, which has a specialized algorithm for these problems. We refer to Appendices C.4 and D.4 for the convex RPT relaxation and problem instances, respectively. The results are illustrated in Table 6.

$\mathcal{X}$	#	RPT-BB			RPT-SDP-BB			BARON		CPLEX	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Opt	Time
	1	394.75	0.35	1	394.75	0.55	1	394.75	0.19	394.75	0.11
	2	884.75	0.31	1	884.75	0.43	1	884.75	0.21	884.75	0.08
	3	6888.78	0.03	0	6888.78	0.05	0	6888.78	0.11	6888.78	0.07
	4	98382.63	2.86	0	98382.63	3.77	0	98382.63	3600*	98382.63	1.32
	5	774482.38	18.14	0	774482.38	67.51	0	774482.38	3600*	774482.38	13.25
	6	1717.80	0.52	0	1717.80	0.16	0	1717.80	0.14		
	7	4507.75	0.22	5	4507.75	0.21	1	4507.75	0.32		
	8	17084.82	0.37	6	17084.82	0.39	1	17084.82	0.33		
	9	52557.99	3600*	5924	46632.36	35.07	26	46632.36	41.64		
	10	76929.21	3600*	1903	52233.70	211.22	51	52233.70	3600*		

Table 6 Results for Problem (36). Instances 1-5 contain only linear constraints, while instances 5-10 contain additional nonconvex quadratic constraints.

From Table 6 we observe that for instances 1-3, 6-8, all approaches find the optimum in less than a second. CPLEX is in general better for instances 1-5. Moreover, for instances 6, 8 BARON is faster, while for instance 7 RPT-SDP-BB is faster. Further, for instances 9 and 10 RPT-SDP-BB achieves the best computation time, while RPT-BB is not able to find the optimum within an hour. Finally, we notice that RPT-SDP-BB is able to solve all instances, while BARON is not able to prove optimality for instances 4, 5, and 10.

8.5. Dike height optimization

Eijgenraam et al. (2017) develop a model 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 D.5. Observe that the feasible region is not compact. However, we can add redundant upper bounds on $\boldsymbol{x}$ such that we obtain a compact feasible region. Moreover, since $C>0$, the conditions for convergence of RPT-BB, as specified in Theorem 3, are satisfied.

The objective of (DHO) is to minimize the sum of investment costs and the total expected cost of flooding, both as a result of heightening dikes, see 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-BB. We compare RPT-BB, RPT-SDP-BB, and BARON. We refer to Appendices C.5 and D.5 for the convex RPT relaxation and problem instances, respectively. The results for the homogeneous dike rings 10, 15 and 16 in the Netherlands are shown in Table 7.

$\boldsymbol{t}$	#	RPT-BB			RPT-SDP-BB			BARON	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time
$\boldsymbol{t}_{\text{ir}}$	10	61.98	3600*	1373	61.98	0.41	0	61.98	110.70
	15	608.74	3600*	2231	608.74	0.29	0	608.74	3600*
	16	1268.11	3600*	922	1268.11	0.29	0	1268.11	3600*
$\boldsymbol{t}_{25}$	10	61.31	3600*	2197	61.31	0.49	0	61.31	1680.36
	15	609.92	3600*	1308	609.92	1.15	0	609.92	3600*
	16	1269.63	3600*	1897	1269.63	0.88	0	1269.63	3600*
$\boldsymbol{t}_{50}$	10	55.50	3600*	6680	55.50	0.16	0	55.50	1.32
	15	545.23	3600*	724	545.23	0.27	0	545.23	1.81
	16	1100.07	3600*	1225	1100.07	0.22	0	1100.07	3600*

Table 7 Results for Problem (DHO), for dike rings 10, 15, and 16 in the Netherlands.

From Table 7 we observe that RPT-SDP-BB outperforms RPT-BB and BARON, since it finds the global optimal solution for every instance in the root node, in about a second. On the other hand, RPT-BB finds the optimal solution for every case, but is not able to prove optimality. Moreover, BARON is not able to prove optimality in each case for dike ring 16 and for 2 out of the 3 cases for dike ring 15. Finally, for dike ring 16, BARON is able to achieve a smaller optimality gap than RPT-BB, as can be seen in Table 8.

$\boldsymbol{t}$	#	RPT-BB			BARON		
		UB	LB	Time	UB	LB	Time
	$\boldsymbol{t}_{\text{ir}}$	1268.11	1255.58	3600*	1268.11	1266.79	3600*
	$\boldsymbol{t}_{25}$	1269.63	1256.84	3600*	1269.63	1267.98	3600*
	$\boldsymbol{t}_{50}$	1100.07	1090.56	3600*	1100.07	1100	3600*

Table 8 Upper and lower bounds obtained for RPT-BB and BARON, within one hour for dikerings 16. We set the maximum time limit equal to 3600 seconds, hence if the computation time equals 3600*, the optimum cannot be found within 3600 seconds and all approaches return the best upper and lower bounds they can obtain within 3600 seconds.

9. Discussion and conclusion

In summary, we develop a method for globally solving nonconvex optimization problems involving SLC functions. We introduce the RPT framework, which enables us to obtain a convex relaxation of the original nonconvex problem, while introducing additional variables and constraints. We then incorporate it in spatial branch and bound in order to solve the initial problem to optimality by sequentially partitioning the feasible region in smaller regions. In the numerical experiments, we demonstrate that the proposed method stands well against the current state of the art global optimization methods. Overall, we observe that for the considered problem instances, RPT-BB and RPT-SDP-BB are able to solve most problems by generating a few hyperplanes. This, together with the efficiency of MOSEK for solving conic optimization problems, is what drives the speed of the method.

A key limitation of the RPT-BB method is its reduced tractability when applied to problems with a large number of variables and constraints. This issue arises because the method involves squaring the number of variables and performing pairwise multiplications across all linear and convex constraints within the feasible region. To enhance scalability, future research could explore methodological adaptations designed to manage larger problem instances more efficiently. Potential adaptations might include employing partial constraint multiplications and selectively generating variable products. An iterative constraint generation approach could also be beneficial. This would involve initially solving the RPT relaxation with a subset of constraints, then progressively incorporating additional constraints based on whether the solutions violate any remaining constraints. Assessing the efficiency of this iterative approach, considering it requires multiple resolutions of the RPT relaxation, would be essential.

Finally, it would be interesting to investigate the potential of RPT-BB beyond the scope discussed in this paper, by applying RPT-BB also to nonconvex optimization problems in other fields, such as mixed integer nonlinear optimization, robust optimization, adaptive robust optimization, distributionally robust optimization, polynomial optimization, and bilevel optimization.

Acknowledgements

We would like to thank Bram Gorissen for the suggestion to consider fractional optimization from an RPT lens. We thank Etienne de Klerk for providing us reference Hartman (1959). Finally, we thank Kurt Anstreicher and Daniel Kuhn for their valuable comments on a draft version of this paper. The second author of this paper is funded by NWO grant 406.18.EB.003.

References

T. Achterberg. Scip: solving constraint integer programs. *Mathematical Programming Computation*, 1:1–41, 2009.

F.A. Al-Khayyal, C. Larsen, and T. Van Voorhis. A relaxation method for nonconvex quadratically constrained quadratic programs. *Journal of Global Optimization*, 6(3):215–230, 1995.

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

K.M. Anstreicher. Solving two-trust-region subproblems using semidefinite optimization with eigenvector branching. *Journal of Optimization Theory and Applications*, 202(1):303–319, 2022.

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 and E. Roos. An algorithm for maximizing a convex function based on its minimum. *INFORMS Journal on Computing*, 34(6):3200–3214, 2022.

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.

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

J. Chen and S. Burer. Globally solving nonconvex quadratic programming problems via completely positive programming. *Mathematical Programming Computation*, 4(1):33–52, 2012.

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.

J.E. Falk and R.M. Soland. An algorithm for separable nonconvex programming problems. *Management Science*, 15(9):550–569, 1969.

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.

Gurobi Optimization. Gurobi optimizer reference manual 8.1.1, 2019. URL http://www.gurobi.com.

P. Hartman. On functions representable as a difference of convex functions. *Pacific Journal of Mathematics*, 9(3):707–713, 1959.

R. Horst. An algorithm for nonconvex programming problems. *Mathematical Programming*, 10:312–321, 1976.

ILOG,Inc. *ILOG CPLEX 12.7, User Manual*, 2017. URL https://www.ibm.com/support/knowledgecenter/SSSA5P_12.7.0/ilog.odms.studio.help/Optimization_Studio/topics/COS_home.html.

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.

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.

H.S. Ryoo and N.V. Sahinidis. A branch-and-reduce approach to global optimization. *Journal of Global Optimization*, 8:107–138, 1996.

N.V. Sahinidis. Baron: A general purpose global optimization software package. *Journal of Global Optimization*, 8(2):201–205, 1996.

A. Saxena, P. Bonami, and J. Lee. Convex relaxations of non-convex mixed integer quadratically constrained programs: extended formulations. *Mathematical Programming*, 124(1):383–411, 2010.

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

A. Selvi, A. Ben-Tal, R. Brekelmans, and D. den Hertog. Convex maximization via adjustable robust optimization. *INFORMS Journal on Computing*, 34(4):2091–2105, 2022.

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 Solving Discrete and Continuous Nonconvex Problems*, volume 31. Springer Science & Business Media, 1999.

H. Sherali and W. Adams. A hierarchy of relaxations leading to the convex hull representation for general discrete optimization problems. *Annals of Operations Research*, 140:21–47, 2005.

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, 2009.

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.

A. Waechter, C. Laird, F. Margot, and Y. Kawajir. Introduction to ipopt: A tutorial for downloading, installing, and using ipopt. *Revision*, 2009. URL https://coin-or.github.io/Ipopt/.

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 optimization: A two-stage robust optimization perspective. *INFORMS Journal on Computing*, 34(5):2410–2427, 2022.

J. Zhen, D. Kuhn, and W. Wiesemann. A unified theory of robust and distributionally robust optimization via the primal-worst-equals-dual-best principle. *Operations Research*, 2023. URL https://pubsonline.informs.org/doi/10.1287/opre.2021.0268.

Appendix

A. When infimum is not attained

If the infimum of (3) is not attained, we assume that (1) satisfies the following regularity condition.

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

Note that Assumption 3 implies that $\boldsymbol{x}^{\mathrm{S}}$ resides in the sets $\cap_{\substack{k\in\mathcal{K}_0\\ i\in\mathcal{I}_0}} \mathrm{ri}(\mathrm{dom}(c_{ik}))$ and $\cap_{j\in\mathcal{J}_0}\mathrm{ri}(\mathrm{dom}(h_j))$ thanks to Proposition 2.42 in Rockafellar and Wets (2009), and thus, $\boldsymbol{x}^{\mathrm{S}}$ is a strict Slater point of (1). Furthermore, there exists a $(\boldsymbol{\tau}^{\mathrm{S}}, \boldsymbol{X}^{\mathrm{S}}, \boldsymbol{V}^{\mathrm{S}})$ such that $(\boldsymbol{x}^{\mathrm{S}}, \boldsymbol{\tau}^{\mathrm{S}}, \boldsymbol{X}^{\mathrm{S}}, \boldsymbol{V}^{\mathrm{S}})$ is a strict Slater point of the corresponding RPT relaxation (8) of (1) with $f_k(\boldsymbol{x}) \le 0$ is replaced by

$$\begin{cases} c_{0k}(\boldsymbol{x}) - \sup\limits_{y\in\mathrm{dom}(c^*_{1k})} \{\boldsymbol{x}^\top \boldsymbol{y} - c^*_{1k}(\boldsymbol{y})\} \le 0 \\ \boldsymbol{y} \in \mathrm{dom}(c^*_{1k}). \end{cases} \tag{37}$$

Finally, thanks to Remark 1 and the proof of Theorem 6(iii) of Zhen et al. (2023), the inf operator in the constraint of (8) 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 (8).

B. Mountain climbing procedure

We use a mountain climbing (MC) procedure based on the algorithm from Tao and An (1997), to find an upper bound for problems involving the biconjugate. The MC procedure takes as input $\mathcal{X}''$, the list of candidate vectors obtained from the solution of the RPT relaxation (see Section 3.4) and returns a local optimum. The procedure is summarized in Algorithm 3, for the problem of maximizing a function $f(\boldsymbol{x}, \boldsymbol{y})$ over $\mathcal{X} \times \mathcal{Y}$, where $\mathcal{X}$ and $\mathcal{Y}$ are disjoint sets. For $\mathcal{X} = \mathcal{X}_3$ we only apply it for the candidate vectors that are feasible. Note that it is possible that $\mathcal{X}'' = \emptyset$, in which case MC cannot be applied. Moreover, for $\mathcal{X} = \mathcal{X}_2$ and $\mathcal{X} = \mathcal{X}_3$ in the numerical experiments we alternate between maximizing for $\boldsymbol{x}, \boldsymbol{z} \in \mathcal{X}$ and maximizing for $\boldsymbol{y} \in \mathcal{Y}$ and vice versa.

C. RPT-SDP formulations of the numerical experiments

Throughout the experiments we consider five cases of the feasible set $\mathcal{X}$, those are $\mathcal{X} = \mathcal{X}_1$, $\mathcal{X} = \mathcal{X}_2$, $\mathcal{X} = \mathcal{X}_3$, $\mathcal{X} = \mathcal{X}_4$, and $\mathcal{X} = \mathcal{X}_5$, where

$$\begin{aligned}
\mathcal{X}_1 &= \{\boldsymbol{x} \in \mathbb{R}^{n_x}_+ \mid \boldsymbol{D}^\top \boldsymbol{x} \le \boldsymbol{d}\} \\
\mathcal{X}_2 &= \left\{ \boldsymbol{x} \in \mathcal{X}_1 \;\middle|\; \log\left(\sum_{i=1}^{n_x} \exp(x_i)\right) \le a \right\} \\
\mathcal{X}_3 &= \left\{ \boldsymbol{x} \in \mathcal{X}_1 \;\middle|\; \|\boldsymbol{x}\|_2 + \sum_{i=1}^{n_x} \sqrt{x_i} \le c \right\} \\
\mathcal{X}_4 &= \left\{ \boldsymbol{x} \in \mathcal{X}_1 \middle| \boldsymbol{A}_i^\top \boldsymbol{x} + b_i \ge 0, i \in \{1, \ldots, n_y\} \right\} \\
\mathcal{X}_5 &= \left\{ \boldsymbol{x} \in \mathcal{X}_1 \middle| \boldsymbol{x}^\top \boldsymbol{P}_k \boldsymbol{x} + \boldsymbol{q}_k^\top \boldsymbol{x} + r_k \le 0,\ k \in \mathcal{K}_{\mathrm{C}} \right\}.
\end{aligned}$$

Algorithm 3 Mountain climbing procedure

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

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

We notice that both $\mathcal{X}_2$ and $\mathcal{X}_3$ are not in conic form, but they can be reformulated as such, in the following way. First, for $\mathcal{X}_2$ we observe that $\log\left(\sum_{i=1}^{n_x} \exp(x_i)\right) \leq a \iff \sum_{i=1}^{n_x} \exp(x_i - a) \leq 1$. Using epigraphical variables z_i we obtain the following equivalent form:

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

Regarding $\mathcal{X}_3$ we first reformulate the nonconvex constraint via the biconjugate and obtain the equivalent set

$$\mathcal{X}_3 = \left\{ \boldsymbol{x} \in \mathcal{X}_1 \,\middle|\, \|\boldsymbol{x}\|_2 + \sum_{i=1}^{n_x} \frac{1}{4z_i} + \boldsymbol{x}^\top \boldsymbol{z} \leq 0 \right\}.$$

We introduce epigraphical variables for the convex component of the SLC constraint. Since the convex component of the SLC constraint consists of a sum of two basic cone functions we introduce an epigraphical variable for each basic cone function. Subsequently, we reformulate every convex constraint in terms of one of the basic cone constraints. Next, we convexify the SLC constraint, such that we obtain the following relaxed set of constraints

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

We choose c to be large enough such that (29) with $\mathcal{X} = \mathcal{X}_3$ satisfies Assumption 2.

In the formulations for the numerical experiments we encounter several products of variables. These are linearized as follows: We linearize $\boldsymbol{x}\boldsymbol{x}^\top$ by $\boldsymbol{X}$, $\boldsymbol{y}\boldsymbol{y}^\top$ by $\boldsymbol{Y}$, $\boldsymbol{z}\boldsymbol{z}^\top$ by $\boldsymbol{Z}$, $\boldsymbol{w}\boldsymbol{w}^\top$ by $\boldsymbol{W}$, $\boldsymbol{t}\boldsymbol{t}^\top$ by $\boldsymbol{T}$, $\boldsymbol{x}\boldsymbol{y}^\top$ by $\boldsymbol{U}$, $\boldsymbol{x}\boldsymbol{z}^\top$ by $\boldsymbol{V}$, $\boldsymbol{x}\boldsymbol{w}^\top$ by $\boldsymbol{Q}$, $\boldsymbol{x}\boldsymbol{t}^\top$ by $\boldsymbol{F}$, $\boldsymbol{y}\boldsymbol{z}^\top$ by $\boldsymbol{R}$, $\boldsymbol{y}\boldsymbol{w}^\top$ by $\boldsymbol{P}$, $\boldsymbol{y}\boldsymbol{t}^\top$ by $\boldsymbol{G}$, $\boldsymbol{z}\boldsymbol{w}^\top$ by $\boldsymbol{K}$, $\boldsymbol{z}\boldsymbol{t}^\top$ by $\boldsymbol{H}$, $s\boldsymbol{x}$ by $\boldsymbol{\alpha}$, $s\boldsymbol{y}$ by $\boldsymbol{\beta}$, $s\boldsymbol{z}$ by $\boldsymbol{\gamma}$, $s\boldsymbol{t}$ by $\boldsymbol{\phi}$, s^2 by σ, $p\boldsymbol{x}$ by $\boldsymbol{\lambda}$, $p\boldsymbol{y}$ by $\boldsymbol{\mu}$, $p\boldsymbol{z}$ by $\boldsymbol{\nu}$, $p\boldsymbol{t}$ by $\boldsymbol{\psi}$, ps by ρ and p^2 by π.

C.1. RPT-SDP formulation of Problem (28)

Replacing the objective function with the biconjugate function in (28) 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 $\mathcal{Y}$ is given by

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

$\boldsymbol{\mathcal{X} = \mathcal{X}_1}$. The RPT-SDP formulation is given by

$$\begin{aligned}
\max_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y}} \quad & \operatorname{Tr}(\boldsymbol{U}\boldsymbol{A}) + \boldsymbol{b}^\top \boldsymbol{y} \\
\text{s.t.} \quad & \boldsymbol{D}^\top \boldsymbol{X}_i - \boldsymbol{d} x_i \le \boldsymbol{0}, && i \in [n_x], && \text{(38a)} \\
& \boldsymbol{D}^\top \boldsymbol{U}_j - \boldsymbol{d} y_j \le \boldsymbol{0}, && j \in [n_y], && \text{(38b)} \\
& \boldsymbol{d}\boldsymbol{x}^\top \boldsymbol{D} + \boldsymbol{D}^\top \boldsymbol{x}\boldsymbol{d}^\top \le \boldsymbol{D}^\top \boldsymbol{X}\boldsymbol{D} + \boldsymbol{d}\boldsymbol{d}^\top, && && \text{(38c)} \\
& \sum_{j \in \mathcal{J}_\ell} y_j = 1, && \ell \in \mathcal{L}, && \text{(38d)} \\
& \sum_{j \in \mathcal{J}_\ell} \boldsymbol{U}_j - \boldsymbol{x} = \boldsymbol{0}, && \ell \in \mathcal{L}, && \text{(38e)} \\
& \sum_{j \in \mathcal{J}_\ell} \boldsymbol{Y}_j - \boldsymbol{y} = \boldsymbol{0}, && \ell \in \mathcal{L}, && \text{(38f)} \\
& \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{U} \ge \boldsymbol{0}, && && \text{(38g)} \\
& \begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} & \boldsymbol{x} \\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{y} \\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & 1 \end{pmatrix} \succeq \boldsymbol{0}, && && \text{(38h)}
\end{aligned}$$

where constraints (38a) - (38g) result from pairwise multiplication of the linear constraints and constraint (38h) results from the additional SDP relaxation.

Observe that $\boldsymbol{x} \in \mathcal{X}_1$ is redundant. The constraint $\boldsymbol{D}^\top \boldsymbol{x} \le \boldsymbol{d}$ is redundant by (38e), (38d) and (38b):

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

The non-negativity constraint $\boldsymbol{x} \ge \boldsymbol{0}$ is redundant by (38e) and (38g). Moreover, the non-negativity constraint $\boldsymbol{y} \ge \boldsymbol{0}$ is redundant by (38f) and (38g). Hence, these constraints are not included in the above formulation.

In the RPT-SDP formulation we hence obtain $n_x|\mathcal{L}| + n_y|\mathcal{L}| + n_x n_y + L' n_y + n_x(n_x+1)/2 + L'(L'+1)/2 + n_x L' + n_y(n_y+1)/2$ additional linear constraints, one additional SDP constraint and $n_x(n_x+1)/2 + n_y(n_y+1)/2 + n_x n_y$ extra variables.

$\boldsymbol{\mathcal{X} = \mathcal{X}_2}$. The RPT-SDP formulation is given by

$$\begin{aligned}
\max_{\substack{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z} \\ \boldsymbol{U},\boldsymbol{V},\boldsymbol{R} \\ \boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}}} \quad & \mathrm{Tr}(\boldsymbol{U}\boldsymbol{A}) + \boldsymbol{b}^\top \boldsymbol{y} \\
\text{s.t.} \quad & (38a) - (38g) \\
& \sum_{i=1}^{n_x} z_i \le 1, && && (39a)\\
& \sum_{i=1}^{n_x} \boldsymbol{V}_i - \boldsymbol{x} \le \boldsymbol{0}, && && (39b)\\
& \sum_{i=1}^{n_x} \boldsymbol{R}_i - \boldsymbol{y} \le \boldsymbol{0}, && && (39c)\\
& \boldsymbol{D}^\top \boldsymbol{x} - \boldsymbol{D}^\top \sum_{i=1}^{n_x} \boldsymbol{V}_i \le \boldsymbol{d}(1 - \sum_{i=1}^{n_x} z_i), && && (39d)\\
& \sum_{i,j=1}^{n_x} Z_{ij} - 2\sum_{i=1}^{n_x} z_i + 1 \ge 0, && && (39e)\\
& \exp(x_i - a) \le z_i, && i \in [n_x], && (39f)\\
& x_j \exp\left(\frac{X_{ij} - a x_j}{x_j} \right) \le V_{ji}, && i,j \in [n_x], && (39g)\\
& y_j \exp\left(\frac{U_{ij} - a y_j}{y_j} \right) \le R_{ji}, && i \in [n_x], j \in [n_y], && (39h)\\
& (d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}) \exp\left(\frac{d_\ell x_i - a d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{X}_i + a \boldsymbol{D}_\ell^\top \boldsymbol{x}}{d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}} \right) \le d_\ell z_i - \boldsymbol{D}_\ell^\top \boldsymbol{V}_i, && i \in [n_x], \ell \in [L'], && (39i)\\
& \left(1 - \sum_{j=1}^{n_x} z_j \right) \exp\left(\frac{x_i - a - \sum_{j=1}^{n_x} V_{ij} + a \sum_{j=1}^{n_x} z_j}{1 - \sum_{j=1}^{n_x} z_j} \right) \le z_i - \sum_{j=1}^{n_x} Z_{ji}, && i \in [n_x], && (39j)\\
& \exp\left(x_i + x_j - 2a \right) \le Z_{ij}, && i \le j \in [n_x], && (39k)\\
& z_j \exp\left(\frac{V_{ij} - a z_j}{z_j} \right) \le Z_{ij}, && i \le j \in [n_x], && (39l)
\end{aligned}$$

$$\sum_{j\in\mathcal{J}_k} \boldsymbol{R}_j - \boldsymbol{z} = \boldsymbol{0}, \qquad k\in\mathcal{K}, \tag{39m}$$

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} & \boldsymbol{V} & \boldsymbol{x} \\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{R} & \boldsymbol{y} \\ \boldsymbol{V}^\top & \boldsymbol{R}^\top & \boldsymbol{Z} & \boldsymbol{z} \\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & \boldsymbol{z}^\top & 1 \end{pmatrix} \succeq \boldsymbol{0}, \tag{39n}$$

where constraints (39b) - (39e) result from pairwise multiplication of the new linear constraint over $\boldsymbol{z}$ with the previous linear constraints (39g) - (39l) result from pairwise multiplication of the exponential constraint with the linear inequalities and itself, constraint (39m) results from pairwise multiplication of the initial linear constraint over $\boldsymbol{y}$ with $\boldsymbol{z}$ and constraint (39n) results from the additional SDP relaxation.

In the RPT-SDP formulation we hence obtain $2n_x|\mathcal{L}| + n_y|\mathcal{L}| + n_xn_y + L'n_y + n_x(n_x+1)/2 + L'(L'+1)/2 + n_xL' + n_y(n_y+1)/2 + n_y + n_x + L' + 1$ additional linear constraints, $n_xn_y + n_x^2 + L'n_x + n_x(n_x+1) + n_x$ additional exponential constraints, one additional SDP constraint and $n_x(n_x+1) + n_y(n_y+1)/2 + 2n_xn_y + n_x^2$ extra variables.

$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_3$.

The pairwise multiplication of the linear constraints gives us the following constraints:

(38a) – (38g)

$$s + p + \sum_{i=1}^{n_x} V_{ii} \le c, \tag{40a}$$

$$\sum_{i=1}^{n_x} t_i \le p, \tag{40b}$$

$$\boldsymbol{D}^\top \boldsymbol{V}_i \le z_i \boldsymbol{d}, \qquad i\in[n_x], \tag{40c}$$

$$\boldsymbol{D}^\top \boldsymbol{F}_i \le t_i \boldsymbol{d}, \qquad i\in[n_x], \tag{40d}$$

$$\boldsymbol{D}^\top \boldsymbol{\lambda} \le p\boldsymbol{d}, \tag{40e}$$

$$\sum_{j\in\mathcal{J}_\ell} \boldsymbol{r}_j - \boldsymbol{z} = \boldsymbol{0}, \qquad \ell\in\mathcal{L}, \tag{40f}$$

$$\sum_{j\in\mathcal{J}_\ell} \boldsymbol{g}_j - \boldsymbol{t} = \boldsymbol{0}, \qquad \ell\in\mathcal{L}, \tag{40g}$$

$$\sum_{j\in\mathcal{J}_\ell} \gamma_j - s = 0, \qquad \ell\in\mathcal{L}, \tag{40h}$$

$$\sum_{j\in\mathcal{J}_\ell} \mu_j - p = 0, \qquad \ell\in\mathcal{L}, \tag{40i}$$

$$\sum_{i=1}^{n_x} \boldsymbol{F}_i \le \boldsymbol{\lambda}, \tag{40j}$$

$$\sum_{i=1}^{n_x} \boldsymbol{G}_i \le \boldsymbol{\mu}, \tag{40k}$$

$$\sum_{i=1}^{n_x} \boldsymbol{H}_i \le \boldsymbol{\nu}, \tag{40l}$$

$$\sum_{i=1}^{n_x} \boldsymbol{T}_i \leq \boldsymbol{\psi}, \tag{40m}$$

$$\sum_{i=1}^{n_x} \psi_i \leq \pi, \tag{40n}$$

$$\sum_{i=1}^{n_x} d_\ell t_i - \boldsymbol{D}_\ell^\top \boldsymbol{F}_i \leq p d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{\lambda}, \qquad \ell \in [L'], \tag{40o}$$

$$\pi - 2\sum_{i=1}^{n_x} \psi_i + \sum_{i,j=1}^{n_x} T_{ij} \geq 0, \tag{40p}$$

$$\boldsymbol{V}, \boldsymbol{F}, \boldsymbol{R}, \boldsymbol{G}, \boldsymbol{H}, \boldsymbol{Z}, \boldsymbol{T} \geq \boldsymbol{0}, \tag{40q}$$

$$\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{\psi}, \sigma, \pi \geq \boldsymbol{0}. \tag{40r}$$

Further, the pairwise multiplications of the non-linear ones result in the following constraints:

$$\|\boldsymbol{x}\|_2 \leq s, \tag{41a}$$

$$\left\|(z_i - t_i, 1)^\top\right\|_2 \leq z_i + t_i, \qquad i \in [n_x], \tag{41b}$$

$$\|\boldsymbol{X}_i\|_2 \leq \alpha_i, \qquad i \in [n_x], \tag{41c}$$

$$\|\boldsymbol{U}_i\|_2 \leq \beta_i, \qquad \imath \in [n_y], \tag{41d}$$

$$\|\boldsymbol{V}_i\|_2 \leq \gamma_i, \qquad i \in [n_x], \tag{41e}$$

$$\|\boldsymbol{F}_i\|_2 \leq \phi_i, \qquad i \in [n_x], \tag{41f}$$

$$\|\boldsymbol{\lambda}\|_2 \leq \rho, \tag{41g}$$

$$\|d_\ell \boldsymbol{x} - \boldsymbol{X}\boldsymbol{D}_\ell\|_2 \leq s d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{\alpha}, \qquad \ell \in [L'], \tag{41h}$$

$$\left\|\boldsymbol{\lambda} - \sum_{i=1}^{n_x} \boldsymbol{F}_i\right\|_2 \leq \rho - \sum_{i=1}^{n_x} \phi_i, \tag{41i}$$

$$\|\boldsymbol{X}\|_F \leq \sigma, \tag{41j}$$

$$\|(V_{ji} - F_{ji}, x_j)\|_2 \leq V_{ji} + F_{ji}, \qquad i, j \in [n_x], \tag{41k}$$

$$\|(R_{ji} - G_{ji}, y_j)\|_2 \leq R_{ji} + G_{ji}, \qquad i \in [n_x], j \in [n_y], \tag{41l}$$

$$\|(Z_{ji} - H_{ji}, z_j)\|_2 \leq Z_{ji} + H_{ji}, \qquad i, j \in [n_x], \tag{41m}$$

$$\|(H_{ij} - T_{ij}, t_j)\|_2 \leq H_{ij} + T_{ij}, \qquad i, j \in [n_x], \tag{41n}$$

$$\|(\nu_i - \psi_i, p)\|_2 \leq \nu_i + \psi_i, \qquad i \in [n_x], \tag{41o}$$

$$\left\|\left(d_\ell(z_i - t_i) + \boldsymbol{D}_\ell^\top(\boldsymbol{F}_i - \boldsymbol{V}_i), d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}\right)\right\|_2 \leq d_\ell(z_i + t_i) - \boldsymbol{D}_\ell^\top(\boldsymbol{F}_i + \boldsymbol{V}_i), \qquad \ell \in [L'], \tag{41p}$$

$$\|(\boldsymbol{V}_i - \boldsymbol{F}_i, \boldsymbol{x})\|_2 \leq \gamma_i + \phi_i, \qquad i \in [n_x], \tag{41q}$$

$$\left\|\left(\nu_i - \psi_i - \sum_{j=1}^{n_x} H_{ij} + \sum_{j=1}^{n_x} T_{ij}, p - \sum_{j=1}^{n_x} t_j\right)\right\|_2 \leq \nu_i + \psi_i - \sum_{j=1}^{n_x} H_{ij} - \sum_{j=1}^{n_x} T_{ij}, \qquad i \in [n_x], \tag{41r}$$

$$\left\|\begin{pmatrix} Z_{ij} - H_{ij} - H_{ji} + T_{ij} & z_i - t_i \\ z_j - t_j & 1 \end{pmatrix}\right\|_2 \leq Z_{ij} + H_{ij} + H_{ji} + T_{ij}, \qquad i, j \in [n_x], \tag{41s}$$

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} & \boldsymbol{V} & \boldsymbol{F} & \boldsymbol{\alpha} & \boldsymbol{\lambda} & \boldsymbol{x} \\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{R} & \boldsymbol{G} & \boldsymbol{\beta} & \boldsymbol{\mu} & \boldsymbol{y} \\ \boldsymbol{V}^\top & \boldsymbol{R}^\top & \boldsymbol{Z} & \boldsymbol{H} & \boldsymbol{\gamma} & \boldsymbol{\nu} & \boldsymbol{z} \\ \boldsymbol{F}^\top & \boldsymbol{G}^\top & \boldsymbol{H}^\top & \boldsymbol{T} & \boldsymbol{\phi} & \boldsymbol{\psi} & \boldsymbol{t} \\ \boldsymbol{\alpha}^\top & \boldsymbol{\beta}^\top & \boldsymbol{\gamma}^\top & \boldsymbol{\phi}^\top & \sigma & \rho & s \\ \boldsymbol{\lambda}^\top & \boldsymbol{\mu}^\top & \boldsymbol{\nu}^\top & \boldsymbol{\psi}^\top & \rho & \pi & p \\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & \boldsymbol{z}^\top & \boldsymbol{t}^\top & s & p & 1 \end{pmatrix} \succeq \boldsymbol{0}. \tag{41t}$$

In the RPT-SDP formulation we hence obtain $(3|\mathcal{L}|+2L+7)n_x + (|\mathcal{L}|+1)n_y + 3n_x n_y + L' n_y + 3n_x(n_x+1)/2 + 3n_x^2 + L'(L'+1)/2 + n_x L' + n_y(n_y+1)/2 + 2|\mathcal{L}| + 2L' + 6$ additional linear constraints, $5n_x^2 + n_x n_y + 5n_x + n_y + 2L' + 3$ additional second order cone constraints, one additional SDP constraint, and $3n_x(n_x+1)/2 + 3n_x^2 + n_y(n_y+1)/2 + 3n_x n_y + 6n_x + 2n_y + 3$ extra variables.

C.2. RPT-SDP formulation of Problem (31)

Replacing the objective function with the biconjugate function in (31) we obtain the following equivalent maximization problem

$$\max_{\substack{\boldsymbol{x}\in\mathcal{X}_1 \\ (\boldsymbol{y},\boldsymbol{w})\in\mathcal{Y}}} \quad \boldsymbol{x}^\top \boldsymbol{y} + \sum_{i=1}^{n_x} w_i, \tag{LSEM$_B$}$$

where $\mathcal{Y}$ is given by

$$\mathcal{Y} = \left\{ \boldsymbol{y} \in \mathbb{R}_+^{n_x},\, \boldsymbol{w} \in \mathbb{R}^{n_x} \;\middle|\; y_i \exp\left(\frac{w_i}{y_i}\right) \le 1,\, i \in [n_x],\, \sum_{i=1}^{n_x} y_i = 1 \right\}.$$

The RPT-SDP formulation is given by

$$\begin{aligned}
\max_{\substack{\boldsymbol{x},\boldsymbol{y},\boldsymbol{w}, \\ \boldsymbol{X},\boldsymbol{Y},\boldsymbol{W}, \\ \boldsymbol{U},\boldsymbol{Q},\boldsymbol{P}}} \quad & \mathrm{Tr}(\boldsymbol{U}) + \sum_{i=1}^{n_x} w_i \\
\text{s.t.} \quad & \boldsymbol{x} \in \mathcal{X}_1, && \text{(42a)} \\
& (\boldsymbol{y},\boldsymbol{w}) \in \mathcal{Y}, && \text{(42b)} \\
& \boldsymbol{X},\boldsymbol{Y},\boldsymbol{U} \ge \boldsymbol{0}, && \text{(42c)} \\
& \boldsymbol{D}^\top \boldsymbol{X}_i - \boldsymbol{d} x_i \le \boldsymbol{0}, \quad i \in [n_x], && \text{(42d)} \\
& \boldsymbol{D}^\top \boldsymbol{U}_i - \boldsymbol{d} y_i \le \boldsymbol{0}, \quad i \in [n_x], && \text{(42e)} \\
& \boldsymbol{d}\boldsymbol{x}^\top \boldsymbol{D} + \boldsymbol{D}^\top \boldsymbol{x} \boldsymbol{d}^\top \le \boldsymbol{D}^\top \boldsymbol{X} \boldsymbol{D} + \boldsymbol{d}\boldsymbol{d}^\top, && \text{(42f)} \\
& \sum_{i\in[n_x]} \boldsymbol{U}_i = \boldsymbol{x}, && \text{(42g)} \\
& \sum_{i\in[n_x]} \boldsymbol{Y}_i = \boldsymbol{y}, && \text{(42h)} \\
& \sum_{i\in[n_x]} (\boldsymbol{P})_i^\top = \boldsymbol{w}, && \text{(42i)} \\
& U_{ji} \exp\left(\frac{Q_{ji}}{U_{ji}}\right) \le x_j, \quad i,j \in [n_x], && \text{(42j)} \\
& Y_{ij} \exp\left(\frac{P_{ji}}{Y_{ij}}\right) \le y_j, \quad i,j \in [n_x], && \text{(42k)}
\end{aligned}$$

$$\left(d_\ell y_i - \boldsymbol{D}_\ell^\top \boldsymbol{U}_i\right) \exp\left(\frac{d_\ell w_i - \boldsymbol{D}_\ell^\top \boldsymbol{V}_i}{d_\ell y_i - \boldsymbol{D}_\ell^\top \boldsymbol{U}_i}\right) \le d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}, \qquad i \in [n_x],\ \ell \in \mathcal{L}, \tag{42l}$$

$$Y_{ij} \exp\left(\frac{P_{ji} + P_{ij}}{Y_{ij}}\right) \le 1, \qquad i \le j \in [n_x], \tag{42m}$$

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} & \boldsymbol{Q} & \boldsymbol{x} \\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{P} & \boldsymbol{y} \\ \boldsymbol{Q}^\top & \boldsymbol{P}^\top & \boldsymbol{W} & \boldsymbol{w} \\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & \boldsymbol{w}^\top & 1 \end{pmatrix} \succeq \boldsymbol{0}, \tag{42n}$$

where constraints (42c) - (42i) result from pairwise multiplication of the linear constraints. Note that we only multiply the linear equality constraint with the variables (see Theorem 1). Constraints (42j) - (42l) result from pairwise multiplication of the linear inequality constraints with the exponential cone constraints, constraint (42m) results from pairwise multiplication of the exponential cone constraints with each other, and constraint (42n) results from the additional SDP relaxation.

In the RPT-SDP formulation we hence obtain $n_x(n_x+1) + n_x^2 + (2L+3)n_x + L(L+1)/2$ additional linear constraints, $2n_x^2 + Ln_x + n_x(n_x+1)/2$ additional exponential cone constraints, one additional SDP constraint and $3n_x(n_x+1)/2 + 3n_x^2$ additional variables.

Observe that we could exclude the non-negativity constraints from the above reformulation, since from constraints (42b), (42g), and (42h) it follows that they are redundant.

C.3. RPT-SDP formulation of Problem (33)

Replacing the objective function with the biconjugate function in (33) we obtain the following equivalent maximization problem

$$\max_{\substack{\boldsymbol{x} \in \mathcal{X} \\ (\boldsymbol{y}, \boldsymbol{w}) \in \mathcal{Y}}} \quad -(\boldsymbol{A}^\top \boldsymbol{x} + \boldsymbol{b})^\top \boldsymbol{y} - \sum_{i \in [n_y]} w_i, \tag{CM$_B$}$$

where $\mathcal{Y}$ is given by

$$\mathcal{Y} = \left\{ \boldsymbol{y} \in \mathbb{R}_+^{n_y}, \boldsymbol{w} \in \mathbb{R}^{n_y} \;\middle|\; \exp(-w_i - 1) \le y_i,\ i \in [n_y] \right\}.$$

The RPT-SDP formulation is given by

$$\begin{aligned}
\max_{\substack{\boldsymbol{x},\boldsymbol{y},\boldsymbol{w}, \\ \boldsymbol{X},\boldsymbol{Y},\boldsymbol{W}, \\ \boldsymbol{U},\boldsymbol{Q},\boldsymbol{P}}} \quad & \mathrm{Tr}(\boldsymbol{U}\boldsymbol{A}) + \boldsymbol{b}^\top y + \sum_{i=1}^{n_y} w_i \\
\text{s.t.} \quad & \boldsymbol{D}^\top \boldsymbol{x} \le \boldsymbol{d}, && \text{(43a)} \\
& \boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{U} \ge \boldsymbol{0}, && \text{(43b)} \\
& \boldsymbol{D}^\top \boldsymbol{X}_i - \boldsymbol{d} x_i \le \boldsymbol{0}, \qquad i \in [n_x], && \text{(43c)} \\
& \boldsymbol{D}^\top \boldsymbol{U}_j - \boldsymbol{d} y_j \le \boldsymbol{0}, \qquad j \in [n_y], && \text{(43d)} \\
& \boldsymbol{d}\boldsymbol{x}^\top \boldsymbol{D} + \boldsymbol{D}^\top \boldsymbol{x}\boldsymbol{d}^\top \le \boldsymbol{D}^\top \boldsymbol{X}\boldsymbol{D} + \boldsymbol{d}\boldsymbol{d}^\top, && \text{(43e)} \\
& \exp(-w_i - 1) \le y_i, \qquad i \in [n_y], && \text{(43f)}
\end{aligned}$$

$$x_j \exp\left(\frac{-Q_{ji}-x_j}{x_j}\right) \le U_{ji}, \qquad i \in [n_y],\, j \in [n_x], \quad (43g)$$

$$y_j \exp\left(\frac{-P_{ji}-y_j}{y_j}\right) \le Y_{ij}, \qquad i,j \in [n_y], \quad (43h)$$

$$(d_j - \boldsymbol{D}_j^\top \boldsymbol{x}) \exp\left(\frac{\boldsymbol{D}_j^\top \boldsymbol{x} - d_j - w_i d_j + \boldsymbol{D}_j^\top \boldsymbol{Q}_i}{d_j - \boldsymbol{D}_j^\top \boldsymbol{x}}\right) \le d_j y_i - \boldsymbol{D}_j^\top \boldsymbol{U}_i, \qquad i \in [n_y],\, j \in [L], \quad (43i)$$

$$\exp(-w_i - w_j - 2) \le Y_{ij}, \qquad i \le j \in [n_y], \quad (43j)$$

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} & \boldsymbol{Q} & \boldsymbol{x} \\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{P} & \boldsymbol{y} \\ \boldsymbol{Q}^\top & \boldsymbol{P}^\top & \boldsymbol{W} & \boldsymbol{w} \\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & \boldsymbol{w}^\top & 1 \end{pmatrix} \succeq \boldsymbol{0}, \quad (43k)$$

where constraints (43b) - (43d) result from pairwise multiplication of the linear constraints, constraints (43g) - (43j) result from pairwise multiplication of the exponential constraints with the linear and constraint (43k) results from the additional SDP relaxation.

In the RPT-SDP formulation we hence obtain $n_x(n_x+1)/2 + n_y(n_y+1)/2 + n_x n_y + L(n_x+n_y) + L(L+1)/2$ additional linear constraints, $n_x n_y + n_y^2 + L n_y + n_y(n_y+1)/2$ additional exponential cone constraints, one additional SDP constraint and $n_x(n_x+1)/2 + n_y(n_y+1) + 2n_x n_y + n_y^2$ additional variables.

C.4. RPT-SDP formulation of Problem (36)

The convex quadratic constraints ($\mathcal{C}$) are reformulated as second order cone constraints, that is $\left\|\boldsymbol{P}_i^{1/2}\boldsymbol{x}\right\|_2 \le -r_i - \boldsymbol{q}_i^\top \boldsymbol{x}$. The nonconvex quadratic constraints ($\mathcal{NC}$) are linearized as follows: $tr(\boldsymbol{P}_i\boldsymbol{X}) + q_i^\top \boldsymbol{x} + r_i \le 0$. The RPT-SDP formulation is given by

$$\min_{\boldsymbol{x},\boldsymbol{X}} \quad \mathrm{Tr}(\boldsymbol{P}_0\boldsymbol{X}) + \boldsymbol{q}_0^\top x + r_0$$

$$\text{s.t.} \quad \boldsymbol{x} \in \mathcal{X}_1, \quad (44a)$$

$$\boldsymbol{D}^\top \boldsymbol{X}_i - \boldsymbol{d} x_i \le \boldsymbol{0}, \qquad i \in [n_x], \quad (44b)$$

$$\boldsymbol{d}\boldsymbol{x}^\top \boldsymbol{D} + \boldsymbol{D}^\top \boldsymbol{x}\boldsymbol{d}^\top \le \boldsymbol{D}^\top \boldsymbol{X}\boldsymbol{D} + \boldsymbol{d}\boldsymbol{d}^\top, \quad (44c)$$

$$\mathrm{Tr}(\boldsymbol{P}_k\boldsymbol{X}) + \boldsymbol{k}_0^\top x + r_k \le 0, \qquad k \in \mathcal{NC}, \quad (44d)$$

$$\left\|\boldsymbol{P}_k^{1/2}\boldsymbol{x}\right\|_2 \le -r_k - \boldsymbol{q}_k^\top \boldsymbol{x}, \qquad k \in \mathcal{C}, \quad (44e)$$

$$\left\|\boldsymbol{P}_i^{1/2}\boldsymbol{X}\boldsymbol{P}_j^{1/2}\right\|_2 \le r_i r_j + r_i \boldsymbol{q}_j^\top x + r_j \boldsymbol{q}_i^\top \boldsymbol{x} + \boldsymbol{q}_i^\top \boldsymbol{X}\boldsymbol{q}_j, \qquad i,j \in \mathcal{C}, \quad (44f)$$

$$\left\|d_\ell \boldsymbol{P}_k^{1/2}\boldsymbol{x} - \boldsymbol{P}_k^{1/2}\boldsymbol{X}\boldsymbol{D}_\ell\right\|_2 \le -r_k d_\ell + r_k \boldsymbol{D}_\ell^\top \boldsymbol{x} - d_\ell \boldsymbol{q}_k^\top \boldsymbol{x} + \boldsymbol{q}_k^\top \boldsymbol{X}\boldsymbol{D}_\ell, \qquad k \in \mathcal{C},\, \ell \in \mathcal{L}, \quad (44g)$$

$$\left\|\boldsymbol{P}_k^{1/2}\boldsymbol{X}_j\right\| \le -r_k x_j - \boldsymbol{q}_k^\top \boldsymbol{X}_j, \qquad k \in \mathcal{C},\, j \in [n_x], \quad (44h)$$

$$\boldsymbol{X} \ge \boldsymbol{0}, \quad (44i)$$

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^\top & 1 \end{pmatrix} \succeq \boldsymbol{0}, \quad (44j)$$

where constraints (44b) - (44c) and (44i) result from pairwise multiplication of the linear constraints, constraints (44f) - (44h) result from pairwise multiplication of the convex quadratic constraints with the linear constraints and each other and constraint (44j) results from the additional SDP relaxation.

In the RPT-SDP formulation we hence obtain $Ln_x + n_x(n_x+1)/2 + L(L+1)/2 + |\mathcal{NC}|$ additional linear constraints, $(n_x + L + 1)|\mathcal{C}| + |\mathcal{C}|(|\mathcal{C}|+1)/2$ additional second order cone constraints, one additional SDP constraint and $n_x(n_x+1)/2$ additional variables.

C.5. RPT-SDP formulation of Problem (DHO)

We introduce the following epigraphical variables: We use z_k for the nonconvex terms in the objective $(C + bx_k)\exp\left(\lambda\sum_{i=0}^k x_i - \delta t_k\right)$, and w_k for $\exp\left(-\theta\sum_{i=0}^k x_i\right)$. The RPT-SDP formulation is given by

$$
\begin{aligned}
\min_{\substack{\boldsymbol{x},\boldsymbol{u},\boldsymbol{v},\boldsymbol{z},\\ \boldsymbol{q},\boldsymbol{X},\boldsymbol{V},\\ \boldsymbol{S},w,W}} \quad & \sum_{k\in\mathcal{K}} z_k + \sum_{k\in\mathcal{K}_0} \frac{S_0}{\beta_\delta}\left(\exp\left(\beta_\delta t_{k+1}\right) - \exp\left(\beta_\delta t_k\right)\right) w_k + \frac{S_0}{\delta}\exp\left(\beta_\delta T\right) w_K \\
\text{s.t.} \quad & (C + bx_k)\exp\left(\frac{\lambda C h_k + \lambda b\sum_{i=0}^k X_{ik} - \delta t_k(C + bx_k)}{C + bx_k}\right) \le z_k, && k\in\mathcal{K}_0, && \text{(45a)}\\
& \boldsymbol{D}^\top \boldsymbol{x} \le \boldsymbol{d}, && && \text{(45b)}\\
& \exp\left(-\theta\sum_{i=0}^k x_i\right) \le w_k, && k\in\mathcal{K}_0, && \text{(45c)}\\
& \boldsymbol{D}^\top \boldsymbol{X}_k \le \boldsymbol{x}_k d, && k\in\mathcal{K}_0, && \text{(45d)}\\
& \boldsymbol{d}\boldsymbol{x}^\top \boldsymbol{D} + \boldsymbol{D}^\top \boldsymbol{x}\boldsymbol{d}^\top \le \boldsymbol{D}^\top \boldsymbol{X}\boldsymbol{D} + \boldsymbol{d}\boldsymbol{d}^\top, && && \text{(45e)}\\
& (d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x})\exp\left(\frac{-\theta d_\ell\sum_{i=0}^k x_i + \theta\sum_{i=0}^k \boldsymbol{D}_\ell^\top \boldsymbol{X}_i}{d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}}\right) \le d_\ell w_k - \boldsymbol{D}_\ell^\top \boldsymbol{Q}_k, && k\in\mathcal{K}_0, \ell\in\mathcal{L}, && \text{(45f)}\\
& x_j \exp\left(\frac{-\theta\sum_{i=0}^k X_{ij}}{x_j}\right) \le Q_{jk}, && k,j\in\mathcal{K}_0, && \text{(45g)}\\
& \exp\left(-\theta\sum_{i=1}^k x_i - \theta\sum_{i=1}^j x_i\right) \le W_{jk}, && j,k\in\mathcal{K}_0, && \text{(45h)}\\
& w_j \exp\left(\frac{-\theta\sum_{i=1}^k Q_{kj}}{w_j}\right) \le W_{jk}, && j,k\in\mathcal{K}_0, && \text{(45i)}\\
& \boldsymbol{x}, \boldsymbol{X} \ge \boldsymbol{0}, && && \text{(45j)}\\
& \begin{pmatrix} \boldsymbol{X} & \boldsymbol{Q} & \boldsymbol{x} \\ \boldsymbol{Q}^\top & \boldsymbol{W} & \boldsymbol{w} \\ \boldsymbol{x}^\top & \boldsymbol{w}^\top & 1 \end{pmatrix} \succeq \boldsymbol{0}, && && \text{(45k)}
\end{aligned}
$$

where constraint (45b) represents the upper bounds on $\boldsymbol{x}$ such that we obtain a compact feasible region, (45d) - (45f) result from pairwise multiplication of (45b) with all other constraints, (45g) results from pairwise multiplication of (45c) with the nonnegativity constraint $\boldsymbol{x} \geq 0$, (45h) - (45i) result from pairwise multiplication of the exponential constraint with itself, (45j) results from pairwise multiplication of the nonnegativity constraints, and constraint (45k) results from the additional SDP relaxation.

In the RPT-SDP formulation we hence obtain $|\mathcal{K}_0| + L(L+1)/2 + |\mathcal{K}_0|(|\mathcal{K}_0|+1)/2$ additional linear inequalities, $(L+1)|\mathcal{K}_0| + 2|\mathcal{K}_0|^2$ additional exponential cone inequalities, and one additional LMI.

D. Data generation of numerical experiments

D.1. Data generation of numerical experiments of Problem (28)

We use the data generated by Selvi et al. (2022, Appendix F.5). Instances 1 - 5 refer to the instances 1, 2, 3, 7, and 11 in Selvi et al. (2022, Appendix F.5) respectively. In every problem, every max-term has the same number of elements, i.e., $|\mathcal{J}_\ell| = |\mathcal{J}_{\ell'}|$ for every $\ell, \ell' \in \mathcal{L}$.

Problem instance 1: $A_{ij} \sim [-5,5], b_j = 0, \boldsymbol{D}^\top = \boldsymbol{I}, d_i = n_x/i,$

Problem instance 2: $A_{ij} \sim [-5,5], b_j = 0, \boldsymbol{D}^\top = \boldsymbol{I}, d_i = n_x/i,$

Problem instance 3: $A_{ij} \sim [-5,5], b_j = 0, \boldsymbol{D}^\top = \boldsymbol{I}, d_i = n_x/i,$

Problem instance 4: $A_{ij} \sim [-5,5], b_j \sim [-10,10], D_{ij} \sim [0,1], d_i \sim [5,15],$

Problem instance 5: $A_{ij} \sim [-5,10], b_j \sim [-10,10],$ 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. The values for the parameters of each distinct problem instance are given in Table 9.

Instance	n_x	$\lvert\mathcal{L}\rvert$	$\lvert\mathcal{J}_\ell\rvert$	a	c	M
1	5	1	5	3	6	100
2	5	10	5	3	6	100
3	20	10	10	11	25	1000
4	10	2	5	3	7	1000
5	20	10	10	5	30	1000

Table 9 Problem (29) parameters for each instance. n_x refers to the number of variables, $|\mathcal{L}|$ to the number of max linear terms, $|\mathcal{J}_\ell|$ to the number of elements within a max-term, a to the parameter used in $\mathcal{X}_2$, c to the parameter used in $\mathcal{X}_3$ and M to the big M parameter used in Problem (30).

D.2. Data generation of numerical experiments of Problem (31)

The problem instances are adopted from Selvi et al. (2022) and can be summarized as follows: In instances 1 and 2 the linear constraints are defined as

$$-\frac{i}{n} \leq x_i \leq \frac{i}{n},$$

in instance 3 as

$$x_i \leq 8,\ \ x_i + x_j \leq u_{ij},$$

where $u_{ij} \sim [5, 15]$. Finally, for the last two we have

$$\textbf{Problem instance 4: } D_{ij} \sim [0,1], d_i \sim [10,30],$$

$$\textbf{Problem instance 5: } D_{ij} \sim [0,1], d_i \sim [20,60].$$

The parameters describing each instance are summarized in Table 10.

Table 10 Problem (31) parameters for each instance. n_x refers to the number of variables and L to the number of linear constraints.

Instance	n_x	L
1	10	20
2	40	80
3	10	100
4	20	20
5	50	50

D.3. Data generation of numerical experiments of Problem (33)

The problem instances were generated in the same way as in BARON (Ryoo and Sahinidis, 1996). Namely, the constraint coefficients were generated as $D_{ij} \sim [-100, 0], d_i \sim [-100, 0]$ and the linear terms as $A_{ij} \sim [0, 10], b_i \sim [0, 10]$. The parameters describing each instance are summarized in Table 11.

Table 11 Problem (33) parameters for each instance. n_x refers to the number of variables, L to the number of linear constraints, and n_y to the number of linear multiplications in the objective.

Instance	n_x	L	n_y
1	5	5	5
2	7	7	10
3	10	10	9
4	20	20	8
5	40	40	4

D.4. Data generation of numerical experiments of Problem (36)

The first 5 problem instances are adopted from Selvi et al. (2022) and can be summarized as follows: In instances 1 and 2 the objectives are $f(\boldsymbol{x}) = -\frac{1}{2}\sum_{i=1}^{20}(x_i-2)^2$ and $f(\boldsymbol{x}) = -\frac{1}{2}\sum_{i=1}^{20}(x_i+5)^2$ respectively and the linear constraints are as in instance 11 for problem (28). Instances 3, 4 and 5 are defined by the linear constraints $\boldsymbol{D}^\top \boldsymbol{x} \leq \boldsymbol{d}, \quad \boldsymbol{x} \leq x_u \boldsymbol{e}$, where

$$\textbf{Problem instance 3: } D_{ij} \sim [0,1], d_i \sim [20,60], x_u = 5,$$

$$\textbf{Problem instance 4: } D_{ij} \sim [0,1], d_i \sim [30,60], x_u = 3,$$

$$\textbf{Problem instance 5: } D_{ij} \sim [0,1], d_i \sim [80,120], x_u = 2.$$

Instances 6, 7, 8, 9 and 10 were adopted from Al-Khayyal et al. (1995). Each matrix $\boldsymbol{P}_i \in \mathbb{R}^{n_x \times n_x}$ in both the objective and the constraints has integer entries uniformly at random between -10 and 10 and further in each row, half of the entries are randomly set to 0. Each vector $\boldsymbol{q}_i \in \mathbb{R}^{n_x}$ is also generated with integer entries between -10 and 10 and each r_i is set to 0. The parameters describing each instance are summarized in Table 12.

Table 12 Problem (36) parameters for each instance. n_x refers to the number of variables, L to the number of linear constraints and nc-q to the number of nonconvex quadratic constraints.

Instance	n_x	L	$nc-q$
1	20	10	0
2	20	10	0
3	10	15	0
4	50	62	0
5	100	130	0
6	8	8	4
7	12	12	6
8	16	16	8
9	30	30	15
10	40	40	20

D.5. Data generation of numerical experiments of Problem (DHO)

The linear constraints are defined as $x_i \leq 300$. Moreover, the time periods in each instance are:

$$\boldsymbol{t}_{25} = (0,25,50,75,100,125,150,175,200,225,250,275)^\top,$$
$$\boldsymbol{t}_{50} = (0,50,100,150,200,250)^\top,$$
$$\boldsymbol{t}_{ir} = (0,20,50,90,130,155,180,210,255,270)^\top.$$

In each instance, the number of variables n_x is equal to the number of time periods. The parameters describing each instance are summarized in Table 13. Moreover, we have $\theta = \alpha - \zeta,\ \beta_\delta = \alpha\eta + \gamma - \delta$.

Table 13 Problem (DHO) parameters for each instance.

Instance	α	C	b	λ	ζ	η	S_0	γ	δ	T
10	0.033027	16.6939	0.6258	0.0014	0.003774	0.32	0.68938	0.02	0.04	300
15	0.0502	125.6422	1.1268	0.0098	0.003764	0.76	16.2008	0.02	0.04	300
16	0.0574	324.6287	2.1304	0.01	0.002032	0.76	25.0071	0.02	0.04	300