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 written in this way, such as concave minimization problems, difference of convex problems, and fractional optimization problems. Our approach exploits 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 develop a spatial Branch and Bound scheme, leveraging RPT, to obtain a global optimal solution. Numerical experiments on four different convex maximization problems, 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 problem instancdes of larger dimension our approach outperforms both BARON and SCIP on computation time for most problem instances, while for smaller dimension BARON overall performs better on computation time.

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

---

1. Introduction

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 written in this way, such as nonconvex quadratic optimization

* Corresponding author.

problems, mixed binary linear optimization problems, concave minimization (or equivalently, convex maximization) problems, difference of convex problems, and fractional optimization problems.

For nonconvex quadratic optimization problems and mixed binary linear optimization problems, hierarchical convex approximations can be obtained from the Reformulation-Linearization Technique (RLT). RLT was introduced in Sherali and Adams (1990), and improved by many authors (Sturm and Zhang, 2003; Anstreicher, 2009, 2012, 2017; Bao et al., 2011; Yang and Burer, 2016; Jiang and Li, 2016, 2019). RLT is also applicable to mixed binary polynomial and to continuous, nonconvex optimization problems (Sherali and Adams, 1999), and has been extended to mixed binary semi-infinite and convex optimization problems (Sherali and Adams, 1994a). 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 propose a spatial branch and bound scheme, levering RPT, 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. More recently, Chen and Burer (2012) develop a branch and bound method, utilizing co-positive programming, addressing nonconvex quadratic problems over linear constraints. 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. Later, Achterberg (2009) develops the global optimization algorithm SCIP, addressing nonconvex problems, with an emphasis on integer problems, utilizing branch and bound.

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 present a novel way for obtaining convex relaxations as well as a novel partitioning scheme leveraging the solution of the latter. We refer to our approach by 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. (2020) develop an algorithm based on adjustable robust optimization and Ben-Tal and Roos (2022) develop an algorithm, called CoMax, based on gradient ascent. The latter is also applicable in integer optimization problems, where the feasible set is a polytope. While both methods can find high quality bounds on the optimal solution, there is no guarantee about convergence to the global optimum. Moreover, both methods cannot handle nonconvex constraints.

Our main contributions can be summarized as follows:

1. We extend the existing RLT approach to a broader class of nonlinear optimization problems. 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. The proposed RPT-BB approach can obtain the global optimal solution of nonconvex optimization problems in which the nonconvex components are sums of products of convex functions.
3. We demonstrate the effectiveness of the proposed RPT-BB approach, by conducting numerical experiments on four different convex maximization problems, a quadratic constraint quadratic optimization problem, and a dike height optimization problem. We show that the RPT-BB 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 find the global optimal solution and 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. 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.*, $\mathcal{I}=\{1,\dots,I\}$ etc. The subscript 0 for an index set indicates that the set additionally includes 0, *i.e.*, $\mathcal{I}_0=\{0,\dots,I\}$ 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 $\mathcal{V}\subseteq\mathbb{R}^{n_\nu}$. The *domain* of a function $f:\mathbb{R}^{n_\nu}\to[-\infty,+\infty]$ is defined as $\mathrm{dom}(f)=\{\boldsymbol{\nu}\in\mathbb{R}^{n_\nu}\mid f(\boldsymbol{\nu})<+\infty\}$. The function f is *proper* if $f(\boldsymbol{\nu})>-\infty$ for all $\boldsymbol{\nu}\in\mathbb{R}^{n_\nu}$ and $f(\boldsymbol{\nu})<+\infty$ for at least one $\boldsymbol{\nu}\in\mathbb{R}^{n_\nu}$, 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 $\boldsymbol{\nu}\in\mathbb{R}^{n_\nu}$ or $f(\boldsymbol{\nu})=-\infty$ for all $\boldsymbol{\nu}\in\mathbb{R}^{n_\nu}$. The *conjugate* of a function $f:\mathbb{R}^{n_\nu}\to[-\infty,+\infty]$ is the function $f^*:\mathbb{R}^{n_\nu}\to[-\infty,+\infty]$ 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* $\delta_{\mathcal{V}}:\mathbb{R}^{n_\nu}\to[-\infty,+\infty]$ of a set $\mathcal{V}\subseteq\mathbb{R}^{n_\nu}$ 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* $\delta^*_{\mathcal{V}}:\mathbb{R}^{n_\nu}\to[-\infty,+\infty]$ of a set $\mathcal{V}\subseteq\mathbb{R}^{n_\nu}$ 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,

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

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

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

where $\boldsymbol{A} \in \mathbb{R}^{L_1 \times n_x}$, $\boldsymbol{P} \in \mathbb{R}^{L_2 \times n_x}$, $\boldsymbol{b} \in \mathbb{R}^{L_1}$, $\boldsymbol{s} \in \mathbb{R}^{L_2}$, $\boldsymbol{h}(\boldsymbol{x}) = [h_0(\boldsymbol{x})\ h_1(\boldsymbol{x})\ \cdots\ h_J(\boldsymbol{x})]^\top \subseteq [-\infty, +\infty]^{J+1}$, and $h_j : \mathbb{R}^{n_x} \to (-\infty, +\infty]$ is proper, closed and convex for every $j \in \mathcal{J}_0$. 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 $r_{ik}(\boldsymbol{x}) \geq 0$ and $c_{ik}(\boldsymbol{x}) \geq 0$ 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 $r_{ik}(\boldsymbol{x}) \geq 0$ 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:

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

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,

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

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). An important class of SLC representable functions are difference of convex (DC) functions. For example, every twicely differentiable continuous function has a DC decomposition (see Hartman (1959)) and can therefore be written as an SLC function. If the constraint in (1) contains a difference of convex function, that is, $f_k(\boldsymbol{x}) = c_{0k}(\boldsymbol{x}) - c_{1k}(\boldsymbol{x}) \leq 0$, where c_{0k}, $c_{1k} : \mathbb{R}^{n_x} \to (-\infty, +\infty]$ are proper, closed and convex for some $k \in \mathcal{K}_0$, then we can reformulate the corresponding constraint function into an SLC function using the biconjugate reformulation (Rockafellar, 1970) and obtain

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

as long as the infimum is attained, since we then can delete the inf operator. In case the infimum is not attained we refer to Appendix A. Observe that if the DC function appears in the objective instead of in the constraints, then using the biconjugate reformulation we can rewrite the objective as a bilinear function analogously.

For many important classes of convex functions, their conjugates and domains are readily available from the literature. Table 1 lists several convex functions and their conjugates and domains.

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

#	f	$\mathrm{dom}(f^*)$	f^*
1	$f(\boldsymbol{x},\bar{x}) = \|\boldsymbol{x}\|_2 - \bar{x}$	$\{(\boldsymbol{y},\bar{y}) : \|\boldsymbol{y}\|_2 \leq 1,\ \bar{y}=1\}$	$f^*(\boldsymbol{y},\bar{y}) = 0$
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	$f(\boldsymbol{x}) = \max_i x_i$	$\{\boldsymbol{y} : \boldsymbol{y} \geq 0, \sum_i y_i = 1, \forall k\}$	$f^*(\boldsymbol{y}) = 0$
6	$f(\boldsymbol{x}) = \sum_k \max_{i\in I_k} x_i$	$\{\boldsymbol{y}_k : \boldsymbol{y}_k \geq 0,\ \sum_i y_{ki} = 1, \forall k\}$	$f^*(\boldsymbol{w}) = 0$
7	$f(\boldsymbol{x}) = \log(\sum_i \exp(x_i))$	$\{\boldsymbol{y} : \boldsymbol{y} \geq 0, \sum_i y_i = 1\}$	$f^*(\boldsymbol{y}) = \sum_i y_i \log(y_i)$
8	$f(\boldsymbol{x}) = \sum_i f_i(\boldsymbol{x})$	$\{\{\boldsymbol{y}_i\}_i : \sum_i \boldsymbol{y}_i = \boldsymbol{y}, \boldsymbol{y}_i \in \mathrm{dom}(f_i^*), \forall i\}$	$f^*(\boldsymbol{y}) = \min_{\{\boldsymbol{y}_i\}_i} \sum_i f_i^*(\boldsymbol{y}_i)$

In the case that the constraint functions do not admit a closed form conjugate, often one can write the conjugate function as an infimum over some additional variables, see for more detail Roos et al. (2020).

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, the approach proposed in this paper can deal with Problems (1) containing (sum of) such nonconvex components.

□

Table 2 Examples of SLC representable functions.

#	f	c	$(q-\boldsymbol{d}^\top\boldsymbol{x})$	Perspectification	Assumptions
1	$-x\ln x$	$-\ln x$	x	$-x\ln(u/x)$	$x\geq 0$
2	$\sqrt{x}$	$x^{-1/2}$	x	$x\sqrt{x/u}$	$x\geq 0$
3	x^θ	$x^{\theta-1}$	x	$x(u/x)^{\theta-1}$	$\theta\in[0,1]$ & $x\geq 0$
4	$-x^\theta$	$-x^{\theta-1}$	x	$-x(u/x)^{\theta-1}$	$\theta\in[1,2]$ & $x\geq 0$
5	$-x_1\ln x_2$	$-\ln x_2$	x_1	$-x_1\ln(u_{12}/x_1)$	$x_1,x_2\geq 0$
6	$\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}$	$(\boldsymbol{Q}\boldsymbol{x})_i$	x_i	$\mathrm{Tr}(\boldsymbol{U}\boldsymbol{Q})$	-
7	$(q-\boldsymbol{d}^\top\boldsymbol{x})\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}$	$\boldsymbol{x}^\top\boldsymbol{Q}\boldsymbol{x}$	$(q-\boldsymbol{d}^\top\boldsymbol{x})$	$\frac{(q\boldsymbol{x}-\boldsymbol{U}\boldsymbol{d})^\top\boldsymbol{Q}(q\boldsymbol{x}-\boldsymbol{U}\boldsymbol{d})}{(q-\boldsymbol{d}^\top\boldsymbol{x})}$	$\boldsymbol{d}^\top\boldsymbol{x}\leq q$ & $\boldsymbol{Q}\succeq\boldsymbol{0}$

3. 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 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 every convex component in the nonconvex SLC functions of (1), and from (2) we have

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

where $\mathcal{T} = \{(\boldsymbol{x},\boldsymbol{\tau}) \in \mathbb{R}^{n_x} \times \mathbb{R}^{K+1} \mid \boldsymbol{x} \in \mathcal{X},\ \boldsymbol{c}_0(\boldsymbol{x}) \leq \boldsymbol{\tau}\}$, and $\boldsymbol{c}_0(\boldsymbol{x}) = [c_{00}(\boldsymbol{x})\ c_{01}(\boldsymbol{x})\ \cdots\ c_{0K}(\boldsymbol{x})]^\top \subseteq (-\infty, +\infty]^{K+1}$. As we will see later, 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. 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 $\boldsymbol{U} \in \mathbb{S}^{n_x}$ to obtain the following sum of perspective functions:

$$\sum_{i\in\mathcal{I}} (q_i - \boldsymbol{d}_i^\top \boldsymbol{x}) c_i \left(\frac{q_i \boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}_i}{q_i - \boldsymbol{d}_i^\top \boldsymbol{x}} \right), \tag{4}$$

which is jointly convex in $(\boldsymbol{x}, \boldsymbol{U})$ 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., $q_i - \boldsymbol{d}_i^\top \boldsymbol{x} \leq 0$ 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:

$$\begin{aligned} \min_{\boldsymbol{x},\boldsymbol{\tau},\boldsymbol{U}} \quad & \tau_0 + \sum_{i\in\mathcal{I}} (q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}) c_{i0} \left(\frac{q_{i0} \boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}_{i0}}{q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}} \right) \\ \text{s.t.} \quad & \tau_k + \sum_{i\in\mathcal{I}} (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) c_{ik} \left(\frac{q_{ik} \boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}} \right) \leq 0 \qquad k \in \mathcal{K} \\ & (\boldsymbol{x}, \boldsymbol{\tau}) \in \mathcal{T}. \end{aligned} \tag{5}$$

By pairwise multiplying inequalities in the set $\mathcal{T}$, we can obtain additional redundant SLC constraints which can then be convexified in a similar manner as described above. The convexified SLC constraints are then not redundant anymore and actually serve as bounds on the newly introduced variables for the product terms. We can pairwise multiply the linear inequality constraints in the set $\mathcal{T}$, similarly as in RLT, to obtain bounds on the newly introduced variables $\boldsymbol{U} \in \mathbb{S}^{n_x}$. However, with RPT we can adaptively improve this approximation by also considering pairwise multiplication of the linear and convex constraints in the set $\mathcal{T}$ and subsequently pairwise multiplication of the convex inequalities in the set $\mathcal{T}$. To be more precise, by considering the following cases of pairwise multiplying the constraints in the set $\mathcal{T}$, we adaptively improve the convex RPT approximation of (1):

Linear inequality × Linear inequality. This is well-known in RLT: we multiply the constraints $\boldsymbol{Ax} \leq \boldsymbol{b}$ of (1) with $\boldsymbol{Ax} \leq \boldsymbol{b}$, and obtain $L_1(L_1+1)/2$ redundant constraints:

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

since the (i,j)-th constraint is exactly the (j,i)-th constraint. Hence, we only consider the upper triangular of the matrix equations; so $L_1(L_1+1)/2$ constraints instead of L_1^2. Next, the nonlinear quadratic terms $\boldsymbol{x}\boldsymbol{x}^\top$ in (1) and the additional redundant constraints are linearized by substituting them with $\boldsymbol{U} \in \mathbb{S}^{n_x}$. We then obtain the convex relaxation

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

Linear inequality × Convex inequality. By multiplying each ℓ-th linear inequality $\boldsymbol{a}_\ell^\top \boldsymbol{x} \leq b_\ell$ of (1) with the convex constraints $\boldsymbol{c}_0(\boldsymbol{x}) \leq \boldsymbol{\tau}$ and $\boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0}$, we obtain $L_1(J+K+2)$ redundant SLC constraints of the form

$$(b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{c}_0(\boldsymbol{x}) \leq (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{\tau} \quad \text{and} \quad (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0} \qquad \ell \in \mathcal{L}_1.$$

Next, the redundant SLC constraints can be reformulated into:

$$(b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0} \quad \Longleftrightarrow \quad (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{h}\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}} \right) \leq \boldsymbol{0} \quad \text{and}$$

$$(b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{c}(\boldsymbol{x}) \leq (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{\tau} \quad \Longleftrightarrow \quad (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{c}\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{x}\boldsymbol{x}^\top \boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}} \right) \leq (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{\tau}.$$

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

$$(b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{h}\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}} \right) \leq \boldsymbol{0} \quad \text{and} \quad (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{c}\left(\frac{b_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}} \right) \leq b_\ell \boldsymbol{\tau} - \boldsymbol{V}^\top \boldsymbol{a}_\ell.$$

Moreover, we include the additional constraints

$$u_{ii} \geq 0, \quad i \in [n_x],$$

since $x_i^2 \geq 0$ for all $i \in [n_x]$.

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

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 $(q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{q\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) = 0$ is redundant to $\left\{(\boldsymbol{x}, \boldsymbol{U}) \mid \boldsymbol{d}^\top \boldsymbol{x} = q,\ \boldsymbol{U}\boldsymbol{d} = q\boldsymbol{x}\right\}$.*

Proof. Since $\boldsymbol{d}^\top \boldsymbol{x} = q$, $\boldsymbol{U}\boldsymbol{d} = q\boldsymbol{x}$, 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

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

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 $L_2 n_x + L_2 n_\tau$ redundant SLC constraints of the form

$$(s_\ell - \boldsymbol{p}_\ell^\top \boldsymbol{x})\boldsymbol{x} = \boldsymbol{0} \quad \text{and} \quad (s_\ell - \boldsymbol{p}_\ell^\top \boldsymbol{x})\boldsymbol{\tau} = \boldsymbol{0} \qquad \ell \in \mathcal{L}_2.$$

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

$$\begin{array}{lll}
\min\limits_{\boldsymbol{x},\boldsymbol{\tau},\boldsymbol{U},\boldsymbol{V}} & \tau_0 + \sum\limits_{i \in \mathcal{I}} (q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}) c_{i0}\left(\dfrac{q_{i0}\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}_{i0}}{q_{i0} - \boldsymbol{d}_{i0}^\top \boldsymbol{x}}\right) & \\
\text{s.t.} & \tau_k + \sum\limits_{i \in \mathcal{I}} (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) c_{ik}\left(\dfrac{q_{ik}\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}}\right) \leq 0, & k \in \mathcal{K}, \\
& \boldsymbol{b}\boldsymbol{x}^\top \boldsymbol{A}^\top + \boldsymbol{A}\boldsymbol{x}\boldsymbol{b}^\top \leq \boldsymbol{A}\boldsymbol{U}\boldsymbol{A}^\top + \boldsymbol{b}\boldsymbol{b}^\top, & \\
& u_{ii} \geq 0, & i \in [n_x] \\
& (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{h}\left(\dfrac{b_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}}\right) \leq \boldsymbol{0}, & \ell \in \mathcal{L}_1, \\
& (b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x})\boldsymbol{c}_0\left(\dfrac{b_\ell \boldsymbol{x} - \boldsymbol{U}\boldsymbol{a}_\ell}{b_\ell - \boldsymbol{a}_\ell^\top \boldsymbol{x}}\right) \leq b_\ell \boldsymbol{\tau} - \boldsymbol{V}^\top \boldsymbol{a}_\ell, & \ell \in \mathcal{L}_1, \\
& s_\ell \boldsymbol{x} - \boldsymbol{U} p_\ell = \boldsymbol{0}, & \ell \in \mathcal{L}_2, \\
& s_\ell \boldsymbol{\tau} - \boldsymbol{V} p_\ell = \boldsymbol{0}, & \ell \in \mathcal{L}_2, \\
& (\boldsymbol{x}, \boldsymbol{\tau}) \in \mathcal{T}. &
\end{array} \tag{7}$$

Note that there are $L_1(J+1+\frac{L_1+1}{2})+(L_1+1)(K+1)+(L_2+1)n_x+L_2 n_\tau$ additional constraints and $n_x^2+(n_x+1)(K+1)$ additional variables in (7) compared to (1).

Remark 1. We remark that introducing epigraph variables for the nonlinear convex components in the prepocessing step tightens the RPT approximation. To demonstrate this, consider the following convex maximization problem:

$$\begin{array}{ll} \max\limits_{x} & x_1 x_2 - \frac{(x_2-2)^2}{4} \\ \text{s.t.} & 0 \le x_1 \le 1. \end{array} \tag{8}$$

In the following we compare two convex relaxations of (8) obtained from (i) applying RPT without the epigraphical reformulation, and (ii) applying RPT with the epigraphical reformulation, respectively,

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

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

Remark 2. If the objective or a constraint function is a sum of several nonlinear convex functions, we can introduce epigraphical variables for each convex function. In this way, we can pairwise multiply these epigraphical constraints with each other and all other constraints in the set $\mathcal{T}$ to obtain tighter bounds on the newly introduced variables. □

Convex inequality × Convex inequality. Just multiplying a nonlinear convex constraint $h_j(\boldsymbol{x}) \le 0$ with another nonlinear convex constraint $h_{j'}(\boldsymbol{x}) \le 0$ results in a constraint $-h_j(\boldsymbol{x})h_{j'}(\boldsymbol{x}) \le 0$ for which the constraint function is not an 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:

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

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:

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

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, 2016).

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

$$\begin{pmatrix} \boldsymbol{U} & \boldsymbol{x} \\ \boldsymbol{x}^\top & 1 \end{pmatrix} \succeq 0. \tag{9}$$

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

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

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 $u_{ii} \ge 0,\ i \in [n_x]$, are redundant to the LMI.

Remark 3. Several convex reformulations and relaxations of several classes of nonconvex problems derived in the literature can also be obtained via RPT or RPT including the SDP relaxation (RPT-SDP). In Appendix B this is shown for disjunctive optimization, generalized linear optimization, the approximate S-lemma for quadratically constrained quadratic optimization, and fractional optimization. □

3.4. Obtaining upper bounds

Since RPT is a conservative approximation, it yields a lower bound for the optimal objective value for problem (1). We can also obtain upper bounds in the following way. Let $(\boldsymbol{x}^*, \boldsymbol{\tau}^*, \boldsymbol{U}^*, \boldsymbol{V}^*)$ be the solution of the convex RPT relaxation. Then we can construct the set $\mathcal{X}' = \left\{\boldsymbol{x}^*, \boldsymbol{x}_1^U, \ldots, \boldsymbol{x}_{n_x}^U, \boldsymbol{x}_1^V, \ldots, \boldsymbol{x}_{n_\tau}^V\right\}$ of candidate solutions for (1), where

$$\boldsymbol{x}_i^U = \begin{cases} \boldsymbol{x}^* & \text{if } x_i^* = 0, \\ \frac{\boldsymbol{U}_i^*}{x_i^*} & \text{otherwise,} \end{cases} \quad \text{and} \quad \boldsymbol{x}_j^V = \begin{cases} \boldsymbol{x}^* & \text{if } \tau_j^* = 0, \\ \frac{\boldsymbol{V}_j^*}{\tau_j^*} & \text{otherwise,} \end{cases} \quad \text{for all } i \in [n_x], j \in [n_\tau].$$

Observe that if the nonconvexity is only in the objective, i.e., there are no SLC constraints in (1), we have $\boldsymbol{x}^* \in \mathcal{X}$. If $\boldsymbol{x}$ is assumed to be nonnegative, i.e., the constraint $\boldsymbol{x} \geq \boldsymbol{0}$ is included in the original set of constraints, then $\mathcal{X}' \subseteq \mathcal{X}$, as is shown by the following lemma:

Lemma 2. *Let $\mathcal{X} = \{\boldsymbol{x} \in \mathbb{R}_+^{n_x} \mid \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \boldsymbol{P}\boldsymbol{x} = \boldsymbol{s}, \boldsymbol{h}(\boldsymbol{x}) \leq \boldsymbol{0}\}$, where $\boldsymbol{A} \in \mathbb{R}^{L_1 \times n_x}$, $\boldsymbol{P} \in \mathbb{R}^{L_2 \times n_x}$, $\boldsymbol{b} \in \mathbb{R}^{L_1}$, $\boldsymbol{s} \in \mathbb{R}^{L_2}$, $\boldsymbol{h}(\boldsymbol{x}) = [h_0(\boldsymbol{x}) \ h_1(\boldsymbol{x}) \ \cdots \ h_J(\boldsymbol{x})]^\top \subseteq (-\infty, +\infty]^{J+1}$, and $h_j : \mathbb{R}^{n_x} \to (-\infty, +\infty]$ is proper, closed and convex for every $j \in \mathcal{J}_0$. Then the set of candidate solutions $\mathcal{X}'$ is contained in the original feasible set $\mathcal{X}$.*

Proof. Let $g_\ell(\boldsymbol{x}) = \boldsymbol{b}_\ell - \boldsymbol{A}_\ell^\top \boldsymbol{x}$ for $\ell \in \mathcal{L}_1$. If $\boldsymbol{x} > 0$, we have $\boldsymbol{U} > \boldsymbol{0}$ and thus $\frac{\boldsymbol{U}_i^*}{x_i^*} > \boldsymbol{0}$. Moreover, we have

$$x_i^* \boldsymbol{g}_\ell\left(\frac{\boldsymbol{U}_i^*}{x_i^*}\right) \leq 0 \implies \boldsymbol{g}_\ell\left(\frac{\boldsymbol{U}_i^*}{x_i^*}\right) \leq 0, \quad \ell \in \mathcal{L}_1 \quad \text{and} \quad x_i^* \boldsymbol{h}\left(\frac{\boldsymbol{U}_i^*}{x_i^*}\right) \leq 0 \implies \boldsymbol{h}\left(\frac{\boldsymbol{U}_i^*}{x_i^*}\right) \leq 0.$$

Finally, we also have

$$x_i^* \boldsymbol{P} \frac{\boldsymbol{U}_i^*}{x_i^*} = \boldsymbol{s} x_i^* \implies \boldsymbol{P} \frac{\boldsymbol{U}_i^*}{x_i^*} = \boldsymbol{s}.$$

Observe that if $x_i^* = 0$ we have $\boldsymbol{x}_i^U = x^*$, hence $x^* \in \mathcal{X}$. Therefore $\boldsymbol{x}_i^U \in \mathcal{X}$ for every $i \in [n_x]$. In a similar way we can prove $\boldsymbol{x}_j^V \in \mathcal{X}$ for every $j \in [n_\tau]$. This concludes the proof. □

Hence, by substituting the candidate solutions in the original problem (1) we obtain upper bounds corresponding to each candidate solution and we can 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 $\mathcal{X}' \nsubseteq \mathcal{X}$, since for the latter the solution from the RPT relaxation might not be feasible for (1). Therefore, to compute the best upper bound, we select the set of candidate solutions in $\mathcal{X}'$ that satisfy all constraints in the set $\mathcal{X}$, such that we obtain the finite set

$$\mathcal{X}'' = \left\{ \boldsymbol{x} \in \mathcal{X}' \;\middle|\; x \in \mathcal{X} \right\}.$$

Observe that if we have nonconvexity in the constraints, it is possible that $\mathcal{X}'' = \emptyset$, in which case we cannot find a feasible solution.

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 written 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 C.

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 an upper and lower bound to problem (1). If the lower bound does not equal the upper bound, we search for a hyperplane that separates the candidate solutions for N_0 as described in Section 3.4, that is, we search for a hyperplane $H = \left\{\boldsymbol{x} \mid \boldsymbol{f}^\top \boldsymbol{x} = l\right\}$ such that $\boldsymbol{f}^\top \boldsymbol{x}_i \leq l$ for $i \in \mathcal{I}_1$, and $\boldsymbol{f}^\top \boldsymbol{x}_i \geq l$ for $i \in \mathcal{I}_2$, where $\mathcal{I}_1$, $\mathcal{I}_2 \subseteq [n_x]$ and $\mathcal{I}_1 \cup \mathcal{I}_2 = \mathcal{X}'$. Ideally, we want this hyperplane to cut the feasible region into two equally spaced sub regions. Hence, we require the hyperplane to pass through the analytic center, i.e., $\boldsymbol{f}^\top \boldsymbol{x}_{\text{ac}} = l$. The analytic center is defined as a point that maximizes the product of the slacks of the constraints. Assuming a generic optimization problem with convex constraints $g_i(\boldsymbol{x}) \leq 0$, the analytic center can be computed as the solution of the following problem:

$$\begin{aligned} \boldsymbol{x}_{ac} \quad = \quad & \underset{\boldsymbol{x}}{\text{argmax}} \quad \sum_i \ln(-g_i(\boldsymbol{x})) \\ & \text{s.t.} \qquad g_i(\boldsymbol{x}) \leq 0. \end{aligned} \tag{10}$$

Although the analytic center is not the geometric center of the feasible region, we choose the former as it is easier to calculate.

Ideally, the sets $\mathcal{I}_1$ and $\mathcal{I}_2$ contain about the same number of candidate solutions, since if we would split the region such that all candidate solutions are in one of the two subregions it might take longer to improve the bounds. This is demonstrated in for example Zhen et al. (2022). Here, a disjoint bilinear problem over a polyhedral feasible set is reformulated as a two-stage robust optimization problem and a convex relaxation is obtained by imposing linear decision rules. It is shown in this paper that one can equivalently obtain a convex relaxation using RLT. The critical points in this case are the worst-case scenarios. If the region was split such that all worst-case scenarios were in one of the two subregions, it was possible to end up with the same worst-case scenarios and thus no improvement. Therefore, it makes sense to try to split the critical points, which in our case are the candidate solutions, as much as possible over the two regions. Hence, we search for a hyperplane that is of maximum margin, that is, it separates the points as much as possible. Since the distance of a point $\boldsymbol{x}$ from the hyperplane equals $|\boldsymbol{f}^\top \boldsymbol{x} - l|/\|\boldsymbol{f}\|$, we can find a maximum margin hyperplane that passes through the analytic center by solving the following problem

$$\begin{array}{lll} \min\limits_{\boldsymbol{f},l,\boldsymbol{z}} & \|\boldsymbol{f}\| & \\ \text{s.t.} & \boldsymbol{f}^\top \boldsymbol{x}_{\text{ac}} = l, & \\ & \boldsymbol{f}^\top \boldsymbol{x}_i \leq l + M(1 - z_i), & i \in \mathcal{I}, \\ & \boldsymbol{f}^\top \boldsymbol{x}_i \geq l - M z_i, & i \in \mathcal{I}, \\ & \sum\limits_{i \in \mathcal{I}} z_i = |\mathcal{I}_1|, & \\ & \sum\limits_{i \in [n_x]} f_i \geq 1, & \\ & z_i \in \{0, 1\}, & \end{array} \tag{11}$$

where $\|\cdot\|$ denotes any norm. We further add the constraint $\sum_{i \in [n_x]} f_i \geq 1$ to avoid the trivial hyperplane $(\boldsymbol{f}, l) = (\boldsymbol{0}, 0)$. Hence, the hyperplane H is given by $H_0 = \left\{\boldsymbol{x} \mid \boldsymbol{f}_0^\top \boldsymbol{x} = l_0\right\}$, where $(\boldsymbol{f}_0, l_0)$ represents the optimal solution of problem (11). If problem (11) is infeasible we can reduce the set $\mathcal{I}_1$ and increase the set $\mathcal{I}_2$ or the other way around. We remark that in case the set of candidate solutions is not contained in the original feasible set, i.e., $\mathcal{X}' \nsubseteq \mathcal{X}$, some of the candidate solutions might not be inside the feasible region. Nevertheless, since the hyperplane passes through the analytic center, the hyperplane still cuts the feasible region further.

Next, 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.,

$$\mathcal{X}_l^0 = \left\{\boldsymbol{x} \in \mathcal{X} \mid \boldsymbol{f}_0^\top \boldsymbol{x} \leq l_0\right\},$$
$$\mathcal{X}_r^0 = \left\{\boldsymbol{x} \in \mathcal{X} \mid \boldsymbol{f}_0^\top \boldsymbol{x} \geq l_0\right\}.$$

Subsequently, we apply RPT to each childnode and obtain a lower and 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 can search for a maximum margin hyperplane H_k that passes through the analytic center and separates the candidate solutions of N_k to create again two new child nodes, and so on.

A key element of the branch and bound algorithm is pruning parts of the tree in order to speed up the method. The main condition that we use for pruning is when a lower bound is greater than the current best upper bound. In problems where we use the mountain climbing algorithm for obtaining upper bounds, we also prune nodes that result in worst upper bounds than the root node as they lead towards worst local optima. Another important aspect is which node to select from the unexplored ones. We propose picking the one with the smallest lower bound. In Algorithm 1, we summarize the described spatial branch and bound procedure via RPT to obtain a global solution to problem (1).

REMARK 4. For minimization problems containing concave or difference of convex functions, i.e., $f(\boldsymbol{x}) = c_0(\boldsymbol{x}) - c_1(\boldsymbol{x})$, where c_0 and c_1 proper, closed, and convex, and $c_0(\boldsymbol{x}) = 0$ in case of a concave function, either in the objective or in the constraints, we use the biconjugate reformulation in order to obtain a minimization problem in generic form (1). We then subsequently obtain a bilinear objective or constraint, i.e., we obtain a term $\boldsymbol{x}^\top \boldsymbol{y}$ in the objective or constraint, where $\boldsymbol{x} \in \mathcal{X}$ and $\boldsymbol{y} \in \text{dom}(c_1^*)$. We remark that in this case we only generate hyperplanes by separating the candidate solutions in the set $\mathcal{X}'$, i.e., we only generate hyperplanes in the $\boldsymbol{x}$-space. If we also generate hyperplanes in the $\boldsymbol{y}$-space, we obtain much more constraints in each branch and bound iteration, increasing the computation time in each successive childnode. On the other hand, generating hyperplanes in the $\boldsymbol{y}$-space might reduce the number of hyperplanes that need to be generated, which could also reduce the computation time. We leave the question if you could benefit from also generating hyperplanes in the $\boldsymbol{y}$-space to future research. □

4. A Simple Example

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

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

Let $\mathcal{X}_\text{T}$ 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:

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

Algorithm 1 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} and N_{j_2} by solving problem (11)
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**

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 u_{11}, u_{22}, u_{33}, u_{12}, u_{13}, and $u_{23} \in \mathbb{R}$ respectively to obtain the following convex relaxation:

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

The solution of (13) appears to be

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

with objective value 3, which constitutes a lower bound on the optimal value of (12). Since $\mathcal{X}$ 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 $\mathcal{X}_{\text{TLL}}$ denote the feasible set of (13). We pairwise multiply the linear with the nonlinear constraints and obtain the SLC constraints

$$\begin{aligned}
&(x_1+x_2+1)\exp(x_2-x_3) \le (x_1+x_2+1)x_1, \\
&(x_1+x_2+1)2\exp\left(\frac{-x_1}{2}\right)+(x_1+x_2+1)2\exp\left(\frac{-x_2}{2}\right) \le (x_1+x_2+1)(2+\exp(-1)), \\
&(10-x_i)\exp(x_2-x_3) \le (10-x_i)x_1, && i\in\{1,2,3\}, \\
&(10-x_i)2\exp\left(\frac{-x_1}{2}\right)+(10-x_i)2\exp\left(\frac{-x_2}{2}\right) \le (10-x_i)(2+\exp(-1)), && i\in\{1,2,3\}.
\end{aligned}$$

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

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

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 u_{11}, u_{22}, u_{33}, u_{12}, u_{13}, and u_{23} respectively. The convex relaxation that results from the RPT approach is therefore:

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

The solution of (14) is

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

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:

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

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 $\text{Lb}^0 = 19.778$. The analytic center of the feasible region is $\boldsymbol{x}_{ac} = (6.59, 2.47, 2.70)$. The maximum margin hyperplane that passes through the analytic center is given by $H = \{\boldsymbol{x} \in \mathbb{R}^3 \mid -0.28x_1 + 0.70x_2 + 0.57x_3 = 1.46\}$. 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 $\mathcal{X}_l^0 = \{\boldsymbol{x} \in \mathcal{X}_\text{T} \mid -0.28x_1 + 0.70x_2 + 0.57x_3 \leq 1.46\}$ and N_2 represents problem (12) in which the feasible region $\mathcal{X}_\text{T}$ is intersected with $\mathcal{X}_r^0 = \{\boldsymbol{x} \in \mathcal{X}_\text{T} \mid -0.28x_1 + 0.70x_2 + 0.57x_3 \geq 1.46\}$. We apply steps 1-4 on N_1 and N_2 and obtain:

$$\text{Lb}_1 = 23.202, \; \text{Ub}_1 = 23.202 \;, \; \text{Lb}_2 = 19.785, \; \text{Ub}_2 = 19.798.$$

We set $\text{Ub} = \min\{\text{Ub}_0, \text{Ub}_1, \text{Ub}_2\} = 19.798$. Moreover, node N_2 becomes active, node N_1 remains inactive, since $\text{Lb}_1 > \text{Ub}$, and we delete node N_0 from the list of active nodes, i.e., $\text{ACTIVE} = \{N_2\}$. We set $\text{Lb} = \min\{\text{Lb}_1, \text{Lb}_2\} = 19.785$.

Since Ub - Lb $= 0.013 > \delta$, we select node N_2 from the list of active nodes. The analytic center of the new feasible region is $\boldsymbol{x}_{ac} = (5.71, 5.14, 5.84)$. The maximum margin hyperplane that passes through the analytic center is given by $H = \{\boldsymbol{x} \in \mathbb{R}^3 \mid 2.24x_1 + 3.65x_2 - 4.89x_3 = 2.97\}$. Hence we create two child nodes N_3 and N_4 from N_2 such that N_3 represents problem (12) in which the feasible region $\mathcal{X}$ is intersected with $\mathcal{X}_l^1 = \{\boldsymbol{x} \in \mathcal{X}_l^0 \mid 2.24x_1 + 3.65x_2 - 4.89x_3 \leq 2.97\}$ and N_4 represents problem (12) in which the feasible set $\mathcal{X}$ is intersected with $\mathcal{X}_r^1 = \{\boldsymbol{x} \in \mathcal{X}_l^0 \mid 2.24x_1 + 3.65x_2 - 4.89x_3 \geq 2.97\}$. We apply steps 1-4 on N_3 and N_4 and obtain:

$$\text{Lb}_3 = 19.946, \; \text{Ub}_3 = 19.946 \;, \; \text{Lb}_4 = 19.787, \; \text{Ub}_4 = 19.787.$$

We set $\mathrm{Ub} = 19.787$. Then, node N_3 becomes inactive, since $\mathrm{Lb}_3 > \mathrm{Ub}$. We set $\mathrm{Lb} = 19.787$ and therefore obtain the optimal solution

$$\boldsymbol{x}' = \begin{bmatrix} 1.18 \\ 0.92 \\ 0.75 \end{bmatrix} \quad \text{and} \quad \boldsymbol{U}' = \begin{bmatrix} u'_{11} & u'_{12} & u'_{13} \\ u'_{21} & u'_{22} & u'_{23} \\ u'_{31} & u'_{32} & u'_{33} \end{bmatrix} = \begin{bmatrix} 1.41 & 1.09 & 0.89 \\ 1.09 & 0.85 & 0.69 \\ 0.89 & 0.69 & 0.57 \end{bmatrix},$$

with optimal objective value 19.787.

5. Strenghtening the RPT-BB approach

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

5.1. Adding 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.

Theorem 1. *If 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* $q \in \mathbb{R}$*, is redundant to* $\{\boldsymbol{x} \mid \boldsymbol{B}\boldsymbol{x} \le \boldsymbol{p}\} \neq \emptyset$*, where* $\boldsymbol{B} \in \mathbb{R}^{L \times n_x}$ *and* $\boldsymbol{p} \in \mathbb{R}^L$*, then the constraints* $\boldsymbol{d}^\top \boldsymbol{x} \le q$*,* $2q\boldsymbol{d}^\top \boldsymbol{x} \le \boldsymbol{d}^\top \boldsymbol{U}\boldsymbol{d} + q^2$ *and* $(q - \boldsymbol{d}^\top \boldsymbol{x}) f\left(\frac{q\boldsymbol{x} - \boldsymbol{U}\boldsymbol{d}}{q - \boldsymbol{d}^\top \boldsymbol{x}}\right) \le \boldsymbol{0}$ *are redundant to*

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

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

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

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

coincide and both are nonnegative thanks to the strong duality of linear optimization and the redundancy of $\boldsymbol{d}^\top \boldsymbol{x} \le q$ to $\{\boldsymbol{x} \mid \boldsymbol{B}\boldsymbol{x} \le \boldsymbol{p}\}$, 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 $\boldsymbol{b}_y = \boldsymbol{B}^\top \boldsymbol{y} = \boldsymbol{d}$ 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 $\boldsymbol{U} \in \mathbb{S}^{n_x}$ we have

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

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

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

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

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

Thus, the claim follows. □

Note that adding linear constraints that are redundant to existing nonlinear constraints might be useful, as is demonstrated in Example 4.

Example 4. 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. Removing redundant constraints when using the biconjugate reformulation

In this subsection, we consider the specific case in which we have a minimization problem with concave or difference of convex objective and convex constraints. As described in Example 1, we can equivalently write this problem as the following disjoint bilinear optimization problem using the biconjugate reformulation:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & c_0(\boldsymbol{x}) - \boldsymbol{x}^\top \boldsymbol{y} + c_1^*(\boldsymbol{x}) \\ \text{s.t.} & \boldsymbol{x} \in \mathcal{X} \\ & \boldsymbol{y} \in \operatorname{dom}(c_1^*), \end{array} \tag{16}$$

where c_0 and c_1 are proper, closed, and convex, and $c_0 = 0$ in case of a concave objective. Observe, that by definition of a convex function, $\operatorname{dom}(c_1^*)$ is convex. Let $\operatorname{dom}(c_1^*)$ be given by

$$\operatorname{dom}(c_1^*) = \{\boldsymbol{y} : \boldsymbol{A}_y \boldsymbol{y} \le \boldsymbol{b}_y, \boldsymbol{P}_y \boldsymbol{y} = \boldsymbol{s}_y,\ g_j(\boldsymbol{y}) \le 0, j \in \mathcal{J}\},$$

where $\boldsymbol{A}_y \in \mathbb{R}^{L_y \times n_y}$, $\boldsymbol{P}_y \in \mathbb{R}^{L_y \times n_y}$, $\boldsymbol{b}_y \in \mathbb{R}^{L_y}$, $\boldsymbol{s}_y \in \mathbb{R}^{L_y}$, and $g_j : \mathbb{R}^{n_x} \to (-\infty, +\infty]$ is proper, closed and convex for every $j \in \mathcal{J}$.

The following Lemma shows 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 $\operatorname{dom}(c_1^*)$, are redundant as long as the LMI in Step 3 is not incuded.

Lemma 3. *The additional constraints in the RPT relaxation of* (16) *resulting from pairwise multiplication in $\mathcal{X} \times \mathcal{X}$ and $dom(c_1^*) \times dom(c_1^*)$ are redundant as long as the additional constraints resulting from pairwise multiplication in $\mathcal{X} \times dom(c_1^*)$ are included and the following LMI*

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

where $\boldsymbol{x}\boldsymbol{y}^\top$ and $\boldsymbol{y}\boldsymbol{y}^\top$ are linearized by $\boldsymbol{W} \in \mathbb{R}^{n_x \times n_y}$ 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 $\mathcal{X} \times \operatorname{dom}(c_1^*)$. On other other hand, if a solution is feasible for the problem involving only the multiplications in $\mathcal{X} \times \operatorname{dom}(c_1^*)$, since we are not using the LMI, we can take $\boldsymbol{U} = \boldsymbol{x}\boldsymbol{x}^\top$, $\boldsymbol{Y} = \boldsymbol{y}\boldsymbol{y}^\top$ 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 $\operatorname{dom}(c_1^*) \times \operatorname{dom}(c_1^*)$ are redundant.

5.3. Directly applying RPT versus biconjugate reformulation

It can happen that a function has both an SLC representation as well as a biconjugate reformulation resulting in a different SLC representation. In this case, the question arises whether it is better to apply RPT directly on the SLC function or apply RPT to the biconjugate reformulation. The following example shows that directly applying RPT in this case might be better than applying RPT to the biconjugate reformulation.

EXAMPLE 5. Consider the following nonconvex optimization problem

$$\begin{aligned} \max_{x} \quad & \sqrt{x} \\ \text{s.t.} \quad & 1 \leq x \leq 2. \end{aligned} \tag{17}$$

Observe that $\sqrt{x}$ can be written as $x\frac{1}{\sqrt{x}}$. Since $\frac{1}{\sqrt{x}}$ is convex and its domain is given by $\mathbb{R}_{++}$, $\sqrt{x}$ has a SLC representation, see case 1 in Table 2. However, it can also be equivalently written as $\inf_{y\geq 0}\{-xy + \frac{1}{4y}\}$ using the biconjugate reformulation. In the following we compare two convex relaxations of (17) obtained from (1) applying RPT to the SLC representation, and (2) applying RPT to the biconjugate reformulation, respectively,

$$\begin{aligned} \max_{x,u} \quad & x\sqrt{\tfrac{x}{u}} \\ \text{s.t.} \quad & 1 \leq x_1 \leq 2, \\ & u - 2x + 1 \geq 0, \\ & u - 4x + 4 \geq 0, \\ & 3x - u - 2 \geq 0, \end{aligned} \qquad \text{and} \qquad \begin{aligned} \max_{x,u,v,t} \quad & -v + t \\ \text{s.t.} \quad & y \geq 0 \\ & 0 \leq x_1 \leq 1, \\ & v - y \geq 0, \\ & 2y - v \geq 0, \\ & \left\|(y - t, 1)^\top\right\|_2 \geq y + t, \\ & \left\|(v - \theta - y + t, x - 1)^\top\right\|_2 \leq v + \theta - y + t, \\ & \left\|(2y - 2t - v + \theta, 2 - x)^\top\right\|_2 \leq 2y + 2t - v - \theta, \end{aligned}$$

where in (2) we use an epigraph variable t such that $\frac{1}{4y} \leq t$ and subsequently write it as a second order cone constraint. Moreover, observe that because of Lemma 3 we only pairwise multiply constraints containing an x variable with constraints containing a y variable. The optimal objective value of the convex relaxation in (1) is 1, which is also the optimal solution to (17), while in (2) v is unbounded below, hence the optimal objective value of the problem in (2) is $-\infty$.

We could not find an example for which the biconjugate reformulation yields better RPT results than using the SLC representation. However, we could not prove that this is always the case. When considering RPT-BB, in Section 7.2, where we consider a Euclidean norm maximization over convex constraints, we show that for some instances RPT-BB on the biconjugate reformulation has a lower computation time than RPT-BB in which we use the SLC representation.

5.4. Strengthening upper bounds with eigenvectors of the optimal solution

At optimality we will always have $\boldsymbol{U} = \boldsymbol{x}\boldsymbol{x}^\top$. If we multiply both sides with $\boldsymbol{x}$ we obtain $\boldsymbol{U}\boldsymbol{x} = (\boldsymbol{x}^\top \boldsymbol{x})\boldsymbol{x}$. Therefore, we notice that $\boldsymbol{x}$ is an eigenvector of $\boldsymbol{U}$ with corresponding eigenvalue $\boldsymbol{x}^\top \boldsymbol{x}$. Hence, we can add the eigenvectors of $\boldsymbol{U}$ 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 $\boldsymbol{U}$.

Example 6. Consider the following toy problem:

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & (x_1 + x_2 + 1)\exp(x_2) \\ \text{s.t.} & x_1 + x_2 + 1 \geq 0, \\ & x_1 x_2 \geq -1. \end{array} \tag{18}$$

The optimal solution of (18) is $(-1, 0)$ with optimal value 0. After applying RPT we obtain the following relaxation

$$\begin{array}{ll} \min\limits_{\boldsymbol{x}} & (x_1 + x_2 + 1)\exp\left(\frac{u_{12}+u_{22}+x_2}{x_1+x_2+1}\right) \\ \text{s.t.} & x_1 + x_2 + 1 \geq 0, \\ & u_{12} \geq -1, \\ & u_{11} + u_{22} + 2u_{12} + 2x_1 + 2x_2 + 1 \geq 0. \end{array} \tag{19}$$

The optimal solution of (19) is

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

with optimal value 0 which gives us a lower bound. We have the following candidate vectors

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

The eigenvectors of $\boldsymbol{U}^*$ 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, $\boldsymbol{x}_1^U$ is infeasible as $(x_1^U)_1 (x_1^U)_2 = -5.44 < -1$ and $\boldsymbol{x}_3^U$ is infeasible as $(x_3^U)_1 + (x_1^U)_2 + 1 = -0.26 < 0$. Finally, we notice that $\boldsymbol{x}_2^U$ 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 $\boldsymbol{U}^*$. □

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 through the analytic center decreases the volume of the feasible region, but 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 1.

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 $x_i^{max} = \arg\max x_i$ and $x_i^{min} = \arg\min x_i$, for all $i \in [n_x]$, subject to the feasible region of this leaf. We then separate the feasible region of this leaf by adding the hyperplane $\boldsymbol{x} = \frac{1}{2}(\boldsymbol{x}^{max} + \boldsymbol{x}^{min})$, 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 2. *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 1 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:

$$\{\boldsymbol{x} : \|\boldsymbol{x} - \boldsymbol{\alpha}\|_\infty \leq \epsilon\}. \tag{20}$$

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 $j \to \infty$.

Step 2) In this step we show that

$$|U_{ij} - x_i x_j| \leq 4\epsilon(|\alpha_j| + \epsilon). \tag{21}$$

To prove this, first observe that for $i, j \in [n_x]$, we have

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

$$x_i - \alpha_i \geq -\epsilon \tag{23}$$

$$x_j - \alpha_j \geq -\epsilon \tag{24}$$

$$x_j - \alpha_j \leq \epsilon. \tag{25}$$

Clearly, Constraints (22) – (25) are redundant with respect to the corresponding feasible region of the leaf of the B&B tree. It follows from Theorem 1 that the inequalities that are obtained after multiplying these redundant constraints and perspectification are also redundant. We multiply the inequalities (23) and (24) and apply perspectification to obtain:

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

From this inequality we obtain

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

By multiplying (23) and (25) we can prove in a similar way:

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

Hence, we obtain (21).

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

$$\begin{aligned} \tilde{f}_k(\boldsymbol{x}, \boldsymbol{U}_1) - \tilde{f}_k(\boldsymbol{x}, \boldsymbol{U}_2) &= \sum_i (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) \left(c_{ik}\left(\frac{q_{ik}\boldsymbol{x} - \boldsymbol{U}_1 \boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}} \right) - c_{ik}\left(\frac{q_{ik}\boldsymbol{x} - \boldsymbol{U}_2 \boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}} \right) \right) \\ &\leq \sum_i (q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}) L_{ik} \| \frac{q_{ik}\boldsymbol{x} - \boldsymbol{U}_1 \boldsymbol{d}_{ik} - q_{ik}\boldsymbol{x} + \boldsymbol{U}_2 \boldsymbol{d}_{ik}}{q_{ik} - \boldsymbol{d}_{ik}^\top \boldsymbol{x}} \| \\ &= \sum_i L_{ik} \| (\boldsymbol{U}_1 - \boldsymbol{U}_2) \boldsymbol{d}_{ik} \| \\ &\leq \sum_i L_{ik} \| \boldsymbol{U}_1 - \boldsymbol{U}_2 \| \| \boldsymbol{d}_{ik} \| \\ &\leq \tilde{L_k} \| \boldsymbol{U}_1 - \boldsymbol{U}_2 \|, \end{aligned}$$

where $\tilde{L_k} = \sum_i L_{ik} \| \boldsymbol{d}_{ik} \|$. In particular if we take $U_1 = U$ and $U_2 = xx^\top$, and using (21) we get

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

where $\tilde{\alpha} = \max_j |\alpha_j|$. This means that $\tilde{f}_k(\boldsymbol{x}, \boldsymbol{U}) \to \tilde{f}_k(\boldsymbol{x}, xx^\top)$ 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:

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

The perspective approximation is:

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

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

We emphasize that if the constraints in (1) contains difference of convex functions, the feasible region of the biconjugate reformulation, i.e., Problem (B), may not be bounded (see Example 1) and Theorem 2 does not apply. The following proposition shows that the adapted Algorithm 1 converges if the sets $\text{dom}(c_{ik}^*)$ obtained from the biconjugate reformulation are bounded.

PROPOSITION 1. *If for each $i \in \mathcal{I}$ and $k \in \mathcal{K}_0$, the function c_{ik} and its corresponding recession function $\delta^*_{\text{dom}(c_{ik}^*)}$ are Lipschitz continuous, and the set $\text{dom}(c_{ik}^*)$ obtained from the biconjugate reformulation is bounded, then the spatial B&B Algorithm 1 with Adaptation A converges to a global optimal solution of Problem (B).*

Proof. The proof consists of three steps. Since Steps 1) and 3) are identical to those of Theorem 2, we omit them here. For the remainder of this proof, it suffices to show that

$$\text{Tr}\,(V) \to \boldsymbol{x}^\top \boldsymbol{y},$$

where the variable $\boldsymbol{y}$ is due to the biconjugate reformulation. Because the set $\text{dom}(c_{ik}^*)$ is bounded, for every $i \in [n_x]$, there exists a $K \in \mathbb{R}$ such that

$$y_i \leq K \tag{26}$$

$$y_i \geq -K. \tag{27}$$

It follows from Theorem 1 that the inequalities that are obtained after multiplying (22) and (23) with (26) and (27) and perspectification are redundant. Multiplying constraints (22) and (27) and perspectification yields:

$$0 \leq (\epsilon - x_i + \alpha_i)(y_i + K) = K(\epsilon - x_i + \alpha_i) + y_i\epsilon - V_{ii} + \alpha_i y_i.$$

Starting from this inequality we obtain

$$V_{ii} - \alpha_i y_i \leq y_i\epsilon + K(\epsilon - x_i + \alpha_i) \leq K\epsilon + K(\epsilon + \epsilon) = 3K\epsilon,$$

in which the last inequality follows from (23) and (26). Multiplying constraints (22) and (26) and perspectification yields:

$$0 \leq (\epsilon - x_i + \alpha_i)(K - y_i) = K(\epsilon - x_i + \alpha_i) - y_i\epsilon + V_{ii} - \alpha_i y_i.$$

Starting from this inequality we obtain

$$V_{ii} - \alpha_i y_i \geq y_i \epsilon + K(-\epsilon + x_i - \alpha_i) \geq -K\epsilon + K(-\epsilon - \epsilon) = -3K\epsilon,$$

in which the last inequality follows from (23) and (27). Combining these results and using (22), (23), and (26) finally yields:

$$\begin{aligned} |V_{ii} - x_i y_i| &= |V_{ii} - \alpha_i y_i + \alpha_i y_i - x_i y_i| \\ &\leq |V_{ii} - \alpha_i y_i| + |(\alpha_i - x_i) y_i| \\ &\leq 3K\epsilon + \epsilon K \\ &= 4K\epsilon. \end{aligned}$$

It follows from triangle inequality that

$$|\mathrm{Tr}\,(V) - \boldsymbol{x}^\top \boldsymbol{y}| \leq 4nK\epsilon.$$

Finally, because $\epsilon \to 0$ as $j \to \infty$, we have that $\mathrm{Tr}\,(V) \to \boldsymbol{x}^\top \boldsymbol{y}$. □

7. 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 written 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, for BARON are conducted with BARON version 20.10.16 Sahinidis (1996) implemented using the Python package `pyomo` version 6.4.1, and for CPLEX are conducted with CPLEX version 22.1.0 ILOG,Inc. (2017) 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. Finally, we note that in all branch and bound implementations the optimality gap was set to 10^{-4}.

7.1. 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. (2020):

$$\max_{\boldsymbol{x}\in\mathcal{X}} \quad \sum_{k\in\mathcal{K}} \max_{j\in\mathcal{J}_k}\{\boldsymbol{A}_j^\top \boldsymbol{x} + b_j\}, \tag{28}$$

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

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

Here $\boldsymbol{D}\in\mathbb{R}^{n_x\times L}$ and $\boldsymbol{d}\in\mathbb{R}^L$. Since the objective of (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.,

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

Observe that $n_y = |\mathcal{J}|$. 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

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

We solve problem (30) with Gurobi for $\mathcal{X}_1$ and Mosek for $\mathcal{X}_2$. We refer to Appendices D.1 and E.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. All approaches find

$\mathcal{X}$	#	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	0.69	0.4	211.87	1.47	0	211.87	1.08	211.87	0.06
	3	1081.62	94.02	18	1081.62	42.32	0	1081.62	3600*	1081.62	1.63
	3a	1159.90	75.11	10.4	1159.90	286.51	0.6	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	6.31	9.6	83.78	0.19	0	83.78	4.09	83.78	0.02
	5	3002.44	65.39	6	3002.44	566.65	1	3002.44	3600*	3002.44	647.18
	5a	2898.05	63.27	7.2	2898.05	1278.26	3.3	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	4.46	1	136.22	3.43	0	136.22	3600*	136.22	778.65
	2a	122.21	12.02	1.8	122.21	25.41	1.3	122.21	3600*	122.21	977.05
	3	837.94	25.79	1	837.94	201.34	0	837.94	3600*	837.94	3600*
	3a	890.07	62.72	3.2	890.07	618.84	0.9	890.07	3600*	886.41	3600*
	4	33.73	4.12	7	33.73	6.77	5	33.73	75.44	33.73	0.09
	4a	31.81	1.08	1.5	31.81	2.71	0.8	31.81	39.33	31.81	0.11
	5	1610.69	14.84	2	1610.69	492.74	1	1610.69	3600*	1610.69	3600*
	5a	1670.92	31.79	2.8	1670.92	700.99	1.6	1670.92	3600*	1670.92	3600*
$\mathcal{X}_3$	1	13.44	2.97	1	13.44	5.12	1	13.44	0.31		
	1a	15.02	8.62	2.5	15.02	14.32	6.3	15.02	0.13		
	2	140.89	21.96	1	140.89	59.31	1	140.89	8.48		
	2a	129.31	145.92	6.5	129.31	83.31	1.9	129.31	417.41		
	3	768.96	909.69	4	768.96	1788.87	1	768.96	3600*		
	3a	805.95	3600*	5.2	805.95	3585.16	2.4	805.95	3600*		
	4	45.34	1.08	1	45.34	7.48	1	45.34	114.87		
	4a	43.79	18.72**	1.9	44.97	35.41	2.2	44.97	97.11		
	5	2273.67	3600*	35	1711.03***	3600*	4	1700.68	3600*		
	5a	2307.39	1953.18**	43.6	1735.74***	3600*	4.3	2000.09	3600*		

Table 3 Comparison for the sum-of-max-linear-terms maximization problem of RPT-BB, RPT-SDP-BB, BARON, and the exact mixed integer reformulation (Gurobi for $\mathcal{X} = \mathcal{X}_1$ and MOSEK for $\mathcal{X} = \mathcal{X}_2$) over the feasible regions $\mathcal{X}_1, \mathcal{X}_2$, and $\mathcal{X}_3$, on problem instances 1, 2, 3, 4, and 5. The results for problem instances 1a, 2a, 3a, 4a, and 5a reflect the average of 10 randomly generated instances corresponding to problems 1, 2, 3, 4, and 5 respectively. Opt represents the optimal value, Gen Hyp represents the total number of hyperplanes generated during branch and bound and Time represents the computation time. 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 an 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.

the optimum for all instances, except for BARON, which cannot find the optimum for instances 3, 3a, 5, and 5a within the time limit.

For $\mathcal{X} = \mathcal{X}_2$, MIR still finds the optimum within the lowest computation time for instances 1, 1a, 4, and 4a. For instances 2 and 2a, the computation time of MIR is much larger than the computation time of RPT-BB and RPT-SDP-BB. RPT-BB and RPT-SDP-BB find the optimum for all instances within the computation limit, while BARON is unable to find the optimum for instances 2, 2a, 3, 3a, 5, and 5a, and MIR is unable to find the optimum for instances 3, 3a, 5, and 5a within the time limit of one hour.

For $\mathcal{X} = \mathcal{X}_3$, MOSEK cannot solve the instances because of the nonconvex constraint. For instances 1, 1a, 2, and 2a, all approaches find the optimum. For instances 1, 1a, and 2, BARON has the lowest computation time, while for instance 2a RPT-SDP-BB has the lowest computation time. Moreover, RPT-SDP-BB also finds the optimum for instances 3, 3a, 4, and 4a, whereas RPT-BB could only find the optimum for instances 3 and 4a, and BARON could only find the optimum for 4, and 4a within a computation time of one hour. For these instances, RPT-SDP-BB has the lowest computation time, except for instance 4, for which RPT-BB finds the optimum faster. Note that for instance 3a and 4a, we hence need the SDP relaxation to solve the instances to optimality. For instance 5, all approaches could not find the optimum within the time limit of one hour, and we observe that RPT-BB finds a better lower bound than BARON. For instance 5a only RPT-BB could solve some of the instances within an hour.

7.2. Euclidean norm maximization over convex constraints

We consider the following generic euclidean norm maximization problem subject to convex constraints:

$$\max_{\boldsymbol{x}\in\mathcal{X}} \quad \|\boldsymbol{x}\|_2. \tag{31}$$

We consider two cases of $\mathcal{X}$, that is, $\mathcal{X} = \mathcal{X}_1$ and $\mathcal{X} = \mathcal{X}_2$, where $\mathcal{X}_1$ and $\mathcal{X}_2$ are defined as in Section 7.1. Observe that the objective of (31) has two different SLC representations. The first SLC representation results from rewriting the objective of (31) as $\sum_{i=1}^{n_x} U_{ii}$, where $U_{ii} = x_i^2$ for $i \in [n_x]$. The second SLC representation results from using the biconjugate reformulation. 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:

$$\max_{\substack{\boldsymbol{x}\in\mathcal{X}\\ \boldsymbol{y}\in\mathcal{Y}}} \quad \boldsymbol{x}^\top \boldsymbol{y}, \tag{32}$$

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

$$\mathcal{Y} = \left\{ \boldsymbol{y} \in \mathbb{R}^{n_x}_+ \,\middle|\, \|\boldsymbol{y}\|_2 \le 1 \right\}.$$

We compare the proposed approach applied either on the SLC representation resulting from the biconjugate reformulation (RPT-BB, RPT-SDP-BB) or on the direct SLC representation (RPT-BB-Dir, RPT-SDP-BB-Dir), with BARON and SCIP. We refer to Sections H.2 and I.2 in the Appendix for the convex RPT relaxation and problem instances respectively. The results are illustrated in Table 4.

$\mathcal{X}$	#	RPT-BB			RPT-SDP-BB			RPT-BB-Dir			RPT-SDP-BB-Dir			BARON		SCIP	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Opt	Time
$\mathcal{X}_1$	1	29.67	0.86	4.7	29.67	0.86	1.6	29.67	0.02	0	29.67	0.08	0	29.67	0.08	29.67	0.38
	2	22.83	2.30	4.8	22.83	3.40	4.1	22.83	1.51	3.5	22.83	0.34	0.3	22.83	0.07	22.83	0.18
	3	21.79	6.75	3.1	21.79	15.88	3.1	21.79	2.98	1.2	21.79	1.39	0.2	21.79	0.11	21.79	0.41
	4	49.66	89.73	7.0	49.66	238.37	5.1	49.66	39.87	2.1	49.66	41.63	1.5	49.66	0.59	49.66	0.73
	5	47.93	1027.97	45.8	47.93	817.27	5.2	47.93	175.95	4.1	47.93	248.49	3.8	47.93	1.84	47.93	1.25
$\mathcal{X}_2$	1	11.96	2.21	4.8	11.96	3.97	5.7	11.96	0.84	2.1	11.96	0.77	1.9	11.96	0.25	11.96	361.51
	2	13.30	5.69	2.9	13.30	10.60	4.1	13.30	2.35	2.1	13.30	2.12	1.7	13.30	2.74	13.30	34.12
	3	17.93	7.85	4.1	17.93	15.98	1.1	17.93	19.31	5.8	17.93	12.01	2.7	17.93	1.07	17.90	3.37
	4	37.20	115.95	10.3	37.20	692.73	2.1	37.20	122.55	2.2	37.20	227.27	1.9	37.20	3600*	35.61	3600*
	5	39.12	118.73	2.2	39.12	1602.32	2.1	39.12	909.75	6.1	39.12	1541.06	5.8	39.12	3600*	11.07	3600*

Table 4 Comparison of RPT-BB, RPT-SDP-BB, BARON, and SCIP for the Euclidean norm maximization problem over the feasible regions $\mathcal{X} = \mathcal{X}_1$ and $\mathcal{X} = \mathcal{X}_2$, for problem instances 1, 2, 3, 4 and 5 which reflect the average of 10 randomly generated instances. Opt represents the optimal value, Gen Hyp represents the total number of hyperplanes generated during branch and bound and Time represents the computation time. 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.

From Table 4 we observe that for $\mathcal{X} = \mathcal{X}_1$, all approaches find the optimum and BARON outperforms the other approaches on computation time, except for instance 1, for which RPT-BB-Dir has the lowest computation time and instance 5 for which SCIP has the lowest computation time. On all instances RPT-BB-Dir and RPT-SDP-BB-Dir have a lower computation time than RPT-BB and RPT-SDP-BB.

For $\mathcal{X} = \mathcal{X}_2$ all approaches find the optimum for all instances, except for instances 4 and 5, for which BARON cannot prove optimality and SCIP cannot find the optimum within the time limit of one hour. For instances 1 and 3, BARON outperforms the other approaches on computation time, while for instance 2, RPT-SDP-BB-Dir has the lowest computation time. Moreover, we observe that using the biconjugate reformulation we find the optimum faster than using the direct SLC representation for the larger instances 3, 4 and 5, while for the smaller instances we have a lower computation time using the direct SLC representation.

Finally, we observe that when using the biconjugate reformulation, adding the SDP relaxation results in a larger computation time for all instances, except instance 5 for $\mathcal{X} = \mathcal{X}_1$. When using the direct SLC representation, the SDP relaxation reduces the computational time for both $\mathcal{X} = \mathcal{X}_1$ and $\mathcal{X} = \mathcal{X}_2$ in the smaller scale instances, however it increases it in the larger scale ones, those are instances 4 and 5.

7.3. Log-sum-exp maximization over linear constraints

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

$$\max_{\boldsymbol{x} \in \mathcal{X}_1} \quad \log \left(\sum_{i \in [n_x]} \exp(x_i) \right), \tag{33}$$

where $\mathcal{X}_1$ is defined as in Section 7.1. Since the objective of (33) is a closed convex function, we can replace it by its biconjugate function and obtain the following equivalent maximization problem:

$$\max_{\substack{\boldsymbol{x}\in\mathcal{X}_1 \\ (\boldsymbol{y},\boldsymbol{w})\in\mathcal{Y}}} \quad \boldsymbol{x}^\top \boldsymbol{y} + \sum_{i\in[n_x]} w_i, \tag{34}$$

where $\mathcal{Y}$ equals the domain of the conjugate function of the objective of (33), i.e.,

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

Observe that here we make use of case 7 in Table 1 and introduce epigraph variables w_i for every $i\in[n_x]$.

We compare RPT-BB, RPT-SDP-BB, BARON, and SCIP. We refer to Sections D.3 and E.3 in the Appendix for the convex RPT relaxation and problems instances respectively. The results are illustrated in Table 5.

$\mathcal{X}$	#	RPT-BB			RPT-SDP-BB			BARON		SCIP	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Opt	Time
	1	10.01	0.14	0	10.01	0.48	0	10.01	0.06	10.01	0.07
	2	40.00	1.51	0	40.00	158.44	0	40.00	360.65	3.69	3600*
	3	6.09	2.89	0.9	6.09	23.02	0.8	6.09	0.12	6.09	0.17
	4	21.96	63.65	23.8	21.96	13.51	0.9	21.78	1082.38**	19.31	2520.37**
	5	34.76	61.11	0.7	34.76	368.69	0	34.76	1450.84*	15.43	3600*

Table 5 Comparison of RPT-BB, RPT-SDP-BB, BARON, and SCIP for the log-sum-exp maximization problem over the feasible region $\mathcal{X}=\mathcal{X}_1$, for problem instances 1, 2, 3, 4, and 5, which reflect the average of 10 randomly generated instances. Opt represents the optimal value, Gen Hyp represents the total number of hyperplanes generated during branch and bound and Time represents the computation time. 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 an hour while others were not and returned the best value obtained.

From Table 5 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 the time limit. Observe that while BARON 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 performs best on computation time, while for instance 2, RPT-BB has the lowest computation time.

When we compare RPT-BB and RPT-SDP-BB, we observe that RPT-BB has the lowest computation time for all instances except for instance 4, since for the latter the number of generated hyperplanes is much higher for RPT-BB. Moreover for instances 1 and 2 both methods find the optimum at the root node, and for instance 5, RPT-SDP-BB also finds the optimum at the root node.

7.4. Linear multiplicative optimization

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

$$\min_{\boldsymbol{x}\in\mathcal{X}} \quad \prod_{i\in[n_y]} \boldsymbol{A}_i^\top \boldsymbol{x} + b_i \tag{35}$$

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 (35) can be equivalently reformulated as

$$\min_{\boldsymbol{x}\in\mathcal{X}} \quad \sum_{i\in[n_y]} \ln\left(\boldsymbol{A}_i^\top\boldsymbol{x}+b_i\right). \tag{36}$$

Since the objective of (36) is a closed concave function, we can replace it by its biconjugate function and obtain the following equivalent maximization problem:

$$\max_{\substack{\boldsymbol{x}\in\mathcal{X}\\(\boldsymbol{y},\boldsymbol{w})\in\mathcal{Y}}} \quad -(\boldsymbol{A}^\top\boldsymbol{x}+\boldsymbol{b})^\top\boldsymbol{y}-\sum_{i\in[n_y]} w_i, \tag{37}$$

where $\mathcal{Y}$ equals the domain of the conjugate function of the objective of (36), i.e.,

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

Observe that here we make use of case 3 in Table 1 and introduce epigraph variables w_i for every $i\in[n_y]$. Moreover, observe that $\mathcal{Y}$ is not bounded. Hence, the conditions for convergence of RPT-BB are not satisfied. Nevertheless, we compare RPT-BB, RPT-SDP-BB, BARON, and SCIP. We refer to sections D.4 and E.4 in the Appendix for the convex RPT relaxation and problem instances respectively. The results are illustrated in Table 6.

From Table 6 we observe that RPT-BB has an edge over BARON on problems involving a large number of linear multiplications in the objective. Namely, on instances 3, 4 and 5, corresponding to 8, 20, and 30 linear multiplications in the objective respectively, we observe that RPT-BB is able to find the global optimum in seconds, while BARON is unable to prove optimality within one hour. Ryoo and Sahinidis (1996) mention that BARON can handle LMP problems up to 5 linear terms in the objective. Although there has been much development in the software from then, it is still unable to solve problems including more than 10 linear multiplications in the objective. On the other hand, we observe that RPT-BB can handle problems including up to at least 30 linear multiplications in the objective.

$\mathcal{X}$	#	RPT-BB			RPT-SDP-BB			BARON		SCIP	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Opt	Time
	1	11.39	7.74	23.5	11.39	3.86	7.1	11.39	0.24	11.39	1.06
	2	5.38	1798.90**	146.4	-	3600*	0	5.38	2.34	5.38	10.12
	3	22.84	909.15	137.6	22.84	136.39	57.2	22.84	3600*	22.84	3600*
	4	44.82	305.15	47.2	44.82	1426.03	46.9	44.82	3600*	-	3600*
	5	66.16	55.29	5.5	66.16	2323.97	8.1	66.16	3600*	-	3600*

Table 6 Comparison of RPT-BB, RPT-SDP-BB, BARON and SCIP for the linear multiplicative optimization problem instances 1, 2, 3, 4, and 5, which reflect the average of 10 randomly generated instances. Opt represents the optimal value, Gen Hyp represents the total number of hyperplanes generated during branch and bound and Time represents the computation time. 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 an hour while others were not and returned the best value obtained A - indicates that no solution was returned after one hour.

However, there is a trade-off between the number of variables and constraints that each method can handle. We observe that BARON and SCIP are able to solve instance 2, corresponding to 200 variables, 200 linear constraints, and 3 linear multiplications in the objective, in seconds, while RPT-BB is unable to find the global optimum for some instances within one hour and RPT-SDP-BB is unable to solve the problem at the root node. The total number of variables and constraints resulting after the multiplications grows significantly when considering a large number of numerical variables, hence making the problem intractable. We conclude that RPT-BB can be advantageous in LMP problems involving a large number of linear multiplications in the objective, while BARON and SCIP can be advantageous in LMP problems with more variables and constraints and less linear multiplications in the objective.

7.5. Quadratic constraint quadratic optimization

We consider the following quadratic constraint quadratic optimization problem:

$$\begin{aligned} \min_{\boldsymbol{x}\in\mathcal{X}_1} \quad & \boldsymbol{x}^\top \boldsymbol{P}_0 \boldsymbol{x} + \boldsymbol{q}_0^\top \boldsymbol{x} + r_0 \\ \text{s.t.} \quad & \boldsymbol{x}^\top \boldsymbol{P}_k \boldsymbol{x} + \boldsymbol{q}_k^\top \boldsymbol{x} + r_k \le 0, \qquad k\in\mathcal{K}, \end{aligned} \tag{38}$$

where $\boldsymbol{P}_k$, $k\in\mathcal{K}_0$, are not necessarily positive semi-definite. Hence, problem (38) is not necessarily convex. However, the nonconvex quadratic functions are SLC, hence we can apply RPT. We compare RPT-BB, RPT-SDP-BB, and BARON. Moreover, on the first five instances involving nonconvex QPs over linear constraints we also compare with CPLEX. We refer to Appendices D.5 and E.5 for the convex RPT relaxation and problem instances respectively. The results are illustrated in Table 7.

From Table 7 we observe that for instances 1-5, all approaches find the optimum for all instances, except for BARON, which is not able to prove optimality within the computation limit for instance

$\mathcal{X}$	#	RPT-BB			RPT-SDP-BB			BARON		CPLEX	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time	Opt	Time
	1	394.75	0.63	4	394.75	0.55	1	394.75	0.19	394.75	0.11
	2	884.75	1.19	5	884.75	0.39	1	884.75	0.21	884.75	0.08
	3	6888.78	0.03	0	6888.78	0.03	0	6888.78	0.11	6888.78	0.07
	4	98,382.63	2.86	0	98,382.63	3.77	0	98,382.63	3600*	98,382.63	1.32
	5	774,482.38	12.53	0	774,482.38	67.51	0	774,482.38	3600*	774,482.38	13.25
	6	3415.62	0.16	3	3415.62	0.41	2	3415.62	0.06		
	7	16,805.89	0.83	8	16,805.89	0.44	1	16,805.89	0.11		
	8	15,433.13	0.45	2	15,433.13	0.09	0	15,433.13	0.17		

Table 7 Comparison of RPT-BB, RPT-SDP-BB, BARON and CPLEX for the quadratic constraint quadratic optimization problem over the feasible region $\mathcal{X} = \mathcal{X}_5$, for problem instances 1, 2, 3, 4, 5, 6, 7 and 8. Opt represents the optimal value, Gen Hyp represents the total number of hyperplanes generated during branch and bound and Time represents the computation time. 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.

4 and 5. For instances 1, 2, and 4, CPLEX has the lowest computation time, whereas for instances 3 and 5, RPT-BB has the lowest computation time. For instances 6-8, containing also nonconvex QP constraints, all three approaches are able to find the optimum within one second, with BARON having the lowest computation time for problem instances 6 and 7, and RPT-SDP-BB for instance 8.

7.6. 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):

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

where $\boldsymbol{t}$ is the vector of all moments in time at which the dike height is increased, $t_0 = 0$, $\boldsymbol{x}$ is the vector of all increases in dike height, where x_k is the increment of the dike height at time t_k, h_k is the increase in dike height after t_k years, i.e., $h_k = \sum_{i=0}^{k} x_i$, $h_K = \sum_{k\in\mathcal{K}_0} x_k$ and $\beta_\delta, \delta, \theta, \lambda, b, C, T$ and S_0 are constants, which are explained in more detail in Appendix E.6. 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 2, 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 D.6 and E.6 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 8.

$\boldsymbol{t}$	#	RPT-BB			RPT-SDP-BB			BARON	
		Opt	Time	Gen Hyp	Opt	Time	Gen Hyp	Opt	Time
$\boldsymbol{t}_{\text{ir}}$	10	61.98	5.88	7	61.98	0.15	0	61.98	110.70
	15	608.74	3600*	1769	608.74	0.29	0	608.74	3600*
	16	1268.11	3600*	2054	1268.11	0.29	0	1268.11	3600*
$\boldsymbol{t}_{25}$	10	61.31	34.78	12	61.31	0.49	0	61.31	1680.36
	15	609.92	5.81	3	609.92	1.15	0	609.92	3600*
	16	1269.63	3600*	1151	1269.63	0.88	0	1269.63	3600*
$\boldsymbol{t}_{50}$	10	55.50	62.12	58	55.50	0.16	0	55.50	1.32
	15	545.23	2.94	6	545.23	0.27	0	545.23	1.81
	16	1100.07	3600*	1256	1100.07	0.22	0	1100.07	3600*

Table 8 Comparison of RPT-BB, RPT-SDP-BB, and BARON for the dike height optimization problem, for dike rings 10, 15, and 16 in the Netherlands. Opt represents the optimal value, Gen Hyp represents the total number of hyperplanes generated during branch and bound and Time represents the computation time. 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.

From Table 8 we observe that RPT-SDP-BB outperforms RPT-BB and BARON on both the number of global optimal solutions found and computation time, since it finds the global optimal solution for every instance in the root node and within less than a second. Both RPT-BB and BARON find the global optimal solution in each case for dike ring 10 and are not able to prove optimality in each case for dike ring 16. For dike ring 16, BARON is able to achieve a smaller optimality gap than RPT-BB as can be seen from Table 9.

Moreover, from Table 8 we observe that RPT-BB can find the global optimal solution for dike ring 15 for $\boldsymbol{t}_{25}$ and $\boldsymbol{t}_{50}$, while BARON is only able to find the global optimal solution for $\boldsymbol{t}_{50}$. For the instances that can be solved by both RPT-BB and BARON, BARON has much lower computation time in the case of $\boldsymbol{t}_{50}$, while RPT-BB finds the global optimum much faster in case of $\boldsymbol{t}_{\text{ir}}$ and $\boldsymbol{t}_{25}$.

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

Table 9 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.

8. 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. Since we are multiplying all constraints in the feasible region, the method becomes less tractable in problems involving a large number of variables and constraints. However, one can consider to not pairwise multiply every constraint in order to reduce computational effort.

So far we have only considered pairwise multiplication of convex constraints and reformulated any SLC constraints with RPT. However, in future work one could further consider multiplying the SLC constraints and convexify the resulting constraints by introducing additional variables.

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(2–3):471–484, 2009.

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

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

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

A. Ben-Tal 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. Convex relaxations with second order cone constraints for nonconvex quadratically constrained quadratic programming. *arXiv*, 2016. URL https://arxiv.org/pdf/1608.02096.pdf.

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

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

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.

K. Roos, M. Balvert, B.L. Gorissen, and D. den Hertog. A universal and structured way to derive dual optimization problem formulations. *INFORMS Journal on Optimization*, 2(4):229–255, 2020.

H.S. Ryoo and N.V. Sahinidis. A branch-and-reduce approach to global optimization. *Journal of Global Optimization*, 8:107–138, 1996. doi: https://doi.org/10.1007/BF00138689.

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

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. *Optimization Online*, 2020. URL http://www.optimization-online.org/DB_HTML/2020/07/7881.html.

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

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

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

H. Sherali and W. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, volume 31. Springer Science & Business Media, 1999.

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.

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

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.

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* $k \in \mathcal{K}$, $\boldsymbol{A}\boldsymbol{x} < \boldsymbol{b}$ *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 $(\boldsymbol{\tau}^{\mathrm{S}}, \boldsymbol{U}^{\mathrm{S}}, \boldsymbol{V}^{\mathrm{S}})$ such that $(\boldsymbol{x}^{\mathrm{S}}, \boldsymbol{\tau}^{\mathrm{S}}, \boldsymbol{U}^{\mathrm{S}}, \boldsymbol{V}^{\mathrm{S}})$ is a strict Slater point of the corresponding RPT relaxation (7) of (1) with $f_k(\boldsymbol{x}) \le 0$ is replaced by

$$\begin{cases} c_{0k}(\boldsymbol{x}) - \sup\limits_{y\in\mathrm{dom}(c^*_{1k})}\{\boldsymbol{x}^\top\boldsymbol{y} - c^*_{1k}(\boldsymbol{y})\} \le 0 \\ \boldsymbol{y} \in \mathrm{dom}(c^*_{1k}). \end{cases} \tag{39}$$

Finally, thanks to Remark 1 and the proof of Theorem 6(iii) of Zhen et al. (2021), the inf operator in the constraint of (7) 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 (7).

B. Known convex reformulations and relaxations obtained via RPT

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

B.1. Disjunctive optimization

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

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

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

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

This convex relaxation is exact according to Gorissen et al. (2014, Lemma 1), which applies because $\mathcal{X}_k$, $k \in \mathcal{K}$, are nonempty, compact and convex sets. We now use this observation to derive convex relaxation for disjunctive optimization problems with general convex sets. In Sherali and

Adams (1994b, Section 4), the authors derive similar result for disjunctive optimization problems with a linear objective function and polyhedral sets $\mathcal{X}_k$, $k \in \mathcal{K}$. Consider a generic disjunctive optimization problem

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

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.

B.2. Generalized linear optimization

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

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

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

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

The convex 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., 2021, Lemma 6). While for a special case where $\mathcal{X}_k$, $k \in \mathcal{K}$, are (nonempty) boxes, the corresponding linear relaxation of (GLP) is exact due to Dantzig (1963).

B.3. Approximate $\mathcal{S}$-Lemma for quadratically constrained quadratic optimization

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

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

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

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

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

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

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

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

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:

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

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).

B.4. Fractional optimization

Consider the following generic fractional optimization problem

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

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:

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

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

Appendix C: 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, for example problem (16). 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 2, 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 $\mathcal{X}''=$, 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.

Appendix D. 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

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

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:

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

Algorithm 2 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**
    $\text{Lb} \leftarrow f(\boldsymbol{x}, \boldsymbol{y})$
    $\boldsymbol{x} \leftarrow \arg\max_{\boldsymbol{x} \in \mathcal{X}} f(\boldsymbol{x}, \boldsymbol{y})$
    $\boldsymbol{y} \leftarrow \arg\max_{\boldsymbol{y} \in \mathcal{Y}} f(\boldsymbol{x}, \boldsymbol{y})$
    $\text{Lb}_{\boldsymbol{x}} \leftarrow f(\boldsymbol{x}, \boldsymbol{y})$
    $\varepsilon \leftarrow \text{Lb}_{\boldsymbol{x}} - \text{Lb}$
  end while
  $\mathcal{L} \leftarrow \mathcal{L} \cup \{(\boldsymbol{x}, \boldsymbol{y})\}$
end for
$(\boldsymbol{x}^\star, \boldsymbol{y}^\star) \leftarrow \arg\max_{(\boldsymbol{x}, \boldsymbol{y}) \in \mathcal{L}} f(\boldsymbol{x}, \boldsymbol{y})$
$\text{Lb}^\star = f(\boldsymbol{x}^\star, \boldsymbol{y}^\star)$
return $(\text{Lb}^\star, \boldsymbol{x}^\star, \boldsymbol{y}^\star)$

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 epigraph 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 epigraph 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

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

We choose c to be large enough such that (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.

D.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

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

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

$$
\begin{aligned}
\max_{\boldsymbol{x},\boldsymbol{y},\boldsymbol{U},\boldsymbol{X},\boldsymbol{Y}} \quad & \mathrm{Tr}(\boldsymbol{U}\boldsymbol{A})+\boldsymbol{b}^\top\boldsymbol{y} \\
\text{s.t.} \quad & \boldsymbol{D}^\top\boldsymbol{X}_i-\boldsymbol{d}x_i\le\boldsymbol{0}, && i\in[n_x], && \text{(40a)}\\
& \boldsymbol{D}^\top\boldsymbol{U}_j-\boldsymbol{d}y_j\le\boldsymbol{0}, && j\in[n_y], && \text{(40b)}\\
& \boldsymbol{d}\boldsymbol{x}^\top\boldsymbol{D}+\boldsymbol{D}^\top\boldsymbol{x}\boldsymbol{d}^\top\le\boldsymbol{D}^\top\boldsymbol{X}\boldsymbol{D}+\boldsymbol{d}\boldsymbol{d}^\top, && && \text{(40c)}\\
& \sum_{j\in\mathcal{J}_k} y_j=1, && k\in\mathcal{K}, && \text{(40d)}\\
& \sum_{j\in\mathcal{J}_k} \boldsymbol{U}_j-\boldsymbol{x}=\boldsymbol{0}, && k\in\mathcal{K}, && \text{(40e)}\\
& \sum_{j\in\mathcal{J}_k} \boldsymbol{Y}_j-\boldsymbol{y}=\boldsymbol{0}, && k\in\mathcal{K}, && \text{(40f)}\\
& \boldsymbol{X},\boldsymbol{Y},\boldsymbol{U}\ge\boldsymbol{0}, && && \text{(40g)}\\
& \begin{pmatrix}\boldsymbol{X} & \boldsymbol{U} & \boldsymbol{x}\\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{y}\\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & 1\end{pmatrix}\succeq\boldsymbol{0}, && && \text{(40h)}
\end{aligned}
$$

where constraints (40a) - (40g) result from pairwise multiplication of the linear constraints and constraint (40h) 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 (40e), (40d) and (40b):

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

The non-negativity constraint $\boldsymbol{x}\ge\boldsymbol{0}$ is redundant by (40e) and (40g). Moreover, the non-negativity constraint $\boldsymbol{y}\ge\boldsymbol{0}$ is redundant by (40f) and (40g). Hence, these constraints are not included in the above formulation.

In the RPT-SDP formulation we hence obtain $n_x|\mathcal{K}| + n_y|\mathcal{K}| + n_x n_y + Ln_y + n_x(n_x+1)/2 + L(L+1)/2 + n_x L + n_y(n_y+1)/2$ 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.

$\boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{X}}_2$**.** 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 & (40a) - (40g) \\
& \sum_{i=1}^{n_x} z_i \le 1, && (41a)\\
& \sum_{i=1}^{n_x} \boldsymbol{V}_i - \boldsymbol{x} \le \boldsymbol{0}, && (41b)\\
& \sum_{i=1}^{n_x} \boldsymbol{R}_i - \boldsymbol{y} \le \boldsymbol{0}, && (41c)\\
& \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), && (41d)\\
& \sum_{i,j=1}^{n_x} Z_{ij} - 2\sum_{i=1}^{n_x} z_i + 1 \ge 0, && (41e)\\
& \exp(x_i - a) \le z_i, \qquad i \in [n_x], && (41f)\\
& x_j \exp\left(\frac{X_{ij} - a x_j}{x_j}\right) \le V_{ji}, \qquad i,j \in [n_x], && (41g)\\
& y_j \exp\left(\frac{U_{ij} - a y_j}{y_j}\right) \le R_{ji}, \qquad i \in [n_x], j \in [n_y], && (41h)\\
& \left(d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}\right) \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, \qquad i \in [n_x], \ell \in [L], && (41i)\\
& \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}, \qquad i \in [n_x], && (41j)\\
& \exp\left(x_i + x_j - 2a\right) \le Z_{ij}, \qquad i \le j \in [n_x], && (41k)\\
& z_j \exp\left(\frac{V_{ij} - a z_j}{z_j}\right) \le Z_{ij} \qquad i \le j \in [n_x] && (41l)\\
& \sum_{j \in \mathcal{J}_k} \boldsymbol{R}_j - \boldsymbol{z} = \boldsymbol{0}, \qquad k \in \mathcal{K}, && (41m)\\
& \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}, && (41n)
\end{aligned}
$$

where constraints (41b) - (41e) result from pairwise multiplication of the new linear constraint over $\boldsymbol{z}$ with the previous linear constraints (41g) - (41l) result from pairwise multiplication of the exponential constraint with the linear inequalities and itself, constraint (41m) results from pairwise multiplication of the initial linear constraint over $\boldsymbol{y}$ with $\boldsymbol{z}$ and constraint (41n) results from the additional SDP relaxation.

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

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

The RPT-SDP formulation is given by

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

$$
\begin{aligned}
&(40a) - (40g) && \\
&s + p + \sum_{i=1}^{n_x} V_{ii} \le c, && && (42a)\\
&\sum_{i=1}^{n_x} t_i \le p, && && (42b)\\
&\boldsymbol{D}^\top \boldsymbol{V}_i \le z_i \boldsymbol{d}, && i \in [n_x], && (42c)\\
&\boldsymbol{D}^\top \boldsymbol{F}_i \le t_i \boldsymbol{d}, && i \in [n_x], && (42d)\\
&\boldsymbol{D}^\top \boldsymbol{\lambda} \le p\boldsymbol{d}, && && (42e)\\
&\sum_{j \in \mathcal{J}_k} \boldsymbol{r}_j - \boldsymbol{z} = \boldsymbol{0}, && k \in \mathcal{K}, && (42f)\\
&\sum_{j \in \mathcal{J}_k} \boldsymbol{g}_j - \boldsymbol{t} = \boldsymbol{0}, && k \in \mathcal{K}, && (42g)\\
&\sum_{j \in \mathcal{J}_k} \gamma_j - s = 0, && k \in \mathcal{K}, && (42h)\\
&\sum_{j \in \mathcal{J}_k} \mu_j - p = 0, && k \in \mathcal{K}, && (42i)\\
&\sum_{i=1}^{n_x} \boldsymbol{F}_i \le \boldsymbol{\lambda}, && && (42j)\\
&\sum_{i=1}^{n_x} \boldsymbol{G}_i \le \boldsymbol{\mu}, && && (42k)\\
&\sum_{i=1}^{n_x} \boldsymbol{H}_i \le \boldsymbol{\nu}, && && (42l)\\
&\sum_{i=1}^{n_x} \boldsymbol{T}_i \le \boldsymbol{\psi}, && && (42m)\\
&\sum_{i=1}^{n_x} \psi_i \le \pi, && && (42n)\\
&\sum_{i=1}^{n_x} d_\ell t_i - \boldsymbol{D}_\ell^\top \boldsymbol{F}_i \le p d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{\lambda}, && \ell \in \mathcal{L}, && (42o)\\
&\pi - 2\sum_{i=1}^{n_x} \psi_i + \sum_{i,j=1}^{n_x} T_{ij} \ge 0, && && (42p)\\
&\boldsymbol{V}, \boldsymbol{F}, \boldsymbol{R}, \boldsymbol{G}, \boldsymbol{H}, \boldsymbol{Z}, \boldsymbol{T} \ge \boldsymbol{0}, && && (42q)\\
&\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{\psi}, \sigma, \pi \ge \boldsymbol{0}. && && (42r)
\end{aligned}
$$

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

$$\|\boldsymbol{x}\|_2 \le s, \tag{43a}$$

$$\left\|(z_i - t_i, 1)^\top\right\|_2 \le z_i + t_i, \qquad i \in [n_x], \tag{43b}$$

$$\|\boldsymbol{X}_i\|_2 \le \alpha_i, \qquad i \in [n_x], \tag{43c}$$

$$\|\boldsymbol{U}_i\|_2 \le \beta_i, \qquad \imath \in [n_y], \tag{43d}$$

$$\|\boldsymbol{V}_i\|_2 \le \gamma_i, \qquad i \in [n_x], \tag{43e}$$

$$\|\boldsymbol{F}_i\|_2 \le \phi_i, \qquad i \in [n_x], \tag{43f}$$

$$\|\boldsymbol{\lambda}\|_2 \le \rho, \tag{43g}$$

$$\|d_\ell \boldsymbol{x} - \boldsymbol{X}\boldsymbol{D}_\ell\|_2 \le s d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{\alpha}, \qquad \ell \in \mathcal{L}, \tag{43h}$$

$$\|\boldsymbol{\lambda} - \sum_{i=1}^{n_x} \boldsymbol{F}_i\|_2 \le \rho - \sum_{i=1}^{n_x} \phi_i, \tag{43i}$$

$$\|\boldsymbol{X}\|_F \le \sigma, \tag{43j}$$

$$\|\,(V_{ji} - F_{ji}, x_j)\,\|_2 \le V_{ji} + F_{ji}, \qquad i, j \in [n_x], \tag{43k}$$

$$\|\,(R_{ji} - G_{ji}, y_j)\,\|_2 \le R_{ji} + G_{ji}, \qquad i \in [n_x], j \in [n_y], \tag{43l}$$

$$\|\,(Z_{ji} - H_{ji}, z_j)\,\|_2 \le Z_{ji} + H_{ji}, \qquad i, j \in [n_x], \tag{43m}$$

$$\|\,(H_{ij} - T_{ij}, t_j)\,\|_2 \le H_{ij} + T_{ij}, \qquad i, j \in [n_x], \tag{43n}$$

$$\|\,(\nu_i - \psi_i, p)\,\|_2 \le \nu_i + \psi_i, \qquad i \in [n_x], \tag{43o}$$

$$\|\,\left(d_\ell(z_i - t_i) + \boldsymbol{D}_\ell^\top (\boldsymbol{F}_i - \boldsymbol{V}_i), d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}\right)\,\|_2 \le d_\ell(z_i + t_i) - \boldsymbol{D}_\ell^\top(\boldsymbol{F}_i + \boldsymbol{V}_i), \qquad \ell \in \mathcal{L}, \tag{43p}$$

$$\|\,(\boldsymbol{V}_i - \boldsymbol{F}_i, \boldsymbol{x})\,\|_2 \le \gamma_i + \phi_i, \qquad i \in [n_x], \tag{43q}$$

$$\|\left(\nu_i - \psi_i - \sum_{j=1}^{n_x} H_{ij} + \sum_{j=1}^{n_x} T_{ij}, p - \sum_{j=1}^{n_x} t_j\right)\|_2 \le \nu_i + \psi_i - \sum_{j=1}^{n_x} H_{ij} - \sum_{j=1}^{n_x} T_{ij}, \qquad i \in [n_x], \tag{43r}$$

$$\left\|\begin{pmatrix} Z_{ij} - H_{ij} - H_{ji} + T_{ij} & z_i - t_i \\ z_j - t_j & 1 \end{pmatrix}\right\|_2 \le Z_{ij} + H_{ij} + H_{ji} + T_{ij}, \qquad i, j \in [n_x], \tag{43s}$$

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} & \boldsymbol{V} & \boldsymbol{F} & \boldsymbol{\alpha} & \boldsymbol{\lambda} & \boldsymbol{x} \\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{R} & \boldsymbol{G} & \boldsymbol{\beta} & \boldsymbol{\mu} & \boldsymbol{y} \\ \boldsymbol{V}^\top & \boldsymbol{R}^\top & \boldsymbol{Z} & \boldsymbol{H} & \boldsymbol{\gamma} & \boldsymbol{\nu} & \boldsymbol{z} \\ \boldsymbol{F}^\top & \boldsymbol{G}^\top & \boldsymbol{H}^\top & \boldsymbol{T} & \boldsymbol{\phi} & \boldsymbol{\psi} & \boldsymbol{t} \\ \boldsymbol{\alpha}^\top & \boldsymbol{\beta}^\top & \boldsymbol{\gamma}^\top & \boldsymbol{\phi}^\top & \sigma & \rho & s \\ \boldsymbol{\lambda}^\top & \boldsymbol{\mu}^\top & \boldsymbol{\nu}^\top & \boldsymbol{\psi}^\top & \rho & \pi & p \\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & \boldsymbol{z}^\top & \boldsymbol{t}^\top & s & p & 1 \end{pmatrix} \succeq \boldsymbol{0}. \tag{43t}$$

In the RPT-SDP formulation we hence obtain $(3|\mathcal{K}| + 2L + 7)n_x + (|\mathcal{K}| + 1)n_y + 3n_x n_y + L n_y + 3n_x(n_x + 1)/2 + 3n_x^2 + L(L + 1)/2 + n_x L + n_y(n_y + 1)/2 + 2|\mathcal{K}| + 2L + 6$ 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.

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

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

$$\max_{\substack{\boldsymbol{x}\in\mathcal{X}\\ \boldsymbol{y}\in\mathcal{Y}}} \quad \boldsymbol{x}^\top \boldsymbol{y}, \tag{ENM$_B$}$$

where $\mathcal{Y}$ is given by

$$\mathcal{Y} = \left\{\boldsymbol{y}\in\mathbb{R}^{n_x},\ \left\|\boldsymbol{y}\right\|_2 \le 1\right\}.$$

$\mathcal{X} = \mathcal{X}_1$

We notice that since we are maximizing $\boldsymbol{x}^\top\boldsymbol{y}$ with $\boldsymbol{x}\ge\boldsymbol{0}$, at the optimal solution we will always have $\boldsymbol{y}\ge\boldsymbol{0}$. Therefore we can pairwise multiply every basic cone constraint with y_j to further tighten the RPT relaxation. The formulation of the RPT-SDP relaxation is given by:

$$\begin{aligned}
\max_{\substack{\boldsymbol{x},\boldsymbol{y}\\ \boldsymbol{X},\boldsymbol{Y},\boldsymbol{U}}} \quad & \mathrm{Tr}(\boldsymbol{U}) \\
\text{s.t.} \quad & \boldsymbol{x}\in\mathcal{X}_1, && && \text{(44a)}\\
& \boldsymbol{y}\in\mathcal{Y}, && && \text{(44b)}\\
& \boldsymbol{X},\boldsymbol{Y},\boldsymbol{U}\ge\boldsymbol{0} && && \text{(44c)}\\
& \boldsymbol{D}^\top\boldsymbol{X}_i - \boldsymbol{d}x_i \le \boldsymbol{0}, && i\in[n_x], && \text{(44d)}\\
& \boldsymbol{D}^\top\boldsymbol{U}_i - \boldsymbol{d}y_i \le \boldsymbol{0}, && i\in[n_x], && \text{(44e)}\\
& \boldsymbol{d}\boldsymbol{x}^\top\boldsymbol{D} + \boldsymbol{D}^\top\boldsymbol{x}\boldsymbol{d}^\top \le \boldsymbol{D}^\top\boldsymbol{X}\boldsymbol{D} + \boldsymbol{d}\boldsymbol{d}^\top, && && \text{(44f)}\\
& \left\|(\boldsymbol{U})_i^\top\right\|_2 \le x_i, && i\in[n_x], && \text{(44g)}\\
& \left\|\boldsymbol{Y}_i\right\|_2 \le y_i, && i\in[n_x], && \text{(44h)}\\
& \left\|d_\ell\boldsymbol{y} - \boldsymbol{U}^\top\boldsymbol{D}_\ell\right\|_2 \le d_\ell - \boldsymbol{D}_\ell^\top\boldsymbol{x}, && \ell\in\mathcal{L}, && \text{(44i)}\\
& \left\|\boldsymbol{Y}\right\|_F \le 1, && && \text{(44j)}\\
& \begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} & \boldsymbol{x}\\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{y}\\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & 1\end{pmatrix} \succeq \boldsymbol{0}, && && \text{(44k)}
\end{aligned}$$

where constraints (44c) - (44f) result from pairwise multiplication of the linear constraints, constraints (44g) - (44i) result from pairwise multiplication of the linear constraints with the second order cone constraint and constraint (44k) results from the additional SDP relaxation.

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

Observe that we could exclude the constraints (44a) and the non-negativity constraint $\boldsymbol{y}\ge\boldsymbol{0}$, since from constraints (44g) - (44i) it follows that both are redundant.

$\mathcal{X} = \mathcal{X}_2$

The formulation of the RPT-SDP relaxation is given by:

$$
\begin{aligned}
\max_{\substack{\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}\\ \boldsymbol{X},\boldsymbol{Y},\boldsymbol{Z}\\ \boldsymbol{U},\boldsymbol{V},\boldsymbol{R}}} \quad & \mathrm{Tr}(\boldsymbol{U}) \\
\text{s.t.} \quad & (44a) - (44j), \\
& \sum_{i=1}^{n_x} z_i \le 1, && && (45a)\\
& \exp(x_i - a) \le z_i, && i \in [n_x], && (45b)\\
& \Big\|\boldsymbol{y} - \sum_{i=1}^{n_x} \boldsymbol{R}_i\Big\|_2 \le 1 - \sum_{i=1}^{n_x} z_i, && && (45c)\\
& \sum_{i=1}^{n_x} \boldsymbol{V}_i - \boldsymbol{x} \le \boldsymbol{0}, && && (45d)\\
& \sum_{i=1}^{n_x} \boldsymbol{R}_i - \boldsymbol{y} \le \boldsymbol{0}, && && (45e)\\
& \boldsymbol{D}^\top \boldsymbol{x} - \boldsymbol{D}^\top \sum_{i=1}^{n_x} \boldsymbol{V}_i \le \boldsymbol{d}\left(1 - \sum_{i=1}^{n_x} z_i\right), && && (45f)\\
& \sum_{i,j=1}^{n_x} Z_{ij} - 2\sum_{i=1}^{n_x} z_i + 1 \ge 0, && && (45g)\\
& x_j \exp\left(\frac{X_{ij} - a x_j}{x_j}\right) \le V_{ji}, && i,j \in [n_x], && (45h)\\
& y_j \exp\left(\frac{U_{ij} - a y_j}{y_j}\right) \le R_{ji}, && i,j \in [n_x], && (45i)\\
& (d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}) \exp\left(\frac{d_\ell(x_i - a) - \boldsymbol{D}_\ell^\top \boldsymbol{X}_i + a\boldsymbol{D}_\ell \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 \mathcal{L}, && (45j)\\
& \left(1 - \sum_{j=1}^{n_x} z_j\right) \exp\left(\frac{x_i - \sum_{j=1}^{n_x} \boldsymbol{V}_{ij} - a + 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_{ij}, && i \in [n_x], && (45k)\\
& \exp(x_i + x_j - 2a) \le Z_{ij}, && i \le j \in [n_x], && (45l)\\
& \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}, && && (45m)
\end{aligned}
$$

where constraints (45c) - (45g) result from pairwise multiplication of the linear constraint over $\boldsymbol{z}$, with the initial constraints and itself, constraints (45h) - (45l) result from pairwise multiplication of the exponential constraint with the linear constraints and itself and constraint (45m) results from the additional SDP relaxation. Note that the multiplication of the exponential constraint with the second order cone constraint yields a trivial constraint, as the left-hand side of the latter is 1.

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

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

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

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

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{(46a)}\\
& (\boldsymbol{y},\boldsymbol{w})\in\mathcal{Y} && \text{(46b)}\\
& \boldsymbol{X},\boldsymbol{Y},\boldsymbol{U}\ge\boldsymbol{0} && \text{(46c)}\\
& \boldsymbol{D}^\top\boldsymbol{X}_i-\boldsymbol{d}x_i\le\boldsymbol{0}, \qquad i\in[n_x], && \text{(46d)}\\
& \boldsymbol{D}^\top\boldsymbol{U}_i-\boldsymbol{d}y_i\le\boldsymbol{0}, \qquad i\in[n_x], && \text{(46e)}\\
& \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{(46f)}\\
& \sum_{i\in[n_x]}\boldsymbol{U}_i=\boldsymbol{x}, && \text{(46g)}\\
& \sum_{i\in[n_x]}\boldsymbol{Y}_i=\boldsymbol{y}, && \text{(46h)}\\
& \sum_{i\in[n_x]}(\boldsymbol{P})_i^\top=\boldsymbol{w}, && \text{(46i)}\\
& U_{ji}\exp\left(\frac{Q_{ji}}{U_{ji}}\right)\le x_j, \qquad i,j\in[n_x], && \text{(46j)}\\
& Y_{ij}\exp\left(\frac{P_{ji}}{Y_{ij}}\right)\le y_j, \qquad i,j\in[n_x], && \text{(46k)}\\
& \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}, && \text{(46l)}\\
& Y_{ij}\exp\left(\frac{P_{ji}+P_{ij}}{Y_{ij}}\right)\le 1, \qquad i\le j\in[n_x], && \text{(46m)}\\
& \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}, && \text{(46n)}
\end{aligned}$$

where constraints (46c) - (46i) result from pairwise multiplication of the linear constraints. Note that we only multiply the linear equality constraint with the variables (see Theorem 1). Constraints (46j) - (46l) result from pairwise multiplication of the linear inequality constraints with the exponential cone constraints, constraint (46m) results from pairwise multiplication of the exponential cone constraints with each other, and constraint (46n) 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 (46b), (46g), and (46h) it follows that they are redundant.

D.4. RPT-SDP formulation of Problem (35)

Replacing the objective function with the biconjugate function in (35) 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

$$\begin{aligned}
\max_{\substack{\boldsymbol{x},\boldsymbol{y},\boldsymbol{w},\\ \boldsymbol{X},\boldsymbol{Y},\boldsymbol{W},\\ \boldsymbol{U},\boldsymbol{Q},\boldsymbol{P}}} \quad & \operatorname{Tr}(\boldsymbol{U}\boldsymbol{A})+\boldsymbol{b}^\top y+\sum_{i=1}^{n_y}w_i \\
\text{s.t.} \quad & \boldsymbol{D}^\top\boldsymbol{x}\le\boldsymbol{d}, && && \text{(47a)}\\
& \boldsymbol{X},\boldsymbol{Y},\boldsymbol{U}\ge\boldsymbol{0}, && && \text{(47b)}\\
& \boldsymbol{D}^\top\boldsymbol{X}_i-\boldsymbol{d}x_i\le\boldsymbol{0}, && i\in[n_x], && \text{(47c)}\\
& \boldsymbol{D}^\top\boldsymbol{U}_j-\boldsymbol{d}y_j\le\boldsymbol{0}, && j\in[n_y], && \text{(47d)}\\
& \boldsymbol{d}\boldsymbol{x}^\top\boldsymbol{D}+\boldsymbol{D}^\top\boldsymbol{x}\boldsymbol{d}^\top\le\boldsymbol{D}^\top\boldsymbol{X}\boldsymbol{D}+\boldsymbol{d}\boldsymbol{d}^\top, && && \text{(47e)}\\
& \exp(-w_i-1)\le y_i, && i\in[n_y], && \text{(47f)}\\
& x_j\exp\left(\frac{-Q_{ji}-x_j}{x_j}\right)\le U_{ji}, && i\in[n_y],\, j\in[n_x], && \text{(47g)}\\
& y_j\exp\left(\frac{P_{ji}-y_j}{y_j}\right)\le Y_{ij}, && i,j\in[n_y], && \text{(47h)}\\
& (d_j-\boldsymbol{D}_j^\top\boldsymbol{x})\exp\left(\frac{\boldsymbol{D}_j^\top\boldsymbol{x}-d_j-w_id_j+\boldsymbol{D}_j^\top\boldsymbol{Q}_i}{d_j-\boldsymbol{D}_j^\top\boldsymbol{x}}\right)\le d_jy_i-\boldsymbol{D}_j^\top\boldsymbol{U}_i, && i\in[n_y],\, j\in[L], && \text{(47i)}\\
& \exp(-w_i-w_j-2)\le Y_{ij}, && i\le j\in[n_y], && \text{(47j)}\\
& \begin{pmatrix}\boldsymbol{X} & \boldsymbol{U} & \boldsymbol{Q} & \boldsymbol{x}\\ \boldsymbol{U}^\top & \boldsymbol{Y} & \boldsymbol{P} & \boldsymbol{y}\\ \boldsymbol{Q}^\top & \boldsymbol{P}^\top & \boldsymbol{W} & \boldsymbol{w}\\ \boldsymbol{x}^\top & \boldsymbol{y}^\top & \boldsymbol{w}^\top & 1\end{pmatrix}\succeq\boldsymbol{0}, && && \text{(47k)}
\end{aligned}$$

where constraints (47b) - (47d) result from pairwise multiplication of the linear constraints, constraints (47g) - (47j) result from pairwise multiplication of the exponential constraints with the linear and constraint (47k) results from the additional SDP relaxation.

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

D.5. RPT-SDP formulation of Problem (10)

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

$$
\begin{aligned}
\min_{\boldsymbol{x},\boldsymbol{X}} \quad & \mathrm{Tr}(\boldsymbol{P}_0\boldsymbol{X}) + \boldsymbol{q}_0^\top x + r_0 \\
\text{s.t.} \quad & \boldsymbol{x} \in \mathcal{X}_1, && && \text{(48a)}\\
& \boldsymbol{D}^\top \boldsymbol{X}_i - \boldsymbol{d}x_i \leq \boldsymbol{0}, && i \in [n_x], && \text{(48b)}\\
& \boldsymbol{d}\boldsymbol{x}^\top \boldsymbol{D} + \boldsymbol{D}^\top \boldsymbol{x}\boldsymbol{d}^\top \leq \boldsymbol{D}^\top \boldsymbol{X}\boldsymbol{D} + \boldsymbol{d}\boldsymbol{d}^\top, && && \text{(48c)}\\
& \mathrm{Tr}(\boldsymbol{P}_k\boldsymbol{X}) + \boldsymbol{k}_0^\top x + r_k \leq 0, && k \in \mathcal{NC}, && \text{(48d)}\\
& \left\|\boldsymbol{P}_k^{1/2}\boldsymbol{x}\right\|_2 \leq -r_k - \boldsymbol{q}_k^\top \boldsymbol{x}, && k \in \mathcal{C}, && \text{(48e)}\\
& \left\|\boldsymbol{P}_i^{1/2}\boldsymbol{X}\boldsymbol{P}_j^{1/2}\right\|_2 \leq r_i r_j + r_i \boldsymbol{q}_j^\top x + r_j \boldsymbol{q}_i^\top \boldsymbol{x} + \boldsymbol{q}_i^\top \boldsymbol{X}\boldsymbol{q}_j, && i,j \in \mathcal{C}, && \text{(48f)}\\
& \left\|d_\ell \boldsymbol{P}_k^{1/2}\boldsymbol{x} - \boldsymbol{P}_k^{1/2}\boldsymbol{X}\boldsymbol{D}_\ell\right\|_2 \leq -r_k d_\ell + r_k \boldsymbol{D}_\ell^\top \boldsymbol{x} - d_\ell \boldsymbol{q}_k^\top \boldsymbol{x} + \boldsymbol{q}_k^\top \boldsymbol{X}\boldsymbol{D}_\ell, && k \in \mathcal{C},\, \ell \in \mathcal{L}, && \text{(48g)}\\
& \left\|\boldsymbol{P}_k^{1/2}\boldsymbol{X}_j\right\| \leq -r_k x_j - \boldsymbol{q}_k^\top \boldsymbol{X}_j, && k \in \mathcal{C},\, j \in [n_x], && \text{(48h)}\\
& \boldsymbol{X} \geq \boldsymbol{0}, && && \text{(48i)}\\
& \begin{pmatrix} \boldsymbol{X} & \boldsymbol{x} \\ \boldsymbol{x}^\top & 1 \end{pmatrix} \succeq \boldsymbol{0}, && && \text{(48j)}
\end{aligned}
$$

where constraints (48b) - (48c) and (48i) result from pairwise multiplication of the linear constraints, constraints (48f) - (48h) result from pairwise multiplication of the convex quadratic constraints with the linear constraints and each other and constraint (48j) 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.

D.6. 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

$$
\begin{aligned}
\min_{\substack{\boldsymbol{x},\boldsymbol{u},\boldsymbol{v},\boldsymbol{z},\\ \boldsymbol{q},\boldsymbol{X},\boldsymbol{V},\\ \boldsymbol{S},w,W}} \quad & \sum_{k\in\mathcal{K}} z_k + \sum_{k\in\mathcal{K}_0} \frac{S_0}{\beta_\delta}\left(\exp(\beta_\delta t_{k+1}) - \exp(\beta_\delta t_k)\right) w_k + \frac{S_0}{\delta}\exp(\beta_\delta T)\, w_K \\
\text{s.t.} \quad & (C + bx_k)\exp\left(\frac{\lambda C h_k + \lambda b \sum_{i=0}^k U_{ik} - \delta t_k (C + bx_k)}{C + bx_k}\right) \leq z_k, && k \in \mathcal{K}_0, && \text{(49a)}\\
& \boldsymbol{D}^\top \boldsymbol{x} \leq \boldsymbol{d} && && \text{(49b)}\\
& \exp\left(-\theta \sum_{i=0}^k x_i\right) \leq w_k, && k \in \mathcal{K}_0, && \text{(49c)}\\
& \boldsymbol{D}^\top \boldsymbol{X}_k \leq \boldsymbol{x}_k d, && k \in \mathcal{K}_0, && \text{(49d)}\\
& \boldsymbol{d}\boldsymbol{x}^\top \boldsymbol{D} + \boldsymbol{D}^\top \boldsymbol{x}\boldsymbol{d}^\top \leq \boldsymbol{D}^\top \boldsymbol{X}\boldsymbol{D} + \boldsymbol{d}\boldsymbol{d}^\top, && && \text{(49e)}
\end{aligned}
$$

$$(d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}) \exp\left(\frac{-\theta d_\ell \sum_{i=0}^k x_i + \theta \sum_{i=0}^k \boldsymbol{D}_\ell^\top \boldsymbol{X}_i}{d_\ell - \boldsymbol{D}_\ell^\top \boldsymbol{x}}\right) \leq d_\ell w_k - \boldsymbol{D}_\ell^\top \boldsymbol{Q}_k, \qquad k \in \mathcal{K}_0, \ell \in \mathcal{L}, \tag{49f}$$

$$x_j \exp\left(\frac{-\theta \sum_{i=0}^k X_{ij}}{x_j}\right) \leq Q_{jk}, \qquad k, j \in \mathcal{K}_0, \tag{49g}$$

$$\exp\left(-\theta \sum_{i=1}^k x_i - \theta \sum_{i=1}^j x_i\right) \leq W_{jk} \qquad j, k \in \mathcal{K}_0 \tag{49h}$$

$$w_j \exp\left(\frac{-\theta \sum_{i=1}^k Q_{kj}}{w_j}\right) \leq W_{jk} \qquad j, k \in \mathcal{K}_0 \tag{49i}$$

$$\boldsymbol{x}, \boldsymbol{X} \geq \boldsymbol{0}, \tag{49j}$$

$$\begin{pmatrix} \boldsymbol{X} & \boldsymbol{Q} & \boldsymbol{x} \\ \boldsymbol{Q}^\top & \boldsymbol{W} & \boldsymbol{w} \\ \boldsymbol{x}^\top & \boldsymbol{w}^\top & 1 \end{pmatrix} \succeq \boldsymbol{0}, \tag{49k}$$

where constraint (49b) represents the upper bounds on $\boldsymbol{x}$ such that we obtain a compact feasible region, (49d) - (49f) result from pairwise multiplication of (49b) with all other constraints, (49g) results from pairwise multiplication of (49c) with the nonnegativity constraint $\boldsymbol{x} \geq 0$, (49h) - (49i) result from pairwise multiplication of the exponential constraint with itself, (49j) results from pairwise multiplication of the nonnegativity constraints, and constraint (49k) 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.

Appendix E. Data generation of numerical experiments

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

We use the data generated by Selvi et al. (2020, Appendix F.5). Instances 1 - 5 refer to the instances 1, 2, 3, 7, and 11 in Selvi et al. (2020, Appendix F.5) respectively. In every problem, every max-term has the same number of elements, i.e., $|\mathcal{J}_k| = |\mathcal{J}_{k'}|$ for every $k, k' \in \mathcal{K}$.

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 10.

Instance	n_x	$\lvert\mathcal{K}\rvert$	$\lvert\mathcal{J}_k\rvert$	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 10 Problem (29) parameters for each instance. n_x refers to the number of variables, $|\mathcal{K}|$ to the number of max linear terms, $|\mathcal{J}_k|$ 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).

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

The data for each problem instance are generated as follows:

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

Problem instance 2: $D_{ij} \sim [0,1], d_i \sim [10,30]$,

Problem instance 3: $D_{ij} \sim [0,1], d_i \sim [10,30]$,

Problem instance 4: $D_{ij} \sim [0,1], d_i \sim [30,70]$,

Problem instance 5: $D_{ij} \sim [0,1], d_i \sim [30,70]$.

The parameters describing each instance are summarized in Table 11.

Instance	n_x	L	a
1	6	3	8
2	10	10	8
3	20	20	13
4	40	40	20
5	50	50	25

Table 11 Problem (31) parameters for each instance. n_x refers to the number of variables, L to the number of linear constraints and a to the parameter used in $\mathcal{X}_2$.

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

The problem instances are adopted from Selvi et al. (2020) 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 12.

Table 12 Problem (33) 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

E.4. Data generation of numerical experiments of Problem (35)

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 13.

Table 13 Problem (35) 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	200	200	3
3	20	20	8
4	30	20	20
5	40	10	30

E.5. Data generation of numerical experiments of Problem (10)

The first 5 problem instances are adopted from Selvi et al. (2020) 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 and 8 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 14.

Table 14 Problem (10) 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

E.6. 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 15. Moreover, we have $\theta = \alpha - \zeta,\ \beta_\delta = \alpha\eta + \gamma - \delta$.

Table 15 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