

Reformulation-Perspectification Technique for 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

University of Chinese Academy of Sciences, Beijing, China, zhenjianzhe@ucas.ac.cn

In this paper, we propose a new global optimization approach for solving nonconvex optimization problems in which the nonconvex components are sums of products of convex functions. A broad class of nonconvex problems can be formulated in this way, such as concave minimization problems, problems with difference of convex functions in the objective and constraints, and fractional optimization problems. Our approach leverages two techniques: first, we introduce a new technique, called the Reformulation-Perspectification Technique (RPT), to obtain a convex approximation of the considered nonconvex continuous optimization problem. Next, we employ a spatial branch and bound scheme that combines RPT with eigenvector branching to obtain a global optimum solution. Numerical experiments on sum-of-max-linear-terms maximization, log-sum-exp maximization, linear multiplicative optimization, quadratically constrained quadratic optimization, and dike height optimization demonstrate that the proposed approach is competitive with state-of-the-art global optimization solvers and can substantially outperform them on selected challenging instances, particularly when strengthened with semidefinite constraints.

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

1. Introduction

We introduce a new global optimization approach for nonconvex optimization problems where the nonconvex elements are sums of products of convex functions. This formulation covers a wide range of optimization problems, including nonconvex quadratic optimization, mixed binary linear optimization, concave minimization, difference of convex programming, and fractional optimization.

For nonconvex quadratic optimization problems and mixed binary linear optimization problems, hierarchical convex approximations can be obtained from the Reformulation-Linearization Technique

(RLT) (Sherali and Adams, 1999). RLT was introduced in Sherali and Adams (1990), and improved by many authors (Sturm and Zhang, 2003; Anstreicher, 2009, 2012, 2017; Bao et al., 2011; Yang and Burer, 2016; Jiang and Li, 2019). RLT is also applicable to mixed binary polynomial and to continuous nonconvex polynomial optimization problems (Sherali and Adams, 1999), and has been extended to mixed binary semi-infinite and convex optimization problems (Sherali and Adams, 2009). RLT consists of two steps: 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 the 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 handling more types of nonconvexity than RLT as well as potentially obtaining tighter approximations than RLT-based methods.

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

Another state-of-the-art method that utilizes branch and bound, addressing nonconvex problems with an emphasis on integer problems, is the global optimization algorithm SCIP, developed by [Achterberg \(2009\)](#).

Although eigenvector branching has been used before in global optimization, the novelty of RPT-BB lies in combining it with the RPT relaxation. The branching inequalities and the bounds on the branching direction are multiplied with the other constraints of the node. This yields child relaxations that are at least as tight as those of [Anstreicher \(2022\)](#), and it extends the rule to relaxations in which the semidefinite constraint linking the lifted and original variables is not imposed. Section 6 proves that, under the stated assumptions, the obtained lower bounds converge to the global optimum up to the accuracy with which the node relaxations are solved.

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

Our main contributions can be summarized as follows:

1. We extend the existing RLT approach to a broader class of nonconvex optimization problems, namely optimization problems in which the nonconvex components are sums of products of convex functions. The proposed RPT approach can handle multiplication of constraints that are neither linear nor quadratic, and thereby potentially obtains tighter approximations than RLT. Moreover, it can also handle more types of nonconvexity than RLT.
2. We introduce RPT-BB, a new global optimization approach that combines the new RPT relaxation with a modification of the eigenvector branching strategy of [Anstreicher \(2022\)](#). Our modification extends the branching strategy to settings in which the lifted rank-one residual need not be positive semidefinite. Under the stated assumptions, we establish convergence of the lower bounds to the global optimum.
3. We provide several theoretical insights for our approach. We show that using epigraphical variables for the nonlinear convex components in the nonconvex objective and/or constraints, yields an RPT relaxation that is at least as tight as without the introduction of epigraphical variables. Further, we show that adding linear constraints that are redundant to existing linear constraints does not tighten the RPT relaxation, while adding linear constraints that are redundant to existing nonlinear constraints can be useful.

4. We conduct a computational evaluation of the proposed RPT-based branch and bound methods on five classes of problems: sum-of-max-linear-terms maximization, log-sum-exp maximization, linear multiplicative optimization, quadratically constrained quadratic optimization, and dike height optimization. In addition, we study alternative branching implementations in Appendix A.

This paper is organized as follows. Section 2 introduces the generic nonconvex optimization problem we consider. Section 3 presents the RPT-BB approach. Section 4 illustrates the approach on a simple example. Section 5 discusses additional strengthening strategies. Section 6 provides the convergence analysis. Section 7 shows how several known convex reformulations and relaxations can also be obtained via RPT. Section 8 reports numerical experiments. Section 9 concludes.

Notation. We generally use bold faced characters such as $\mathbf{a} \in \mathbb{R}^n$ and $\mathbf{A} \in \mathbb{R}^{m \times n}$ to represent vectors and matrices, respectively, a_i to denote the i -th element of the vector $\mathbf{a}$, $\mathbf{A}_i \in \mathbb{R}^m$ to denote the i -th column of matrix $\mathbf{A}$, and A_{ij} to denote the entry of $\mathbf{A}$ in the i -th row and j -th column, unless specified otherwise. The calligraphic letters $\mathcal{I}, \mathcal{J}, \mathcal{K}, \mathcal{L}$ and the corresponding capital Roman letters I, J, K, L are reserved for finite index sets and their respective cardinalities, *i.e.*, $\mathcal{I} = \{1, \dots, I\}$ 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 $\text{ri}(\mathcal{V})$ to denote the relative interior of a set $\mathcal{V} \subseteq \mathbb{R}^{n_\nu}$. The *domain* of a function $f : \mathbb{R}^{n_\nu} \rightarrow [-\infty, +\infty]$ is defined as $\text{dom}(f) = \{\boldsymbol{\nu} \in \mathbb{R}^{n_\nu} \mid f(\boldsymbol{\nu}) < +\infty\}$. 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 $\text{dom}(f) \neq \emptyset$. 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 $f : \mathbb{R}^{n_\nu} \rightarrow [-\infty, +\infty]$ is the function $f^* : \mathbb{R}^{n_\nu} \rightarrow [-\infty, +\infty]$ defined through $f^*(\mathbf{w}) = \sup_{\boldsymbol{\nu}} \{\boldsymbol{\nu}^\top \mathbf{w} - f(\boldsymbol{\nu})\}$. The conjugate $(f^*)^*$ of f^* is called the *biconjugate* of f and is abbreviated as f^{**} . The *indicator function* $\delta_{\mathcal{V}} : \mathbb{R}^{n_\nu} \rightarrow [-\infty, +\infty]$ 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* $\delta_{\mathcal{V}}^* : \mathbb{R}^{n_\nu} \rightarrow [-\infty, +\infty]$ of a set $\mathcal{V} \subseteq \mathbb{R}^{n_\nu}$ is defined through $\delta_{\mathcal{V}}^*(\mathbf{w}) = \sup_{\boldsymbol{\nu} \in \mathcal{V}} \{\boldsymbol{\nu}^\top \mathbf{w}\}$. The *perspective function* of a proper, closed and convex function $f : \mathbb{R}^{n_\nu} \rightarrow (-\infty, +\infty]$ is defined as $h(\boldsymbol{\nu}, t) = tf(\boldsymbol{\nu}/t)$ if $t > 0$, and $h(\boldsymbol{\nu}, 0) = \delta_{\text{dom}(f^*)}^*(\boldsymbol{\nu})$ for all $\boldsymbol{\nu} \in \mathbb{R}^{n_\nu}$. For ease of exposition, we use $tf(\boldsymbol{\nu}/t)$ to denote the perspective function $h(\boldsymbol{\nu}, t)$ for the rest of this paper.

2. Generic problem formulation

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

$$\begin{aligned} \min_{\mathbf{x}} \quad & f_0(\mathbf{x}) \\ \text{s.t.} \quad & f_k(\mathbf{x}) \leq 0, \quad k \in \mathcal{K}, \\ & \mathbf{x} \in \mathcal{X}, \end{aligned} \tag{1}$$

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

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

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

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

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

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

ASSUMPTION 2. *If r_{ik} and c_{ik} are both nonlinear, then $r_{ik}(\mathbf{x}) \geq 0$ and $c_{ik}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathcal{X}$, for every $i \in \mathcal{I}$ and $k \in \mathcal{K}_0$. If r_{ik} is linear and c_{ik} is nonlinear, then $r_{ik}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathcal{X}$, 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:

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

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,

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

and $q_{ik} \in \mathbb{R}$, $\mathbf{d}_{ik} \in \mathbb{R}^{n_x}$, and $c_{0k}, c_{ik} : \mathbb{R}^{n_x} \rightarrow (-\infty, +\infty]$ are proper, closed and convex for every $i \in \mathcal{I}_k$, and $k \in \mathcal{K}_0$. Throughout, (2) is read with the convention $0 \cdot c_{ik}(\mathbf{x}) = 0$.

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

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

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

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

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

#	f	$\text{dom}(f^*)$	f^*
1	$f(\mathbf{x}, \bar{x}) = \ \mathbf{x}\ _2 - \bar{x}$	$\{(\mathbf{y}, \bar{y}) : \ \mathbf{y}\ _2 \leq 1, \bar{y} = -1\}$	$f^*(\mathbf{y}, \bar{y}) = 0$
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	$f(\mathbf{x}) = \max_i x_i$	$\{\mathbf{y} : \mathbf{y} \geq 0, \sum_i y_i = 1\}$	$f^*(\mathbf{y}) = 0$
6	$f(\mathbf{x}) = \sum_i \max_{k \in \mathcal{K}_i} x_{ik}$	$\{\{\mathbf{y}_i\}_i : \mathbf{y}_i \geq 0, \sum_{k \in \mathcal{K}_i} y_{ik} = 1, \forall i\}$	$f^*(\mathbf{y}) = 0$
7	$f(\mathbf{x}) = \log(\sum_i \exp(x_i))$	$\{\mathbf{y} : \mathbf{y} \geq 0, \sum_i y_i = 1\}$	$f^*(\mathbf{y}) = \sum_i y_i \log(y_i)$
8	$f(\mathbf{x}) = \sum_i g_i(\mathbf{x})$	$\{\sum_i \mathbf{y}_i : \mathbf{y}_i \in \text{dom}(g_i^*), \forall i\}$	$f^*(\mathbf{y}) = \min_{\sum_i \mathbf{y}_i = \mathbf{y}} \sum_i g_i^*(\mathbf{y}_i)$

EXAMPLE 2 (FRACTIONAL OPTIMIZATION). Consider the following fractional function

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

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

EXAMPLE 3 (SOME EXAMPLES OF SLC FUNCTIONS). Table 2 lists several examples of SLC functions that are generally nonconvex. These functions are SLC-representable provided that Assumption 2 holds. Hence, our proposed approach can deal with Problem (1) containing (sums of) such nonconvex components. $\square$

#	$f(\mathbf{x})$	$c_1(\mathbf{x})$	$(q - \mathbf{d}^\top \mathbf{x})$	Assumptions
1	$\sqrt{q - \mathbf{d}^\top \mathbf{x}}$	$(q - \mathbf{d}^\top \mathbf{x})^{-1/2}$	$q - \mathbf{d}^\top \mathbf{x}$	—
2	$(q - \mathbf{d}^\top \mathbf{x})^\theta$	$(q - \mathbf{d}^\top \mathbf{x})^{\theta-1}$	$q - \mathbf{d}^\top \mathbf{x}$	$\theta \in [0, 1]$
3	$-(q - \mathbf{d}^\top \mathbf{x})^\theta$	$-(q - \mathbf{d}^\top \mathbf{x})^{\theta-1}$	$q - \mathbf{d}^\top \mathbf{x}$	$\theta \in [1, 2]$
4	$-(q_1 - \mathbf{d}_1^\top \mathbf{x}) \ln(q_2 - \mathbf{d}_2^\top \mathbf{x})$	$-\ln(q_2 - \mathbf{d}_2^\top \mathbf{x})$	$q_1 - \mathbf{d}_1^\top \mathbf{x}$	—
5	$(q - \mathbf{d}^\top \mathbf{x}) \mathbf{x}^\top \mathbf{Q} \mathbf{x}$	$\mathbf{x}^\top \mathbf{Q} \mathbf{x}$	$q - \mathbf{d}^\top \mathbf{x}$	$\mathbf{Q} \succeq 0$

Table 2 Selected SLC-representable functions and their defining components.

3. Reformulation-Perspectification Technique and Branch and Bound

In this section, we describe our new approach, called RPT-BB, to solve (1) to global optimality. Our approach comprises five steps:

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

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

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

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

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

3.1. Preprocessing step

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

$$\begin{aligned} \min_{\mathbf{x}, \boldsymbol{\tau}} \quad & \tau_0 + \sum_{i \in \mathcal{I}_0} (q_{i0} - \mathbf{d}_{i0}^\top \mathbf{x}) c_{i0}(\mathbf{x}) \\ \text{s.t.} \quad & \tau_k + \sum_{i \in \mathcal{I}_k} (q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}) c_{ik}(\mathbf{x}) \leq 0, \quad k \in \mathcal{K}, \\ & (\mathbf{x}, \boldsymbol{\tau}) \in \mathcal{T}, \end{aligned} \tag{4}$$

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

3.2. Reformulation and perspectification

Now we are ready to explain the core idea of RPT. We first consider the univariate case. The intuition is as follows: Consider the convex function $c : \mathbb{R} \rightarrow \mathbb{R}$ and the nonconvex constraint set $\mathcal{X} = \{x \in \mathbb{R}_+ : x c(x) \leq 0\}$. The constraint set can be written as $\mathcal{X} =$

$\left\{ (x, x') \in \mathbb{R}_+ \times \mathbb{R}_+ : x c\left(\frac{x'}{x}\right) \leq 0, x' = x^2 \right\}$. Observe that the new set $\mathcal{X}$ is also nonconvex due to the constraint $x' = x^2$. However, observe that for $x \geq 0$, the function $h(x, x') = x c\left(\frac{x'}{x}\right)$ is the perspective function of $c(\cdot)$ and therefore it is jointly convex in x and x' . Thus, by either completely relaxing $x' = x^2$ and adding the hitherto redundant $x' \geq 0$, or by relaxing it as $x' \geq x^2$ to preserve convexity (and nonnegativity), one obtains a convex outer approximation of $\mathcal{X}$.

REMARK 1 (RELATIONSHIP BETWEEN RLT AND RPT). Observe that the RLT approach is a subcase of RPT as described above, assuming the function $c(\cdot)$ is linear. Since the perspective of a linear function is the function itself, the perspectification step is not needed in RLT. However, for nonlinear functions this is not the case and therefore the perspectification step is necessary.

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

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

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

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

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

$$\begin{aligned} \min_{\mathbf{x}, \boldsymbol{\tau}, \mathbf{X}} \quad & \tau_0 + \sum_{i \in \mathcal{I}_0} (q_{i0} - \mathbf{d}_{i0}^\top \mathbf{x}) c_{i0} \left(\frac{q_{i0} \mathbf{x} - \mathbf{X} \mathbf{d}_{i0}}{q_{i0} - \mathbf{d}_{i0}^\top \mathbf{x}} \right) \\ \text{s.t.} \quad & \tau_k + \sum_{i \in \mathcal{I}_k} (q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}) c_{ik} \left(\frac{q_{ik} \mathbf{x} - \mathbf{X} \mathbf{d}_{ik}}{q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}} \right) \leq 0, \quad k \in \mathcal{K}, \\ & (\mathbf{x}, \boldsymbol{\tau}) \in \mathcal{T}. \end{aligned} \quad (6)$$

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

considering the pairwise multiplication of the linear and convex constraints in the set $\mathcal{T}$, followed by the pairwise multiplication of the convex inequalities in the set $\mathcal{T}$.

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

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

$$\mathbf{b}\mathbf{x}^\top \mathbf{A} + \mathbf{A}^\top \mathbf{x}\mathbf{b}^\top \leq \mathbf{A}^\top \mathbf{x}\mathbf{x}^\top \mathbf{A} + \mathbf{b}\mathbf{b}^\top.$$

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

$$\mathbf{b}\mathbf{x}^\top \mathbf{A} + \mathbf{A}^\top \mathbf{x}\mathbf{b}^\top \leq \mathbf{A}^\top \mathbf{X} \mathbf{A} + \mathbf{b}\mathbf{b}^\top. \quad (7)$$

Moreover, we include the additional constraints

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

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

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

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

Next, the redundant SLC constraints can be reformulated into:

$$\begin{aligned} (b_\ell - \mathbf{A}_\ell^\top \mathbf{x})\mathbf{h}(\mathbf{x}) \leq \mathbf{0} &\iff (b_\ell - \mathbf{A}_\ell^\top \mathbf{x})\mathbf{h}\left(\frac{b_\ell \mathbf{x} - \mathbf{x}\mathbf{x}^\top \mathbf{A}_\ell}{b_\ell - \mathbf{A}_\ell^\top \mathbf{x}}\right) \leq \mathbf{0} \quad \text{and} \\ (b_\ell - \mathbf{A}_\ell^\top \mathbf{x})\mathbf{c}_0(\mathbf{x}) \leq (b_\ell - \mathbf{A}_\ell^\top \mathbf{x})\boldsymbol{\tau} &\iff (b_\ell - \mathbf{A}_\ell^\top \mathbf{x})\mathbf{c}_0\left(\frac{b_\ell \mathbf{x} - \mathbf{x}\mathbf{x}^\top \mathbf{A}_\ell}{b_\ell - \mathbf{A}_\ell^\top \mathbf{x}}\right) \leq (b_\ell - \mathbf{A}_\ell^\top \mathbf{x})\boldsymbol{\tau}. \end{aligned}$$

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

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

Table 3 complements Table 2 in Example 3 by showing the perspectified forms of the SLC-representable functions.

#	$f(\mathbf{x})$	Perspectification	Assumptions
1	$\sqrt{q - \mathbf{d}^\top \mathbf{x}}$	$(q - \mathbf{d}^\top \mathbf{x}) \sqrt{\frac{q - \mathbf{d}^\top \mathbf{x}}{q^2 - 2q\mathbf{d}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{X} \mathbf{d}}}$	—
2	$(q - \mathbf{d}^\top \mathbf{x})^\theta$	$(q - \mathbf{d}^\top \mathbf{x}) \left(\frac{q^2 - 2q\mathbf{d}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{X} \mathbf{d}}{q - \mathbf{d}^\top \mathbf{x}} \right)^{\theta-1}$	$\theta \in [0, 1]$
3	$-(q - \mathbf{d}^\top \mathbf{x})^\theta$	$-(q - \mathbf{d}^\top \mathbf{x}) \left(\frac{q^2 - 2q\mathbf{d}^\top \mathbf{x} + \mathbf{d}^\top \mathbf{X} \mathbf{d}}{q - \mathbf{d}^\top \mathbf{x}} \right)^{\theta-1}$	$\theta \in [1, 2]$
4	$-(q_1 - \mathbf{d}_1^\top \mathbf{x}) \ln(q_2 - \mathbf{d}_2^\top \mathbf{x})$	$-(q_1 - \mathbf{d}_1^\top \mathbf{x}) \ln \left(\frac{q_1 q_2 - q_1 \mathbf{d}_2^\top \mathbf{x} - q_2 \mathbf{d}_1^\top \mathbf{x} + \mathbf{d}_1^\top \mathbf{X} \mathbf{d}_2}{q_1 - \mathbf{d}_1^\top \mathbf{x}} \right)$	—
5	$(q - \mathbf{d}^\top \mathbf{x}) \mathbf{x}^\top \mathbf{Q} \mathbf{x}$	$\frac{(q \mathbf{x} - \mathbf{X} \mathbf{d})^\top \mathbf{Q} (q \mathbf{x} - \mathbf{X} \mathbf{d})}{q - \mathbf{d}^\top \mathbf{x}}$	$\mathbf{Q} \succeq \mathbf{0}$

Table 3 Perspectified forms of the SLC-representable functions in Table 2.

Linear equality $\times$ Convex inequality. When multiplying a linear equality constraint with a convex inequality constraint, the denominator and coefficient of the resulting perspective function are zero. Fortunately, all additional nonlinear constraints resulting from multiplying a linear equality constraint with a convex inequality constraint are redundant as long as we consider the pairwise multiplication of the linear equality constraints with all variables (see Lemma 1). For quadratic problems, a similar observation was first mentioned by [Sherali and Adams \(1999, Remark 8.1\)](#). Before we formally prove Lemma 1, we first define redundant constraints.

DEFINITION 1 (REDUNDANT EQUALITY CONSTRAINTS). An equality constraint $f(\mathbf{x}) = 0$, where $f: \mathbb{R}^{n_x} \rightarrow (-\infty, +\infty]$, is *redundant* to the nonempty set $\mathcal{X} \subseteq \mathbb{R}^{n_x}$, if $\mathcal{X} \subseteq \{\mathbf{x} \mid f(\mathbf{x}) = 0\}$.

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

Proof. Fix $(\mathbf{x}, \mathbf{X})$ such that $\mathbf{d}^\top \mathbf{x} = q$ and $\mathbf{X}\mathbf{d} = q\mathbf{x}$. Then $q - \mathbf{d}^\top \mathbf{x} = 0$ and $q\mathbf{x} - \mathbf{X}\mathbf{d} = \mathbf{0}$. By the standard closure formula for the perspective of a proper, closed and convex function,

$$(q - \mathbf{d}^\top \mathbf{x}) f\left(\frac{q\mathbf{x} - \mathbf{X}\mathbf{d}}{q - \mathbf{d}^\top \mathbf{x}}\right) = \delta_{\text{dom}(f^*)}^*(\mathbf{0}).$$

Since f is proper, closed, and convex, its conjugate f^* is proper ([Rockafellar, 1970, Thm. 12.2](#)). By definition of properness, $\text{dom}(f^*) \neq \emptyset$, hence

$$\delta_{\text{dom}(f^*)}^*(\mathbf{0}) = \sup_{\mathbf{y} \in \text{dom}(f^*)} \langle \mathbf{0}, \mathbf{y} \rangle = 0.$$

Therefore, $(q - \mathbf{d}^\top \mathbf{x}) f\left(\frac{q\mathbf{x} - \mathbf{X}\mathbf{d}}{q - \mathbf{d}^\top \mathbf{x}}\right) = 0$ is satisfied by every $(\mathbf{x}, \mathbf{X})$ in the set, i.e., it is redundant. $\square$

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

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

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

$$\begin{aligned} \min_{\mathbf{x}, \boldsymbol{\tau}, \mathbf{X}, \mathbf{V}} \quad & \tau_0 + \sum_{i \in \mathcal{I}_0} (q_{i0} - \mathbf{d}_{i0}^\top \mathbf{x}) c_{i0} \left(\frac{q_{i0} \mathbf{x} - \mathbf{X} \mathbf{d}_{i0}}{q_{i0} - \mathbf{d}_{i0}^\top \mathbf{x}} \right) \\ \text{s.t.} \quad & \tau_k + \sum_{i \in \mathcal{I}_k} (q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}) c_{ik} \left(\frac{q_{ik} \mathbf{x} - \mathbf{X} \mathbf{d}_{ik}}{q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}} \right) \leq 0, \quad k \in \mathcal{K}, \\ & \mathbf{b}\mathbf{x}^\top \mathbf{A} + \mathbf{A}^\top \mathbf{x}\mathbf{b}^\top \leq \mathbf{A}^\top \mathbf{X} \mathbf{A} + \mathbf{b}\mathbf{b}^\top, \\ & X_{ii} \geq 0, \quad i \in \{1, \dots, n_x\}, \\ & (b_\ell - \mathbf{A}_\ell^\top \mathbf{x}) \mathbf{h} \left(\frac{b_\ell \mathbf{x} - \mathbf{X} \mathbf{A}_\ell}{b_\ell - \mathbf{A}_\ell^\top \mathbf{x}} \right) \leq 0, \quad \ell \in \{1, \dots, m_1\}, \\ & (b_\ell - \mathbf{A}_\ell^\top \mathbf{x}) c_0 \left(\frac{b_\ell \mathbf{x} - \mathbf{X} \mathbf{A}_\ell}{b_\ell - \mathbf{A}_\ell^\top \mathbf{x}} \right) \leq b_\ell \boldsymbol{\tau} - \mathbf{V} \mathbf{A}_\ell, \quad \ell \in \{1, \dots, m_1\}, \\ & s_\ell \mathbf{x} - \mathbf{X} \mathbf{P}_\ell = \mathbf{0}, \quad \ell \in \{1, \dots, m_2\}, \\ & s_\ell \boldsymbol{\tau} - \mathbf{V} \mathbf{P}_\ell = \mathbf{0}, \quad \ell \in \{1, \dots, m_2\}, \\ & (\mathbf{x}, \boldsymbol{\tau}) \in \mathcal{T}. \end{aligned} \tag{8}$$

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

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

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

Proof. Applying RPT without the preprocessing step yields the relaxation:

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{X}} \quad & c_{00}(\mathbf{x}) + \sum_{i \in \mathcal{I}_0} (q_{i0} - \mathbf{d}_{i0}^\top \mathbf{x}) c_{i0} \left(\frac{q_{i0} \mathbf{x} - \mathbf{X} \mathbf{d}_{i0}}{q_{i0} - \mathbf{d}_{i0}^\top \mathbf{x}} \right) \\ \text{s.t.} \quad & c_{0k}(\mathbf{x}) + \sum_{i \in \mathcal{I}_k} (q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}) c_{ik} \left(\frac{q_{ik} \mathbf{x} - \mathbf{X} \mathbf{d}_{ik}}{q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}} \right) \leq 0, \quad k \in \mathcal{K}, \\ & \mathbf{b}\mathbf{x}^\top \mathbf{A} + \mathbf{A}^\top \mathbf{x}\mathbf{b}^\top \leq \mathbf{A}^\top \mathbf{X} \mathbf{A} + \mathbf{b}\mathbf{b}^\top, \\ & X_{ii} \geq 0, \quad i \in \{1, \dots, n_x\}, \\ & (b_\ell - \mathbf{A}_\ell^\top \mathbf{x}) \mathbf{h} \left(\frac{b_\ell \mathbf{x} - \mathbf{X} \mathbf{A}_\ell}{b_\ell - \mathbf{A}_\ell^\top \mathbf{x}} \right) \leq 0, \quad \ell \in \{1, \dots, m_1\}, \\ & s_\ell \mathbf{x} - \mathbf{X} \mathbf{P}_\ell = \mathbf{0}, \quad \ell \in \{1, \dots, m_2\}, \\ & \mathbf{x} \in \mathcal{X}. \end{aligned} \tag{9}$$

Compared to (8), problem (9) omits the following two sets of constraints:

$$\begin{aligned} (b_\ell - \mathbf{A}_\ell^\top \mathbf{x}) c_0 \left(\frac{b_\ell \mathbf{x} - \mathbf{X} \mathbf{A}_\ell}{b_\ell - \mathbf{A}_\ell^\top \mathbf{x}} \right) &\leq b_\ell \boldsymbol{\tau} - \mathbf{V} \mathbf{A}_\ell, & \ell \in \{1, \dots, m_1\}, \\ s_\ell \boldsymbol{\tau} - \mathbf{V} \mathbf{P}_\ell &= \mathbf{0}, & \ell \in \{1, \dots, m_2\}. \end{aligned} \quad (10)$$

Let (11) denote the problem obtained from (8) by dropping the two sets of constraints in (10):

$$\begin{aligned} \min_{\mathbf{x}, \boldsymbol{\tau}, \mathbf{X}, \mathbf{V}} \quad & \tau_0 + \sum_{i \in \mathcal{I}_0} (q_{i0} - \mathbf{d}_{i0}^\top \mathbf{x}) c_{i0} \left(\frac{q_{i0} \mathbf{x} - \mathbf{X} \mathbf{d}_{i0}}{q_{i0} - \mathbf{d}_{i0}^\top \mathbf{x}} \right) \\ \text{s.t.} \quad & \tau_k + \sum_{i \in \mathcal{I}_k} (q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}) c_{ik} \left(\frac{q_{ik} \mathbf{x} - \mathbf{X} \mathbf{d}_{ik}}{q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}} \right) \leq 0, & k \in \mathcal{K}, \\ & \mathbf{b} \mathbf{x}^\top \mathbf{A} + \mathbf{A}^\top \mathbf{x} \mathbf{b}^\top \leq \mathbf{A}^\top \mathbf{X} \mathbf{A} + \mathbf{b} \mathbf{b}^\top, \\ & X_{ii} \geq 0, & i \in \{1, \dots, n_x\}, \\ & (b_\ell - \mathbf{A}_\ell^\top \mathbf{x}) h \left(\frac{b_\ell \mathbf{x} - \mathbf{X} \mathbf{A}_\ell}{b_\ell - \mathbf{A}_\ell^\top \mathbf{x}} \right) \leq 0, & \ell \in \{1, \dots, m_1\}, \\ & s_\ell \mathbf{x} - \mathbf{X} \mathbf{P}_\ell = \mathbf{0}, & \ell \in \{1, \dots, m_2\}, \\ & c_{0k}(\mathbf{x}) \leq \tau_k, & k \in \mathcal{K}_0, \\ & \mathbf{x} \in \mathcal{X}. \end{aligned} \quad (11)$$

Since dropping constraints can only enlarge the feasible region, it follows that $\min(8) \geq \min(11)$. Next, observe that after removing (10), the variable $\mathbf{V}$ no longer appears in any constraint and can therefore be eliminated without affecting feasibility or the minimum. Moreover, in (11) each τ_k appears, besides its epigraph constraint $c_{0k}(\mathbf{x}) \leq \tau_k$, only with coefficient +1 in the objective ($k=0$) or in the k -th constraint ($k \in \mathcal{K}$). Hence decreasing τ_k to $c_{0k}(\mathbf{x})$ preserves feasibility and does not increase the objective, so that τ_k can be eliminated by setting it to $c_{0k}(\mathbf{x})$ for every $k \in \mathcal{K}_0$. Observe that, eliminating $\boldsymbol{\tau}$ from (11) recovers exactly (9). Hence, $\min(11) = \min(9)$, which implies that $\min(8) \geq \min(9)$. Therefore, (8) is at least as tight as (9). $\square$

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

EXAMPLE 4 (EFFECTIVENESS OF PREPROCESSING). Consider the following nonconvex optimization problem:

$$\begin{aligned} \max_{\mathbf{x}} \quad & x_1 x_2 - \frac{(x_2 - 2)^2}{4} \\ \text{s.t.} \quad & 0 \leq x_1 \leq 1, \quad -4 \leq x_2 \leq 4. \end{aligned} \quad (12)$$

In the following we compare two convex relaxations of (12) obtained from (i) applying RPT without the epigraphical reformulation, and (ii) applying RPT with the epigraphical reformulation,

respectively,

$$\begin{aligned}
\max_{\mathbf{x}, \mathbf{X}} \quad & X_{12} - \frac{(x_2 - 2)^2}{4} \\
\text{s.t.} \quad & 0 \leq x_1 \leq 1, \quad -4 \leq x_2 \leq 4, \\
& X_{11} - 2x_1 + 1 \geq 0, \\
& x_1 - X_{11} \geq 0, \quad X_{11} \geq 0, \\
& -4x_1 \leq X_{12} \leq 4x_1, \\
& x_2 - 4(1 - x_1) \leq X_{12} \leq x_2 + 4(1 - x_1), \\
& X_{22} + 8x_2 + 16 \geq 0, \quad 16 - X_{22} \geq 0, \\
& X_{22} - 8x_2 + 16 \geq 0,
\end{aligned}
\quad \text{and} \quad
\begin{aligned}
\max_{\substack{\mathbf{x}, \mathbf{X} \\ \tau, \mathbf{v}}} \quad & X_{12} - \tau \\
\text{s.t.} \quad & 0 \leq x_1 \leq 1, \quad -4 \leq x_2 \leq 4, \\
& X_{11} - 2x_1 + 1 \geq 0, \\
& x_1 - X_{11} \geq 0, \quad X_{11} \geq 0, \\
& -4x_1 \leq X_{12} \leq 4x_1, \\
& x_2 - 4(1 - x_1) \leq X_{12} \leq x_2 + 4(1 - x_1), \\
& X_{22} + 8x_2 + 16 \geq 0, \quad 16 - X_{22} \geq 0, \\
& X_{22} - 8x_2 + 16 \geq 0, \\
& \frac{(x_2 - 2)^2}{4} \leq \tau, \\
& \frac{(X_{12} - 2x_1)^2}{4x_1} \leq v_1, \\
& \frac{(x_2 - 2 - X_{12} + 2x_1)^2}{4 - 4x_1} \leq \tau - v_1, \\
& \frac{(2x_2 + X_{22} - 8)^2}{16 + 4x_2} \leq 4\tau + v_2, \\
& \frac{(6x_2 - X_{22} - 8)^2}{16 - 4x_2} \leq 4\tau - v_2.
\end{aligned}$$

Here v_1 and v_2 are the additional variables that linearize the products τx_1 and τx_2 , respectively. The maximum objective value of the convex relaxation without the epigraphical reformulation is then $\frac{13}{4}$, which is finite but strictly larger than the global optimum 3. It is attained at $x_1^* = \frac{7}{8}$, $x_2^* = 3$, $X_{12}^* = \frac{7}{2}$. In contrast, the objective value of the epigraphical reformulation equals the global optimum 3, attained at $(x_1^*, x_2^*, X_{11}^*, X_{12}^*, X_{22}^*, \tau^*, v_1^*, v_2^*) = (1, 4, 1, 4, 16, 1, 1, 4)$. Thus the epigraphical reformulation is exact here, whereas the non-epigraphical relaxation retains a strictly positive gap. In fact, any two-sided bound $-M \leq x_2 \leq M$ with $M \geq 4$ yields the same picture, with non-epigraphical value $\frac{2M+5}{4}$. $\square$

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

LEMMA 2 (Redundant Epigraph Variables). *Introducing epigraphical variables for the linear components in the preprocessing step of RPT does not result in a tighter convex relaxation.*

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

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

Here $(\mathbf{V}^\top)_k$ and $\mathbf{T}_k$ represent $\tau_k \mathbf{x}$ and $\tau_k \boldsymbol{\tau}$, respectively. Write $c_{0k}(\mathbf{x}) = \mathbf{a}^\top \mathbf{x} + b$. The displayed equalities fix $(\mathbf{V}^\top)_k = \mathbf{X} \mathbf{a} + b \mathbf{x}$ and, when $\mathbf{T}$ is introduced, $\mathbf{T}_k = \mathbf{V} \mathbf{a} + b \boldsymbol{\tau}$. Substituting these

identities and $\tau_k = c_{0k}(\mathbf{x})$ into the additional RPT products of this equality reduces them to identities or to constraints already present before introducing τ_k . Thus the new variables can be eliminated without strengthening the projected relaxation. Using the epigraph inequality instead of this equality cannot give a stronger relaxation. $\square$

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

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

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

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

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

A related mechanism arises when a nonlinear convex inequality is replaced by a redundant linear support (e.g., a supporting hyperplane) and then combined with a second nonlinear convex inequality. Although such linear supports do not change the feasible region, they can generate additional valid inequalities after RPT, thereby tightening the relaxation. An analogous observation has already been made in the literature for RLT; see, for example, the SOC–RLT construction in Burer and Anstreicher (2013).

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 $\mathbf{x}\mathbf{x}^\top$ are linearized by a symmetric matrix $\mathbf{X}$. Such a linearization based relaxation for the nonconvex quadratic equality $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ may

be significantly improved by the SDP relaxation $\mathbf{X} \succeq \mathbf{x}\mathbf{x}^\top$, which can be equivalently reformulated as an LMI by using Schur complement (Boyd and Vandenberghe, 2004):

$$\begin{pmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{x}^\top & 1 \end{pmatrix} \succeq \mathbf{0}. \quad (13)$$

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

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

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

3.4. Obtaining upper bounds

Let $(\mathbf{x}^*, \boldsymbol{\tau}^*, \mathbf{X}^*, \mathbf{V}^*)$ be a feasible solution of the convex RPT relaxation (8). We construct the candidate set

$$\mathcal{X}' = \{\mathbf{x}^*, \mathbf{x}_1^X, \dots, \mathbf{x}_{n_x}^X, \mathbf{x}_0^V, \dots, \mathbf{x}_K^V\},$$

where

$$\mathbf{x}_i^X = \begin{cases} \mathbf{x}^*, & x_i^* = 0, \\ \frac{\mathbf{X}_i^*}{x_i^*}, & x_i^* \neq 0, \end{cases} \quad i \in \{1, \dots, n_x\}, \quad \text{and} \quad \mathbf{x}_k^V = \begin{cases} \mathbf{x}^*, & \tau_k^* = 0, \\ \frac{(\mathbf{V}^{*\top})_k}{\tau_k^*}, & \tau_k^* \neq 0, \end{cases} \quad k \in \mathcal{K}_0.$$

Because $(\mathbf{x}^*, \boldsymbol{\tau}^*) \in \mathcal{T}$, we have $\mathbf{x}^* \in \mathcal{X}$.

The construction is motivated by the exact lifted identities. If the relaxation is rank-one exact, then $\mathbf{X}^* = \mathbf{x}^*(\mathbf{x}^*)^\top$, and consequently $\mathbf{X}_i^* = x_i^* \mathbf{x}^*$ for every i . Thus $\mathbf{X}_i^*/x_i^*$ recovers $\mathbf{x}^*$ whenever $x_i^* \neq 0$. When $x_i^* = 0$, we use $\mathbf{x}^*$ itself. Similarly, the exact identity $\mathbf{V}^* = \boldsymbol{\tau}^*(\mathbf{x}^*)^\top$ gives $(\mathbf{V}^{*\top})_k = \tau_k^* \mathbf{x}^*$ and motivates the candidates $\mathbf{x}_k^V$.

LEMMA 3 (Candidate Solution Construction). *If the nonnegativity constraints $x_i \geq 0$, $i \in \{1, \dots, n_x\}$, occur explicitly among the rows of $\mathbf{A}^\top \mathbf{x} \leq \mathbf{b}$, then $\mathbf{x}_i^X \in \mathcal{X}$ for all $i \in \{1, \dots, n_x\}$. Suppose, in addition, that $c_{0k}(\mathbf{x}) \geq 0$ for all $\mathbf{x} \in \mathcal{X}$ and $k \in \mathcal{K}_0$, so that the inequalities $\boldsymbol{\tau} \geq \mathbf{0}$ are valid for $\mathcal{T}$. If these inequalities are included in $\mathcal{T}$ and their RPT products with the constraints defining $\mathcal{X}$ are included in the relaxation, that is, if the relaxation contains the linearized constraints*

$$\mathbf{A}^\top (\mathbf{V}^\top)_k \leq \tau_k \mathbf{b}, \quad \tau_k \mathbf{h} \left(\frac{(\mathbf{V}^\top)_k}{\tau_k} \right) \leq \mathbf{0}, \quad k \in \mathcal{K}_0, \quad (14)$$

then also $\mathbf{x}_k^V \in \mathcal{X}$ for all $k \in \mathcal{K}_0$, and hence $\mathcal{X}' \subseteq \mathcal{X}$.

Proof. Fix $i \in \{1, \dots, n_x\}$. Since $\mathbf{x}^* \in \mathcal{X}$ and $x_i \geq 0$ is a row of $\mathbf{A}^\top \mathbf{x} \leq \mathbf{b}$, we have $x_i^* \geq 0$. Hence, if $x_i^* \neq 0$, then $x_i^* > 0$. Moreover, because $x_i \geq 0$ is a row of $\mathbf{A}^\top \mathbf{x} \leq \mathbf{b}$, its RPT products with the linear inequalities, linear equalities, and convex inequalities defining $\mathcal{X}$ are among the constraints of (8), and at the relaxation solution they read

$$\mathbf{A}^\top \mathbf{X}_i^* \leq x_i^* \mathbf{b}, \quad \mathbf{P}^\top \mathbf{X}_i^* = x_i^* \mathbf{s}, \quad x_i^* \mathbf{h}\left(\frac{\mathbf{X}_i^*}{x_i^*}\right) \leq \mathbf{0}.$$

Dividing by $x_i^* > 0$ preserves the inequality directions and gives

$$\mathbf{A}^\top \frac{\mathbf{X}_i^*}{x_i^*} \leq \mathbf{b}, \quad \mathbf{P}^\top \frac{\mathbf{X}_i^*}{x_i^*} = \mathbf{s}, \quad \mathbf{h}\left(\frac{\mathbf{X}_i^*}{x_i^*}\right) \leq \mathbf{0}.$$

These are precisely the defining conditions of $\mathcal{X}$, so $\mathbf{x}_i^X \in \mathcal{X}$. If $x_i^* = 0$, then $\mathbf{x}_i^X = \mathbf{x}^* \in \mathcal{X}$ by definition. Thus $\mathbf{x}_i^X \in \mathcal{X}$ for all $i \in \{1, \dots, n_x\}$.

For the second part, fix $k \in \mathcal{K}_0$. The definition of $\mathcal{T}$ gives $c_{0k}(\mathbf{x}) \leq \tau_k$, which together with $c_{0k}(\mathbf{x}) \geq 0$ on $\mathcal{X}$ shows that $\tau_k \geq 0$ is valid for $\mathcal{T}$. In particular, $\tau_k^* \geq 0$ because $(\mathbf{x}^*, \boldsymbol{\tau}^*) \in \mathcal{T}$, so if $\tau_k^* \neq 0$, then $\tau_k^* > 0$. Evaluating the two relations in (14) and the equality constraints $\mathbf{s}_\ell \boldsymbol{\tau} - \mathbf{V} \mathbf{P}_\ell = \mathbf{0}$ of (8) at the relaxation solution and dividing by $\tau_k^* > 0$ gives

$$\mathbf{A}^\top \frac{(\mathbf{V}^{*\top})_k}{\tau_k^*} \leq \mathbf{b}, \quad \mathbf{P}^\top \frac{(\mathbf{V}^{*\top})_k}{\tau_k^*} = \mathbf{s}, \quad \mathbf{h}\left(\frac{(\mathbf{V}^{*\top})_k}{\tau_k^*}\right) \leq \mathbf{0},$$

and hence $\mathbf{x}_k^V = (\mathbf{V}^{*\top})_k / \tau_k^* \in \mathcal{X}$. If $\tau_k^* = 0$, then $\mathbf{x}_k^V = \mathbf{x}^* \in \mathcal{X}$ by definition. Thus $\mathbf{x}_k^V \in \mathcal{X}$ for all $k \in \mathcal{K}_0$, and consequently $\mathcal{X}' \subseteq \mathcal{X}$. $\square$

Under Assumption 1, the nonnegativity requirement in Lemma 3 is without loss of generality. Indeed, compactness of $\mathcal{X}$ allows us to translate each variable by its minimum over $\mathcal{X}$, yielding an equivalent formulation with explicitly nonnegative variables while preserving convexity and the SLC structure.

Membership in $\mathcal{X}$ is not generally sufficient for feasibility of Problem (1). It is sufficient when $\mathcal{K} = \emptyset$. Otherwise, the inequalities $f_k(\mathbf{x}) \leq 0$, $k \in \mathcal{K}$, must also hold. We therefore test every candidate in $\mathcal{X}'$ by direct substitution into the original problem. Let $\mathcal{X}''$ contain the distinct candidates that are feasible and have finite objective value. This check also covers formulations in which the nonnegativity and product conditions of Lemma 3 are not imposed, so membership in $\mathcal{X}$ is not guaranteed in advance.

We additionally use each candidate in $\mathcal{X}'$ at which the model functions are evaluable as an initial point for a local solver such as Ipopt (Wächter et al., 2009), whether or not that candidate is already feasible. Every feasible point with finite objective value returned by the solver is added to $\mathcal{X}''$. For the disjoint bilinear formulations obtained through the biconjugate representation, the

mountain climbing method of [Tao and An \(1997\)](#) provides another local improvement procedure (Appendix C). If $\mathcal{X}'' \neq \emptyset$, then

$$\min_{\mathbf{x} \in \mathcal{X}''} f_0(\mathbf{x})$$

is a finite valid upper bound on the optimal value of Problem (1). If $\mathcal{X}'' = \emptyset$, this procedure produces no finite upper bound.

3.5. Spatial branch and bound method

We solve Problem (1) to global optimality via a spatial branch and bound scheme. Each node N_j of the search tree is associated with a relaxation region $\mathcal{X}_j \subseteq \mathbb{R}^{n_x} \times \mathbb{S}^{n_x}$ that collects the constraints on $(\mathbf{x}, \mathbf{X})$ inherited along the path to N_j (including $\mathbf{x} \in \mathcal{X}$). At node N_j , we solve the convex RPT relaxation (8) with the additional restriction $(\mathbf{x}, \mathbf{X}) \in \mathcal{X}_j$; this yields a valid lower bound. A feasible solution (and hence an upper bound) is obtained using the techniques in Section 3.4. The algorithm terminates when the global optimality gap (Ub – Lb) is at most δ , a user-specified optimality tolerance.

Before formally defining the branching strategy, we make the following observation. Problem (8) is a relaxation of Problem (1) obtained by omitting the rank-one condition $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$. Any feasible $\mathbf{x}$ of (1) can be lifted to a feasible solution of (8) by setting $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ (and, when epigraph variables are present, choosing $\boldsymbol{\tau}$ accordingly and setting $\mathbf{V} = \boldsymbol{\tau}\mathbf{x}^\top$). Moreover,

$$\{(\mathbf{x}, \mathbf{X}) \mid \mathbf{X} = \mathbf{x}\mathbf{x}^\top\} = \{(\mathbf{x}, \mathbf{X}) \mid \forall \boldsymbol{\nu} \in \mathbb{R}^{n_x} : (\boldsymbol{\nu}^\top \mathbf{x})^2 \leq \boldsymbol{\nu}^\top \mathbf{X} \boldsymbol{\nu}, \boldsymbol{\nu}^\top \mathbf{X} \boldsymbol{\nu} \leq (\boldsymbol{\nu}^\top \mathbf{x})^2\}. \quad (15)$$

Thus, omitting the nonconvex constraints in (15) yields the convex relaxation (8), which can be solved efficiently, but whose solution is often infeasible for Problem (1).

To address violations of (15), let $(\mathbf{x}^*, \boldsymbol{\tau}^*, \mathbf{X}^*, \mathbf{V}^*)$ be the solution of the RPT relaxation at node N_j , which is an optimal solution when the relaxation is solved exactly, let $\mathcal{F}_j$ denote the feasible set of that relaxation, and define

$$\mathbf{R}_j := \mathbf{X}^* - \mathbf{x}^*(\mathbf{x}^*)^\top.$$

If $\mathbf{R}_j \neq \mathbf{0}$, choose a unit eigenvector $\boldsymbol{\nu}_j$ associated with an eigenvalue λ_j of $\mathbf{R}_j$ such that $|\lambda_j| = \|\mathbf{R}_j\|_2$, and set $s_j := \boldsymbol{\nu}_j^\top \mathbf{x}^*$. The rank-one condition requires $\boldsymbol{\nu}_j^\top \mathbf{X} \boldsymbol{\nu}_j = (\boldsymbol{\nu}_j^\top \mathbf{x})^2$, while at the relaxation solution

$$\boldsymbol{\nu}_j^\top \mathbf{X}^* \boldsymbol{\nu}_j - (\boldsymbol{\nu}_j^\top \mathbf{x}^*)^2 = \lambda_j, \quad (16)$$

so $\boldsymbol{\nu}_j$ is a direction of maximal rank-one violation, as in the eigenvector branching of [Anstreicher \(2022\)](#). Next, we compute the range of this direction over the node relaxation,

$$a_j := \inf\{\boldsymbol{\nu}_j^\top \mathbf{x} : (\mathbf{x}, \boldsymbol{\tau}, \mathbf{X}, \mathbf{V}) \in \mathcal{F}_j\} \quad \text{and} \quad b_j := \sup\{\boldsymbol{\nu}_j^\top \mathbf{x} : (\mathbf{x}, \boldsymbol{\tau}, \mathbf{X}, \mathbf{V}) \in \mathcal{F}_j\}, \quad (17)$$

so that $a_j \leq s_j \leq b_j$. Both values are finite (Lemma 5(c) in Appendix F). They need not be attained, because the projection of $\mathcal{F}_j$ onto the $\mathbf{x}$ -space need not be closed when $\mathcal{F}_j$ is unbounded in the lifted variables, but attainment is not needed in what follows. The two convex problems in (17) are independent and can be solved in parallel. Alternative valid bounds are discussed in Remark 3.

Write $t = \boldsymbol{\nu}_j^\top \mathbf{x}$ and $y = \boldsymbol{\nu}_j^\top \mathbf{X} \boldsymbol{\nu}_j$. The left child N_{j_1} is obtained by appending the two linear inequalities $a_j \leq t \leq s_j$ to the constraints of N_j , and the right child N_{j_2} by appending $s_j \leq t \leq b_j$. In each child, the reformulation step of Section 3.2 is applied to the augmented system. The two new inequalities are multiplied with themselves, with each other, and with every other constraint of the node in which only $\mathbf{x}$ appears, and the products are perspectified and linearized as before. Among the generated inequalities, the product $(s_j - t)^2 \geq 0$ of the branching inequality with itself gives

$$y \geq 2s_j t - s_j^2 \quad (18)$$

in both children, the product $(s_j - t)(t - a_j) \geq 0$ of the two new inequalities gives

$$y \leq (a_j + s_j)t - a_j s_j \quad (19)$$

in the left child, and the product $(t - s_j)(b_j - t) \geq 0$ gives

$$y \leq (s_j + b_j)t - s_j b_j \quad (20)$$

in the right child. These are the tangent of t^2 at s_j and its secants on $[a_j, s_j]$ and $[s_j, b_j]$. At the parent solution, $t = s_j$ and $y = s_j^2 + \lambda_j$ by (16), whereas (18) and the applicable secant force $y = s_j^2$ at $t = s_j$. Hence $(\mathbf{x}^*, \mathbf{X}^*)$ is infeasible for both children whenever $\lambda_j \neq 0$. The tangent excludes it when $\lambda_j < 0$ and the secant when $\lambda_j > 0$. The same inequalities are generated in every child, so the sign of λ_j plays no role in the construction. In particular, the rule is applied unchanged when the semidefinite constraint of Section 3.3 is imposed, in which case $\mathbf{R}_j \succeq \mathbf{0}$ and $\lambda_j > 0$ whenever $\mathbf{R}_j \neq \mathbf{0}$.

The construction contains the eigenvector branching of Anstreicher (2022) as a special case. There, the semidefinite constraint is imposed and $\boldsymbol{\nu}_j$ is an eigenvector of the largest eigenvalue of $\mathbf{R}_j$. For bounds $a_j \leq \boldsymbol{\nu}_j^\top \mathbf{x} \leq b_j$ that are valid over the relaxation, the two children are defined by $t \geq s_j$ together with $y \leq (s_j + b_j)t - s_j b_j$, and by $t \leq s_j$ together with $y \leq (a_j + s_j)t - a_j s_j$ (Anstreicher, 2022, Proposition 2.1). These are the split and the secants (19)–(20). Minimizing and maximizing $\boldsymbol{\nu}_j^\top \mathbf{x}$ over the relaxation, as in (17), is one of the ways of obtaining a_j and b_j suggested there. Two features distinguish the present rule. First, the secants are not appended as separate cuts but arise as products of the appended inequalities. These inequalities are also multiplied with every other constraint of the node, with the linear inequalities as in RLT and with the convex inequalities through their perspectives. For a norm constraint, the latter products are the second-order cone

constraints of Burer and Anstreicher (2013). Starting from the same relaxation at N_j and using the same bounds, each child relaxation of the present rule is therefore contained in the corresponding child relaxation of Anstreicher (2022), that is, it is at least as tight. Second, the product of the branching inequality with itself yields the tangent (18). This tangent is implied by the semidefinite constraint when that constraint is imposed, but it is what excludes the parent solution when $\lambda_j < 0$. This is why the rule extends to relaxations in which $\mathbf{R}_j$ need not be positive semidefinite.

Branching at node N_j does not exclude any feasible solution of Problem (1) whose rank-one lifting satisfies the restrictions at N_j . Indeed, let $\mathbf{x}$ be feasible for Problem (1) and suppose $(\mathbf{x}, \mathbf{x}\mathbf{x}^\top) \in \mathcal{X}_j$. Then $\mathbf{x}$ can be lifted to a feasible solution of the node relaxation by setting $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ (and, when epigraph variables are present, choosing $\boldsymbol{\tau}$ accordingly and setting $\mathbf{V} = \boldsymbol{\tau}\mathbf{x}^\top$). The split $\boldsymbol{\nu}_j^\top \mathbf{x} \leq s_j$ or $\boldsymbol{\nu}_j^\top \mathbf{x} \geq s_j$ is exhaustive. The bounds (17) hold at $\mathbf{x}$ because the lifted solution belongs to $\mathcal{F}_j$, and at an exact lift $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ every generated inequality reduces to the square or the product of nonnegative scalar slacks. Hence the lifted solution remains feasible in at least one child (Lemma 5 in Appendix F). Consequently, branching only on $(\mathbf{x}, \mathbf{X})$ is sufficient for correctness: it does not exclude any feasible (in particular, globally optimal) solution of Problem (1). The auxiliary variables $\boldsymbol{\tau}$ and $\mathbf{V}$ need not be branched on and are re-optimized in each child relaxation.

After branching, we solve the RPT relaxation over each child region (optimizing over $(\mathbf{x}, \boldsymbol{\tau}, \mathbf{X}, \mathbf{V})$ subject to $(\mathbf{x}, \mathbf{X}) \in \mathcal{X}_{j_k}$) to obtain lower bounds and update the global upper bound using Section 3.4. The branching rule is summarized in Algorithm 1. We do not use constraints involving $\mathbf{X}$ (including the generated inequalities (18)–(20)) for further constraint multiplications, as doing so would introduce products such as $x_i X_{k\ell}$ (i.e., third-order terms); instead, we keep these constraints in their original form. The new inequalities and the inequalities generated from them are retained by every descendant node. We prune a node whenever its lower bound exceeds the current best upper bound and select the next node using best-bound selection (i.e., the active node with the smallest lower bound). Algorithm 2 summarizes the overall procedure.

REMARK 2. Since the eigenvector $\boldsymbol{\nu}$ corresponding to the eigenvalue with the largest absolute value is typically dense, the resulting branching constraints are dense as well. Dense constraints may increase per-node solve time relative to coordinate branching; however, in our experiments we did not observe a consistent or dominant runtime impact attributable to this density. Moreover, a dense $\boldsymbol{\nu}$ couples many components of $\mathbf{x}$, which can help tighten the relaxation when RPT is applied.

REMARK 3 (ALTERNATIVE DIRECTIONAL BOUNDS). Any bounds $a_j \leq \boldsymbol{\nu}_j^\top \mathbf{x} \leq b_j$ that are valid over $\mathcal{F}_j$ may be used in place of (17). The children remain valid, but the secants become weaker.

REMARK 4 (OPTIONAL SOC STRENGTHENING). When $\lambda_j < 0$, the convex inequality $(\boldsymbol{\nu}_j^\top \mathbf{x})^2 \leq \boldsymbol{\nu}_j^\top \mathbf{X} \boldsymbol{\nu}_j$ is violated at the relaxation solution and may be added to the node before branching,

followed by re-optimization, as in the cutting-plane method of Saxena et al. (2010). This inequality does not partition the node. Appendix A compares this variant with branching alone.

Algorithm 1 Eigenvector branching strategy

Input: A node N_j , its region $\mathcal{X}_j$, and the solution $(\mathbf{x}^*, \boldsymbol{\tau}^*, \mathbf{X}^*, \mathbf{V}^*)$ of the relaxation at N_j .

Output: $(\mathcal{X}_{j1}, \mathcal{X}_{j2})$.

- 1: Set $\mathbf{R}_j \leftarrow \mathbf{X}^* - \mathbf{x}^*(\mathbf{x}^*)^\top$.
 - 2: Compute a unit eigenvector $\boldsymbol{\nu}_j$ of $\mathbf{R}_j$ associated with an eigenvalue λ_j satisfying $|\lambda_j| = \|\mathbf{R}_j\|_2$ (any unit vector if $\mathbf{R}_j = \mathbf{0}$).
 - 3: Set $s_j \leftarrow \boldsymbol{\nu}_j^\top \mathbf{x}^*$.
 - 4: Compute the bounds a_j and b_j of (17).
 - 5: Form the left child $\mathcal{X}_{j1}$ from $\mathcal{X}_j$, $a_j \leq \boldsymbol{\nu}_j^\top \mathbf{x} \leq s_j$, and its standard RPT constraints.
 - 6: Form the right child $\mathcal{X}_{j2}$ from $\mathcal{X}_j$, $s_j \leq \boldsymbol{\nu}_j^\top \mathbf{x} \leq b_j$, and its standard RPT constraints.
-

4. A simple example

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

$$\begin{aligned}
 \min_{x_1, x_2, x_3} \quad & 3x_1 - 3x_2 + 3x_3 + (x_1 + x_2 + 1) \exp(x_1) + (x_1 + x_2 + 1) \exp(x_3) \\
 \text{s.t.} \quad & x_1 + x_2 \geq -1, \\
 & x_i \leq 10, \\
 & \exp(x_2 - x_3) \leq x_1, \\
 & 2 \exp\left(\frac{-x_1}{2}\right) + 2 \exp\left(\frac{-x_2}{2}\right) \leq 2 + \exp(-1).
 \end{aligned} \tag{21}$$

Let $\mathcal{X}_T$ denote the feasible set of toy problem (21), consisting of four linear constraints 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 $\times$ Linear. First, we perspectify the SLC objective, and then we consider adding constraints:

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

Algorithm 2 Branch and bound via RPT**Input:** $(N_0, \mathcal{X}_0, \delta)$, where $\delta \geq 0$ is the absolute optimality tolerance.**Output:** $(\mathbf{x}^*, \text{Lb}, \text{Ub})$.

-
- 1: Solve the RPT relaxation over $\mathcal{X}_0$; obtain its solution $\sigma_0 = (\mathbf{x}_0^*, \tau_0^*, \mathbf{X}_0^*, \mathbf{V}_0^*)$, its lower bound Lb_0 , and an upper bound Ub_0 (possibly $+\infty$) from Section 3.4; if Ub_0 is finite, store the feasible candidate attaining it as the incumbent $\mathbf{x}^*$.
 - 2: $\text{Lb} \leftarrow \text{Lb}_0$; $\text{Ub} \leftarrow \text{Ub}_0$; $\text{ACTIVE} \leftarrow \{N_0\}$.
 - 3: **while** $\text{ACTIVE} \neq \emptyset$ **and** $\text{Ub} - \text{Lb} > \delta$ **do**
 - 4: $N_j \leftarrow \arg \min_{N_i \in \text{ACTIVE}} \text{Lb}_i$; $\text{ACTIVE} \leftarrow \text{ACTIVE} \setminus \{N_j\}$.
 - 5: Apply Algorithm 1 to $(N_j, \mathcal{X}_j, \sigma_j)$ to obtain the child regions $\mathcal{X}_{j_1}$ and $\mathcal{X}_{j_2}$.
 - 6: **for** $k = 1, 2$ **do**
 - 7: $i \leftarrow j_k$. Solve the RPT relaxation over $\mathcal{X}_i$. If it is infeasible, discard N_i ; otherwise obtain σ_i , Lb_i , and Ub_i , and set $\text{Lb}_i \leftarrow \max\{\text{Lb}_i, \text{Lb}_j\}$. {the bound of the parent remains valid for the child}
 - 8: **if** $\text{Ub}_i < \text{Ub}$ **then**
 - 9: $\text{Ub} \leftarrow \text{Ub}_i$ and update the incumbent $\mathbf{x}^*$.
 - 10: **end if**
 - 11: **end for**
 - 12: **for** $k = 1, 2$ **do**
 - 13: $i \leftarrow j_k$. **if** N_i is feasible and $\text{Lb}_i < \text{Ub}$ **then** $\text{ACTIVE} \leftarrow \text{ACTIVE} \cup \{N_i\}$.
 - 14: **end for**
 - 15: $\text{Lb} \leftarrow \min\{\text{Ub}, \min_{N_i \in \text{ACTIVE}} \text{Lb}_i\}$ if $\text{ACTIVE} \neq \emptyset$, and $\text{Lb} \leftarrow \text{Ub}$ otherwise.
 - 16: **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 $X_{11}, X_{22}, X_{33}, X_{12}, X_{13}$, and $X_{23} \in \mathbb{R}$ respectively to obtain the following convex relaxation:

$$\begin{aligned}
\min_{\mathbf{x}, \mathbf{X}} \quad & 3x_1 - 3x_2 + 3x_3 + (x_1 + x_2 + 1) \exp\left(\frac{X_{11} + X_{12} + x_1}{x_1 + x_2 + 1}\right) + (x_1 + x_2 + 1) \exp\left(\frac{X_{13} + X_{23} + x_3}{x_1 + x_2 + 1}\right) \\
\text{s.t.} \quad & \mathbf{x} \in \mathcal{X}_T, \\
& X_{11} + 2X_{12} + X_{22} + 2x_1 + 2x_2 + 1 \geq 0, \\
& X_{ii'} - 10x_i - 10x_{i'} + 100 \geq 0, & i \leq i' \in \{1, 2, 3\}, \\
& 10x_1 + 10x_2 + 10 - X_{i1} - X_{i2} - x_i \geq 0, & i \in \{1, 2, 3\}, \\
& X_{ii} \geq 0, & i \in \{1, 2, 3\}.
\end{aligned} \tag{22}$$

The numerical solution of (22) is

$$\mathbf{x}' = (1.00, 1.11, 1.11)^\top, \quad \mathbf{X}' = \begin{bmatrix} 12.98 & -48.13 & -40.60 \\ -48.13 & 78.08 & -39.55 \\ -40.60 & -39.55 & 0.00 \end{bmatrix}.$$

The relaxation objective is approximately 3.000, which constitutes a lower bound, while the corresponding feasible candidate gives an upper bound of 20.816.

Linear $\times$ Convex. Let $\mathcal{X}_{\text{TLL}}$ denote the feasible set of (22). We pairwise multiply the linear with the nonlinear constraints and obtain the SLC constraints

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

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 $X_{11}, X_{22}, X_{33}, X_{12}, X_{13}$, and X_{23} respectively. The resulting convex relaxation

is:

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

The numerical solution of (23) is

$$\mathbf{x}' = (1.17, 0.93, 0.77)^\top, \quad \mathbf{X}' = \begin{bmatrix} 1.77 & 0.75 & 0.23 \\ 0.75 & 1.16 & 1.31 \\ 0.23 & 1.31 & 0.57 \end{bmatrix}.$$

Its reported lower bound is 19.778 and its feasible candidate gives the upper bound 19.809, improving both bounds obtained from (22).

Set of candidate solutions. The four candidate vectors, computed from the relaxation solution, are

$$\begin{aligned}
\mathbf{x}' &= (1.17, 0.93, 0.77)^\top, & \mathbf{x}_1^X &= (1.51, 0.64, 0.19)^\top, \\
\mathbf{x}_2^X &= (0.81, 1.25, 1.40)^\top, & \mathbf{x}_3^X &= (0.30, 1.70, 0.75)^\top.
\end{aligned}$$

Only $\mathbf{x}'$ passes the feasibility check.

Branch and bound step. We apply the eigenvector rule of Section 3.5, with and without the semidefinite constraint

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

Without the semidefinite constraint. The root bounds are

$$(\text{Lb}_0, \text{Ub}_0) = (19.778, 19.809).$$

The selected eigenpair of $\mathbf{X} - \mathbf{x}\mathbf{x}^\top$ gives

$$\boldsymbol{\nu} = (0.62, -0.54, -0.56)^\top, \quad \lambda = 1.30.$$

Solving both child relaxations yields

$$(\text{Lb}_1, \text{Ub}_1) = (19.790, 19.790), \quad (\text{Lb}_2, \text{Ub}_2) = (19.787, 19.787).$$

Thus the global bounds after one branching iteration are

$$\text{LB} = 19.787, \quad \text{UB} = 19.787,$$

and the incumbent is $\hat{\mathbf{x}} \approx (1.19, 0.92, 0.75)^\top$, which satisfy both absolute and relative stopping criteria.

With the semidefinite constraint. The root bounds are

$$(\text{Lb}_0, \text{Ub}_0) = (19.784, 19.792).$$

The selected eigenpair of $\mathbf{X} - \mathbf{x}\mathbf{x}^\top$ gives

$$\boldsymbol{\nu} = (-0.46, 0.40, 0.79)^\top, \quad \lambda = 0.69.$$

Solving both child relaxations yields

$$(\text{Lb}_1, \text{Ub}_1) = (19.787, 19.787), \quad (\text{Lb}_2, \text{Ub}_2) = (19.787, 19.787).$$

Thus the global bounds after one branching iteration are

$$\text{LB} = 19.787, \quad \text{UB} = 19.787,$$

and the incumbent is $\hat{\mathbf{x}} \approx (1.19, 0.92, 0.75)^\top$, which satisfy both stopping criteria.

REMARK 5. The small search trees in this example are qualitatively consistent with the behavior reported for eigenvector branching in [Anstreicher \(2022\)](#), where many instances are solved with shallow trees and small final gaps.

5. Improving the RPT-BB approach

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

5.1. Multiplying with redundant linear constraints

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

DEFINITION 2 (REDUNDANT INEQUALITY CONSTRAINTS). An inequality constraint $f(\mathbf{x}) \leq 0$, where $f: \mathbb{R}^{n_x} \rightarrow (-\infty, +\infty]$, is *redundant* to a nonempty set $\mathcal{X} \subseteq \mathbb{R}^{n_x}$, if $\mathcal{X} \subseteq \{\mathbf{x} \mid f(\mathbf{x}) \leq 0\}$.

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

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

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

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

$$\begin{array}{ll} \min_{\mathbf{x}} & q - \mathbf{d}^\top \mathbf{x} \\ \text{s.t.} & \mathbf{B}^\top \mathbf{x} \leq \mathbf{p} \end{array} \quad \text{and} \quad \begin{array}{ll} \max_{\mathbf{y} \geq \mathbf{0}} & q - \mathbf{p}^\top \mathbf{y} \\ \text{s.t.} & \mathbf{B}\mathbf{y} = \mathbf{d} \end{array}$$

coincide and both are nonnegative thanks to the strong duality of linear optimization and the redundancy of $\mathbf{d}^\top \mathbf{x} \leq q$ to $\{\mathbf{x} \mid \mathbf{B}^\top \mathbf{x} \leq \mathbf{p}\}$, which implies that there exists a $\mathbf{y} \in \mathbb{R}_+^L$ such that $\mathbf{d}^\top \mathbf{x} \leq q$ is redundant to $\{\mathbf{x} \mid \mathbf{b}_y^\top \mathbf{x} \leq p_y\}$, where $\mathbf{b}_y = \mathbf{B}\mathbf{y} = \mathbf{d}$ and $p_y = \mathbf{p}^\top \mathbf{y} \leq q$. Then, for any $\mathbf{x}$ that satisfies $\mathbf{b}_y^\top \mathbf{x} \leq p_y$ and $f(\mathbf{x}) \leq 0$, we have that

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

Moreover, for any $\mathbf{X} \in \mathbb{S}^{n_x}$, since $\mathbf{b}_y = \mathbf{d}$ and $q - p_y \geq 0$, we have $(q\mathbf{x} - \mathbf{X}\mathbf{d}, q - \mathbf{d}^\top \mathbf{x}) = (p_y\mathbf{x} - \mathbf{X}\mathbf{b}_y, p_y - \mathbf{b}_y^\top \mathbf{x}) + (q - p_y)(\mathbf{x}, 1)$, so that the convexity and positive homogeneity of the perspective function of f , together with $f(\mathbf{x}) \leq 0$, yield

$$(q - \mathbf{d}^\top \mathbf{x})f\left(\frac{q\mathbf{x} - \mathbf{X}\mathbf{d}}{q - \mathbf{d}^\top \mathbf{x}}\right) \leq (p_y - \mathbf{b}_y^\top \mathbf{x})f\left(\frac{p_y\mathbf{x} - \mathbf{X}\mathbf{b}_y}{p_y - \mathbf{b}_y^\top \mathbf{x}}\right) + (q - p_y)f(\mathbf{x}) \leq (p_y - \mathbf{b}_y^\top \mathbf{x})f\left(\frac{p_y\mathbf{x} - \mathbf{X}\mathbf{b}_y}{p_y - \mathbf{b}_y^\top \mathbf{x}}\right).$$

Notice that

$$\begin{aligned} (p_\ell - \mathbf{b}_\ell^\top \mathbf{x})f\left(\frac{p_\ell \mathbf{x} - \mathbf{X}\mathbf{b}_\ell}{p_\ell - \mathbf{b}_\ell^\top \mathbf{x}}\right) \leq 0, \ell \in \mathcal{L} & \implies \sum_{\ell \in \mathcal{L}} \theta_\ell (p_\ell - \mathbf{b}_\ell^\top \mathbf{x})f\left(\frac{p_\ell \mathbf{x} - \mathbf{X}\mathbf{b}_\ell}{p_\ell - \mathbf{b}_\ell^\top \mathbf{x}}\right) \leq 0 \\ & \implies \left(\sum_{\ell \in \mathcal{L}} \theta_\ell (p_\ell - \mathbf{b}_\ell^\top \mathbf{x})\right) f\left(\frac{\sum_{\ell \in \mathcal{L}} \theta_\ell (p_\ell \mathbf{x} - \mathbf{X}\mathbf{b}_\ell)}{\sum_{\ell \in \mathcal{L}} \theta_\ell (p_\ell - \mathbf{b}_\ell^\top \mathbf{x})}\right) \leq 0 \\ & \implies (p_y - \mathbf{b}_y^\top \mathbf{x})f\left(\frac{p_y \mathbf{x} - \mathbf{X}\mathbf{b}_y}{p_y - \mathbf{b}_y^\top \mathbf{x}}\right) \leq 0, \end{aligned}$$

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

$$\left\{ (\mathbf{x}, \mathbf{X}) \mid \mathbf{b}_y^\top \mathbf{x} \leq p_y, f(\mathbf{x}) \leq 0, (p_\ell - \mathbf{b}_\ell^\top \mathbf{x})f\left(\frac{p_\ell \mathbf{x} - \mathbf{X}\mathbf{b}_\ell}{p_\ell - \mathbf{b}_\ell^\top \mathbf{x}}\right) \leq 0, \ell \in \mathcal{L} \right\}.$$

Finally, since $\mathbf{y} \geq \mathbf{0}$, the second family of constraints in the set gives $0 \leq \mathbf{y}^\top (\mathbf{B}^\top \mathbf{X} \mathbf{B} + \mathbf{p} \mathbf{p}^\top - \mathbf{p} \mathbf{x}^\top \mathbf{B} - \mathbf{B}^\top \mathbf{x} \mathbf{p}^\top) \mathbf{y} = \mathbf{d}^\top \mathbf{X} \mathbf{d} + p_y^2 - 2p_y \mathbf{d}^\top \mathbf{x}$, and hence

$$\mathbf{d}^\top \mathbf{X} \mathbf{d} + q^2 - 2q \mathbf{d}^\top \mathbf{x} = \mathbf{d}^\top \mathbf{X} \mathbf{d} + p_y^2 - 2p_y \mathbf{d}^\top \mathbf{x} + (q - p_y) [(q - \mathbf{d}^\top \mathbf{x}) + (p_y - \mathbf{d}^\top \mathbf{x})] \geq 0,$$

because $q - p_y \geq 0$, $q - \mathbf{d}^\top \mathbf{x} \geq 0$ and $p_y - \mathbf{d}^\top \mathbf{x} = \mathbf{y}^\top (\mathbf{p} - \mathbf{B}^\top \mathbf{x}) \geq 0$. This is the constraint $2q \mathbf{d}^\top \mathbf{x} \leq \mathbf{d}^\top \mathbf{X} \mathbf{d} + q^2$. Thus, the claim follows. $\square$

REMARK 6. Theorem 2 parallels a standard observation in the RLT literature: adding a linear inequality that is redundant for the original linear description does not strengthen the RLT relaxation; see, e.g., Proposition 8.2 in [Sherali and Adams \(1999\)](#), and Proposition 3(i) in [Hof and Walter \(2025\)](#) for a recent treatment.

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

EXAMPLE 5. Consider the following nonconvex problem:

$$\begin{aligned} \min_{x_1, x_2} \quad & 2x_1 + 3x_2 - 5x_1x_2 - (x_1 + 2) \ln(x_1 + 2) \\ \text{s.t.} \quad & x_1 + x_2 \leq 1, \\ & \exp(-x_1) + \exp(-x_2) \leq 1 + \exp(-1). \end{aligned} \tag{24}$$

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, namely, $-5x_1x_2$ and $-(x_1 + 2) \ln(x_1 + 2)$. The obtained convex relaxation of Problem (24) from RPT without the optional SDP relaxation has an objective value of -35.17 . The obtained optimal solution $\mathbf{x}' = (1, 0)^\top$ is a feasible solution to (24), and its corresponding objective value is -1.30 , which constitutes an upper bound on the optimal value of (24).

The linear constraints $x_1 \geq -1$ and $x_2 \geq -1$ are redundant to the second inequality. However, adding those constraints to (24) and subsequently applying RPT results in a convex relaxation of Problem (24) with objective value of -4.47 . Again, the obtained optimal solution $\mathbf{x}' = (0.5, 0.5)^\top$ is a feasible solution to (24), and its corresponding objective value is -1.04 , which constitutes an upper bound on the optimal value of (24). Hence, by adding the redundant linear constraints $x_1 \geq -1$ and $x_2 \geq -1$, we obtain a tighter lower bound on the optimal objective value of (24). $\square$

REMARK 7. Example 5 illustrates the strengthening effect discussed in Section 3.2: even linear inequalities that are redundant for the original feasible set can tighten the lifted RPT relaxation, because they generate additional valid product constraints when combined with other (nonlinear) constraints. Related effects are exploited in the SOC-RLT literature for QCQP-type models; see, e.g., [Burer and Anstreicher \(2013\)](#).

5.2. Handling biconvex problems with convex constraints

We consider the following generic instance of (1):

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{y}} \quad & c_0(\mathbf{x}) + \sum_k r_k(\mathbf{y}) c_k(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{X}, \mathbf{y} \in \mathcal{Y}, \end{aligned} \quad (25)$$

where the functions $r_k(\cdot)$ and $c_k(\cdot)$ are assumed to be convex, and $r_k(\mathbf{y}), c_k(\mathbf{x}) \geq 0$. We show that the constraints in the RPT relaxation of (25) resulting from pairwise multiplication in the same set, i.e., either in $\mathcal{X}$ or in $\mathcal{Y}$, are redundant as long as the LMI as given in Section 3.3 is not included.

LEMMA 4 (Redundant Pairwise Multiplications). *The additional constraints in the RPT relaxation of (25) resulting from pairwise multiplication in $\mathcal{X} \times \mathcal{X}$ and $\mathcal{Y} \times \mathcal{Y}$ are redundant, if the additional constraints resulting from pairwise multiplication in $\mathcal{X} \times \mathcal{Y}$ are included, and the following LMI*

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

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

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

Observe that the problem of minimizing a DC function is a special case of (25) where $\mathcal{Y} = \text{dom}(c_1^*)$ and both r_k and c_k are linear, as a result of the biconjugate reformulation as explained in Example 1. Finally, we have the following remark in terms of branching, when we solve Problem (25) with Algorithm 2.

REMARK 8 (LMI IN THE $(\mathbf{x}, \mathbf{y})$ -SPACE). When we apply Algorithm 2 to minimization problems of the form (25) in which we include the LMI, we typically generate hyperplanes by computing the eigenvector corresponding to the largest eigenvalue of $\mathbf{X} - \mathbf{x}\mathbf{x}^\top$, i.e., we only generate hyperplanes in the $\mathbf{x}$ -space. Alternatively, hyperplanes can be generated in the $(\mathbf{x}, \mathbf{y})$ -space by considering the eigenvectors of the matrix

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

where $(\mathbf{x}, \mathbf{y})$ denotes the vector formed by stacking $\mathbf{x}$ and $\mathbf{y}$. We observe that also generating hyperplanes in the $(\mathbf{x}, \mathbf{y})$ -space adds many more constraints in each branch and bound iteration, which can increase the computation time at each successive child node. On the other hand, generating hyperplanes in the $(\mathbf{x}, \mathbf{y})$ -space might reduce the total number of hyperplanes that need to be generated, potentially decreasing the overall computation time. The question of whether generating hyperplanes in the $(\mathbf{x}, \mathbf{y})$ -space could be beneficial is left for future research. Finally, we note that when we apply Algorithm 2 to minimization problems of the form (25) in which we do not include the LMI, one can branch based on left and right directions associated with the matrix $\mathbf{R} = \mathbf{U} - \mathbf{x}\mathbf{y}^\top$. Concretely, let $\mathbf{v}^r$ be a unit eigenvector corresponding to the largest eigenvalue of $\mathbf{R}^\top \mathbf{R}$, i.e.,

$$\mathbf{R}^\top \mathbf{R} \mathbf{v}^r = \lambda_{\max}(\mathbf{R}^\top \mathbf{R}) \mathbf{v}^r.$$

If $\mathbf{R}\mathbf{v}^r \neq \mathbf{0}$, define $\mathbf{v}^\ell = \mathbf{R}\mathbf{v}^r / \|\mathbf{R}\mathbf{v}^r\|$; then $\mathbf{v}^\ell$ is a unit eigenvector corresponding to the largest eigenvalue of $\mathbf{R}\mathbf{R}^\top$, i.e.,

$$\mathbf{R}\mathbf{R}^\top \mathbf{v}^\ell = \lambda_{\max}(\mathbf{R}\mathbf{R}^\top) \mathbf{v}^\ell.$$

□

5.3. Strengthening upper bounds with eigenvectors of the optimal solution

At an exact lift, $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$. If $\mathbf{x} \neq \mathbf{0}$, multiplying both sides by $\mathbf{x}$ gives $\mathbf{X}\mathbf{x} = (\mathbf{x}^\top \mathbf{x})\mathbf{x}$. Therefore, we notice that $\mathbf{x}$ is an eigenvector of $\mathbf{X}$ with corresponding eigenvalue $\mathbf{x}^\top \mathbf{x}$. Hence, we can add the eigenvectors of $\mathbf{X}$ to the candidate set $\mathcal{X}'$, as described in Section 3.4. As for every element of $\mathcal{X}'$, an eigenvector-derived candidate is used to obtain an upper bound only after direct feasibility and finite-objective checks. Example 6 illustrates a case where the tightest upper bound can be obtained from an eigenvector of $\mathbf{X}$.

EXAMPLE 6. Consider the following toy problem:

$$\begin{aligned} \min_{\mathbf{x}} \quad & (x_1 + x_2 + 1) \exp(x_2) \\ \text{s.t.} \quad & x_1 + x_2 + 1 \geq 0, \\ & x_1 x_2 \geq -1. \end{aligned} \tag{26}$$

An optimal solution of (26) is $(-1, 0)$ with optimal value 0. After applying RPT we obtain the following relaxation

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{X}} \quad & (x_1 + x_2 + 1) \exp\left(\frac{X_{12} + X_{22} + x_2}{x_1 + x_2 + 1}\right) \\ \text{s.t.} \quad & x_1 + x_2 + 1 \geq 0, \\ & X_{12} \geq -1, \\ & X_{11} + X_{22} + 2X_{12} + 2x_1 + 2x_2 + 1 \geq 0. \end{aligned} \tag{27}$$

The optimal solution of (27) is

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

with optimal value 0, which gives us a lower bound for the optimal value of (26). We have the following candidate vectors

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

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

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

We observe that $\mathbf{x}^*$ is infeasible as $x_1^*x_2^* = -1.11 < -1$. Moreover, $\mathbf{x}_1^X$ is infeasible as $(x_1^X)_1(x_1^X)_2 = -5.44 < -1$ and $\mathbf{x}_3^{EV}$ is infeasible as $(x_3^{EV})_1 + (x_3^{EV})_2 + 1 = -0.26 < 0$. Finally, we notice that $\mathbf{x}_2^X$ is feasible and gives an upper bound of 1.64, while $\mathbf{x}_4^{EV}$ 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 $\mathbf{X}^*$. $\square$

6. Convergence analysis of the RPT-BB approach

In this section we show that Algorithm 2 converges for Problem (1). Theorem 3 shows that the solution returned by the algorithm is δ -optimal whenever the algorithm terminates, and that the global lower bounds converge to the optimal value, up to the accuracy with which the node relaxations are solved, whenever the algorithm does not terminate. Corollary 1 shows that the algorithm terminates after finitely many iterations under additional conditions. The proofs are given in Appendix F.

Throughout this section we assume that Assumptions 1 and 2 hold, that Problem (1) is feasible, and that its optimal value v^* is finite. Here $\mathcal{X}$ and $\mathbf{x}$ are those of the SLC formulation that is actually solved, including any auxiliary variables introduced by a reformulation.

We use the following notation. For a node N_j , $\mathcal{F}_j$ is the feasible set of its RPT relaxation. Solving this relaxation returns a point $\sigma_j = (\mathbf{x}_j, \boldsymbol{\tau}_j, \mathbf{X}_j, \mathbf{V}_j) \in \mathcal{F}_j$ with relaxation objective value p_j , together with a lower bound ℓ_j on the relaxation objective value of every point of $\mathcal{F}_j$. Algorithm 1 branches on σ_j . The lower bound Lb_j of the node in Algorithm 2 is ℓ_j or, if larger, the bound inherited from the parent. We write $\mathbf{R}_j = \mathbf{X}_j - \mathbf{x}_j\mathbf{x}_j^\top$ and $\rho_j = \|\mathbf{R}_j\|_2$, and we denote the diameter of $\mathcal{X}$ by $D_x := \max_{\mathbf{x}, \mathbf{x}' \in \mathcal{X}} \|\mathbf{x} - \mathbf{x}'\|_2$. As in Algorithm 2, Lb and Ub are the global lower and upper bounds, and $\mathbf{x}^{\text{inc}}$ is the incumbent, that is, the solution $\mathbf{x}^*$ that Algorithm 2 returns. Finally, an infinite nested branch is an infinite sequence of nodes in which every node is a child of the previous one.

ASSUMPTION 3 (Accuracy of the Node Relaxation Solutions). *There is a tolerance $\bar{\varepsilon} \geq 0$ such that, for every node N_j generated by Algorithm 2, solving the relaxation of N_j in finite time either correctly reports that $\mathcal{F}_j = \emptyset$, or returns a point $\sigma_j \in \mathcal{F}_j$ with relaxation objective value p_j and a lower bound $\ell_j \in \mathbb{R}$ on the relaxation objective value of every point of $\mathcal{F}_j$ such that*

$$\ell_j \leq p_j \leq \ell_j + \bar{\varepsilon}. \quad (28)$$

THEOREM 3 (Convergence and Optimality Certificates). *Let Assumptions 1–3 hold, and let Algorithm 2 be run with any $\delta \geq 0$. Then the following statements hold.*

(a) *If the algorithm terminates, then it returns a feasible incumbent $\mathbf{x}^{\text{inc}}$ with*

$$0 \leq f_0(\mathbf{x}^{\text{inc}}) - v^* \leq \text{Ub} - \text{Lb} \leq \delta.$$

In particular, the incumbent is globally optimal when $\delta = 0$.

(b) *If the algorithm does not terminate, then the global lower bounds Lb converge monotonically to a limit L with $v^* - \bar{\varepsilon} \leq L \leq v^*$. Moreover, every accumulation point $\bar{\mathbf{x}}$ of the relaxation solutions $\mathbf{x}_j$ along an infinite nested branch is feasible for Problem (1) and satisfies $f_0(\bar{\mathbf{x}}) \leq L + \bar{\varepsilon} \leq v^* + \bar{\varepsilon}$.*

REMARK 9 (UPPER BOUNDS AND FINITE TERMINATION). Theorem 3 is a result on the lower bounds. It does not state that the upper bounds converge to v^* , and hence it does not guarantee that Algorithm 2 terminates when $\delta > 0$. The proof of Corollary 1 below shows what is needed for termination. Consider an infinite nested branch $\{N_r\}$ and a subsequence of relaxation solutions $\mathbf{x}_{r_\ell} \rightarrow \bar{\mathbf{x}}$. Suppose that the upper-bounding procedure of Section 3.4, applied to the nodes N_{r_ℓ} , produces feasible solutions $\hat{\mathbf{x}}_{r_\ell}$ with $\liminf_\ell f_0(\hat{\mathbf{x}}_{r_\ell}) \leq f_0(\bar{\mathbf{x}})$. Then $\text{Ub} - \text{Lb} \leq f_0(\hat{\mathbf{x}}_{r_\ell}) - \text{Lb}_{r_\ell}$ when N_{r_ℓ} is selected, and the limit inferior of the right-hand side is at most $f_0(\bar{\mathbf{x}}) - L \leq \bar{\varepsilon}$. For $\delta > \bar{\varepsilon}$ this contradicts the stopping criterion, so the algorithm terminates. If $\mathcal{K} = \emptyset$ and f_0 is continuous, then $\hat{\mathbf{x}}_r = \mathbf{x}_r$ has this property. If $\mathcal{K} \neq \emptyset$, then the solutions $\mathbf{x}_r$ may violate the constraints $f_k(\mathbf{x}) \leq 0$ at every node of the branch, even though their limit points satisfy these constraints. In that case, the candidate solutions and local searches of Section 3.4 need not produce feasible solutions whose objective values approach $f_0(\bar{\mathbf{x}})$.

COROLLARY 1 (Finite δ -Optimal Termination When $\mathcal{K} = \emptyset$). *Let Assumptions 1–3 hold with $\bar{\varepsilon} < \delta$. Suppose that $\mathcal{K} = \emptyset$ and that f_0 is finite and continuous on $\mathcal{X}$. Then Algorithm 2 terminates after finitely many iterations and returns a feasible incumbent $\mathbf{x}^{\text{inc}}$ satisfying $0 \leq f_0(\mathbf{x}^{\text{inc}}) - v^* \leq \delta$.*

7. Known convex reformulations and relaxations obtained via RPT

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

7.1. Disjunctive optimization

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

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

where $\mathbf{h}_k(\mathbf{x}) = [h_{1k}(\mathbf{x}) \ h_{2k}(\mathbf{x}) \ \cdots \ h_{Jk}(\mathbf{x})]^\top$, and $h_{jk} : \mathbb{R}^{n_x} \rightarrow (-\infty, +\infty]$ 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 $\lambda_k \mathbf{x}_k$, $k \in \mathcal{K}$. One can apply RPT to obtain the following convex relaxation

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

This convex relaxation is exact according to [Gorissen et al. \(2014, Lemma 1\)](#), which applies because $\mathcal{X}_k$, $k \in \mathcal{K}$, are nonempty, compact and convex sets. We now use this observation to derive convex relaxation for disjunctive optimization problems with general convex sets. In [Sherali and Adams \(2005, Section 4\)](#), the authors derive similar result for disjunctive optimization problems with a linear objective function and polyhedral sets $\mathcal{X}_k$, $k \in \mathcal{K}$. Consider a generic disjunctive optimization problem

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

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

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

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

7.2. Generalized linear optimization

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

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

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

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

The convex problem is in general a convex relaxation of (GLP), which has the same optimal value as (GLP) if one of the following regularity conditions is satisfied: (i) $\mathcal{X}_k$ is bounded for each $k \in \mathcal{K}_0$ (Gorissen et al., 2014, Lemma 1); (ii) there exists a $(\mathbf{y}, \{\mathbf{x}_k\}_k)$ with $\mathbf{y} > \mathbf{0}$ that is feasible for (GLP) (Zhen et al., 2023, Lemma 6). While for a special case where $\mathcal{X}_k$, $k \in \mathcal{K}$, are (nonempty) boxes, the corresponding linear relaxation of (GLP) is exact due to Dantzig (1963).

7.3. Convex hull representation for 0-1 mixed-integer convex programs

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

$$\mathcal{X} = \left\{ \mathbf{x} \mid \begin{array}{ll} h_k(\mathbf{x}) \leq 0, & k \in \mathcal{K}, \\ \mathbf{x}_B \in \{0, 1\}^{n_B}, & \mathbf{x}_C \in [0, 1]^{n_C} \end{array} \right\},$$

where the vector $\mathbf{x} \in \mathbb{R}^{n_x}$ is the concatenation of the binary vector $\mathbf{x}_B$ and the continuous vector $\mathbf{x}_C$, namely, $\mathbf{x} = (\mathbf{x}_B, \mathbf{x}_C)$ with $n_x = n_B + n_C$. Assume $\mathcal{X} \neq \emptyset$. Similarly to Sherali and Adams (2009), for each subset $\mathcal{J} \subseteq \mathcal{B}$, where $\mathcal{B} = \{1, \dots, n_B\}$, we define the polynomial factors

$$F_{\mathcal{J}}(\mathbf{x}_B) = \prod_{j \in \mathcal{J}} (\mathbf{x}_B)_j \prod_{j \in \mathcal{B} \setminus \mathcal{J}} (1 - (\mathbf{x}_B)_j).$$

Observe that these factors satisfy the identity

$$\sum_{\mathcal{J} \subseteq \mathcal{B}} F_{\mathcal{J}}(\mathbf{x}_B) = \prod_{j \in \mathcal{B}} ((\mathbf{x}_B)_j + (1 - (\mathbf{x}_B)_j)) = 1.$$

We linearize each factor $F_{\mathcal{J}}(\mathbf{x}_B)$ by introducing a variable $\lambda_{\mathcal{J}} \in \mathbb{R}_+$. We then impose the simplex constraints

$$\lambda_{\mathcal{J}} \geq 0 \quad \forall \mathcal{J} \subseteq \mathcal{B}, \quad \sum_{\mathcal{J} \subseteq \mathcal{B}} \lambda_{\mathcal{J}} = 1, \quad (29)$$

and

$$(\mathbf{x}_B)_i = \sum_{\mathcal{J} \subseteq \mathcal{B}: i \in \mathcal{J}} \lambda_{\mathcal{J}} \quad \forall i \in \mathcal{B}. \quad (30)$$

Note that, since $\mathbf{x}_B \in \{0, 1\}^{n_B}$, constraints (29) and (30) imply that

$$\lambda_{\mathcal{J}} = \begin{cases} 1, & \text{if } \mathcal{J} = \{i \in \mathcal{B} : (\mathbf{x}_B)_i = 1\}, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, $\lambda_{\mathcal{J}} = F_{\mathcal{J}}(\mathbf{x}_B)$ for all $\mathcal{J} \subseteq \mathcal{B}$.

Next, for each subset $\mathcal{J} \subseteq \mathcal{B}$, we introduce a vector $\mathbf{y}_{\mathcal{J}} \in \mathbb{R}^{n_x}$ to represent the product $\mathbf{x}\lambda_{\mathcal{J}}$ componentwise. For the binary block, this yields the exact identities

$$((\mathbf{y}_{\mathcal{J}})_B)_i = \begin{cases} \lambda_{\mathcal{J}}, & \text{if } i \in \mathcal{J}, \\ 0, & \text{otherwise,} \end{cases} \quad \forall i \in \mathcal{B}, \forall \mathcal{J} \subseteq \mathcal{B}. \quad (31)$$

For the continuous block, we linearize the product $\mathbf{x}_C\lambda_{\mathcal{J}}$ by $(\mathbf{y}_{\mathcal{J}})_C$, and use the bounds $\mathbf{0} \leq \mathbf{x}_C \leq \mathbf{1}$ to impose

$$\mathbf{0} \leq (\mathbf{y}_{\mathcal{J}})_C \leq \mathbf{1}\lambda_{\mathcal{J}} \quad \forall \mathcal{J} \subseteq \mathcal{B}, \quad (32)$$

where $\mathbf{0}, \mathbf{1} \in \mathbb{R}^{n_C}$ denote the all-zeros and all-ones vectors, respectively. Finally, since $\sum_{\mathcal{J} \subseteq \mathcal{B}} \lambda_{\mathcal{J}} = 1$ and $\mathbf{y}_{\mathcal{J}} = \mathbf{x}\lambda_{\mathcal{J}}$, we obtain

$$\mathbf{x} = \sum_{\mathcal{J} \subseteq \mathcal{B}} \mathbf{y}_{\mathcal{J}}. \quad (33)$$

Moreover, (31) and (33) imply (30), so the latter need not be imposed again.

Using these variables, and convexifying each bilinear constraint $\lambda_{\mathcal{J}}h_k(\mathbf{x}) \leq 0$ via the perspective of h_k , we obtain

$$\mathcal{X}_{\text{conv}} = \left\{ \mathbf{x} \mid \exists \lambda, \mathbf{y}_{\mathcal{J}} : \begin{array}{l} \lambda_{\mathcal{J}}h_k\left(\frac{\mathbf{y}_{\mathcal{J}}}{\lambda_{\mathcal{J}}}\right) \leq 0, \quad k \in \mathcal{K}, \mathcal{J} \subseteq \mathcal{B} \\ (29), (31), (32), (33) \end{array} \right\}. \quad (34)$$

The formulation (34) is convex and, as shown in Sherali and Adams (2009), it equals the convex hull of $\mathcal{X}$, that is,

$$\mathcal{X}_{\text{conv}} = \text{conv}(\mathcal{X}).$$

This exact description enumerates all 2^{n_B} subsets $\mathcal{J} \subseteq \mathcal{B}$, that is, all binary assignments.

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

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

$$\begin{array}{ll} \min_{\mathbf{x}} & \mathbf{x}^\top \mathbf{A}_0 \mathbf{x} + 2\mathbf{b}_0^\top \mathbf{x} + c_0 \\ \text{s.t.} & \mathbf{x}^\top \mathbf{A}_1 \mathbf{x} + 2\mathbf{b}_1^\top \mathbf{x} + c_1 \leq 0, \end{array} \quad (\text{QCQP})$$

where $\mathbf{A}_k \in \mathbb{R}^{n_x \times n_x}$, $\mathbf{b}_k \in \mathbb{R}^{n_x}$ 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 $\mathbf{x} \in \mathbb{R}^{n_x}$ with $\mathbf{x}^\top \mathbf{A}_1 \mathbf{x} + 2\mathbf{b}_1^\top \mathbf{x} + c_1 < 0$, then we have

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

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

$$\begin{aligned} \min_{\mathbf{X}, \mathbf{x}} \quad & \text{Tr}(\mathbf{A}_0 \mathbf{X}) + 2\mathbf{b}_0^\top \mathbf{x} + c_0 \\ \text{s.t.} \quad & \text{Tr}(\mathbf{A}_1 \mathbf{X}) + 2\mathbf{b}_1^\top \mathbf{x} + c_1 \leq 0, \\ & \begin{bmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{x}^\top & 1 \end{bmatrix} \succeq 0, \end{aligned}$$

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

$$\begin{aligned} \min_{\mathbf{x}} \quad & \mathbf{x}^\top \mathbf{A}_0 \mathbf{x} + 2\mathbf{b}_0^\top \mathbf{x} + c_0 \\ \text{s.t.} \quad & \mathbf{x}^\top \mathbf{A}_k \mathbf{x} + 2\mathbf{b}_k^\top \mathbf{x} + c_k \leq 0 \quad k \in \mathcal{K}, \end{aligned}$$

where $\mathbf{A}_k \in \mathbb{R}^{n_x \times n_x}$, $\mathbf{b}_k \in \mathbb{R}^{n_x}$ 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:

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

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

7.5. Fractional optimization

Consider the following generic fractional optimization problem

$$\begin{aligned} \min_{\mathbf{x}} \quad & \frac{f(\mathbf{x})}{g(\mathbf{x})} \\ \text{s.t.} \quad & h_k(\mathbf{x}) \leq 0 \quad k \in \mathcal{K}, \end{aligned} \tag{FP}$$

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

$$\begin{aligned} \min_{\mathbf{v}, \tau} \quad & \tau f(\mathbf{v}/\tau) \\ \text{s.t.} \quad & \tau g(\mathbf{v}/\tau) \geq 1, \\ & \tau h_k(\mathbf{v}/\tau) \leq 0, \quad k \in \mathcal{K}. \end{aligned}$$

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

8. Numerical experiments

We evaluate RPT-BB and RPT-SDP-BB on five classes of nonconvex optimization problems: sum-of-max-linear-terms maximization, log-sum-exp maximization, linear multiplicative optimization, quadratically constrained quadratic optimization, and dike height optimization. The first three use the biconjugate reformulation, whereas the latter two are already in the generic form (1).

All numerical experiments in this paper were run on a MacBook Pro with an Apple M3 Pro chip (11 CPU cores: 5 performance and 6 efficiency) and 18 GB RAM. The RPT methods are implemented in Julia 1.12.5 (JuliaLang contributors, 2026) with JuMP.jl 1.29.4 (JuMP developers, 2026). Gurobi 13.0.0 (Gurobi Optimization, LLC, 2026) solves the linear optimization problems and MOSEK 11.0.30 (ApS, 2025) the conic node relaxations. Ipopt is used for the local searches of Section 3.4 in the quadratically constrained quadratic and dike height optimization problems. Both RPT methods in the main tables use the branching rule of Algorithm 1, with an absolute or relative optimality tolerance of 10^{-4} and a time limit of 1800 seconds per problem instance. We compare the proposed approach with BARON 21.1.13 (Sahinidis, 1996; The Optimization Firm, LLC, 2021), called through Pyomo 6.7.2 (Bynum et al., 2021; Hart et al., 2011), with SCIP 10.0.0 (Hojny et al., 2025), called through JuMP.jl, and with CPLEX 22.1.2 (ILOG, Inc., 2017), called through docplex 2.32.264. All solvers are run with one thread, with the same optimality tolerance and time limit as the RPT methods, with a constraint feasibility tolerance of 10^{-6} , and with one solver call per problem instance. We distinguish direct and biconjugate formulations by the suffixes Dir and Bic, where applicable.

The tables report the best feasible objective value returned (Opt), the number of branching hyperplanes (Hyp), the computation time in seconds (Time), and the number of instances that reach the time limit (TO). For the randomly generated groups, Opt, Hyp and Time are averages over the ten instances of the group, and TO is reported out of ten. The time of an instance that reaches the time limit is recorded as 1800 seconds. A solver may also stop before the time limit without proving optimality, so a feasible objective value or a zero timeout count does not by itself establish optimality. We retain feasible objective values even when optimality is not established. Incumbent candidates of all methods are checked against the original constraints with a feasibility tolerance of 10^{-6} . A dash indicates that no feasible solution was returned, and the group mean is withheld if any instance of the group returns none. Appendix A compares branching rules within RPT-BB and RPT-SDP-BB.

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

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

$$\max_{\mathbf{x} \in \mathcal{X}} \sum_{\ell \in \mathcal{L}} \max_{j \in \mathcal{J}_\ell} \{\mathbf{A}_j^\top \mathbf{x} + b_j\}, \quad (35)$$

where $\mathbf{A} \in \mathbb{R}^{n_x \times n_y}$, $\mathbf{b} \in \mathbb{R}^{n_y}$. We consider three cases of $\mathcal{X}$: 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

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

Here $\mathbf{D} \in \mathbb{R}^{n_x \times m}$ and $\mathbf{d} \in \mathbb{R}^m$. Since the objective of (35) is a closed convex function, we can replace it by its biconjugate function and obtain the following equivalent maximization problem:

$$\max_{\substack{\mathbf{x} \in \mathcal{X} \\ \mathbf{y} \in \mathcal{Y}}} (\mathbf{A}^\top \mathbf{x} + \mathbf{b})^\top \mathbf{y}, \quad (36)$$

where $\mathcal{Y}$ equals the domain of the conjugate function of the objective of (35), i.e.,

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

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

$$\begin{aligned} \max_{\substack{\mathbf{x} \in \mathcal{X} \\ \boldsymbol{\lambda}, \mathbf{z}}} & \sum_{\ell \in \mathcal{L}} \lambda_\ell \\ \text{s.t.} & \lambda_\ell \leq \mathbf{A}_j^\top \mathbf{x} + b_j + M(1 - z_j), \quad j \in \mathcal{J}_\ell, \ell \in \mathcal{L}, \\ & \sum_{j \in \mathcal{J}_\ell} z_j = 1, \quad \ell \in \mathcal{L}, \\ & \mathbf{z} \in \{0, 1\}^{n_y}. \end{aligned} \quad (37)$$

We solve Problem (37) with Gurobi for $\mathcal{X}_1$ and MOSEK for $\mathcal{X}_2$. Instances labeled 1–5 correspond to single instances, taken from [Selvi et al. \(2022\)](#), whereas instances labeled 1r–5r report averages over 10 independently generated random instances, generated according to the exact same mechanism. We refer to Appendices D.1 and E.1 for the convex RPT relaxation and problem instances, respectively.

$\mathcal{X}$	#	RPT-BB				RPT-SDP-BB				BARON			MIP		
		Opt	Time	Hyp	TO	Opt	Time	Hyp	TO	Opt	Time	TO	Opt	Time	TO
$\mathcal{X}_1$	1	23.29	0.01	0.0	0/1	23.29	0.01	0.0	0/1	23.29	0.11	0/1	23.29	0.00	0/1
	1r	22.72	0.01	0.0	0/10	22.72	0.01	0.0	0/10	22.72	0.10	0/10	22.72	0.00	0/10
	2	233.94	0.11	0.0	0/1	233.94	0.45	0.0	0/1	233.94	0.31	0/1	233.94	0.02	0/1
	2r	211.87	0.19	0.2	0/10	211.87	0.44	0.0	0/10	211.87	0.74	0/10	211.87	0.02	0/10
	3	1081.62	126.19	78.0	0/1	1081.62	26.50	0.0	0/1	1081.62	1800.00	1/1	1081.62	1.18	0/1
	3r	1159.90	56.70	34.4	0/10	1159.90	59.83	0.3	0/10	1159.90	1800.00	10/10	1159.90	1.06	0/10
	4	113.70	0.03	0.0	0/1	113.70	0.05	0.0	0/1	113.70	2.29	0/1	113.70	0.01	0/1
	4r	83.78	0.16	0.5	0/10	83.78	0.10	0.0	0/10	83.78	2.49	0/10	83.78	0.00	0/10
	5	3002.44	2.63	1.0	0/1	3002.44	118.96	1.0	0/1	3002.44	1800.00	1/1	3002.44	327.52	0/1
	5r	2898.05	9.46	6.0	0/10	2898.05	141.77	1.3	0/10	2898.05	1800.00	10/10	2898.05	1073.88	4/10
$\mathcal{X}_2$	1	14.58	0.02	0.0	0/1	14.58	0.03	0.0	0/1	14.58	0.18	0/1	14.58	0.05	0/1
	1r	14.54	0.02	0.0	0/10	14.54	0.03	0.0	0/10	14.54	0.13	0/10	14.54	0.06	0/10
	2	136.22	0.35	0.0	0/1	136.22	0.96	0.0	0/1	136.22	1800.00	1/1	136.22	1115.66	0/1
	2r	122.21	2.18	3.3	0/10	122.21	3.57	0.7	0/10	122.21	1800.00	10/10	122.21	468.78	0/10
	3	837.94	559.62	76.0	0/1	837.94	92.00	0.0	0/1	837.94	1800.00	1/1	790.75	1800.00	1/1
	3r	890.07	932.13	77.0	4/10	890.07	312.34	0.6	0/10	890.07	1800.00	10/10	843.66	1800.00	10/10
	4	33.73	4.06	15.0	0/1	33.73	3.86	2.0	0/1	33.73	45.78	0/1	33.73	0.13	0/1
	4r	31.81	1.94	4.9	0/10	31.81	1.47	0.8	0/10	31.81	19.76	0/10	31.81	0.11	0/10
	5	1610.69	1800.00	210.0	1/1	1610.69	1800.00	5.0	1/1	1610.69	1800.00	1/1	1588.59	1800.00	1/1
	5r	1670.92	1800.00	140.4	10/10	1670.92	1800.00	5.5	10/10	1671.02	1800.00	10/10	1657.09	1800.00	10/10
$\mathcal{X}_3$	1	13.45	215.12	642.0	0/1	13.41	11.76	19.0	0/1	13.45	0.22	0/1			
	1r	14.99	369.08	432.2	1/10	14.99	11.77	13.1	0/10	15.02	0.14	0/10			
	2	140.87	112.86	179.0	0/1	140.62	41.25	4.0	0/1	140.89	5.59	0/1			
	2r	129.29	148.51	213.9	0/10	129.10	41.29	4.1	0/10	129.31	238.47	0/10			
	3	768.96	801.40	99.0	0/1	768.94	1800.00	2.0	1/1	768.97	1800.00	1/1			
	3r	795.18	1404.88	157.7	7/10	793.83	1800.00	2.0	10/10	805.60	1800.00	10/10			
	4	45.34	1800.00	216.0	1/1	45.34	232.76	51.0	0/1	45.34	74.31	0/1			
	4r	44.79	1492.21	443.6	6/10	44.89	575.90	58.0	1/10	44.97	40.45	0/10			
	5	2267.99	1800.00	195.0	1/1	-	1800.00	2.0	1/1	2125.58	1800.00	1/1			
	5r	2342.89	1800.00	61.6	10/10	-	1800.00	2.1	10/10	2185.23	1800.00	10/10			

Table 4 Results for Problem (36) over the feasible regions $\mathcal{X}_1, \mathcal{X}_2$, and $\mathcal{X}_3$.

From Table 4, we observe that the RPT methods are competitive with MIP over $\mathcal{X}_1$ and $\mathcal{X}_2$, and substantially faster on several instances for which BARON reaches the time limit. The LMI strengthening can reduce computation time when it substantially reduces branching, as on instances 3 and 3r over $\mathcal{X}_2$, but can increase computation time when RPT-BB already needs few hyperplanes. Over $\mathcal{X}_3$, no method is uniformly best. RPT-BB returns the highest objective values on instances 5 and 5r, although all methods reach the time limit, whereas RPT-SDP-BB returns no feasible solution. The lower objective values returned by RPT-SDP-BB on some other instances also indicate numerical difficulties. Thus the LMI strengthening is useful on some instances, but is not consistently advantageous.

8.2. Log-sum-exp maximization over linear constraints

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

$$\max_{\mathbf{x} \in \mathcal{X}_1} \log \left(\sum_{i=1}^{n_x} \exp(x_i) \right). \quad (38)$$

Since the objective of (38) is a closed convex function, we can replace it by its biconjugate function and obtain the following equivalent maximization problem:

$$\max_{\substack{\mathbf{x} \in \mathcal{X}_1 \\ (\mathbf{y}, \mathbf{w}) \in \mathcal{Y}}} \mathbf{x}^\top \mathbf{y} + \sum_{i=1}^{n_x} w_i, \quad (39)$$

where $\mathcal{Y}$ equals the domain of the conjugate function of the objective of (38), i.e.,

$$\mathcal{Y} = \left\{ \mathbf{y} \in \mathbb{R}_+^{n_x}, \mathbf{w} \in \mathbb{R}^{n_x} \mid y_i \exp\left(\frac{w_i}{y_i}\right) \leq 1, i \in \{1, \dots, n_x\}, \sum_{i=1}^{n_x} y_i = 1 \right\}.$$

Observe that here we make use of case 7 in Table 1 and introduce epigraphical variables w_i for every $i \in \{1, \dots, n_x\}$. We refer to Appendices D.2 and E.2 for the convex RPT relaxation and problem instances, respectively. The results are illustrated in Table 5. Each entry corresponds to the average over 10 randomly generated instances.

$\mathcal{X}$	#	RPT-BB				RPT-SDP-BB				BARON-Dir			SCIP-Dir			BARON-Bic			SCIP-Bic		
		Opt	Time	Hyp	TO	Opt	Time	Hyp	TO	Opt	Time	TO	Opt	Time	TO	Opt	Time	TO	Opt	Time	TO
$\mathcal{X}_1$	1	10.01	0.75	0.0	0/10	10.01	0.91	0.0	0/10	10.01	0.07	0/10	10.01	0.00	0/10	10.01	1800.00	10/10	10.01	1800.00	10/10
	2	40.00	1.03	0.0	0/10	40.00	28.74	0.0	0/10	40.00	1800.00	10/10	40.00	0.00	0/10	40.00	1800.00	10/10	-	1800.00	10/10
	3	6.10	2.29	1.0	0/10	6.10	4.06	1.0	0/10	6.10	0.13	0/10	6.10	0.02	0/10	5.71	1800.00	10/10	6.10	1800.00	10/10
	4	21.96	0.28	0.0	0/10	21.96	1.40	0.0	0/10	21.90	1080.07	6/10	21.96	0.11	0/10	14.77	1800.00	10/10	20.89	1800.00	10/10
	5	-	7.54	0.0	0/10	34.76	148.91	0.0	0/10	34.76	1440.05	8/10	34.76	0.21	0/10	15.11	1800.00	10/10	33.40	1800.00	10/10

Table 5 Results for Problem (38).

From Table 5, we observe that SCIP is fastest on the direct formulation, whereas RPT handles the biconjugate formulation much faster than BARON-Bic and SCIP-Bic, both of which reach the time limit on every instance. The benefit of the biconjugate reformulation therefore depends on the relaxation used to solve it. Within RPT, the LMI strengthening does not reduce branching on these instances and increases computation time. RPT-BB is consequently faster, although on instance 5 MOSEK stalls on the root relaxation of one of the ten random instances and RPT-BB returns no solution for it, so the group mean of the objective value is withheld.

8.3. Linear multiplicative optimization

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

$$\min_{\mathbf{x} \in \mathcal{X}} \prod_{i=1}^{n_y} (\mathbf{A}_i^\top \mathbf{x} + b_i), \quad (40)$$

where $\mathbf{A} \in \mathbb{R}^{n_x \times n_y}$, $\mathbf{b} \in \mathbb{R}^{n_y}$, $\mathcal{X} = \{\mathbf{x} \in \mathbb{R}_+^{n_x} \mid \mathbf{D}^\top \mathbf{x} \leq \mathbf{d}, \mathbf{A}_i^\top \mathbf{x} + b_i \geq 0\}$, $\mathbf{D} \in \mathbb{R}^{n_x \times L'}$, and $\mathbf{d} \in \mathbb{R}^{L'}$. Without loss of generality we assume $\mathbf{A}_i^\top \mathbf{x} + b_i > 0$ for all $i \in \mathcal{I}$. Utilizing a log transformation, as in Ryoo and Sahinidis (1996), Problem (40) can be equivalently reformulated as

$$\min_{\mathbf{x} \in \mathcal{X}} \sum_{i=1}^{n_y} \log(\mathbf{A}_i^\top \mathbf{x} + b_i). \quad (41)$$

Since the objective of (41) is a closed concave function, we can replace it by its biconjugate function and obtain the following equivalent maximization problem:

$$\max_{\substack{\mathbf{x} \in \mathcal{X} \\ (\mathbf{y}, \mathbf{w}) \in \mathcal{Y}}} -(\mathbf{A}^\top \mathbf{x} + \mathbf{b})^\top \mathbf{y} - \sum_{i=1}^{n_y} w_i, \quad (42)$$

where $\mathcal{Y}$ equals the domain of the conjugate function of the objective of (41), i.e.,

$$\mathcal{Y} = \{\mathbf{y} \in \mathbb{R}_+^{n_y}, \mathbf{w} \in \mathbb{R}^{n_y} \mid \exp(-w_i - 1) \leq y_i, i \in \{1, \dots, n_y\}\}.$$

Observe that here we make use of case 3 in Table 1 and introduce epigraphical variables w_i for every $i \in \{1, \dots, n_y\}$. Moreover, $\mathcal{Y}$ is not bounded, which is one reason we make no general asymptotic-convergence claim for this formulation. We refer to Appendices D.3 and E.3 for the convex RPT relaxation and problem instances, respectively. The results are illustrated in Table 6. Each entry corresponds to the average over 10 randomly generated instances.

#	RPT-BB				RPT-SDP-BB				BARON-Dir			SCIP-Dir			BARON-Bic			SCIP-Bic		
	Opt	Time	Hyp	TO	Opt	Time	Hyp	TO	Opt	Time	TO	Opt	Time	TO	Opt	Time	TO	Opt	Time	TO
1	12.14	2.27	26.5	0/10	12.14	4.53	18.0	0/10	12.14	0.15	0/10	12.14	0.11	0/10	12.14	25.48	0/10	12.14	2.12	0/10
2	23.47	8.52	53.2	0/10	23.47	24.14	31.1	0/10	23.47	0.29	0/10	23.47	6.05	0/10	23.47	1413.00	6/10	23.47	711.06	2/10
3	20.96	36.65	138.9	0/10	20.96	30.78	28.6	0/10	20.96	0.50	0/10	20.96	266.91	0/10	20.96	1149.62	4/10	20.96	1220.68	6/10
4	18.99	1748.07	1256.5	8/10	18.99	208.97	40.7	0/10	18.99	7.63	0/10	18.99	1800.00	10/10	18.99	1800.00	10/10	18.99	1800.00	10/10
5	8.90	1800.00	291.8	10/10	-	522.11	33.2	0/10	8.90	0.41	0/10	8.90	1416.36	7/10	8.90	52.26	0/10	8.90	110.43	0/10

Table 6 Results for Problem (40).

BARON-Dir and SCIP-Dir are applied to the logarithmic reformulation (41), and BARON-Bic and SCIP-Bic to (42). From Table 6, we observe that the logarithmic reformulation is particularly effective for BARON-Dir, which has the lowest computation time on instances 2–5 and no timeouts. For the biconjugate formulation, RPT-SDP-BB is faster than BARON-Bic and SCIP-Bic on instances 2–4. The LMI strengthening also substantially reduces branching relative to RPT-BB, most clearly on instance 4, but increases computation time on the first two instances. These results illustrate the importance of both the formulation and the relaxation. On instance 5, MOSEK stalls on the root relaxation of one of the ten random instances and RPT-SDP-BB returns no solution for it, so the group mean of the objective value is withheld.

8.4. Quadratically constrained quadratic optimization

We consider the following quadratically constrained quadratic optimization problem:

$$\begin{aligned} \min_{\mathbf{x} \in \mathcal{X}_1} \quad & \mathbf{x}^\top \mathbf{P}_0 \mathbf{x} + \mathbf{q}_0^\top \mathbf{x} + r_0 \\ \text{s.t.} \quad & \mathbf{x}^\top \mathbf{P}_k \mathbf{x} + \mathbf{q}_k^\top \mathbf{x} + r_k \leq 0, \quad k \in \mathcal{K}, \end{aligned} \quad (43)$$

where $\mathbf{P}_k \in \mathbb{R}^{n_x \times n_x}$, $k \in \mathcal{K}_0$, are not necessarily positive semidefinite and as a result Problem (43) is not necessarily convex. However, the nonconvex quadratic functions are SLC, hence we can apply

RPT. CPLEX is run only on instances 1–5, which have linear constraints. It is not run on instances 6–10, which contain nonconvex quadratic constraints. We refer to Appendices D.4 and E.4 for the convex RPT relaxation and problem instances, respectively. The results are illustrated in Table 7.

#	RPT-BB				RPT-SDP-BB				BARON			CPLEX		
	Opt	Time	Hyp	TO	Opt	Time	Hyp	TO	Opt	Time	TO	Opt	Time	TO
1	394.75	6.12	1	0/1	394.75	7.50	1	0/1	394.75	0.38	0/1	394.75	0.07	0/1
2	884.75	0.24	0	0/1	884.75	0.26	0	0/1	884.75	0.27	0/1	884.75	0.04	0/1
3	6888.78	0.02	0	0/1	6888.78	0.02	0	0/1	6888.78	0.07	0/1	6888.78	0.02	0/1
4	98382.63	0.56	0	0/1	98382.63	1.73	0	0/1	98382.63	1800.00	1/1	98382.63	0.71	0/1
5	774482.39	40.25	0	0/1	774482.39	35.70	0	0/1	774482.38	1800.00	1/1	774482.38	8.60	0/1
6	1717.80	0.78	0	0/1	1717.80	0.82	0	0/1	1717.80	0.22	0/1			
7	4507.75	0.29	1	0/1	4507.75	0.33	1	0/1	4507.75	0.21	0/1			
8	17084.82	0.62	3	0/1	17084.82	0.08	0	0/1	17084.82	0.19	0/1			
9	46632.36	1800.00	483	1/1	46632.36	49.32	21	0/1	46632.36	18.12	0/1			
10	52233.70	1800.00	133	1/1	52233.70	1800.00	151	1/1	52233.70	1800.00	1/1			

Table 7 Results for Problem (43). Instances 1–5 contain only linear constraints, while instances 6–10 contain additional nonconvex quadratic constraints.

From Table 7, we observe that all methods return the same objective values within the tolerance. CPLEX has the lowest computation time on instances 1, 2 and 5, and ties with the RPT methods on instance 3 at the reported precision. RPT-BB is fastest on instance 4, whereas BARON reaches the time limit on instances 4 and 5. RPT-BB and RPT-SDP-BB solve instances 1–8 in at most 40.25 seconds with at most three hyperplanes. On instance 9, BARON has the lowest computation time (18.12 seconds), RPT-SDP-BB terminates after 49.32 seconds with 21 hyperplanes, and RPT-BB reaches the time limit after 483 hyperplanes, which shows the effect of the LMI strengthening on the instances with nonconvex quadratic constraints. On instance 10, all methods reach the time limit.

8.5. Dike height optimization

Eijgenraam et al. (2017) develop a model to optimize dike heightening in the Netherlands. The authors show that the optimal solution is periodic, i.e., the dike is heightened by 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 may be desirable to combine heightenings of several dikes. In this section, we propose to use RPT to solve the dike heightening problem in which the years at which heightening takes place are 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):

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

where $\mathbf{t}$ is the vector of all moments in time at which the dike height is increased, $t_0 = 0$, $\mathbf{x}$ 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.5. The feasible region of (DHO) is not compact without upper bounds on $\mathbf{x}$. The constraints $x_i \leq 300$ used in the implementation make it compact, as required by Assumption 1.

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 $\mathbf{t}$ 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.5 and E.5 for the convex RPT relaxation and problem instances, respectively. The results for the homogeneous dike rings 10, 15 and 16 in the Netherlands are shown in Table 8.

$\mathbf{t}$	#	RPT-BB				RPT-SDP-BB				BARON		
		Opt	Time	Hyp	TO	Opt	Time	Hyp	TO	Opt	Time	TO
$\mathbf{t}_{1r}$	10	61.98	1800.00	2935	1/1	61.98	0.08	0	0/1	61.98	4.09	0/1
	15	608.74	1800.00	4335	1/1	608.74	0.10	0	0/1	608.74	213.64	0/1
	16	1268.11	1800.00	3365	1/1	1268.11	0.09	0	0/1	1268.11	55.61	0/1
$\mathbf{t}_{25}$	10	61.31	1800.00	2073	1/1	61.31	8.71	0	0/1	61.31	25.55	0/1
	15	609.92	1800.00	854	1/1	609.92	0.15	0	0/1	609.92	1362.23	0/1
	16	1269.63	1.05	3	0/1	1269.63	0.20	0	0/1	1269.63	220.94	0/1
$\mathbf{t}_{50}$	10	55.50	853.89	2999	0/1	55.50	0.04	0	0/1	55.50	0.73	0/1
	15	545.23	392.32	1176	0/1	545.23	0.04	0	0/1	545.23	1.34	0/1
	16	1100.07	91.43	680	0/1	1100.07	0.05	0	0/1	1100.07	0.45	0/1

Table 8 Results for Problem (DHO), for dike rings 10, 15, and 16 in the Netherlands.

From Table 8, we observe that RPT-SDP-BB solves every instance at the root node, in at most 8.71 seconds, and has the lowest computation time on every instance. RPT-BB solves the three instances with the 50-year schedule and dike ring 16 with the 25-year schedule, but reaches the time limit on the remaining five instances after 854 to 4335 hyperplanes. BARON solves all nine instances, in between 0.45 and 1362.23 seconds. For this problem, the LMI strengthening makes the RPT relaxation exact at the root node.

9. Discussion and conclusion

In summary, we develop a method for globally solving nonconvex optimization problems involving SLC functions. We introduce the RPT framework, which enables us to obtain a convex relaxation of the original nonconvex problem, while introducing additional variables and constraints. We then

incorporate it in spatial branch and bound, combining the new RPT relaxation with a modification of eigenvector branching that also handles the indefinite case. We establish convergence guarantees for the proposed branching scheme, showing that the global lower bounds converge to the optimal value up to a prescribed accuracy. Moreover, when only the objective function is nonconvex, we prove finite termination with a δ -optimal solution under the stated assumptions.

In the numerical experiments, the proposed methods are competitive with BARON, SCIP, and CPLEX, and RPT-SDP-BB is the fastest method on several of the harder instances. Notably, RPT-SDP-BB solves the dike height optimization instances at the root node. Overall, RPT-BB and RPT-SDP-BB solve most instances with few branching hyperplanes. This, together with the efficiency of MOSEK for solving conic optimization problems, drives the speed of the method.

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

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

Acknowledgments

Disclosure of generative AI use. The statements of Theorem 3 and Corollary 1, the proof strategies developed in Appendix F, and most of the initial proof text were generated using OpenAI’s GPT-6 Astra via ChatGPT and Anthropic’s Claude Fable 5.1. These exchanges were directed by one of the authors. The authors checked the proofs, adapted the terminology, and substantially rewrote portions of the generated text. Generative AI was also used to proofread the manuscript during revision. The authors take full responsibility for the accuracy, integrity, and originality of the paper, including the AI-generated material.

References

- T. Achterberg. Scip: solving constraint integer programs. *Mathematical Programming Computation*, 1:1–41, 2009.
- F. 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. Anstreicher. Semidefinite programming versus the reformulation-linearization technique for nonconvex quadratically constrained quadratic programming. *Journal of Global Optimization*, 43(2):471–484, 2009.
- K. Anstreicher. On convex relaxations for quadratically constrained quadratic programming. *Mathematical Programming*, 136(2):233–251, 2012.
- K. Anstreicher. Kronecker product constraints with an application to the two-trust-region subproblem. *SIAM Journal on Optimization*, 27(1):368–378, 2017.
- K. Anstreicher. Solving two-trust-region subproblems using semidefinite optimization with eigenvector branching. *Journal of Optimization Theory and Applications*, 202(1):303–319, 2022.
- MOSEK ApS. *MOSEK Optimization Suite 11.0.30 Documentation*, 2025. URL https://docs.mosek.com/11.0/capi/intro_info.html. MOSEK v11.0.30 used in experiments.
- X. Bao, N. 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.
- S. Burer and K. Anstreicher. Second-order-cone constraints for extended trust-region subproblems. *SIAM Journal on Optimization*, 23(1):432–451, 2013. doi: 10.1137/110826862. URL <https://doi.org/10.1137/110826862>.
- M. Bynum, G. Hackebeil, W. Hart, C. Laird, B. Nicholson, J. Sirola, J.-P. Watson, and D. Woodruff. *Pyomo—Optimization Modeling in Python*, volume 67. Springer, third edition, 2021.
- 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. Falk and R. 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. Grossmann and S. Lee. Generalized convex disjunctive programming: Nonlinear convex hull relaxation. *Computational Optimization and Applications*, 26:83–100, 2003.
- Gurobi Optimization, LLC. Gurobi Optimizer Reference Manual, 2026. URL <https://www.gurobi.com>. Gurobi v13.0.0 used in experiments.
- W. Hart, J.-P. Watson, and D. Woodruff. Pyomo: Modeling and solving mathematical programs in python. *Mathematical Programming Computation*, 3(3):219–260, 2011.
- P. Hartman. On functions representable as a difference of convex functions. *Pacific Journal of Mathematics*, 9(3):707–713, 1959.
- H. Hof and M. Walter. Geometry of the reformulation-linearization-technique: Domination of disjunctions. *CoRR*, abs/2511.13805, 2025. URL <http://arxiv.org/abs/2511.13805>.
- C. Hojny, M. Besançon, K. Bestuzheva, S. Borst, A. Chmiela, J. Dionísio, L. Eifler, M. Ghannam, A. Gleixner, A. Göß, A. Hoen, R. van der Hulst, D. Kamp, T. Koch, K. Kofler, J. Lentz, S. Maher, G. Mexi, E. Mühmer, M. E. Pfetsch, S. Pokutta, F. Serrano, Y. Shinano, M. Turner, S. Vigerske, M. Walter, D. Weninger, and L. Xu. The SCIP Optimization Suite 10.0. Technical report, Optimization Online, November 2025. URL <https://optimization-online.org/2025/11/the-scip-optimization-suite-10-0/>.
- R. Horst. An algorithm for nonconvex programming problems. *Mathematical Programming*, 10:312–321, 1976.
- ILOG, Inc. *ILOG CPLEX 12.7, User Manual*, 2017. URL https://www.ibm.com/support/knowledgecenter/SSSA5P_12.7.0/ilog.odms.studio.help/Optimization_Studio/topics/COS_home.html.
- R. Jiang and D. Li. Second order cone constrained convex relaxations for nonconvex quadratically constrained quadratic programming. *Journal of Global Optimization*, 75:461–494, 2019.
- R. Jiang and D. Li. Semidefinite programming based convex relaxation for nonconvex quadratically constrained quadratic programming. In Hoai An Le Thi, Hoai Minh Le, and Tao Pham Dinh, editors, *Optimization of Complex Systems: Theory, Models, Algorithms and Applications*, pages 213–220, Cham, 2020. Springer International Publishing. ISBN 978-3-030-21803-4.
- JuliaLang contributors. Julia programming language, v1.12.5. <https://julialang.org/downloads/manual-downloads/>, 2026. Accessed: 2026-02-14.
- JuMP developers. Jump.jl v1.29.4 release notes. https://jump.dev/JuMP.jl/stable/release_notes/, 2026. Version 1.29.4 (January 16, 2026); accessed 2026-02-14.
- R. Rockafellar. *Convex Analysis*. Princeton University Press, 1970.
- R. Rockafellar and R. Wets. *Variational Analysis*. Springer, 2009.

- H. Ryoo and N. Sahinidis. A branch-and-reduce approach to global optimization. *Journal of Global Optimization*, 8:107–138, 1996.
- Nikolaos V. Sahinidis. BARON: A general purpose global optimization software package. *Journal of Global Optimization*, 8(2):201–205, 1996. doi: 10.1007/BF00138693.
- A. Saxena, P. Bonami, and J. Lee. Convex relaxations of non-convex mixed integer quadratically constrained programs: extended formulations. *Mathematical Programming*, 124(1):383–411, 2010.
- S. Schaible. Parameter-free convex equivalent and dual programs of fractional programming problems. *Zeitschrift für Operations Research*, 18(5):187–196, 1974.
- A. Selvi, A. Ben-Tal, R. Brekelmans, and D. den Hertog. Convex maximization via adjustable robust optimization. *INFORMS Journal on Computing*, 34(4):2091–2105, 2022.
- H. Serali 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. Serali and W. Adams. *A Reformulation-Linearization Technique for Solving Discrete and Continuous Nonconvex Problems*, volume 31. Springer Science & Business Media, 1999.
- H. Serali and W. Adams. A hierarchy of relaxations leading to the convex hull representation for general discrete optimization problems. *Annals of Operations Research*, 140:21–47, 2005.
- H. Serali and W. Adams. A reformulation-linearization technique for semi-infinite and convex programs under mixed 0-1 and general discrete restrictions. *Discrete Applied Mathematics*, 157(6):1319–1333, 2009.
- J. Sturm and S. Zhang. On cones of nonnegative quadratic functions. *Mathematics of Operations Research*, 28(2):246–267, 2003.
- P. Tao and L. An. Convex analysis approach to d. c. programming: Theory, algorithms and applications. *Acta Mathematica Vietnamica*, 22(1):289–355, 1997.
- The Optimization Firm, LLC. BARON version 21.1.13. Computer software, 2021. URL <https://minlp.com/older-baron-versions>. Release dated January 13, 2021.
- A. Waechter, C. Laird, F. Margot, and Y. Kawajir. Introduction to ipopt: A tutorial for downloading, installing, and using ipopt. *Revision*, 2009. URL <https://coin-or.github.io/Ipopt/>.
- B. Yang and S. Burer. A two-variable approach to the two-trust-region subproblem. *SIAM Journal on Optimization*, 26(1):661–680, 2016.
- J. Zhen, A. Marandi, D. de Moor, D. den Hertog, and L. Vandenberghe. Disjoint bilinear optimization: A two-stage robust optimization perspective. *INFORMS Journal on Computing*, 34(5):2410–2427, 2022.
- J. Zhen, D. Kuhn, and W. Wiesemann. A unified theory of robust and distributionally robust optimization via the primal-worst-equals-dual-best principle. *Operations Research*, 2023. URL <https://pubsonline.informs.org/doi/10.1287/opre.2021.0268>.

Appendix

A. Comparison of branching variants

We compare three branching rules for RPT-BB (U-LP, U-RX and U-RX+SOC) and two for RPT-SDP-BB (U-LP and U-RX) on the five problems of Section 8. All rules branch on the eigenvector associated with the eigenvalue of largest absolute value of the residual $\mathbf{X}^* - \mathbf{x}^*(\mathbf{x}^*)^\top$. U-RX is the rule of Algorithm 1. It computes the bounds a_j and b_j by solving the two problems in (17). U-LP appends only the branching inequality and relies on its products with the linear inequalities of the node, which imply the secants with the bounds obtained from the linear constraints of the node. U-RX+SOC first adds the violated inequality

$$(\boldsymbol{\nu}^\top \mathbf{x})^2 \leq \boldsymbol{\nu}^\top \mathbf{X} \boldsymbol{\nu} \quad (44)$$

and re-solves the node relaxation as long as the selected eigenvalue is negative, subject to the optimality tolerance and the time limit, and then applies U-RX (Remark 4). Under the LMI, negative eigenvalues cannot occur in exact arithmetic, so the SOC variant is not compared for RPT-SDP-BB. The convergence analysis of Section 6 covers U-RX.

The experiments use the same hardware and software as the main tables, with a time limit of 600 seconds per instance. The tables report the objective value (Opt), the computation time in seconds (Time), the number of nodes whose relaxation is solved (Nodes), the number of problems (17) solved (Bnd), the number of re-solves of a node after adding (44) (SOC), and the number of instances that reach the time limit (TO). The conventions are those of Section 8. For the randomly generated groups, the numerical columns are averages over the ten instances of the group. The reported times are the actual computation times, which may exceed the time limit because the limit is checked only between relaxation solves.

A.1. Sum-of-max-linear-terms maximization over convex and nonconvex constraints

Table 9 compares the rules over $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{X}_3$.

RPT-BB. Over $\mathcal{X}_1$, the rules perform similarly on the instances that are solved at the root node. On instances 3 and 3r, U-LP is the fastest rule (90.36 and 39.92 seconds against 120.91 and 52.25 seconds for U-RX). On instance 3, U-RX reduces the number of explored nodes from 181 to 157, but this reduction does not compensate for the 156 additional bound problems. The SOC rounds increase the computation time on these instances and cause one timeout on 3r. Over $\mathcal{X}_2$, the tighter bounds often reduce the tree. On instance 4, U-LP, U-RX and U-RX+SOC explore 33, 31 and 21 nodes in 2.44, 3.82 and 2.77 seconds, respectively, so the smallest tree is not the fastest. All three rules reach the time limit on five of the ten random instances of 3r and on all instances of 5 and 5r.

Sum-of-max-linear-terms maximization over convex and nonconvex constraints																				
$\mathcal{X}$	#	RPT-BB																		
		U-LP						U-RX						U-RX+SOC						
		Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO	
$\mathcal{X}_1$	1	23.29	0.01	1	0	0	0/1	23.29	0.01	1	0	0	0/1	23.29	0.01	1	0	0	0	0/1
	1r	22.72	0.01	1	0	0	0/10	22.72	0.01	1	0	0	0/10	22.72	0.01	1	0	0	0	0/10
	2	233.94	0.11	1	0	0	0/1	233.94	0.08	1	0	0	0/1	233.94	0.09	1	0	0	0	0/1
	2r	211.87	0.17	1.4	0	0	0/10	211.87	0.19	1.4	0.4	0	0/10	211.87	0.32	1.4	0.4	0.2	0	0/10
	3	1081.62	90.36	181	0	0	0/1	1081.62	120.91	157	156	0	0/1	1081.62	162.82	205	193	14	0	0/1
	3r	1159.90	39.92	80.2	0	0	0/10	1159.90	52.25	69.8	68.7	0	0/10	1159.90	108.92	133.2	121.9	9.3	1	1/10
	4	113.70	0.02	1	0	0	0/1	113.70	0.05	1	0	0	0/1	113.70	0.02	1	0	0	0	0/1
	4r	83.78	0.12	2	0	0	0/10	83.78	0.12	2	1	0	0/10	83.78	0.15	2	1	0.2	0	0/10
	5	3002.44	2.01	3	0	0	0/1	3002.44	2.32	3	2	0	0/1	3002.44	2.28	3	2	0	0	0/1
	5r	2898.05	6.03	12.8	0	0	0/10	2898.05	8.44	13	12	0	0/10	2898.05	15.00	12.8	11.4	7.1	0	0/10
$\mathcal{X}_2$	1	14.58	0.02	1	0	0	0/1	14.58	0.02	1	0	0	0/1	14.58	0.02	1	0	0	0	0/1
	1r	14.54	0.02	1	0	0	0/10	14.54	0.02	1	0	0	0/10	14.54	0.02	1	0	0	0	0/10
	2	136.22	0.30	1	0	0	0/1	136.22	0.28	1	0	0	0/1	136.22	0.28	1	0	0	0	0/1
	2r	122.21	1.82	8.4	0	0	0/10	122.21	1.88	7.6	6.6	0	0/10	122.21	1.89	6.6	5.6	0.8	0	0/10
	3	837.94	355.15	195	0	0	0/1	837.94	449.46	153	142	0	0/1	837.94	364.48	121	116	7	0	0/1
	3r	890.07	382.48	220.6	0	0	5/10	890.07	393.21	136.3	127.3	0	5/10	890.07	397.29	133.4	126.6	6.4	5	0/10
	4	33.73	2.44	33	0	0	0/1	33.73	3.82	31	30	0	0/1	33.73	2.77	21	20	3	0	0/1
	4r	31.81	1.25	12.8	0	0	0/10	31.81	1.82	10.8	8.4	0	0/10	31.81	1.08	6.2	4	0.8	0	0/10
	5	1610.69	600.08	386	0	0	1/1	1610.69	600.08	250	140	0	1/1	1610.69	600.08	198	123	55	1	1/1
	5r	1670.92	600.03	395.5	0	0	10/10	1670.92	600.07	252.6	142.2	0	10/10	1670.92	600.08	209	116.5	49.4	10	10/10
$\mathcal{X}_3$	1	13.45	46.01	587	0	0	0/1	13.45	214.65	1285	688	0	0/1	13.45	215.55	1285	688	0	0	0/1
	1r	15.01	260.51	1352.2	0	0	2/10	14.99	248.51	797.4	421.7	0	1/10	14.98	414.13	775	403.3	0.3	4	0/10
	2	140.87	24.59	43	0	0	0/1	140.87	112.63	359	182	0	0/1	140.62	946.20	91	47	0	1	1/1
	2r	129.29	131.51	414.6	0	0	0/10	129.29	147.70	428.8	221.1	0	0/10	129.25	469.52	395.4	203.6	0	3	10/10
	3	768.96	252.97	77	0	0	0/1	768.96	600.42	157	106	0	1/1	768.96	600.11	84	56	0	1	1/1
	3r	804.71	395.35	117.3	0	0	4/10	794.78	558.85	117.8	68.5	0	8/10	794.76	880.18	49.3	29.3	0	10	10/10
	4	45.32	188.94	467	0	0	0/1	45.34	600.01	373	211	0	1/1	45.34	1073.78	108	61	1	1	1/1
	4r	44.60	458.97	699.9	0	0	6/10	44.15	545.61	509.1	267.7	0	9/10	43.50	858.78	244.1	129.4	0.6	10	10/10
	5	2267.32	600.08	218	0	0	1/1	2267.99	600.13	167	113	0	1/1	2273.72	600.00	75	53	31	1	1/1
	5r	2358.87	600.02	202.8	0	0	10/10	2349.45	600.05	154.7	104.7	0	10/10	2342.73	600.09	113	73.9	42.8	10	10/10
$\mathcal{X}$	#	RPT-SDP-BB																		
		U-LP						U-RX												
		Opt	Time	Nodes	Bnd	TO	Opt	Time	Nodes	Bnd	TO									
$\mathcal{X}_1$	1	23.29	0.01	1	0	0/1	23.29	0.01	1	0	0/1									
	1r	22.72	0.01	1	0	0/10	22.72	0.01	1	0	0/10									
	2	233.94	0.45	1	0	0/1	233.94	0.42	1	0	0/1									
	2r	211.87	0.45	1	0	0/10	211.87	0.45	1	0	0/10									
	3	1081.62	26.73	1	0	0/1	1081.62	26.65	1	0	0/1									
	3r	1159.90	43.02	1.6	0	0/10	1159.90	58.70	1.6	0.5	0/10									
	4	113.70	0.04	1	0	0/1	113.70	0.04	1	0	0/1									
	4r	83.78	0.08	1	0	0/10	83.78	0.08	1	0	0/10									
	5	3002.44	80.72	3	0	0/1	3002.44	116.80	3	2	0/1									
	5r	2898.05	92.22	3.6	0	0/10	2898.05	139.43	3.6	2.6	0/10									
$\mathcal{X}_2$	1	14.58	0.03	1	0	0/1	14.58	0.03	1	0	0/1									
	1r	14.54	0.03	1	0	0/10	14.54	0.02	1	0	0/10									
	2	136.22	1.02	1	0	0/1	136.22	0.97	1	0	0/1									
	2r	122.21	2.67	2.4	0	0/10	122.21	3.56	2.4	1.3	0/10									
	3	837.94	92.47	1	0	0/1	837.94	90.75	1	0	0/1									
	3r	890.07	214.31	2.1	0	1/10	890.07	246.28	1.8	0.8	1/10									
	4	33.73	1.29	5	0	0/1	33.73	3.83	5	4	0/1									
	4r	31.81	0.75	2.6	0	0/10	31.81	1.44	2.6	1.6	0/10									
	5	1610.69	1262.88	4	0	1/1	1610.69	605.38	4	4	1/1									
	5r	-	681.05	6.9	0	10/10	1670.92	603.12	4.1	4	10/10									
$\mathcal{X}_3$	1	13.44	3.14	15	0	0/1	13.41	11.73	39	27	0/1									
	1r	14.99	15.43	48	0	0/10	14.99	11.76	27.2	22.5	0/10									
	2	140.62	27.95	9	0	0/1	140.62	41.23	9	5	0/1									
	2r	129.09	25.96	8.6	0	0/10	129.10	41.22	9.2	6.9	0/10									
	3	768.94	604.04	3	0	1/1	768.94	602.88	1	0	1/1									
	3r	793.83	603.46	3	0	10/10	789.16	602.42	1	0	10/10									
	4	45.33	95.63	95	0	0/1	45.34	233.08	103	59	0/1									
	4r	44.84	180.21	105.4	0	1/10	44.57	328.28	81.5	54.3	3/10									
	5	-	600.44	3	0	1/1	-	601.96	1	0	1/1									
	5r	-	605.19	3	0	10/10	-	602.78	1.7	0.7	10/10									

Table 9 Results for Problem (36) over the feasible regions $\mathcal{X}_1$, $\mathcal{X}_2$, and $\mathcal{X}_3$ for the different branching strategies in RPT-BB and RPT-SDP-BB.

Over $\mathcal{X}_3$, the objective values and the timeout counts differ between the rules, and U-LP has fewer timeouts than U-RX on 3r and 4r.

RPT-SDP-BB. U-LP and U-RX perform similarly on the instances that are solved at the root node. When branching occurs, the additional bound problems do not give a consistent advantage.

U-LP is faster on instances 5 and 5r over $\mathcal{X}_1$ and on 4 and 4r over $\mathcal{X}_2$. Overall, no rule is uniformly preferable over the three feasible regions.

A.2. Log-sum-exp maximization over linear constraints

Table 10 shows little difference between the rules on instances 1, 2 and 4, which are solved at the root node by both methods. On instance 3, U-RX+SOC has the lowest computation time for RPT-BB (1.04 seconds against 2.03 for U-LP and 2.28 for U-RX), whereas U-LP is the fastest rule for RPT-SDP-BB. On instance 5, MOSEK stalls on the root relaxation of one of the ten random instances, so RPT-BB returns no solution for it under any of the three rules, whereas both RPT-SDP-BB rules return solutions for all ten. The additional bound problems do not reduce the computation time on this problem.

Log-sum-exp maximization over linear constraints																			
$\mathcal{X}$	#	U-LP						RPT-BB						U-RX+SOC					
								U-RX											
		Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO
$\mathcal{X}_1$	1	10.01	0.76	1	0	0	0/10	10.01	0.74	1	0	0	0/10	10.01	0.75	1	0	0	0/10
	2	40.00	1.05	1	0	0	0/10	40.00	1.03	1	0	0	0/10	40.00	1.03	1	0	0	0/10
	3	6.10	2.03	3.4	0	0	0/10	6.10	2.28	3	1.7	0	0/10	6.10	1.04	1.6	0.6	0.1	0/10
	4	21.96	0.28	1	0	0	0/10	21.96	0.30	1	0	0	0/10	21.96	0.29	1	0	0	0/10
	5	-	7.53	1	0	0	0/10	-	7.48	1	0	0	0/10	-	7.66	1	0	0	0/10

RPT-SDP-BB																			
$\mathcal{X}$	#	U-LP					U-RX												
		Opt	Time	Nodes	Bnd	TO	Opt	Time	Nodes	Bnd	TO								
$\mathcal{X}_1$	1	10.01	0.92	1	0	0/10	10.01	0.91	1	0	0/10								
	2	40.00	28.82	1	0	0/10	40.00	28.40	1	0	0/10								
	3	6.10	2.63	3	0	0/10	6.10	4.04	3	2	0/10								
	4	21.96	1.41	1	0	0/10	21.96	1.39	1	0	0/10								
	5	34.76	154.44	1	0	0/10	34.76	146.82	1	0	0/10								

Table 10 Results for Problem (38) for the different branching strategies in RPT-BB and RPT-SDP-BB.

A.3. Linear multiplicative optimization

Table 11 shows that U-LP has the lowest computation time for RPT-BB on instances 1–3 and reaches the time limit on eight of the ten random instances of instance 4, against ten for U-RX and U-RX+SOC. The SOC rounds increase the computation time considerably. On instance 2, the mean time increases from 8.61 seconds for U-RX to 273.02 seconds for U-RX+SOC. All three rules reach the time limit on all random instances of instance 5. For RPT-SDP-BB, U-LP is also the fastest rule. On instances 3 and 4, it takes 17.16 and 70.79 seconds, against 30.86 and 126.39 seconds for U-RX. On instance 5, MOSEK stalls on the root relaxation of one of the ten random instances, so RPT-SDP-BB returns no solution for it under both rules.

Linear multiplicative optimization																			
$\mathcal{X}$	#	RPT-BB																	
		U-LP						U-RX						U-RX+SOC					
		Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO
	1	12.14	1.77	53.8	0	0	0/10	12.14	2.27	54	48.1	0	0/10	12.14	22.93	178.6	173.5	1.4	0/10
	2	23.47	5.47	106	0	0	0/10	23.47	8.61	107.4	99.9	0	0/10	23.47	273.02	1033.8	985.7	4.9	1/10
	3	20.96	26.51	307.8	0	0	0/10	20.96	36.55	278.8	268.9	0	0/10	20.96	456.90	951.1	923.9	4.4	5/10
	4	18.99	585.60	2043	0	0	8/10	18.99	600.23	1014.2	999.2	0	10/10	18.99	600.69	369.2	348.5	2.7	10/10
	5	8.90	600.84	404.4	0	0	10/10	8.90	600.75	275.8	254.6	0	10/10	8.90	601.51	120	116.2	1.3	10/10
$\mathcal{X}$	#	RPT-SDP-BB																	
		U-LP					U-RX												
		Opt	Time	Nodes	Bnd	TO	Opt	Time	Nodes	Bnd	TO								
	1	12.14	2.91	37	0	0/10	12.14	4.52	37	35.7	0/10								
	2	23.47	13.74	63.4	0	0/10	23.47	24.12	63.2	61.2	0/10								
	3	20.96	17.16	58.2	0	0/10	20.96	30.86	58.2	54.6	0/10								
	4	18.99	70.79	81	0	0/10	18.99	126.39	82.4	79	0/10								
	5	-	298.66	68.2	0	0/10	-	425.36	58.2	56.1	4/10								

Table 11 Results for Problem (40) for the different branching strategies in RPT-BB and RPT-SDP-BB.

A.4. Quadratically constrained quadratic optimization

Quadratically constrained quadratic optimization																			
$\mathcal{X}$	#	RPT-BB																	
		U-LP						U-RX						U-RX+SOC					
		Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO
	1	394.75	6.18	3	0	0	0/1	394.75	6.10	3	2	0	0/1	394.75	6.14	3	2	0	0/1
	2	884.75	0.25	1	0	0	0/1	884.75	0.24	1	0	0	0/1	884.75	0.24	1	0	0	0/1
	3	6888.78	0.02	1	0	0	0/1	6888.78	0.02	1	0	0	0/1	6888.78	0.02	1	0	0	0/1
	4	98382.63	0.81	1	0	0	0/1	98382.63	0.53	1	0	0	0/1	98382.63	0.54	1	0	0	0/1
	5	774482.39	42.00	1	0	0	0/1	774482.39	40.11	1	0	0	0/1	774482.39	40.84	1	0	0	0/1
	6	1717.80	0.90	1	0	0	0/1	1717.80	1.07	1	0	0	0/1	1717.80	0.88	1	0	0	0/1
	7	4507.75	0.21	3	0	0	0/1	4507.75	0.28	3	2	0	0/1	4507.75	0.29	3	2	0	0/1
	8	17084.82	0.60	7	0	0	0/1	17084.82	0.62	7	6	0	0/1	17084.82	1.35	5	4	1	0/1
	9	46632.36	600.22	445	0	0	1/1	46632.36	600.00	302	302	0	1/1	46632.36	600.08	233	232	93	1/1
	10	52233.70	600.01	120	0	0	1/1	52233.70	600.24	85	83	0	1/1	52233.70	600.01	93	92	25	1/1

RPT-SDP-BB											
$\mathcal{X}$	#	U-LP					U-RX				
		Opt	Time	Nodes	Bnd	TO	Opt	Time	Nodes	Bnd	TO
	1	394.75	7.19	3	0	0/1	394.75	7.37	3	2	0/1
	2	884.75	0.29	1	0	0/1	884.75	0.26	1	0	0/1
	3	6888.78	0.02	1	0	0/1	6888.78	0.02	1	0	0/1
	4	98382.63	1.73	1	0	0/1	98382.63	1.74	1	0	0/1
	5	774482.39	36.17	1	0	0/1	774482.39	35.30	1	0	0/1
	6	1717.80	0.86	1	0	0/1	1717.80	0.81	1	0	0/1
	7	4507.75	0.22	3	0	0/1	4507.75	0.33	3	2	0/1
	8	17084.82	0.08	1	0	0/1	17084.82	0.08	1	0	0/1
	9	46632.36	29.02	39	0	0/1	46632.36	59.48	43	42	0/1
	10	52233.70	198.62	101	0	0/1	52233.70	600.70	149	144	1/1

Table 12 Results for the quadratically constrained quadratic optimization problem for the different branching strategies in RPT-BB and RPT-SDP-BB.

For RPT-BB, Table 12 shows similar computation times for the three rules on instances 1–8, which are solved with at most seven explored nodes, and all three rules reach the time limit on instances 9 and 10. For RPT-SDP-BB, the two rules are again close on instances 1–8. U-LP is faster on instance 9 (29.02 seconds against 59.48 seconds for U-RX), and on instance 10 it terminates after 198.62 seconds, whereas U-RX reaches the time limit.

A.5. Dike height optimization

Table 13 shows that the effect of the rule for RPT-BB depends on the dike ring and the heightening schedule. U-LP reaches the time limit on four instances, U-RX on six and U-RX+SOC on eight, and all rules return the same objective values. For RPT-SDP-BB, both rules solve every instance at the root node, so the branching rule plays no role. The LMI strengthening matters more for this problem than the choice of the branching rule.

Dike height optimization																			
t	#	U-LP						RPT-BB						U-RX+SOC					
		Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO	Opt	Time	Nodes	Bnd	SOC	TO
t_{ir}	10	61.98	600.00	4846	0	0	1/1	61.98	600.00	2380	2361	0	1/1	61.98	600.01	1653	1628	733	1/1
	15	608.74	600.02	4732	0	0	1/1	608.74	600.02	3171	3119	0	1/1	608.74	600.00	2641	2606	326	1/1
	16	1268.11	241.48	2063	0	0	0/1	1268.11	600.02	2697	2680	0	1/1	1268.11	600.02	2089	2065	356	1/1
t_{25}	10	61.31	600.21	3504	0	0	1/1	61.31	600.11	1541	1482	0	1/1	61.31	600.05	1276	1247	341	1/1
	15	609.92	600.01	3050	0	0	1/1	609.92	600.04	2070	2055	0	1/1	609.92	600.03	1487	1440	265	1/1
	16	1269.63	509.22	2571	0	0	0/1	1269.63	1.09	7	5	0	0/1	1269.63	600.12	1121	1116	451	1/1
t_{50}	10	55.50	289.94	5279	0	0	0/1	55.50	600.05	4201	4200	0	1/1	55.50	600.02	4131	4129	756	1/1
	15	545.23	92.03	2303	0	0	0/1	545.23	151.17	2353	2350	0	0/1	545.23	600.01	5481	5477	1205	1/1
	16	1100.07	17.94	405	0	0	0/1	1100.07	89.13	1361	1360	0	0/1	1100.07	114.17	1251	1249	287	0/1

RPT-SDP-BB											
t	#	U-LP					U-RX				
		Opt	Time	Nodes	Bnd	TO	Opt	Time	Nodes	Bnd	TO
t_{ir}	10	61.98	0.11	1	0	0/1	61.98	0.08	1	0	0/1
	15	608.74	0.10	1	0	0/1	608.74	0.10	1	0	0/1
	16	1268.11	0.09	1	0	0/1	1268.11	0.09	1	0	0/1
t_{25}	10	61.31	8.72	1	0	0/1	61.31	8.65	1	0	0/1
	15	609.92	0.15	1	0	0/1	609.92	0.15	1	0	0/1
	16	1269.63	0.17	1	0	0/1	1269.63	0.20	1	0	0/1
t_{50}	10	55.50	0.03	1	0	0/1	55.50	0.03	1	0	0/1
	15	545.23	0.04	1	0	0/1	545.23	0.04	1	0	0/1
	16	1100.07	0.05	1	0	0/1	1100.07	0.05	1	0	0/1

Table 13 Results for Problem (DHO).

A.6. Overall comparison of branching rules

Across the five problems, no rule is consistently the fastest. U-LP, which solves no additional optimization problems at a branching, often has the lowest computation time, in particular for linear multiplicative optimization. U-RX can reduce the number of explored nodes, but the two additional problems (17) per branching do not always reduce the total computation time. The SOC rounds give no consistent improvement. When the root relaxation already solves an instance, the branching rule plays no role. We use U-RX in the main tables because it is the rule of Algorithm 1. The comparison shows the trade-off between U-RX and U-LP. Tighter bounds may reduce the number of nodes, but the cost of computing them can outweigh this reduction.

B. When infimum is not attained

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

ASSUMPTION 4. *There exists a vector $\mathbf{x}^S \in \text{ri}(\cap_{k \in \mathcal{K}_0} \text{dom}(f_k))$ such that $f_k(\mathbf{x}^S) < 0$ for all $k \in \mathcal{K}$, $\mathbf{A}^\top \mathbf{x}^S < \mathbf{b}$, $\mathbf{P}^\top \mathbf{x}^S = \mathbf{s}$, and $\mathbf{h}(\mathbf{x}^S) < \mathbf{0}$.*

Note that Assumption 4 implies that $\mathbf{x}^S$ resides in the sets $\cap_{k \in \mathcal{K}_0} \text{ri}(\text{dom}(c_{ik}))$ and $\cap_{j \in \mathcal{J}_0} \text{ri}(\text{dom}(h_j))$ thanks to Proposition 2.42 in Rockafellar and Wets (2009), and thus, $\mathbf{x}^S$ is a strict Slater point of (1). Furthermore, there exists a $(\boldsymbol{\tau}^S, \mathbf{X}^S, \mathbf{V}^S)$ such that $(\mathbf{x}^S, \boldsymbol{\tau}^S, \mathbf{X}^S, \mathbf{V}^S)$ is a strict Slater point of the corresponding RPT relaxation (8) of (1) with $f_k(\mathbf{x}) \leq 0$ is replaced by

$$\exists \mathbf{y} : \begin{cases} c_{0k}(\mathbf{x}) - \sup_{\mathbf{y} \in \text{dom}(c_{1k}^*)} \{\mathbf{x}^\top \mathbf{y} - c_{1k}^*(\mathbf{y})\} \leq 0 \\ \mathbf{y} \in \text{dom}(c_{1k}^*). \end{cases} \quad (45)$$

Finally, thanks to Remark 1 and the proof of Theorem 6(iii) of Zhen et al. (2023), the inf operator in the constraint of (8) can be merged with the inf operator (instead of min operator because the optimal $\mathbf{y}$ may not be obtained) in the objective function without affecting the infimum of (8).

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. The MC procedure takes as starting points the candidates in $\mathcal{X}'$ at which the model functions are evaluable (see Section 3.4) and returns a candidate solution, subject to the direct feasibility and finite-objective checks of Section 3.4. For $\mathcal{X}_3$, a starting point must also pass the feasibility test described in Appendix D. The procedure is summarized in Algorithm 3, for the problem of maximizing a function $f(\mathbf{x}, \mathbf{y})$ over $\mathcal{X} \times \mathcal{Y}$. If no candidate in $\mathcal{X}'$ is usable as a starting point, MC cannot be initialized.

D. RPT-SDP formulations of the numerical experiments

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

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

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 $\log(\sum_{i=1}^{n_x} \exp(x_i)) \leq a \iff \sum_{i=1}^{n_x} \exp(x_i - a) \leq 1$. Using epigraphical variables z_i we obtain the following equivalent form:

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

Algorithm 3 Mountain climbing procedure**Input:** $\mathcal{X}'$, $\mathcal{L} = \emptyset$.

```

1: for  $x \in \mathcal{X}'$  do
2:    $y \leftarrow \arg \max_{y \in \mathcal{Y}} f(x, y)$ 
3:    $\varepsilon \leftarrow 1$ 
4:   while  $\varepsilon > 0.001$  do
5:      $Lb \leftarrow f(x, y)$ 
6:      $x \leftarrow \arg \max_{x \in \mathcal{X}} f(x, y)$ 
7:      $y \leftarrow \arg \max_{y \in \mathcal{Y}} f(x, y)$ 
8:      $Lb_x \leftarrow f(x, y)$ 
9:      $\varepsilon \leftarrow Lb_x - Lb$ 
10:  end while
11:   $\mathcal{L} \leftarrow \mathcal{L} \cup \{(x, y)\}$ 
12: end for
13:  $(x^*, y^*) \leftarrow \arg \max_{(x, y) \in \mathcal{L}} f(x, y)$ 
14:  $Lb^* = f(x^*, y^*)$ 
15: return  $(Lb^*, x^*, y^*)$ 

```

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

$$\mathcal{X}_3 = \left\{ x \in \mathcal{X}_1 \mid \|x\|_2 + \inf_{z \in \mathbb{R}_{++}^{n_x}} \left\{ \sum_{i=1}^{n_x} \frac{1}{4z_i} + x^\top z \right\} \leq c \right\}.$$

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

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

The set $\mathcal{X}_3$ is bounded in the generated instances. The auxiliary representation in z displayed above, however, is not compact, and the infimum over z is not attained when some $x_i = 0$. This lifted

formulation therefore illustrates that compactness and refinement of the original $\mathbf{x}$ -space alone do not establish asymptotic convergence of every auxiliary lifted block (see Section 6).

The mountain climbing procedure is applied as follows. For $\mathcal{X} = \mathcal{X}_3$, it is started only from candidates that pass an explicit feasibility test. For $\mathcal{X}_1$, Lemma 3 guarantees feasibility of $\mathbf{x}^*$ and of the candidates derived from $\mathbf{X}$, whereas the candidates derived from $\mathbf{V}$ are always checked directly. The blocks between which the procedure alternates depend on the formulation. For $\mathcal{X}_1$ it alternates between $\mathbf{x}$ and $\mathbf{y}$, for $\mathcal{X}_2$ between $(\mathbf{x}, \mathbf{z})$ and $\mathbf{y}$, and for $\mathcal{X}_3$ between $\mathbf{x}$ and $(\mathbf{y}, \mathbf{z})$.

In the formulations for the numerical experiments we encounter several products of variables. These are linearized as follows: We linearize $\mathbf{x}\mathbf{x}^\top$ by $\mathbf{X}$, $\mathbf{y}\mathbf{y}^\top$ by $\mathbf{Y}$, $\mathbf{z}\mathbf{z}^\top$ by $\mathbf{Z}$, $\mathbf{w}\mathbf{w}^\top$ by $\mathbf{W}$, $\mathbf{t}\mathbf{t}^\top$ by $\mathbf{T}$, $\mathbf{x}\mathbf{y}^\top$ by $\mathbf{U}$, $\mathbf{x}\mathbf{z}^\top$ by $\mathbf{V}$, $\mathbf{x}\mathbf{w}^\top$ by $\mathbf{Q}$, $\mathbf{x}\mathbf{t}^\top$ by $\mathbf{F}$, $\mathbf{y}\mathbf{z}^\top$ by $\mathbf{R}$, $\mathbf{y}\mathbf{w}^\top$ by $\mathbf{P}$, $\mathbf{y}\mathbf{t}^\top$ by $\mathbf{G}$, $\mathbf{z}\mathbf{w}^\top$ by $\mathbf{K}$, $\mathbf{z}\mathbf{t}^\top$ by $\mathbf{H}$, $s\mathbf{x}$ by α , $s\mathbf{y}$ by β , $s\mathbf{z}$ by γ , $s\mathbf{t}$ by ϕ , s^2 by σ , $p\mathbf{x}$ by λ , $p\mathbf{y}$ by μ , $p\mathbf{z}$ by ν , $p\mathbf{t}$ by ψ , $p\mathbf{s}$ by ρ and p^2 by π .

D.1. RPT-SDP formulation of Problem (35)

Replacing the objective function with the biconjugate function in (35) we obtain the following equivalent maximization problem

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

where $\mathcal{Y}$ is given by

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

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

$$\begin{aligned} \max_{\mathbf{x}, \mathbf{y}, \mathbf{U}, \mathbf{X}, \mathbf{Y}} \quad & \text{Tr}(\mathbf{A}^\top \mathbf{U}) + \mathbf{b}^\top \mathbf{y} \\ \text{s.t.} \quad & \mathbf{D}^\top \mathbf{X}_i - d\mathbf{x}_i \leq \mathbf{0}, \end{aligned} \quad i \in \{1, \dots, n_x\}, \quad (46a)$$

$$\mathbf{D}^\top \mathbf{U}_j - d\mathbf{y}_j \leq \mathbf{0}, \quad j \in \{1, \dots, n_y\}, \quad (46b)$$

$$d\mathbf{x}^\top \mathbf{D} + \mathbf{D}^\top \mathbf{x} d^\top \leq \mathbf{D}^\top \mathbf{X} \mathbf{D} + d d^\top, \quad (46c)$$

$$\sum_{j \in \mathcal{J}_\ell} y_j = 1, \quad \ell \in \mathcal{L}, \quad (46d)$$

$$\sum_{j \in \mathcal{J}_\ell} \mathbf{U}_j - \mathbf{x} = \mathbf{0}, \quad \ell \in \mathcal{L}, \quad (46e)$$

$$\sum_{j \in \mathcal{J}_\ell} \mathbf{Y}_j - \mathbf{y} = \mathbf{0}, \quad \ell \in \mathcal{L}, \quad (46f)$$

$$\mathbf{X}, \mathbf{Y}, \mathbf{U} \geq \mathbf{0}, \quad (46g)$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{U} & \mathbf{x} \\ \mathbf{U}^\top & \mathbf{Y} & \mathbf{y} \\ \mathbf{x}^\top & \mathbf{y}^\top & 1 \end{pmatrix} \succeq \mathbf{0}, \quad (46h)$$

where constraints (46a) - (46g) result from pairwise multiplication of the linear constraints and constraint (46h) results from the additional SDP relaxation.

Observe that $\mathbf{x} \in \mathcal{X}_1$ is redundant. The constraint $\mathbf{D}^\top \mathbf{x} \leq \mathbf{d}$ is redundant by (46e), (46d) and (46b):

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

The nonnegativity constraint $\mathbf{x} \geq \mathbf{0}$ is redundant by (46e) and (46g). Moreover, the nonnegativity constraint $\mathbf{y} \geq \mathbf{0}$ is redundant by (46f) and (46g). Hence, these constraints are not included in the above formulation.

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

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

$$\max_{\substack{\mathbf{x}, \mathbf{y}, \mathbf{z} \\ \mathbf{U}, \mathbf{V}, \mathbf{R} \\ \mathbf{X}, \mathbf{Y}, \mathbf{Z}}} \text{Tr}(\mathbf{A}^\top \mathbf{U}) + \mathbf{b}^\top \mathbf{y}$$

$$\text{s.t. (46a) - (46g)}$$

$$\sum_{i=1}^{n_x} z_i \leq 1, \tag{47a}$$

$$\sum_{i=1}^{n_x} \mathbf{V}_i - \mathbf{x} \leq \mathbf{0}, \tag{47b}$$

$$\sum_{i=1}^{n_x} \mathbf{R}_i - \mathbf{y} \leq \mathbf{0}, \tag{47c}$$

$$\mathbf{D}^\top \mathbf{x} - \mathbf{D}^\top \sum_{i=1}^{n_x} \mathbf{V}_i \leq \mathbf{d} (1 - \sum_{i=1}^{n_x} z_i), \tag{47d}$$

$$\sum_{i,j=1}^{n_x} Z_{ij} - 2 \sum_{i=1}^{n_x} z_i + 1 \geq 0, \tag{47e}$$

$$\exp(x_i - a) \leq z_i, \quad i \in \{1, \dots, n_x\}, \tag{47f}$$

$$x_j \exp\left(\frac{X_{ij} - ax_j}{x_j}\right) \leq V_{ji}, \quad i, j \in \{1, \dots, n_x\}, \tag{47g}$$

$$y_j \exp\left(\frac{U_{ij} - ay_j}{y_j}\right) \leq R_{ji}, \quad i \in \{1, \dots, n_x\}, j \in \{1, \dots, n_y\}, \tag{47h}$$

$$(d_\ell - \mathbf{D}_\ell^\top \mathbf{x}) \exp\left(\frac{d_\ell x_i - ad_\ell - \mathbf{D}_\ell^\top \mathbf{X}_i + a \mathbf{D}_\ell^\top \mathbf{x}}{d_\ell - \mathbf{D}_\ell^\top \mathbf{x}}\right) \leq d_\ell z_i - \mathbf{D}_\ell^\top \mathbf{V}_i, \quad i \in \{1, \dots, n_x\}, \ell \in \{1, \dots, L'\}, \tag{47i}$$

$$\left(1 - \sum_{j=1}^{n_x} z_j\right) \exp\left(\frac{x_i - a - \sum_{j=1}^{n_x} V_{ij} + a \sum_{j=1}^{n_x} z_j}{1 - \sum_{j=1}^{n_x} z_j}\right) \leq z_i - \sum_{j=1}^{n_x} Z_{ji}, \quad i \in \{1, \dots, n_x\}, \tag{47j}$$

$$\exp(x_i + x_j - 2a) \leq Z_{ij}, \quad i \leq j \in \{1, \dots, n_x\}, \tag{47k}$$

$$z_j \exp\left(\frac{V_{ij} - az_j}{z_j}\right) \leq Z_{ij}, \quad i \leq j \in \{1, \dots, n_x\}, \quad (47l)$$

$$\sum_{j \in \mathcal{J}_\ell} (\mathbf{R}^\top)_j - \mathbf{z} = \mathbf{0}, \quad \ell \in \mathcal{L}, \quad (47m)$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{U} & \mathbf{V} & \mathbf{x} \\ \mathbf{U}^\top & \mathbf{Y} & \mathbf{R} & \mathbf{y} \\ \mathbf{V}^\top & \mathbf{R}^\top & \mathbf{Z} & \mathbf{z} \\ \mathbf{x}^\top & \mathbf{y}^\top & \mathbf{z}^\top & 1 \end{pmatrix} \succeq \mathbf{0}, \quad (47n)$$

where constraints (47b) - (47e) result from pairwise multiplication of the new linear constraint over $\mathbf{z}$ with the previous linear constraints (47g) - (47l) result from pairwise multiplication of the exponential constraint with the linear inequalities and itself, constraint (47m) results from pairwise multiplication of the initial linear constraint over $\mathbf{y}$ with $\mathbf{z}$ and constraint (47n) results from the additional SDP relaxation.

Our implementation omits constraint (47l). The displayed formulation therefore contains a valid inequality that is not imposed in the code. The counts below refer to the displayed formulation. In the RPT-SDP formulation we hence obtain $2n_x|\mathcal{L}| + n_y|\mathcal{L}| + n_xn_y + L'n_y + n_x(n_x + 1)/2 + L'(L' + 1)/2 + n_xL' + n_y(n_y + 1)/2 + n_y + n_x + L' + 1$ additional linear constraints, $n_xn_y + n_x^2 + L'n_x + n_x(n_x + 1) + n_x$ additional exponential constraints, one additional SDP constraint and $n_x(n_x + 1) + n_y(n_y + 1)/2 + 2n_xn_y + n_x^2$ extra variables.

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

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

$$(46a) - (46g) \\ s + p + \sum_{i=1}^{n_x} V_{ii} \leq c, \quad (48a)$$

$$\sum_{i=1}^{n_x} t_i \leq p, \quad (48b)$$

$$\mathbf{D}^\top \mathbf{V}_i \leq z_i \mathbf{d}, \quad i \in \{1, \dots, n_x\}, \quad (48c)$$

$$\mathbf{D}^\top \mathbf{F}_i \leq t_i \mathbf{d}, \quad i \in \{1, \dots, n_x\}, \quad (48d)$$

$$\mathbf{D}^\top \boldsymbol{\lambda} \leq p \mathbf{d}, \quad (48e)$$

$$\sum_{j \in \mathcal{J}_\ell} (\mathbf{R}^\top)_j - \mathbf{z} = \mathbf{0}, \quad \ell \in \mathcal{L}, \quad (48f)$$

$$\sum_{j \in \mathcal{J}_\ell} (\mathbf{G}^\top)_j - \mathbf{t} = \mathbf{0}, \quad \ell \in \mathcal{L}, \quad (48g)$$

$$\sum_{j \in \mathcal{J}_\ell} \beta_j - s = 0, \quad \ell \in \mathcal{L}, \quad (48h)$$

$$\sum_{j \in \mathcal{J}_\ell} \mu_j - p = 0, \quad \ell \in \mathcal{L}, \quad (48i)$$

$$\sum_{i=1}^{n_x} \mathbf{F}_i \leq \boldsymbol{\lambda}, \quad (48j)$$

$$\sum_{i=1}^{n_x} \mathbf{G}_i \leq \boldsymbol{\mu}, \quad (48k)$$

$$\sum_{i=1}^{n_x} \mathbf{H}_i \leq \boldsymbol{\nu}, \quad (48l)$$

$$\sum_{i=1}^{n_x} \mathbf{T}_i \leq \boldsymbol{\psi}, \quad (48m)$$

$$\sum_{i=1}^{n_x} \psi_i \leq \pi, \quad (48n)$$

$$\sum_{i=1}^{n_x} (d_\ell t_i - \mathbf{D}_\ell^\top \mathbf{F}_i) \leq p d_\ell - \mathbf{D}_\ell^\top \boldsymbol{\lambda}, \quad \ell \in \{1, \dots, L'\}, \quad (48o)$$

$$\pi - 2 \sum_{i=1}^{n_x} \psi_i + \sum_{i,j=1}^{n_x} T_{ij} \geq 0, \quad (48p)$$

$$\mathbf{V}, \mathbf{F}, \mathbf{R}, \mathbf{G}, \mathbf{H}, \mathbf{Z}, \mathbf{T} \geq \mathbf{0}, \quad (48q)$$

$$\boldsymbol{\lambda}, \boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{\psi}, \sigma, \pi \geq \mathbf{0}. \quad (48r)$$

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

$$\|\mathbf{x}\|_2 \leq s, \quad (49a)$$

$$\|(z_i - t_i, 1)^\top\|_2 \leq z_i + t_i, \quad i \in \{1, \dots, n_x\}, \quad (49b)$$

$$\|\mathbf{X}_i\|_2 \leq \alpha_i, \quad i \in \{1, \dots, n_x\}, \quad (49c)$$

$$\|\mathbf{U}_i\|_2 \leq \beta_i, \quad i \in \{1, \dots, n_y\}, \quad (49d)$$

$$\|\mathbf{V}_i\|_2 \leq \gamma_i, \quad i \in \{1, \dots, n_x\}, \quad (49e)$$

$$\|\mathbf{F}_i\|_2 \leq \phi_i, \quad i \in \{1, \dots, n_x\}, \quad (49f)$$

$$\|\boldsymbol{\lambda}\|_2 \leq \rho, \quad (49g)$$

$$\|d_\ell \mathbf{x} - \mathbf{X} \mathbf{D}_\ell\|_2 \leq s d_\ell - \mathbf{D}_\ell^\top \boldsymbol{\alpha}, \quad \ell \in \{1, \dots, L'\}, \quad (49h)$$

$$\|\boldsymbol{\lambda} - \sum_{i=1}^{n_x} \mathbf{F}_i\|_2 \leq \rho - \sum_{i=1}^{n_x} \phi_i, \quad (49i)$$

$$\|\mathbf{X}\|_F \leq \sigma, \quad (49j)$$

$$\|(V_{ji} - F_{ji}, x_j)\|_2 \leq V_{ji} + F_{ji}, \quad i, j \in \{1, \dots, n_x\}, \quad (49k)$$

$$\|(R_{ji} - G_{ji}, y_j)\|_2 \leq R_{ji} + G_{ji}, \quad i \in \{1, \dots, n_x\}, j \in \{1, \dots, n_y\}, \quad (49l)$$

$$\|(Z_{ji} - H_{ji}, z_j)\|_2 \leq Z_{ji} + H_{ji}, \quad i, j \in \{1, \dots, n_x\}, \quad (49m)$$

$$\|(H_{ij} - T_{ij}, t_j)\|_2 \leq H_{ij} + T_{ij}, \quad i, j \in \{1, \dots, n_x\}, \quad (49n)$$

$$\|(\nu_i - \psi_i, p)\|_2 \leq \nu_i + \psi_i, \quad i \in \{1, \dots, n_x\}, \quad (49o)$$

$$\|(d_\ell(z_i - t_i) + \mathbf{D}_\ell^\top (\mathbf{F}_i - \mathbf{V}_i), d_\ell - \mathbf{D}_\ell^\top \mathbf{x})\|_2 \leq d_\ell(z_i + t_i) - \mathbf{D}_\ell^\top (\mathbf{F}_i + \mathbf{V}_i), \quad \ell \in \{1, \dots, L'\}, \quad (49p)$$

$$\|(\mathbf{V}_i - \mathbf{F}_i, \mathbf{x})\|_2 \leq \gamma_i + \phi_i, \quad i \in \{1, \dots, n_x\}, \quad (49q)$$

$$\left\| \left(\nu_i - \psi_i - \sum_{j=1}^{n_x} H_{ij} + \sum_{j=1}^{n_x} T_{ij}, p - \sum_{j=1}^{n_x} t_j \right) \right\|_2 \leq \nu_i + \psi_i - \sum_{j=1}^{n_x} H_{ij} - \sum_{j=1}^{n_x} T_{ij}, \quad i \in \{1, \dots, n_x\}, \quad (49r)$$

$$\left\| \begin{pmatrix} Z_{ij} - H_{ij} - H_{ji} + T_{ij} & z_i - t_i \\ z_j - t_j & 1 \end{pmatrix} \right\|_2 \leq Z_{ij} + H_{ij} + H_{ji} + T_{ij}, \quad i, j \in \{1, \dots, n_x\}, \quad (49s)$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{U} & \mathbf{V} & \mathbf{F} & \boldsymbol{\alpha} & \boldsymbol{\lambda} & \mathbf{x} \\ \mathbf{U}^\top & \mathbf{Y} & \mathbf{R} & \mathbf{G} & \boldsymbol{\beta} & \boldsymbol{\mu} & \mathbf{y} \\ \mathbf{V}^\top & \mathbf{R}^\top & \mathbf{Z} & \mathbf{H} & \boldsymbol{\gamma} & \boldsymbol{\nu} & \mathbf{z} \\ \mathbf{F}^\top & \mathbf{G}^\top & \mathbf{H}^\top & \mathbf{T} & \boldsymbol{\phi} & \boldsymbol{\psi} & \mathbf{t} \\ \boldsymbol{\alpha}^\top & \boldsymbol{\beta}^\top & \boldsymbol{\gamma}^\top & \boldsymbol{\phi}^\top & \sigma & \rho & s \\ \boldsymbol{\lambda}^\top & \boldsymbol{\mu}^\top & \boldsymbol{\nu}^\top & \boldsymbol{\psi}^\top & \rho & \pi & p \\ \mathbf{x}^\top & \mathbf{y}^\top & \mathbf{z}^\top & \mathbf{t}^\top & s & p & 1 \end{pmatrix} \succeq \mathbf{0}. \quad (49t)$$

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

D.2. RPT-SDP formulation of Problem (38)

Replacing the objective function with the biconjugate function in (38) we obtain the following equivalent maximization problem

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

where $\mathcal{Y}$ is given by

$$\mathcal{Y} = \left\{ \mathbf{y} \in \mathbb{R}_+^{n_x}, \mathbf{w} \in \mathbb{R}^{n_x} \mid y_i \exp\left(\frac{w_i}{y_i}\right) \leq 1, i \in \{1, \dots, n_x\}, \sum_{i=1}^{n_x} y_i = 1 \right\}.$$

The RPT-SDP formulation is given by

$$\begin{aligned} \max_{\substack{\mathbf{x}, \mathbf{y}, \mathbf{w}, \\ \mathbf{X}, \mathbf{Y}, \mathbf{W}, \\ \mathbf{U}, \mathbf{Q}, \mathbf{P}}} \quad & \text{Tr}(\mathbf{U}) + \sum_{i=1}^{n_x} w_i \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{X}_1, \end{aligned} \quad (50a)$$

$$(\mathbf{y}, \mathbf{w}) \in \mathcal{Y}, \quad (50b)$$

$$\mathbf{X}, \mathbf{Y}, \mathbf{U} \geq \mathbf{0}, \quad (50c)$$

$$\mathbf{D}^\top \mathbf{X}_i - d\mathbf{x}_i \leq \mathbf{0}, \quad i \in \{1, \dots, n_x\}, \quad (50d)$$

$$\mathbf{D}^\top \mathbf{U}_i - d\mathbf{y}_i \leq \mathbf{0}, \quad i \in \{1, \dots, n_x\}, \quad (50e)$$

$$d\mathbf{x}^\top \mathbf{D} + \mathbf{D}^\top \mathbf{x} d^\top \leq \mathbf{D}^\top \mathbf{X} \mathbf{D} + d d^\top, \quad (50f)$$

$$\sum_{i \in \{1, \dots, n_x\}} \mathbf{U}_i = \mathbf{x}, \quad (50g)$$

$$\sum_{i \in \{1, \dots, n_x\}} \mathbf{Y}_i = \mathbf{y}, \quad (50h)$$

$$\sum_{i \in \{1, \dots, n_x\}} (\mathbf{P})_i^\top = \mathbf{w}, \quad (50i)$$

$$U_{ji} \exp\left(\frac{Q_{ji}}{U_{ji}}\right) \leq x_j, \quad i, j \in \{1, \dots, n_x\}, \quad (50j)$$

$$Y_{ij} \exp\left(\frac{P_{ji}}{Y_{ij}}\right) \leq y_j, \quad i, j \in \{1, \dots, n_x\}, \quad (50k)$$

$$(d_\ell y_i - \mathbf{D}_\ell^\top \mathbf{U}_i) \exp\left(\frac{d_\ell w_i - \mathbf{D}_\ell^\top \mathbf{Q}_i}{d_\ell y_i - \mathbf{D}_\ell^\top \mathbf{U}_i}\right) \leq d_\ell - \mathbf{D}_\ell^\top \mathbf{x}, \quad i \in \{1, \dots, n_x\}, \ell \in \{1, \dots, L'\}, \quad (50l)$$

$$Y_{ij} \exp\left(\frac{P_{ji} + P_{ij}}{Y_{ij}}\right) \leq 1, \quad i \leq j \in \{1, \dots, n_x\}, \quad (50m)$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{U} & \mathbf{Q} & \mathbf{x} \\ \mathbf{U}^\top & \mathbf{Y} & \mathbf{P} & \mathbf{y} \\ \mathbf{Q}^\top & \mathbf{P}^\top & \mathbf{W} & \mathbf{w} \\ \mathbf{x}^\top & \mathbf{y}^\top & \mathbf{w}^\top & 1 \end{pmatrix} \succeq \mathbf{0}, \quad (50n)$$

where constraints (50c) - (50i) result from pairwise multiplication of the linear constraints. Note that we only multiply the linear equality constraint with the variables (see Theorem 1). Constraints (50j) - (50l) result from pairwise multiplication of the linear inequality constraints with the exponential cone constraints, constraint (50m) results from pairwise multiplication of the exponential cone constraints with each other, and constraint (50n) 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, $2n_x^2 + L'n_x + n_x(n_x + 1)/2$ additional exponential cone constraints, one additional SDP constraint and $3n_x(n_x + 1)/2 + 3n_x^2$ additional variables.

Observe that we could exclude the nonnegativity constraints from the above reformulation, since from constraints (50b), (50g), and (50h) it follows that they are redundant.

D.3. RPT-SDP formulation of Problem (40)

Replacing the objective function with the biconjugate function in (40) we obtain the following equivalent maximization problem

$$\max_{\substack{\mathbf{x} \in \mathcal{X} \\ (\mathbf{y}, \mathbf{w}) \in \mathcal{Y}}} -(\mathbf{A}^\top \mathbf{x} + \mathbf{b})^\top \mathbf{y} - \sum_{i \in \{1, \dots, n_y\}} w_i, \quad (\text{CM}_B)$$

where $\mathcal{Y}$ is given by

$$\mathcal{Y} = \{\mathbf{y} \in \mathbb{R}_+^{n_y}, \mathbf{w} \in \mathbb{R}^{n_y} \mid \exp(-w_i - 1) \leq y_i, i \in \{1, \dots, n_y\}\}.$$

The RPT-SDP formulation is given by

$$\begin{aligned} \max_{\substack{\mathbf{x}, \mathbf{y}, \mathbf{w}, \\ \mathbf{X}, \mathbf{Y}, \mathbf{W}, \\ \mathbf{U}, \mathbf{Q}, \mathbf{P}}} & -\text{Tr}(\mathbf{A}^\top \mathbf{U}) - \mathbf{b}^\top \mathbf{y} - \sum_{i=1}^{n_y} w_i \\ \text{s.t.} & \mathbf{D}^\top \mathbf{x} \leq \mathbf{d}, \end{aligned} \quad (51a)$$

$$\mathbf{X}, \mathbf{Y}, \mathbf{U} \geq \mathbf{0}, \quad (51b)$$

$$\mathbf{D}^\top \mathbf{X}_i - \mathbf{d}x_i \leq \mathbf{0}, \quad i \in \{1, \dots, n_x\}, \quad (51c)$$

$$\mathbf{D}^\top \mathbf{U}_j - \mathbf{d}y_j \leq \mathbf{0}, \quad j \in \{1, \dots, n_y\}, \quad (51d)$$

$$d\mathbf{x}^\top \mathbf{D} + \mathbf{D}^\top \mathbf{x} d^\top \leq \mathbf{D}^\top \mathbf{X} \mathbf{D} + d d^\top, \quad (51e)$$

$$\exp(-w_i - 1) \leq y_i, \quad i \in \{1, \dots, n_y\}, \quad (51f)$$

$$x_j \exp\left(\frac{-Q_{ji} - x_j}{x_j}\right) \leq U_{ji}, \quad i \in \{1, \dots, n_y\}, j \in \{1, \dots, n_x\}, \quad (51g)$$

$$y_j \exp\left(\frac{-P_{ji} - y_j}{y_j}\right) \leq Y_{ij}, \quad i, j \in \{1, \dots, n_y\}, \quad (51h)$$

$$(d_\ell - \mathbf{D}_\ell^\top \mathbf{x}) \exp\left(\frac{\mathbf{D}_\ell^\top (\mathbf{x} + \mathbf{Q}_i) - d_\ell(1 + w_i)}{d_\ell - \mathbf{D}_\ell^\top \mathbf{x}}\right) \leq d_\ell y_i - \mathbf{D}_\ell^\top \mathbf{U}_i, \quad i \in \{1, \dots, n_y\}, \ell \in \{1, \dots, L'\}, \quad (51i)$$

$$\exp(-w_i - w_j - 2) \leq Y_{ij}, \quad i \leq j \in \{1, \dots, n_y\}, \quad (51j)$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{U} & \mathbf{Q} & \mathbf{x} \\ \mathbf{U}^\top & \mathbf{Y} & \mathbf{P} & \mathbf{y} \\ \mathbf{Q}^\top & \mathbf{P}^\top & \mathbf{W} & \mathbf{w} \\ \mathbf{x}^\top & \mathbf{y}^\top & \mathbf{w}^\top & 1 \end{pmatrix} \succeq \mathbf{0}, \quad (51k)$$

where constraints (51b) - (51d) result from pairwise multiplication of the linear constraints, constraints (51g) - (51j) result from pairwise multiplication of the exponential constraints with the linear and constraint (51k) results from the additional SDP relaxation.

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

D.4. RPT-SDP formulation of Problem (43)

For completeness, the SOC-RLT products displayed below apply when the constraints indexed by $\mathcal{C}$ are supplied in the norm-bound form $\|\mathbf{P}_i^{1/2} \mathbf{x}\|_2 \leq -r_i - \mathbf{q}_i^\top \mathbf{x}$, with a nonnegative right-hand side on the feasible region. This form should not be confused with the general convex quadratic constraint $\mathbf{x}^\top \mathbf{P}_i \mathbf{x} + \mathbf{q}_i^\top \mathbf{x} + r_i \leq 0$. The latter has the exact rotated-cone representation $\|\mathbf{P}_i^{1/2} \mathbf{x}\|_2^2 \leq -r_i - \mathbf{q}_i^\top \mathbf{x}$ and requires the corresponding products of that rotated-cone formulation rather than constraints (52e)–(52h). In all instances considered in Section 8, $\mathcal{C} = \emptyset$, so the convex-constraint case is not exercised in our experiments. The nonconvex quadratic constraints ($\mathcal{NC}$) are linearized as follows: $\text{tr}(\mathbf{P}_i \mathbf{X}) + \mathbf{q}_i^\top \mathbf{x} + r_i \leq 0$. The RPT-SDP formulation is given by

$$\begin{aligned} \min_{\mathbf{x}, \mathbf{X}} \quad & \text{Tr}(\mathbf{P}_0 \mathbf{X}) + \mathbf{q}_0^\top \mathbf{x} + r_0 \\ \text{s.t.} \quad & \mathbf{x} \in \mathcal{X}_1, \end{aligned} \quad (52a)$$

$$\mathbf{D}^\top \mathbf{X}_i - d x_i \leq \mathbf{0}, \quad i \in \{1, \dots, n_x\}, \quad (52b)$$

$$d\mathbf{x}^\top \mathbf{D} + \mathbf{D}^\top \mathbf{x} d^\top \leq \mathbf{D}^\top \mathbf{X} \mathbf{D} + d d^\top, \quad (52c)$$

$$\text{Tr}(\mathbf{P}_k \mathbf{X}) + \mathbf{q}_k^\top \mathbf{x} + r_k \leq 0, \quad k \in \mathcal{NC}, \quad (52d)$$

$$\|\mathbf{P}_k^{1/2} \mathbf{x}\|_2 \leq -r_k - \mathbf{q}_k^\top \mathbf{x}, \quad k \in \mathcal{C}, \quad (52e)$$

$$\|\mathbf{P}_i^{1/2} \mathbf{X} \mathbf{P}_j^{1/2}\|_F \leq r_i r_j + r_i \mathbf{q}_j^\top \mathbf{x} + r_j \mathbf{q}_i^\top \mathbf{x} + \mathbf{q}_i^\top \mathbf{X} \mathbf{q}_j, \quad i, j \in \mathcal{C}, \quad (52f)$$

$$\|d_\ell \mathbf{P}_k^{1/2} \mathbf{x} - \mathbf{P}_k^{1/2} \mathbf{X} \mathbf{D}_\ell\|_2 \leq -r_k d_\ell + r_k \mathbf{D}_\ell^\top \mathbf{x} - d_\ell \mathbf{q}_k^\top \mathbf{x} + \mathbf{q}_k^\top \mathbf{X} \mathbf{D}_\ell, \quad k \in \mathcal{C}, \ell \in \{1, \dots, L'\}, \quad (52g)$$

$$\|\mathbf{P}_k^{1/2} \mathbf{X}_j\|_2 \leq -r_k x_j - \mathbf{q}_k^\top \mathbf{X}_j, \quad k \in \mathcal{C}, j \in \{1, \dots, n_x\}, \quad (52h)$$

$$\mathbf{X} \geq \mathbf{0}, \quad (52i)$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{x}^\top & 1 \end{pmatrix} \succeq \mathbf{0}, \quad (52j)$$

where constraints (52b) - (52c) and (52i) result from pairwise multiplication of the linear constraints, constraints (52f) - (52h) result from pairwise multiplication of the convex quadratic constraints with the linear constraints and each other and constraint (52j) results from the additional SDP relaxation.

In the RPT-SDP formulation we hence obtain $L'n_x + n_x(n_x + 1)/2 + L'(L' + 1)/2 + |\mathcal{NC}|$ 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.5. RPT-SDP formulation of Problem (DHO)

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

$$\begin{aligned} \min_{\substack{\mathbf{x}, \mathbf{z}, \mathbf{w}, \\ \mathbf{X}, \mathbf{Q}, \mathbf{W}}} \quad & \sum_{k \in \mathcal{K}_0} z_k + \sum_{k \in \mathcal{K}_0} \frac{S_0}{\beta_\delta} (\exp(\beta_\delta t_{k+1}) - \exp(\beta_\delta t_k)) w_k + \frac{S_0}{\delta} \exp(\beta_\delta T) w_K \\ \text{s.t.} \quad & (C + bx_k) \exp\left(\frac{\lambda C h_k + \lambda b \sum_{i=0}^k X_{ik} - \delta t_k (C + bx_k)}{C + bx_k}\right) \leq z_k, \quad k \in \mathcal{K}_0, \end{aligned} \quad (53a)$$

$$\mathbf{D}^\top \mathbf{x} \leq \mathbf{d}, \quad (53b)$$

$$\exp\left(-\theta \sum_{i=0}^k x_i\right) \leq w_k, \quad k \in \mathcal{K}_0, \quad (53c)$$

$$\mathbf{D}^\top \mathbf{X}_k \leq x_k \mathbf{d}, \quad k \in \mathcal{K}_0, \quad (53d)$$

$$\mathbf{d} \mathbf{x}^\top \mathbf{D} + \mathbf{D}^\top \mathbf{x} \mathbf{d}^\top \leq \mathbf{D}^\top \mathbf{X} \mathbf{D} + \mathbf{d} \mathbf{d}^\top, \quad (53e)$$

$$(d_\ell - \mathbf{D}_\ell^\top \mathbf{x}) \exp\left(\frac{-\theta d_\ell \sum_{i=0}^k x_i + \theta \sum_{i=0}^k \mathbf{D}_\ell^\top \mathbf{X}_i}{d_\ell - \mathbf{D}_\ell^\top \mathbf{x}}\right) \leq d_\ell w_k - \mathbf{D}_\ell^\top \mathbf{Q}_k, \quad k \in \mathcal{K}_0, \ell \in \{1, \dots, L'\}, \quad (53f)$$

$$x_j \exp\left(\frac{-\theta \sum_{i=0}^k X_{ij}}{x_j}\right) \leq Q_{jk}, \quad k, j \in \mathcal{K}_0, \quad (53g)$$

$$\exp\left(-\theta \sum_{i=0}^k x_i - \theta \sum_{i=0}^j x_i\right) \leq W_{jk}, \quad j, k \in \mathcal{K}_0, \quad (53h)$$

$$w_j \exp\left(\frac{-\theta \sum_{i=0}^k Q_{ij}}{w_j}\right) \leq W_{jk}, \quad j, k \in \mathcal{K}_0, \quad (53i)$$

$$\mathbf{x}, \mathbf{X} \geq \mathbf{0}, \quad (53j)$$

$$\begin{pmatrix} \mathbf{X} & \mathbf{Q} & \mathbf{x} \\ \mathbf{Q}^\top & \mathbf{W} & \mathbf{w} \\ \mathbf{x}^\top & \mathbf{w}^\top & 1 \end{pmatrix} \succeq \mathbf{0}, \quad (53k)$$

where constraint (53b) represents the upper bounds on $\mathbf{x}$ such that we obtain a compact feasible region, (53d) - (53f) result from pairwise multiplication of (53b) with all other constraints, (53g) results from pairwise multiplication of (53c) with the nonnegativity constraint $\mathbf{x} \geq 0$, (53h) - (53i) result from pairwise multiplication of the exponential constraint with itself, (53j) results from pairwise multiplication of the nonnegativity constraints, and constraint (53k) 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.

E. Data generation of numerical experiments

E.1. Data generation of numerical experiments of Problem (35)

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

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

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

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

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 $\mathbf{D}$ and $\mathbf{d}$ are given by :

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

respectively. The values for the parameters of each distinct problem instance are given in Table 14.

Instance	n_x	$ \mathcal{L} $	$ \mathcal{J}_\ell $	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 14 Problem (36) parameters for each instance. n_x refers to the number of variables, $|\mathcal{L}|$ to the number of max linear terms, $|\mathcal{J}_\ell|$ to the number of elements within a max-term, a to the parameter used in $\mathcal{X}_2$, c to the parameter used in $\mathcal{X}_3$ and M to the big M parameter used in Problem (37).

E.2. Data generation of numerical experiments of Problem (38)

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

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

in instance 3 as

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

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 15.

Table 15 Problem (38) 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.3. Data generation of numerical experiments of Problem (40)

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 16.

Table 16 Problem (40) parameters for each instance. n_x refers to the number of variables, L' to the number of linear constraints, and n_y to the number of linear multiplications in the objective.

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

E.4. Data generation of numerical experiments of Problem (43)

The first 5 problem instances are adopted from Selvi et al. (2022) and can be summarized as follows: In instances 1 and 2 the objectives are $f(\mathbf{x}) = -\frac{1}{2} \sum_{i=1}^{20} (x_i - 2)^2$ and $f(\mathbf{x}) = -\frac{1}{2} \sum_{i=1}^{20} (x_i + 5)^2$ respectively and the linear constraints are as in instance 5 for problem (35). Instances 3, 4 and 5 are defined by the linear constraints $\mathbf{D}^\top \mathbf{x} \leq \mathbf{d}$, $\mathbf{x} \leq x_u \mathbf{e}$, where

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

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

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

Instances 6, 7, 8, 9 and 10 were adopted from Al-Khayyal et al. (1995). Each matrix $\mathbf{P}_i \in \mathbb{R}^{n_x \times n_x}$ 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 $\mathbf{q}_i \in \mathbb{R}^{n_x}$ is also generated with integer entries between -10 and 10 and each r_i is set to 0. In the data used for

these five instances, $D = I$ and $d = 20e$, so their nonnegative variables satisfy $0 \leq x \leq 20e$. The parameters describing each instance are summarized in Table 17.

Table 17 Problem (43) parameters for each instance. n_x refers to the number of variables, L to the number of linear constraints and nc-q to the number of nonconvex quadratic constraints.

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

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

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

$$t_{25} = (0, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275)^\top,$$

$$t_{50} = (0, 50, 100, 150, 200, 250)^\top,$$

$$t_{ir} = (0, 20, 50, 90, 130, 155, 180, 210, 255, 270)^\top.$$

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

Table 18 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

F. Proofs of the convergence results

Throughout this appendix, we use the assumptions and the notation of Section 6, including Assumption 3. We first state and prove the supporting lemmas, and then prove Theorem 3 and Corollary 1.

F.1. Properties of the node relaxations

Let $\mathcal{F}_j$ be the feasible set of the relaxation at node N_j and let $\sigma_j = (\mathbf{x}_j, \boldsymbol{\tau}_j, \mathbf{X}_j, \mathbf{V}_j) \in \mathcal{F}_j$ be the point returned when this relaxation is solved, on which we branch. When the larger LMI of Section 3.3 is included, $\mathcal{F}_j$ and σ_j denote the corresponding projections with the additional block $\mathbf{T}$ suppressed. For $k \in \mathcal{K}_0$, let

$$\Phi_k(\mathbf{x}, \mathbf{X}) := c_{0k}(\mathbf{x}) + \sum_{i \in \mathcal{I}_k} (q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}) c_{ik} \left(\frac{q_{ik} \mathbf{x} - \mathbf{X} \mathbf{d}_{ik}}{q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}} \right) \quad (54)$$

be the perspectified form of f_k from Section 3.2, where each perspective term follows the convention of Section 1 at a zero multiplier, takes the value $+\infty$ for a negative multiplier, and is replaced by its affine expression in $(\mathbf{x}, \mathbf{X})$ when c_{ik} is affine. Since $(\mathbf{x}, \boldsymbol{\tau}) \in \mathcal{T}$ gives $c_{0k}(\mathbf{x}) \leq \tau_k$, every point of $\mathcal{F}_j$ satisfies $\Phi_k(\mathbf{x}, \mathbf{X}) \leq 0$ for $k \in \mathcal{K}$ and has relaxation objective value at least $\Phi_0(\mathbf{x}, \mathbf{X})$. In particular, $p_j \geq \Phi_0(\mathbf{x}_j, \mathbf{X}_j)$.

LEMMA 5 (Properties of the Node Relaxations). *Let N_j be a node generated by Algorithm 2.*

- (a) *For every $k \in \mathcal{K}_0$, Φ_k is lower semicontinuous on $\mathbb{R}^{n_x} \times \mathbb{S}^{n_x}$, and $\Phi_k(\mathbf{x}, \mathbf{x}\mathbf{x}^\top) = f_k(\mathbf{x})$ for all $\mathbf{x} \in \mathcal{X}$. In particular, $f_0, \dots, f_K$ are lower semicontinuous on $\mathcal{X}$, and v^* is attained.*
- (b) *If $\mathbf{x}$ is feasible for Problem (1) with $f_0(\mathbf{x}) < +\infty$ and satisfies the branching and directional-bound inequalities inherited by N_j , then $(\mathbf{x}, \mathbf{c}_0(\mathbf{x}), \mathbf{x}\mathbf{x}^\top, \mathbf{c}_0(\mathbf{x})\mathbf{x}^\top) \in \mathcal{F}_j$ with relaxation objective value $f_0(\mathbf{x})$. Consequently $\ell_j \leq f_0(\mathbf{x})$, and if N_j is branched, the exact lift of $\mathbf{x}$ belongs to $\mathcal{F}_{j_1}$ or to $\mathcal{F}_{j_2}$.*
- (c) *If $\mathcal{F}_j \neq \emptyset$, then for every unit vector $\boldsymbol{\nu}$ the values $\inf\{\boldsymbol{\nu}^\top \mathbf{x} : (\mathbf{x}, \boldsymbol{\tau}, \mathbf{X}, \mathbf{V}) \in \mathcal{F}_j\}$ and $\sup\{\boldsymbol{\nu}^\top \mathbf{x} : (\mathbf{x}, \boldsymbol{\tau}, \mathbf{X}, \mathbf{V}) \in \mathcal{F}_j\}$ are finite and differ by at most D_x . In particular, the bounds (17) satisfy $a_j \leq s_j \leq b_j$ and $b_j - a_j \leq D_x$.*
- (d) *If $\mathbf{R}_j = \mathbf{0}$, then $\mathbf{x}_j$ is feasible for Problem (1) and $f_0(\mathbf{x}_j) \leq p_j$. Consequently, $\text{Ub} \leq \text{Lb}_j + \bar{\varepsilon}$ from the moment N_j has been solved.*

Proof of Lemma 5. (a) Fix $k \in \mathcal{K}_0$ and $i \in \mathcal{I}_k$ with c_{ik} nonlinear. The function $(t, \mathbf{u}) \mapsto t c_{ik}(\mathbf{u}/t)$ for $t > 0$, extended by $\delta_{\text{dom}(c_{ik}^*)}^*(\mathbf{u})$ at $t = 0$ and by $+\infty$ for $t < 0$, is the closure of the perspective of the proper, closed and convex function c_{ik} and is therefore closed and convex (Rockafellar, 1970, Section 8). Its composition with the affine map $(\mathbf{x}, \mathbf{X}) \mapsto (q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}, q_{ik} \mathbf{x} - \mathbf{X} \mathbf{d}_{ik})$ is lower semicontinuous, as is c_{0k} , and the affine expression used for an affine c_{ik} is continuous. All terms take values in $(-\infty, +\infty]$, so their finite sum Φ_k is lower semicontinuous. Now let $\mathbf{x} \in \mathcal{X}$. For a nonlinear c_{ik} , Assumption 2 gives $q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x} \geq 0$. If $q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x} > 0$, then $(q_{ik} \mathbf{x} - \mathbf{x}\mathbf{x}^\top \mathbf{d}_{ik}) / (q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}) = \mathbf{x}$ and the term equals $(q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}) c_{ik}(\mathbf{x})$. If $q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x} = 0$, then $q_{ik} \mathbf{x} - \mathbf{x}\mathbf{x}^\top \mathbf{d}_{ik} = \mathbf{0}$ and the term equals $\delta_{\text{dom}(c_{ik}^*)}^*(\mathbf{0}) = 0$, as in the proof of Lemma 1, which is the value of $(q_{ik} - \mathbf{d}_{ik}^\top \mathbf{x}) c_{ik}(\mathbf{x})$ under the convention $0 \cdot c_{ik}(\mathbf{x}) = 0$ of (2). For an affine c_{ik} , the affine expression evaluated at $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$

is the product itself. Hence $\Phi_k(\mathbf{x}, \mathbf{x}\mathbf{x}^\top) = f_k(\mathbf{x})$. Since $\mathbf{x} \mapsto (\mathbf{x}, \mathbf{x}\mathbf{x}^\top)$ is continuous, f_k is lower semicontinuous on $\mathcal{X}$. The feasible set of Problem (1) is then a nonempty closed subset of the compact set $\mathcal{X}$, and f_0 attains its minimum v^* on it.

(b) Let $\mathbf{x}$ be as stated and set $\boldsymbol{\tau} := \mathbf{c}_0(\mathbf{x})$, $\mathbf{X} := \mathbf{x}\mathbf{x}^\top$, and $\mathbf{V} := \boldsymbol{\tau}\mathbf{x}^\top$. The vector $\boldsymbol{\tau}$ is finite, because $f_k(\mathbf{x}) \leq 0$ for $k \in \mathcal{K}$ and $f_0(\mathbf{x}) < +\infty$, while every other term of f_k is greater than $-\infty$. We verify the constraints of the relaxation at N_j , which consist of the constraints of (8), the optional semidefinite constraint of Section 3.3, the inherited branching and directional-bound inequalities together with their products with themselves, with each other and with the other constraints in $\mathbf{x}$, and the optional inequalities of Remark 4. The constraints in $\mathbf{x}$ alone hold by hypothesis, and $(\mathbf{x}, \boldsymbol{\tau}) \in \mathcal{T}$ by the choice of $\boldsymbol{\tau}$. By part (a), the perspectified constraint of index $k \in \mathcal{K}$ reads $\tau_k + \Phi_k(\mathbf{x}, \mathbf{X}) - c_{0k}(\mathbf{x}) = f_k(\mathbf{x}) \leq 0$, and the objective equals $\tau_0 + \Phi_0(\mathbf{x}, \mathbf{X}) - c_{00}(\mathbf{x}) = f_0(\mathbf{x})$. The diagonal inequalities $X_{ii} \geq 0$ reduce to $x_i^2 \geq 0$. Every generated product of two linear inequalities, including products involving the inherited inequalities, reduces at $\mathbf{X} = \mathbf{x}\mathbf{x}^\top$ to the product of two nonnegative slacks. For a linear inequality $\mathbf{A}_\ell^\top \mathbf{x} \leq b_\ell$ (or an inherited inequality) and a convex constraint $h_r(\mathbf{x}) \leq 0$, the perspectified product equals $(b_\ell - \mathbf{A}_\ell^\top \mathbf{x})h_r(\mathbf{x}) \leq 0$ when $b_\ell - \mathbf{A}_\ell^\top \mathbf{x} > 0$ and $\delta_{\text{dom}(h_r^*)}^*(\mathbf{0}) = 0$ when $b_\ell - \mathbf{A}_\ell^\top \mathbf{x} = 0$, by the computation in part (a). The products with $\mathbf{c}_0(\mathbf{x}) \leq \boldsymbol{\tau}$ reduce in the same way to $(b_\ell - \mathbf{A}_\ell^\top \mathbf{x})\mathbf{c}_0(\mathbf{x}) \leq (b_\ell - \mathbf{A}_\ell^\top \mathbf{x})\boldsymbol{\tau} = b_\ell\boldsymbol{\tau} - \mathbf{V}\mathbf{A}_\ell$. The products with the equality constraints hold because $s_\ell \mathbf{x} - \mathbf{X}\mathbf{P}_\ell = (s_\ell - \mathbf{P}_\ell^\top \mathbf{x})\mathbf{x} = \mathbf{0}$ and $s_\ell \boldsymbol{\tau} - \mathbf{V}\mathbf{P}_\ell = (s_\ell - \mathbf{P}_\ell^\top \mathbf{x})\boldsymbol{\tau} = \mathbf{0}$. Finally, $\begin{pmatrix} \mathbf{X} & \mathbf{x} \\ \mathbf{x}^\top & 1 \end{pmatrix} = \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \begin{pmatrix} \mathbf{x}^\top \\ 1 \end{pmatrix} \succeq \mathbf{0}$ (if the larger LMI of Section 3.3 is included, its block $\mathbf{T}$ is set to $\boldsymbol{\tau}\boldsymbol{\tau}^\top$, and the full matrix equals $(\mathbf{x}^\top, \boldsymbol{\tau}^\top, 1)^\top (\mathbf{x}^\top, \boldsymbol{\tau}^\top, 1) \succeq \mathbf{0}$), and $(\boldsymbol{\nu}^\top \mathbf{x})^2 = \boldsymbol{\nu}^\top \mathbf{X} \boldsymbol{\nu}$ for every $\boldsymbol{\nu}$, which verifies the optional inequalities of Remark 4. Hence the lift belongs to $\mathcal{F}_j$ with objective value $f_0(\mathbf{x})$, and $\ell_j \leq f_0(\mathbf{x})$ by Assumption 3. If N_j is branched, then $\boldsymbol{\nu}_j^\top \mathbf{x} \leq s_j$ or $\boldsymbol{\nu}_j^\top \mathbf{x} \geq s_j$, and $a_j \leq \boldsymbol{\nu}_j^\top \mathbf{x} \leq b_j$ by (17), because the lift belongs to $\mathcal{F}_j$. Thus $\mathbf{x}$ satisfies the inequalities of at least one child, and the first part of the proof applied to that child shows that the lift belongs to its feasible set.

(c) The $\mathbf{x}$ -component of every point of $\mathcal{F}_j$ lies in $\mathcal{X}$, because $(\mathbf{x}, \boldsymbol{\tau}) \in \mathcal{T}$. Hence, for a unit vector $\boldsymbol{\nu}$, the infimum and the supremum of $\boldsymbol{\nu}^\top \mathbf{x}$ over $\mathcal{F}_j$ lie between $\min_{\mathbf{x} \in \mathcal{X}} \boldsymbol{\nu}^\top \mathbf{x}$ and $\max_{\mathbf{x} \in \mathcal{X}} \boldsymbol{\nu}^\top \mathbf{x}$, which are finite, and they differ by at most $\max\{\boldsymbol{\nu}^\top (\mathbf{x} - \mathbf{x}') : \mathbf{x}, \mathbf{x}' \in \mathcal{X}\} \leq D_x$ by the Cauchy–Schwarz inequality. Since $\sigma_j \in \mathcal{F}_j$ and $s_j = \boldsymbol{\nu}_j^\top \mathbf{x}_j$, the bounds (17) satisfy $a_j \leq s_j \leq b_j$.

(d) If $\mathbf{R}_j = \mathbf{0}$, then $\mathbf{X}_j = \mathbf{x}_j \mathbf{x}_j^\top$, and part (a) gives $f_k(\mathbf{x}_j) = \Phi_k(\mathbf{x}_j, \mathbf{X}_j) \leq 0$ for $k \in \mathcal{K}$, so $\mathbf{x}_j$ is feasible, and $f_0(\mathbf{x}_j) = \Phi_0(\mathbf{x}_j, \mathbf{X}_j) \leq p_j$. In particular, $f_0(\mathbf{x}_j)$ is finite. The point $\mathbf{x}_j$ belongs to the candidate set of Section 3.4 and is evaluated when N_j is solved, so that from then on $\text{Ub} \leq f_0(\mathbf{x}_j) \leq p_j \leq \ell_j + \bar{\varepsilon} \leq \text{Lb}_j + \bar{\varepsilon}$ by (28), where the last inequality holds because Lb_j is ℓ_j or an inherited larger bound. $\square$

F.2. Consistency along infinite branches

LEMMA 6 (**Directional Descendant Estimate**). *Let N_k be branched with direction $\boldsymbol{\nu}_k$, level $s_k = \boldsymbol{\nu}_k^\top \mathbf{x}_k$, and bounds $a_k \leq s_k \leq b_k$, and let N_q be a proper descendant of N_k whose relaxation is feasible. Then*

$$|\boldsymbol{\nu}_k^\top \mathbf{R}_q \boldsymbol{\nu}_k| \leq (b_k - a_k) |\boldsymbol{\nu}_k^\top (\mathbf{x}_q - \mathbf{x}_k)| \leq D_x \|\mathbf{x}_q - \mathbf{x}_k\|_2. \quad (55)$$

Proof of Lemma 6. Write $t = \boldsymbol{\nu}_k^\top \mathbf{x}_q$, $y = \boldsymbol{\nu}_k^\top \mathbf{X}_q \boldsymbol{\nu}_k$, and $d = \boldsymbol{\nu}_k^\top (\mathbf{x}_q - \mathbf{x}_k) = t - s_k$, so that $\boldsymbol{\nu}_k^\top \mathbf{R}_q \boldsymbol{\nu}_k = y - t^2$. The descendant N_q inherits the branching inequalities and the generated inequalities of the child of N_k from which it descends, and σ_q satisfies them because $\sigma_q \in \mathcal{F}_q$. The tangent (18) gives $y - t^2 \geq 2s_k t - s_k^2 - t^2 = -d^2$. If N_q descends from the left child, then $a_k \leq t \leq s_k$ and (19) gives $y - t^2 \leq (a_k + s_k)t - a_k s_k - t^2 = (t - a_k)(s_k - t) \leq (b_k - a_k)|d|$. If N_q descends from the right child, then $s_k \leq t \leq b_k$ and (20) gives $y - t^2 \leq (t - s_k)(b_k - t) \leq (b_k - a_k)|d|$. Since t and s_k both lie in $[a_k, b_k]$, also $d^2 \leq (b_k - a_k)|d|$, which proves the first inequality in (55). The second follows from $b_k - a_k \leq D_x$ (Lemma 5(c)) and from $|d| \leq \|\mathbf{x}_q - \mathbf{x}_k\|_2$, which holds by the Cauchy-Schwarz inequality because $\|\boldsymbol{\nu}_k\|_2 = 1$. $\square$

PROPOSITION 1 (**Vanishing Residual Along Infinite Branches**). *If $\{N_r\}_{r \geq 1}$ is an infinite nested branch of nodes, each branched by Algorithm 1, then $\rho_r \rightarrow 0$.*

Proof of Proposition 1. Suppose some subsequence r_j satisfies $\rho_{r_j} \geq \varepsilon > 0$, and set $\mathbf{S}_{r_j} := \mathbf{R}_{r_j} / \rho_{r_j}$, a symmetric matrix of spectral norm one. Compactness of $\mathcal{X}$, of the unit sphere, and of the set of symmetric matrices with spectral norm one yields a further subsequence, not relabelled, such that $\mathbf{x}_{r_j} \rightarrow \bar{\mathbf{x}}$, $\boldsymbol{\nu}_{r_j} \rightarrow \bar{\boldsymbol{\nu}}$, and $\mathbf{S}_{r_j} \rightarrow \bar{\mathbf{S}}$. Node $N_{r_{j+1}}$ is a descendant of N_{r_j} , and $a_{r_j} \leq s_{r_j} \leq b_{r_j}$ by Lemma 5(c), so Lemma 6 applies with $k = r_j$ and $q = r_{j+1}$. Dividing (55) by $\rho_{r_{j+1}} \geq \varepsilon$ gives

$$|\boldsymbol{\nu}_{r_j}^\top \mathbf{S}_{r_{j+1}} \boldsymbol{\nu}_{r_j}| \leq \frac{D_x}{\varepsilon} \|\mathbf{x}_{r_{j+1}} - \mathbf{x}_{r_j}\|_2 \rightarrow 0,$$

because $\mathbf{x}_{r_j}$ and the shifted sequence $\mathbf{x}_{r_{j+1}}$ converge to the same limit $\bar{\mathbf{x}}$. On the other hand, the selection rule of Algorithm 1 gives $|\boldsymbol{\nu}_{r_{j+1}}^\top \mathbf{S}_{r_{j+1}} \boldsymbol{\nu}_{r_{j+1}}| = |\lambda_{r_{j+1}}| / \rho_{r_{j+1}} = 1$ for every j . By continuity of $(\boldsymbol{\nu}, \mathbf{S}) \mapsto \boldsymbol{\nu}^\top \mathbf{S} \boldsymbol{\nu}$, both left-hand sides converge to $|\bar{\boldsymbol{\nu}}^\top \bar{\mathbf{S}} \bar{\boldsymbol{\nu}}|$, which would have to equal both 0 and 1. Hence no such subsequence exists and $\rho_r \rightarrow 0$. $\square$

LEMMA 7 (**Rank-one Consistency**). *Let $\{N_r\}_{r \geq 1}$ be an infinite nested branch with $\rho_r \rightarrow 0$, and let $\bar{\mathbf{x}}$ be the limit of a subsequence $\{\mathbf{x}_{r_\ell}\}$ of the relaxation solutions. Then $\bar{\mathbf{x}}$ is feasible for Problem (1) and $f_0(\bar{\mathbf{x}}) \leq \liminf_{\ell \rightarrow \infty} p_{r_\ell}$.*

Proof of Lemma 7. We have $\bar{\mathbf{x}} \in \mathcal{X}$ by closedness, and $\mathbf{X}_{r_\ell} = \mathbf{x}_{r_\ell} \mathbf{x}_{r_\ell}^\top + \mathbf{R}_{r_\ell} \rightarrow \bar{\mathbf{x}} \bar{\mathbf{x}}^\top$ because $\rho_r \rightarrow 0$. Every point of $\mathcal{F}_r$ satisfies $\Phi_k(\mathbf{x}_r, \mathbf{X}_r) \leq 0$ for $k \in \mathcal{K}$, so lower semicontinuity and Lemma 5(a) give

$$f_k(\bar{\mathbf{x}}) = \Phi_k(\bar{\mathbf{x}}, \bar{\mathbf{x}} \bar{\mathbf{x}}^\top) \leq \liminf_{\ell \rightarrow \infty} \Phi_k(\mathbf{x}_{r_\ell}, \mathbf{X}_{r_\ell}) \leq 0,$$

so $\bar{\mathbf{x}}$ is feasible for Problem (1). For the objective, $p_r \geq \Phi_0(\mathbf{x}_r, \mathbf{X}_r)$, and the same argument gives $f_0(\bar{\mathbf{x}}) = \Phi_0(\bar{\mathbf{x}}, \bar{\mathbf{x}}\bar{\mathbf{x}}^\top) \leq \liminf_\ell \Phi_0(\mathbf{x}_{r_\ell}, \mathbf{X}_{r_\ell}) \leq \liminf_\ell p_{r_\ell}$. $\square$

F.3. Proofs of the main guarantees

Proof of Theorem 3. We first record two facts about the bounds. Every bound Lb_i stored by Algorithm 2 is a lower bound on the relaxation objective value of every point of $\mathcal{F}_i$. For ℓ_i this is Assumption 3, and an inherited bound is such a bound for the parent, whose relaxation contains that of the child. Every incumbent is feasible by construction (Section 3.4), so $\text{Ub} \geq v^*$ throughout.

Let $\mathbf{x}^\circ$ be a global optimizer, which exists by Lemma 5(a). By Lemma 5(b), the root node contains the exact lift of $\mathbf{x}^\circ$, and whenever a node containing this lift is branched, one of its children contains it as well. Call such a node an optimal node. Its relaxation is feasible, so it is not reported infeasible, and its bound is at most v^* , because the exact lift of $\mathbf{x}^\circ$ belongs to its feasible set with relaxation objective value $f_0(\mathbf{x}^\circ) = v^*$. An optimal node is discarded in one way only, namely when it is not added to the active list because its bound is at least Ub at that time. Then $\text{Ub} \leq v^*$, hence $\text{Ub} = v^*$, from then on. Hence, at initialization and after every completed iteration, either an optimal node is active, so that $\text{Lb} \leq v^*$, or $\text{Ub} = v^*$. In both cases $\text{Lb} \leq v^*$, at initialization because $\text{Lb}_0 = \ell_0 \leq v^*$, and afterwards because the update rule of Algorithm 2 takes the minimum with Ub .

(a) If the active list is empty at termination, then the preceding argument and the update rule give $\text{Lb} = \text{Ub} = v^*$. Otherwise, the stopping test gives $\text{Ub} - \text{Lb} \leq \delta$, and Ub is finite because Lb is. Since the incumbent is feasible, $\text{Ub} = f_0(\mathbf{x}^{\text{inc}})$, and in either case

$$0 \leq f_0(\mathbf{x}^{\text{inc}}) - v^* \leq \text{Ub} - \text{Lb} \leq \delta.$$

In particular, the incumbent is globally optimal when $\delta = 0$.

(b) Let L_s be the value of Lb immediately before the s -th node selection. While the run continues, $\text{Ub} - L_s > \delta \geq 0$, so L_s is the smallest active bound, that is, the bound of the node selected at step s , and the previous paragraph gives $L_s \leq v^*$. The children of the selected node inherit its bound, so their bounds are at least L_s , the bounds of the other active nodes are at least L_s as well, and taking the minimum with $\text{Ub} \geq v^* \geq L_s$ cannot decrease it. Hence L_s is nondecreasing and converges to some $L \leq v^*$.

Suppose $L < v^* - \bar{\varepsilon}$. Every selected node is branched, so an infinite run generates an infinite binary tree of selected nodes, which is closed under taking ancestors, and König's lemma yields an infinite nested branch $\{N_r\}$ of selected nodes. Node N_r is selected at some step s_r with $s_r \rightarrow \infty$, and its bound is $\text{Lb}_r = L_{s_r} \rightarrow L$. By compactness of $\mathcal{X}$, $\{\mathbf{x}_r\}$ has a subsequence converging to some $\bar{\mathbf{x}}$, which by Proposition 1 and Lemma 7 is feasible with

$$f_0(\bar{\mathbf{x}}) \leq \liminf_\ell p_{r_\ell} \leq \liminf_\ell (\ell_{r_\ell} + \bar{\varepsilon}) \leq \liminf_\ell (\text{Lb}_{r_\ell} + \bar{\varepsilon}) = L + \bar{\varepsilon} < v^*,$$

where we used (28) and $\ell_r \leq \text{Lb}_r$. This contradicts the optimality of v^* . Hence $L \geq v^* - \bar{\varepsilon}$.

Finally, let $\{N_r\}$ be an infinite nested branch. Each of its nodes is branched and therefore selected, at some step $s_r \rightarrow \infty$, with $\text{Lb}_r = L_{s_r} \rightarrow L$. By Proposition 1 and Lemma 7, every accumulation point $\bar{\mathbf{x}}$ of $\{\mathbf{x}_r\}$ is feasible and, as above, satisfies $f_0(\bar{\mathbf{x}}) \leq L + \bar{\varepsilon} \leq v^* + \bar{\varepsilon}$. $\square$

Proof of Corollary 1. Suppose, to the contrary, that the run is infinite, and let L_s and U_s denote the global lower and upper bounds immediately before the s -th node selection. The stopping criterion gives $U_s - L_s > \delta$ for every s , and Theorem 3(b) gives $L_s \rightarrow L$. As in the proof of Theorem 3(b), the run contains an infinite nested branch $\{N_r\}$ of selected nodes with $\text{Lb}_r = L_{s_r} \rightarrow L$, and by compactness of $\mathcal{X}$, Proposition 1 and Lemma 7, the relaxation solutions along this branch have a subsequence $\mathbf{x}_{r_\ell} \rightarrow \bar{\mathbf{x}}$ with $\bar{\mathbf{x}}$ feasible and $f_0(\bar{\mathbf{x}}) \leq L + \bar{\varepsilon}$.

Since $\mathcal{K} = \emptyset$, each relaxation solution $\mathbf{x}_j \in \mathcal{X}$ is feasible for the original problem and has finite objective value. It belongs to the candidate set of Section 3.4 and is evaluated when its node is solved, before the node is selected. Consequently, $U_{s_{r_\ell}} \leq f_0(\mathbf{x}_{r_\ell})$, and

$$\delta < U_{s_{r_\ell}} - L_{s_{r_\ell}} \leq f_0(\mathbf{x}_{r_\ell}) - \text{Lb}_{r_\ell} \rightarrow f_0(\bar{\mathbf{x}}) - L \leq \bar{\varepsilon},$$

where continuity of f_0 gives $f_0(\mathbf{x}_{r_\ell}) \rightarrow f_0(\bar{\mathbf{x}})$. This contradicts $\bar{\varepsilon} < \delta$. The run therefore terminates, and δ -optimality follows from Theorem 3(a). $\square$