Convexification of a Separable Function over a Polyhedral Ground Set

Santanu S. Dey* Burak Kocuk†

October 18, 2025

Abstract

In this paper, we study the set $\mathcal{S}^\kappa = \{(x,y) \in \mathcal{G} \times \mathbb{R}^n : y_j = x_j^\kappa, j = 1,\dots,n\}$, where $\kappa > 1$ and the *ground set* $\mathcal{G}$ is a nonempty polytope contained in $[0,1]^n$. This nonconvex set is closely related to separable standard quadratic programming and appears as a substructure in potential-based network flow problems from gas and water networks. Our aim is to obtain the convex hull of $\mathcal{S}^\kappa$ or its tight outer-approximation for the special case when the ground set $\mathcal{G}$ is the standard simplex. We propose power cone, second-order cone and semidefinite programming relaxations for this purpose, which are further strengthened by the Reformulation-Linearization Technique and the Reformulation-Perspectification Technique. For $\kappa = 2$, we obtain the convex hull of $\mathcal{S}^\kappa$ in the low-dimensional setting. For general κ, we give approximation guarantees for the power cone representable relaxation, the weakest relaxation we consider. We prove that this weakest relaxation is tight with probability one as $n \to \infty$ when a uniformly generated linear objective is optimized over it. Finally, we provide the results of our extensive computational experiments comparing the empirical strength of several conic programming relaxations that we propose.

Keywords: convexification, conic programming, power cone, semidefinite programming, reformulation-linearization technique, reformulation-perspectification technique

1 Introduction

Consider the set

$$\mathcal{S}^\kappa = \{(x,y) \in \mathcal{G} \times \mathbb{R}^n : y_j = x_j^\kappa, j = 1,\dots,n\},$$

where $\kappa > 1$ and the *ground set* $\mathcal{G}$ is the following nonempty polytope

$$\mathcal{G} = \{x \in [0,1]^n : \ Ax = b, \ Cx \le d\}.$$

Here, $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$, $C \in \mathbb{R}^{k\times n}$ and $d \in \mathbb{R}^k$. Notice that this set is closely related to the convexification of the separable function $\sum_{j=1}^n(\alpha_j x_j + \beta_j x_j^\kappa)$ over a polyhedral ground set $\mathcal{G}$ and the optimization problem $\min_{(x,y)\in\mathcal{S}^\kappa}\{\sum_{j=1}^n(\alpha_j x_j + \beta_j y_j)\}$. This substructure appears in many applications and we provide two motivating examples to study the set $\mathcal{S}^\kappa$ below.

The first example arises from a special case of the well-studied standard quadratic programming problem [6, 8, 22, 16, 15, 19], called the *separable* standard quadratic programming [7], which is an optimization problem of the form

$$\min_{x\in\Delta^n} \left\{ \sum_{j=1}^n (\alpha_j x_j + \beta_j x_j^2) \right\}, \tag{1}$$

*H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology (santanu.dey@isye.gatech.edu).

†Industrial Engineering Program, Sabancı University (burakkocuk@sabanciuniv.edu).

where $\alpha \in \mathbb{R}^n$, $\beta \in \mathbb{R}^n$ and the set $\Delta^n := \left\{x \in \mathbb{R}^n_+ : \sum_{j=1}^n x_j = 1\right\}$ is the standard simplex in $\mathbb{R}^n$. In our notation, problem (1) is equivalent to $\min_{(x,y)\in\mathcal{S}^2}\{\alpha^T x + \beta^T y\}$, with the ground set $\mathcal{G} = \Delta^n$. Therefore, the set we study is a generalization of the extended formulation of the feasible region of an important problem.

The second example arises from potential-based flow networks (e.g., gas and water networks [17, 27, 20, 13, 25, 21, 9]). Let us consider a "node-based" substructure, where we have a node set $\{0, 1, \ldots, n\}$ and an edge set $\{(1, 0), \ldots, (n, 0)\}$; see Figure 1 for an illustration with $n = 4$ (we also refer the reader to [12] for a similar substructure studied in the context of power systems).

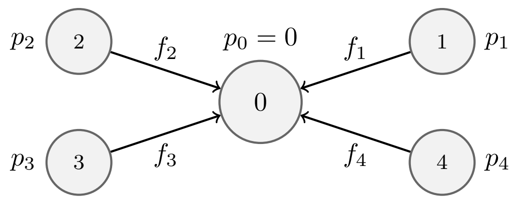


Figure 1: An example node-based structure with $n = 4$ adjacent edges to node 0.

We will think of node 0 as a demand node with a positive demand δ and the others as supply nodes. Each node j has a nonnegative *potential*, denoted by p_j, $0 = 1, \ldots, n$, and we will assume that the demand node 0 has zero potential, i.e., $p_0 = 0$. Each edge has a nonnegative *flow*, denoted by f_j, which is upper bounded by $\bar{f}_j$, $j = 1, \ldots, n$. The relationship between the potential differences and the flow value is governed by the law of physics and is of the form $p_j - p_0 = \chi_j f_j^\kappa$, where χ_j is a positive number representing a certain physical characteristic of edge j. In addition, the relation $\sum_{j=1}^n f_j = \delta$ models the demand constraint. By defining $x_j = \frac{f_j}{\bar{f}_j}$ and $y_j = \frac{p_j}{\chi_j \bar{f}_j^\kappa}$, we can model this situation in the form of $\mathcal{S}^\kappa$, where the ground set is $\mathcal{G} = \{x \in [0, 1]^n : \sum_{j=1}^n \bar{f}_j x_j = \delta\}$. Note that in the special case where we have $\chi_j = \bar{f}_j = \delta = 1$ for $j = 1, \ldots, n$, the ground set is again the standard simplex, that is, $\mathcal{G} = \Delta^n$. As opposed to our illustration above, the directions of flows on the edges are typically not fixed in general. However, the analysis of the set $\mathcal{S}^\kappa$ is still useful as one can deal with a more general case in which the flow directions are variables via a disjunctive formulation that involves several sets of the form $\mathcal{S}^\kappa$.

In this study, our aim is to find $\mathrm{conv}(\mathcal{S}^\kappa)$ or a close outer-approximation of $\mathrm{conv}(\mathcal{S}^\kappa)$ in the space of x and y variables or in an extended space. Although the convexification of functions is an active research area (see, e.g., [30, 26, 28, 31, 4, 24, 3, 2, 32, 1, 18, 33]), the specific substructure we are interested in is less explored. Arguably, references [10], [7] and [15] are most closely related to our study although they only consider the case where the exponent is $\kappa = 2$. For instance, the results in [10] imply that $\mathrm{conv}(\mathcal{S}^2)$ can be obtained as a completely positive cone representable set although this representation is not tractable in general. It is proven in [7] that the optimizing a linear function over $\mathcal{S}^2$ can be done is polynomial-time although an explicit convex hull description is not provided. In [15], it is shown that optimizing a linear function over $\mathcal{S}^2$ can be performed by solving a doubly nonnegative relaxation when β vector is either nonnegative or nonpositive.

In this paper, we provide several conic representable outer-approximations for the set $\mathcal{S}^\kappa$ or equivalently conic programming relaxations for the optimization problem $\min_{(x,y)\in\mathcal{S}^\kappa}\{\sum_{j=1}^n(\alpha_j x_j + \beta_j y_j)\}$. In particular, we propose a power cone relaxation (hereafter abbreviated as the **P** relaxation) obtained as a *single row relaxation*, that is, the nonconvex relation $y_j = x_j^\kappa$ over $x_j \in [0, 1]$ is simply relaxed as $x_j^\kappa \le y_j \le x_j$. We also construct stronger relaxations using reformulation-linearization technique (RLT), second-order cone programming and semidefinite programming. Then, we specialize our analysis to the important special case when the ground set is the standard simplex. For this setting, we provide convex hull results in low-dimension for $\kappa = 2$, and approximation guarantees and a probabilistic tightness analysis for general κ. We further improve our relaxations using the reformulation-perspectification technique (RPT). Finally, we run an extensive set of computational experiments focusing on the case when the ground set is the standard simplex and provide empirical evidence of the comparative strength of the nine relaxations we propose.

Below, we list our main contributions and key results from each section:

- In Section 2, we prove that optimizing a linear function over $\mathcal{S}^\kappa$ is NP-hard in Proposition 1 and propose several conic programming relaxations.
- In Section 3, we analyze the case where the ground set is the standard simplex.
  - For the special case $\kappa = 2$, we provide two convex hull results in low-dimensional setting (Theorem 1 for $n = 2$ and Corollary 2 for $n = 3$) and give some sufficient conditions under which a doubly nonnegative relaxation of the problem $\min_{(x,y)\in\mathcal{S}^\kappa}\{\sum_{j=1}^n(\alpha_j x_j + \beta_j y_j)\}$ is exact (Theorem 3) for general n.
  - For general κ, we provide both distance-based (Propositions 8 and 9) and objective-function based (Proposition 12) approximation results for the $\mathbf{P}$ relaxation. Interestingly, this weakest relaxation we consider is tight with high probability when the dimension n is sufficiently large, as proven in Theorem 4.
- In Section 4, we report the results of our computational experiments and verify several key observations from the previous sections empirically.

Notation: Throughout the paper, $\mathbb{R}^n$, $\mathbb{R}^n_+$, $\mathbb{S}^n$, $\mathbb{S}^n_+$, $\mathbb{CP}^n$ are respectively the set of n-dimensional real vectors, the set of n-dimensional nonnegative real vectors, the set of $n \times n$ real symmetric matrices, the set of $n \times n$ real positive semidefinite matrices, the set of $n \times n$ completely positive matrices, and e is the all-ones vector of appropriate dimension. Also, let e^j be the j-th standard unit vector. Given a square matrix X, we denote the vector of diagonal of X by $\mathrm{diag}(X)$. For a set $\mathcal{S} \subseteq \mathbb{R}^n$, we denote its convex hull and the set of extreme points as $\mathrm{conv}(\mathcal{S})$ and $\mathrm{extr}(\mathcal{S})$, respectively.

2 General Ground Sets

In this section, we focus on a general ground set $\mathcal{G}$. We first prove that it is NP-Hard to optimize a linear function over the set $\mathcal{S}^\kappa$ in Section 2.1. Then, we propose several conic relaxations for the set $\mathcal{S}^\kappa$ in Section 2.2.

2.1 Complexity

Proposition 1. *Optimizing a linear function over $\mathcal{S}^\kappa$ is NP-Hard for $\kappa > 1$.*

Proof. Consider the SUBSET-SUM Problem, which is known to be NP-Hard [14]: Given $a \in \mathbb{Z}^n_+$ and $b \in \mathbb{Z}_+$, does there exist $J \subseteq \{1, \ldots, n\}$ such that $\sum_{j\in J} a_j = b$?

Consider the following problem, which minimizes a linear function over the set $\mathcal{S}^\kappa$:

$$(Q): \min_{x\in[0,1]^n, y\in\mathbb{R}^n} \left\{ \sum_{j=1}^n (x_j - y_j) : a^T x = b,\ y_j = x_j^\kappa, j = 1, \ldots, n \right\}.$$

Notice that a SUBSET-SUM instance is feasible if and only if the optimal value of (Q) is zero. Hence, the result follows. □

Although optimizing a linear function over the set $\mathcal{S}^\kappa$ is NP-Hard in general, there is a known polynomially solvable case due to [7] when $\kappa = 2$ and $(a, b) = (e, 1)$, that is, $\mathcal{G} = \Delta^n$. We note that the algorithm proposed in this reference can be adapted to any $\kappa > 1$.

2.2 Relaxations

In this section, we propose several conic relaxations for the set $\mathcal{S}^\kappa$. We start with a definition.

Definition 1 (Power cone)**.** *The power cone in* $\mathbb{R}^3$ *parameterized with* $\gamma \in (0,1)$ *is defined as*

$$\mathcal{C}^\gamma := \{x \in \mathbb{R}^2_+ \times \mathbb{R} : x_1^\gamma x_2^{1-\gamma} \geq |x_3|\}.$$

For example, the set defined by inequalities $y \geq x^\kappa, y \geq 0$ is equivalent to $(y,1,x) \in \mathcal{C}^{1/\kappa}$ for $\kappa > 1$. In the special case with $\kappa = 2$, $\mathcal{C}^{1/2}$ is the rotated second-order cone in $\mathbb{R}^3$.

The simplest relaxation we construct for the set $\mathcal{S}^\kappa$ involves applying a *single row relaxation* for each nonconvex constraint separately and it utilizes the power cone as defined above to convexify each constraint. We will call this relaxation as *the power cone relaxation* (or the **P** relaxation, in short) and define it as below:

$$\mathcal{S}^\kappa_P = \{(x,y) \in \mathcal{G} \times \mathbb{R}^n : x_j \geq y_j \geq x_j^\kappa, j = 1, \dots, n\}.$$

The **P** relaxation can be further strengthened using other approaches once a lifted matrix variable $X = xx^T \in \mathbb{S}^n$ is used. Below, we will introduce three of them.

The first approach to strengthen the **P** relaxation utilizes RLT [29]. For this purpose, let us consider the following set of constraints obtained by using the first level in the RLT hierarchy

$$\max\{x_i + x_j - 1, 0\} \leq X_{ij} \leq \min\{x_i, x_j\} \qquad 1 \leq i \leq j \leq n \tag{2a}$$

$$AX - bx^T = 0, \quad Axe^T - AX - be^T + bx^T = 0 \tag{2b}$$

$$CX - dx^T \leq 0, \quad Cxe^T - CX - de^T + dx^T \leq 0 \tag{2c}$$

$$AXA^T - Axb^T - bx^TA^T + bb^T = 0 \tag{2d}$$

$$CXC^T - Cxd^T - dx^TC^T + dd^T \geq 0 \tag{2e}$$

$$AXC^T - Axd^T - bx^TC^T + bd^T = 0. \tag{2f}$$

The second approach to strengthen the **P** relaxation utilizes semidefinite programming and involves the following constraints:

$$X \geq 0, \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \succeq 0. \tag{3}$$

The third approach to strengthen the **P** relaxation utilizes second-order cone programming and involves adding the positive semidefinite conditions for the 2×2 minors of the matrix $\begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix}$ only. For this purpose, we consider the constraints

$$X \geq 0, \; x_j^2 \leq X_j \leq x_j, j = 1, \dots, n, \; X_{ij}^2 \leq X_{ii}X_{jj}, \; 1 \leq i < j \leq n. \tag{4}$$

In order to make the connection between the **P** relaxation and the three sets of constraints defined above (i.e., constraints (2), (3) and (4)) stronger, we make a simple observation: Note that we have $X_{jj} = x_j^2$ and $y_j = x_j^\kappa$, implying that $X_{jj} = y_j^{2/\kappa}$. This observation allows us to relate y_j variables with X_{jj} variables through the following constraints:

$$\begin{cases} y_j \geq X_{jj} \geq y_j^{2/\kappa}, j = 1, \dots, n & \text{if } \kappa < 2 \\ X_{jj} = y_j, j = 1, \dots, n & \text{if } \kappa = 2 \; . \\ X_{jj} \geq y_j \geq X_{jj}^{\kappa/2}, j = 1, \dots, n & \text{if } \kappa > 2 \end{cases} \tag{5}$$

We are ready to formally introduce the six relaxations we define for the set $\mathcal{S}^\kappa$:

- The **P** relaxation: $\mathcal{S}^\kappa_P$

- The **PR** relaxation: $\mathcal{S}^{\kappa}_{P,R} := \{(x, y) \in \mathcal{S}^{\kappa}_{P} : \exists X \in \mathbb{S}^n : (2), (5)\}$
- The **PS** relaxation: $\mathcal{S}^{\kappa}_{P,S} := \{(x, y) \in \mathcal{S}^{\kappa}_{P} : \exists X \in \mathbb{S}^n : (3), (5)\}$
- The **Ps** relaxation: $\mathcal{S}^{\kappa}_{P,s} := \{(x, y) \in \mathcal{S}^{\kappa}_{P} : \exists X \in \mathbb{S}^n : (4), (5)\}$
- The **PRS** relaxation: $\mathcal{S}^{\kappa}_{P,R,S} := \{(x, y) \in \mathcal{S}^{\kappa}_{P} : \exists X \in \mathbb{S}^n : (2), (3), (5)\}$
- The **PRs** relaxation: $\mathcal{S}^{\kappa}_{P,R,s} := \{(x, y) \in \mathcal{S}^{\kappa}_{P} : \exists X \in \mathbb{S}^n : (2), (4), (5)\}$

Let us now discuss some basic properties of these relaxations. The lemma below characterizes the extreme points of the **P** relaxation:

Lemma 1. *Let $(\hat{x}, \hat{y}) \in \operatorname{extr}(\mathcal{S}^{\kappa}_{P})$. Then, we have $\hat{y}_j \in \{\hat{x}_j, \hat{x}^{\kappa}_j\}$, $j = 1, \ldots, n$.*

Interestingly, adding the SDP constraint (3) to the **P** relaxation does not further strengthen this relaxation as proven below.

Proposition 2. *Let $\kappa > 1$. Then, $\mathcal{S}^{\kappa}_{P} = \mathcal{S}^{\kappa}_{P,S}$.*

Proof. Note that $\mathcal{S}^{\kappa}_{P} \supseteq \mathcal{S}^{\kappa}_{P,S}$. Suppose that $(\hat{x}, \hat{y}) \in \operatorname{extr}(\mathcal{S}^{\kappa}_{P})$. In order to prove the assertion of the proposition, it suffices to show that $(\hat{x}, \hat{y}) \in \mathcal{S}^{\kappa}_{P,S}$. Due to Lemma 1, we know that $\hat{y}_j \in \{\hat{x}_j, \hat{x}^{\kappa}_j\}$, $j = 1, \ldots, n$. Now, consider the matrix $\hat{X} := \hat{x}\hat{x}^T + D$, where $D \in \mathbb{S}^n$ is a diagonal matrix with entries

$$D_{jj} = \begin{cases} 0 & \text{if } \hat{y}_j = \hat{x}^{\kappa}_j \\ \hat{y}_j - \hat{x}^2_j & \text{if } \hat{y}_j = \hat{x}_j \text{ and } \kappa \leq 2 \\ \hat{y}^{2/\kappa}_j - \hat{x}^2_j & \text{if } \hat{y}_j = \hat{x}_j \text{ and } \kappa > 2 \end{cases}.$$

Observe that both $\hat{x}\hat{x}^T$ and D are doubly nonnegative matrices and $(\hat{x}, \hat{y}, \hat{X})$ satisfies (5). Hence, the result follows. ☐

Corollary 1. *Let $\kappa > 1$. Then, $\mathcal{S}^{\kappa}_{P} = \mathcal{S}^{\kappa}_{P,s}$.*

Due to Proposition 2 and Corollary 1, we will not study the **PS** relaxation and the **Ps** relaxation.

Unlike the SDP constraint (3), adding the RLT constraints (2) strengthens the **P** relaxation. Perhaps more interestingly, the addition of the SDP constraint (3) and the RLT constraints (2) together further strengthens the **PR** relaxation. The proposition below formalizes this statement.

Proposition 3. *We have $\mathcal{S}^{\kappa}_{P} \supseteq \mathcal{S}^{\kappa}_{P,R} \supseteq \mathcal{S}^{\kappa}_{P,R,s} \supseteq \mathcal{S}^{\kappa}_{P,R,S}$, and all the set containment relations can be strict.*

We prove this proposition in Section 3.1 with suitable examples with $\mathcal{G} = \Delta^n$ and $\kappa = 2$.

3 The Standard Simplex as the Ground Set

Let $\mathcal{T}$ be a relaxation of the nonconvex set $\mathcal{S}^{\kappa}$. We will use two metrics to measure the quality of the relaxation $\mathcal{T}$ for the nonconvex set $\mathcal{S}^{\kappa}$:

- Distance-based:
$$D_{\mathcal{S}^{\kappa},\mathcal{T}} := \max_{(\hat{x},\hat{y}) \in \mathcal{T}} \min_{(x,y) \in \mathcal{S}^{\kappa}} \|(x - \hat{x}, y - \hat{y})\|.$$
This measure quantifies the distance of the farthest point in $\mathcal{T}$ from $\mathcal{S}^{\kappa}$.
- Objective function-based: Let $(\alpha, \beta) \in \mathbb{R}^n \times \mathbb{R}^n$.
$$O_{\mathcal{S}^{\kappa},\mathcal{T}}(\alpha, \beta) := \min_{(x,y) \in \mathcal{S}^{\kappa}} \{\alpha^T x + \beta^T y\} - \min_{(x,y) \in \mathcal{T}} \{\alpha^T x + \beta^T y\}.$$
This measure quantifies the additive gap between optimizing over $\mathcal{T}$ and $\mathcal{S}^{\kappa}$.

Remark 1. *Notice that if $O_{\mathcal{S}^\kappa,\mathcal{T}}(\alpha,\beta) = 0$ for all (α,β), then we conclude that $\mathcal{T} = \mathrm{conv}(\mathcal{S}^\kappa)$.*

A simple result that we will frequently use in the rest of the paper is the following proposition.

Proposition 4. $\min_{(x,y)\in\mathcal{S}_P^\kappa}\{\alpha^T x + \beta^T y\} = \min_{x\in\mathcal{G}}\{\sum_{j:\beta_j>0}(\alpha_j x_j + \beta_j x_j^\kappa) + \sum_{j:\beta_j\le 0}(\alpha_j + \beta_j)x_j\}$.

Note that the proof of Proposition 4 immediately follows from the extreme point description of the set $\mathrm{extr}(\mathcal{S}_P^\kappa)$ (the feasible region of the **P** relaxation) given in Lemma 1.

The remainder of this section is organized as follows. In Section 3.1, we prove Proposition 3 for the case $\mathcal{G} = \Delta^n$. Section 3.2 presents several cases where the exact convex hull can be obtained when $\kappa = 2$. In Section 3.3, we establish results that assess the strength of $\mathcal{S}_P^\kappa$ in approximating $\mathcal{S}^\kappa$, with our main result showing that $\mathcal{S}_P^\kappa$ provides an increasingly tight relaxation as $n \to \infty$. Section 3.4 introduces another class of conic-representable inequalities for $\mathcal{S}^\kappa$, derived using RPT.

3.1 Set containment examples

We now prove the strict set inclusion relations stated in Proposition 3. We start with a lemma, which simplifies the RLT constraints when the ground set is the standard simplex.

Lemma 2. *Let $\mathcal{G} = \Delta^n$. Then,*

$$\{x \in \Delta^n, X \in \mathbb{S}^n : (2)\} = \{x \in \Delta^n, X \in \mathbb{S}^n : (6)\},$$

where

$$X \ge 0,\ \ Xe = x. \tag{6}$$

Moreover, if $n \le 3$ and $\kappa = 2$, we have

$$\mathcal{S}_{P,R}^\kappa = \left\{(x,y) \in \mathcal{S}_P^\kappa :\ x_j - y_j \le \sum_{i\neq j}(x_i - y_i),\ j = 1,\dots,n\right\}.$$

Proof. For $\mathcal{G} = \Delta^n$, we have $A = e^T$, $b = 1$ and $C = d = 0$.

In order to prove the first result, notice that equations (2b) give the equation $Xe = x$. Together with the condition $x \in \Delta^n$, we observe that the McCormick envelopes given in (2a) are redundant except $X \ge 0$. Therefore, the first result follows.

In order to prove the second result for $n = 2$, observe that we have

$$\begin{aligned}
\mathcal{S}_{P,R}^2 &= \{(x,y) \in \mathcal{S}_P^2 : \exists X \in \mathbb{S}^2 : X \ge 0,\ Xe = x,\ \mathrm{diag}(X) = y\}\\
&= \{(x,y) \in \mathcal{S}_P^2 : \exists X \in \mathbb{S}^2 : X \ge 0,\ y_1 + X_{12} = x_1, X_{12} + x_2 = y_2\}\\
&= \{(x,y) \in \mathcal{S}_P^2 : x_1 - y_1 = x_2 - y_2\}.
\end{aligned}$$

In order to prove the second result for $n = 3$, observe that we have

$$\begin{aligned}
\mathcal{S}_{P,R}^2 &= \{(x,y) \in \mathcal{S}_P^2 : \exists X \in \mathbb{S}^3 :\ X \ge 0,\ Xe = x, \mathrm{diag}(X) = y\}\\
&= \{(x,y) \in \mathcal{S}_P^2 : \exists X \in \mathbb{S}^3 :\ X \ge 0,\ y_1 + X_{12} + X_{13} = x_1,\\
&\qquad\qquad X_{12} + y_2 + X_{23} = x_2, X_{13} + X_{23} + y_3 = x_3\}\\
&= \{(x,y) \in \mathcal{S}_P^2 : (x_3 - y_3) - (x_1 - y_1) - (x_2 - y_2) \le 0,\\
&\qquad\qquad (x_2 - y_2) - (x_1 - y_1) - (x_3 - y_3) \le 0,\\
&\qquad\qquad (x_1 - y_1) - (x_2 - y_2) - (x_3 - y_3) \le 0\},
\end{aligned}$$

where the last equality follows from the fact that the values X_{12}, X_{13} and X_{23} are uniquely determined given x and y. □

Remark 2. *By virtue of the first result in Lemma 2, we can replace the generic RLT constraints* (2) *with constraints* (6) *when* $\mathcal{G} = \Delta^n$ *for* ***PR***, ***PRS*** *and* ***PRs*** *relaxations. More precisely, we have the following:*

- $\mathcal{S}^{\kappa}_{P,R} := \{(x, y) \in \mathcal{S}^{\kappa}_{P} : \exists X \in \mathbb{S}^n : (6), (5)\}$
- $\mathcal{S}^{\kappa}_{P,R,S} := \{(x, y) \in \mathcal{S}^{\kappa}_{P} : \exists X \in \mathbb{S}^n : (6), (3), (5)\}$
- $\mathcal{S}^{\kappa}_{P,R,s} := \{(x, y) \in \mathcal{S}^{\kappa}_{P} : \exists X \in \mathbb{S}^n : (6), (4), (5)\}$

The following result shows that the **PR** relaxation can be strictly stronger than the **P** relaxation.

Proposition 5. *The relation* $\mathcal{S}^{\kappa}_{P} \supseteq \mathcal{S}^{\kappa}_{P,R}$ *can be strict.*

Proof. Let $\kappa = 2$ and consider $\mathcal{G} = \Delta^2$. Consider the point $(\hat{x}, \hat{y})$, where

$$\hat{x} = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix} \text{ and } \hat{y} = \begin{bmatrix} 1/2 \\ 1/4 \end{bmatrix}.$$

Due to Lemma 2, we deduce that $(\hat{x}, \hat{y}) \in \mathcal{S}^{\kappa}_{P}$ but $(\hat{x}, \hat{y}) \notin \mathcal{S}^{\kappa}_{P,R}$. □

The following result shows that the **PRs** relaxation can be strictly stronger than the **PR** relaxation.

Proposition 6. *The relation* $\mathcal{S}^{\kappa}_{P,R} \supseteq \mathcal{S}^{\kappa}_{P,R,s}$ *can be strict.*

Proof. Let $\kappa = 2$ and consider $\mathcal{G} = \Delta^3$. Consider the point $(\hat{x}, \hat{y})$, where

$$\hat{x} = \begin{bmatrix} 1/3 \\ 1/3 \\ 1/3 \end{bmatrix} \text{ and } \hat{y} = \begin{bmatrix} 1/9 \\ 1/9 \\ 1/3 \end{bmatrix}.$$

Due to Lemma 2, we deduce that this point belongs to $\mathcal{S}^{\kappa}_{P,R}$ with the following *unique* selection of $\hat{X}$:

$$\hat{X} = \begin{bmatrix} 1/9 & 2/9 & 0 \\ 2/9 & 1/9 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}.$$

However, $(\hat{x}, \hat{y}) \notin \mathcal{S}^{\kappa}_{P,R,s}$ since the inequality $X_{11}X_{22} \geq X_{12}^2$ is violated. □

The following result shows that the **PRS** relaxation can be strictly stronger than the **PRs** relaxation.

Proposition 7. *The relation* $\mathcal{S}^{\kappa}_{P,R,s} \supseteq \mathcal{S}^{\kappa}_{P,R,S}$ *can be strict.*

Proof. Let $\kappa = 2$ and consider $\mathcal{G} = \Delta^3$. Consider the point $(\hat{x}, \hat{y})$, where

$$\hat{x} = \begin{bmatrix} 1/2 \\ 1/3 \\ 1/6 \end{bmatrix} \text{ and } \hat{y} = \begin{bmatrix} 1/4 \\ 1/8 \\ 1/30 \end{bmatrix}.$$

Due to Lemma 2, we deduce that this point belongs to $\mathcal{S}^{\kappa}_{P,R}$ with the following *unique* selection of $\hat{X}$:

$$\hat{X} = \frac{1}{240} \begin{bmatrix} 60 & 39 & 21 \\ 39 & 30 & 11 \\ 21 & 11 & 8 \end{bmatrix}.$$

In addition, its principal 2×2 minors, which are $\frac{279}{240^2}$, $\frac{39}{240^2}$, $\frac{119}{240^2}$, are all nonnegative, hence, $(\hat{x}, \hat{y}) \in \mathcal{S}^{\kappa}_{P,R,s}$. However, its determinant is $-\frac{1}{240^2} < 0$, hence, $\hat{X}$ is not positive semidefinite. Therefore, $(\hat{x}, \hat{y}) \notin \mathcal{S}^{\kappa}_{P,R,S}$. □

3.2 Exactness Results for $\kappa = 2$

We now provide some exactness results for $\kappa = 2$. Theorem 1 states that the convex hull of $\mathcal{S}^\kappa$ is second-order cone representable for $n = 2$, which is obtained from the **PR** relaxation.

Theorem 1. *Let $\mathcal{G} = \Delta^2$ and $\kappa = 2$. Then,* $\text{conv}(\mathcal{S}^\kappa) = \mathcal{S}^\kappa_{P,R}$.

Proof. It suffices to show that $\text{conv}(\mathcal{S}^\kappa) \supseteq \mathcal{S}^\kappa_{P,R}$. Let $(\hat{x}, \hat{y}) \in \mathcal{S}^\kappa_{P,R}$. Then, we have $e^T\hat{x} = 1$, $\hat{x}_j \geq \hat{y}_j \geq \hat{x}_j^2$ and $\hat{x}_1 - \hat{y}_1 = \hat{x}_2 - \hat{y}_2$, where the last equality follows from Lemma 2.
Case 1: $\hat{y}_1 = 0$. In this case, we conclude that

$$\hat{y}_1 = 0 \implies \hat{x}_1 = 0 \implies \hat{x}_2 = 1 \implies \hat{y}_2 = 1.$$

Notice that $(\hat{x}, \hat{y}) \in \mathcal{S}^\kappa$, hence, $(\hat{x}, \hat{y}) \in \text{conv}(\mathcal{S}^\kappa)$.
Case 2: $\hat{y}_1 > 0$. In this case, we also have $\hat{x}_1 > 0$. We claim that

$$\begin{bmatrix} \hat{x}_1 \\ \hat{x}_2 \\ \hat{y}_1 \\ \hat{y}_2 \end{bmatrix} = (1-\lambda) \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} + \lambda \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \tilde{y}_1 \\ \tilde{y}_2 \end{bmatrix},$$

for some $(\tilde{x}, \tilde{y}) \in \mathcal{S}^\kappa$ with $\lambda = \frac{\hat{x}_1^2}{\hat{y}_1} \in (0, 1)$. In fact, we have $\tilde{y}_1 = \frac{\hat{y}_1^2}{\hat{x}_1^2}$, $\tilde{y}_2 = \frac{(\hat{y}_1 - \hat{x}_1)^2}{\hat{x}_1^2}$, $\tilde{x}_1 = \frac{\hat{y}_1}{\hat{x}_1}$ and $\tilde{x}_2 = \frac{\hat{x}_1 - \hat{y}_1}{\hat{x}_1}$. Then, it is straightforward to check that $(\tilde{x}, \tilde{y}) \in \mathcal{S}^\kappa$, hence, we conclude that $(\tilde{x}, \tilde{y}) \in \text{conv}(\mathcal{S}^\kappa)$. □

For $n > 2$, the convex hull of $\mathcal{S}^\kappa$ is harder to characterize. The following fact provides the convex hull using the intractable completely positive cone due to [10].

Theorem 2. *Let $\mathcal{G} = \Delta^n$ and $\kappa = 2$. Then, we have*

$$\text{conv}(\mathcal{S}^\kappa) = \left\{ (x, y) \in \Delta^n \times \mathbb{R}^n : \exists X \in \mathbb{S}^n : \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \in \mathbb{CP}^{n+1},\ (6) \right\}.$$

In our paper, we look for tractable relaxations, for example, the **PRS** relaxation, which is a doubly nonnegative relaxation. Since we analyze the case where $\mathcal{G} = \Delta^n$ and $\kappa = 2$ in this part, observe that optimizing a linear function $\alpha^T x + \beta^T y$ over the **PRS** relaxation is equivalent to the following optimization problem, which we call as the **PRS'** relaxation:

$$\min_{x \in \Delta^n, X \in \mathbb{S}^n} \left\{ \alpha^T x + \beta^T \text{diag}(X) : \begin{bmatrix} X & x \\ x^T & 1 \end{bmatrix} \in \mathbb{S}^{n+1}_+,\ (6) \right\}.$$

The following corollary is a consequence of Theorem 2.

Corollary 2. *The following hold true:*

(i) Let $\mathcal{G} = \Delta^3$ and $\kappa = 2$. Then, $\text{conv}(\mathcal{S}^\kappa) = \mathcal{S}^\kappa_{P,R,S}$.

*(ii) Consider an optimal solution $(\hat{x}, \hat{X})$ to the **PRS'** relaxation and assume that the matrix $\begin{bmatrix} \hat{X} & \hat{x} \\ \hat{x}^T & 1 \end{bmatrix} \in \mathbb{CP}^{n+1}$. Then, $\mathcal{O}_{\mathcal{S}^\kappa, \mathcal{S}^\kappa_{P,R,S}}(\alpha, \beta) = 0$.*

Note that part (i) of Corollary 2 follows from the fact that up to 4×4 matrices, doubly non-negative matrices are also completely positive [23]. By the same token, we have $\mathcal{G} = \Delta^2$ and $\kappa = 2$, then $\text{conv}(\mathcal{S}^\kappa) = \mathcal{S}^\kappa_{P,R,S}$. Theorem 1 is a strengthening, since it says that we do not need the PSD constraints to obtain the convex hull when $n = 2$.

Next, in Theorem 3, we present three sufficient conditions on β so that we obtain $\mathcal{O}_{\mathcal{S}^\kappa,\mathcal{S}^\kappa_{P,R,S}}(\alpha,\beta) = 0$.

Theorem 3. *Let $\mathcal{G} = \Delta^n$ and $\kappa = 2$. Then, we have $O_{\mathcal{S}^\kappa,\mathcal{S}^\kappa_{P,R,S}}(\alpha,\beta) = 0$ under any of the following cases:*

- *Case 1: $|\{j = 1,\dots,n : \beta_j \le 0\}| = n$.*
- *Case 2: $|\{j = 1,\dots,n : \beta_j \ge 0\}| = n$.*
- *Case 3: $|\{j = 1,\dots,n : \beta_j > 0\}| = 1$.*

Proof. Consider an optimal solution $(\hat{x},\hat{X})$ to the **PRS'** relaxation. The assertion of the theorem follows if the set $J_< = \{j : \hat{X}_{jj} > \hat{x}_j^2\} = \emptyset$ since in this case $\hat{X} = \hat{x}\hat{x}^T$. In the rest of the proof, let us assume that $J_< \neq \emptyset$ for every optimal solution to the **PRS'** relaxation.

We first prove two preliminary results:

Claim 1. *For $i \notin J_<$, we have $\hat{X}_{ij} = \hat{x}_i\hat{x}_j$ for any $j = 1,\dots,n$.*

Proof. Consider positive semidefinite matrix $\hat{X} - \hat{x}\hat{x}$. Note that if a diagonal entry of this matrix is zero, i.e., $\hat{X}_{ii} - \hat{x}_i^2 = 0$ for some i, then we have that the off-diagonal entries $\hat{X}_{ij} - \hat{x}_i\hat{x}_j = 0$ for every $j = 1,\dots,n$. Hence, the result follows. $\square$

Claim 2. *If there exist $i,j \in J_<$ such that $\beta_i + \beta_j < 0$, then $\hat{X}_{ij} = 0$.*

Proof. Suppose not. Then, we can construct a new solution $\tilde{X}$ that agrees with $\hat{X}$ on each entry except:

$$\begin{bmatrix} \tilde{X}_{ii} & \tilde{X}_{ij} \\ \tilde{X}_{ij} & \tilde{X}_{jj} \end{bmatrix} = \begin{bmatrix} \hat{X}_{ii} & \hat{X}_{ij} \\ \hat{X}_{ij} & \hat{X}_{jj} \end{bmatrix} + \begin{bmatrix} \hat{X}_{ij} & -\hat{X}_{ij} \\ -\hat{X}_{ij} & \hat{X}_{ij} \end{bmatrix}.$$

Notice that $\tilde{X}$ is obtained from $\hat{X}$ by a diagonally dominant shift, hence, it is positive semidefinite. In addition, it maintains the row sums being the same as $\hat{X}$. Therefore, it is a feasible solution. However, the objective function difference between $(\hat{x},\hat{X})$ and $(\hat{x},\tilde{X})$ is $(\beta_i\hat{X}_{ii} + \beta_j\hat{X}_{jj}) - (\beta_i\tilde{X}_{ii} + \beta_j\tilde{X}_{jj}) = -(\beta_i + \beta_j)\hat{X}_{ij} > 0$. However, this is a contradiction to $(\hat{x},\hat{X})$ being an optimal solution. $\square$

In the remainder of the proof, we will assume that $\hat{X}_{ij} = 0$ for i,j such that $\beta_i + \beta_j \le 0$ (notice that if $\beta_i + \beta_j = 0$, we can find a solution that satisfies this property).

Now, we will prove each of the three sufficient conditions separately.

- Case 1: Due to Claim 2, we have that $\hat{X}_{ij} = 0$ for $i \neq j$, implying that $\hat{X}_{jj} = \hat{x}_j$. Then, observe that we have
$$(\hat{x},\hat{X}) = \sum_{j\in J_<} \hat{x}_j(e^j,(e^j)(e^j)^T).$$
In other words, $(\hat{X},\hat{x})$ is a convex combination of $|J_<|$ many feasible solutions. Therefore, at least one of these solutions must have an objective function value at least as good as $(\hat{X},\hat{x})$, which is a contradiction.

- Case 2: Let us consider a new feasible solution $(\hat{x},\hat{x}\hat{x}^T)$ to the **PRS'** relaxation. Observe that the objective function difference between $(\hat{x},\hat{X})$ and $(\hat{x},\hat{x}\hat{x}^T)$ is $\sum_{i\in J_<}\beta_i(\hat{X}_{ii} - \hat{x}_i^2) \ge 0$. This implies that $(\hat{x},\tilde{X})$ is also an optimal solution, which is a contradiction.

- Case 3: Let $J_<^- := \{j \in J_< : \hat{X}_{jj} > \hat{x}_j^2,\ \beta_j \le 0\}$, $J_=^- := \{j : \hat{X}_{jj} = \hat{x}_j^2 > 0,\ \beta_j \le 0\}$, $J_0^- := \{j : \hat{X}_{jj} = \hat{x}_j^2 = 0,\ \beta_j \le 0\}$. If $J_<^- = \emptyset$, then the statement is trivially true as we reduce to Case 2 with a single variable having a positive β_j coefficient. Suppose, without loss of generality, that $\beta_n > 0$, $J_<^- = \{1, \ldots, k\}$, $J_=^- = \{k+1, \ldots, k'\}$ and $J_0^- = \{k'+1, \ldots, n-1\}$. In the rest of the proof, we will consider the following submatrix of $\begin{bmatrix} \hat{X} & \hat{x} \\ \hat{x}^T & 1 \end{bmatrix}$, which is obtained by deleting the identically zero rows and columns in J_0^-:

$$\begin{bmatrix} \hat{X}_{11} & & & \hat{X}_{1n} & \hat{x}_1 \\ & \ddots & & \vdots & \vdots \\ & & \hat{X}_{k',k'} & \hat{X}_{k',n} & \hat{x}_{k'} \\ \hat{X}_{1n} & \cdots & \hat{X}_{k',n} & \hat{X}_{nn} & \hat{x}_n \\ \hat{x}_1 & \cdots & \hat{x}_{k'} & \hat{x}_n & 1 \end{bmatrix}. \tag{7}$$

  Notice that the Schur complement of this matrix with respect to the first k' rows and columns is obtained as

$$0 \preceq \begin{bmatrix} \hat{X}_{nn} - \sum_{j=1}^{k'} \frac{\hat{X}_{jn}^2}{\hat{X}_{jj}} & \hat{x}_n - \sum_{j=1}^{k'} \frac{\hat{x}_j \hat{X}_{jn}}{\hat{X}_{jj}} \\ \hat{x}_n - \sum_{j=1}^{k'} \frac{\hat{x}_j \hat{X}_{jn}}{\hat{X}_{jj}} & 1 - \sum_{j=1}^{k'} \frac{\hat{x}_j^2}{\hat{X}_{jj}} \end{bmatrix} = \left(1 - \sum_{j=1}^{k'} \frac{\hat{x}_j^2}{\hat{X}_{jj}}\right) \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},$$

  where the equality follows as a consequence of $\hat{X}e = \hat{x}$. Since we have $1 - \sum_{j=1}^{k'} \frac{\hat{x}_j^2}{\hat{X}_{jj}} \ge 0$, we conclude that $|J_=^-| \le 1$. Now, we have two subcases:

  - $|J_=^-| = 1$. In this case, we must have $k = 0$ and $k' = 1$, meaning that $J_<^- = \emptyset$. In this situation, we again reduce to Case 2 with a single variable having a positive β_j coefficient.
  - $|J_=^-| = 0$. In this case, we must have $k = k'$. Since we have $1 - \sum_{j=1}^{k} \frac{\hat{x}_j^2}{\hat{X}_{jj}} \ge 0$, we can write an explicit completely positive decomposition of the solution matrix in equation (7) as (recall Corollary 2)

$$\sum_{j=1}^{k} \frac{\hat{x}_j^2}{\hat{X}_{jj}} \left(\frac{\hat{X}^j}{\hat{x}_j}\right)\left(\frac{\hat{X}^j}{\hat{x}_j}\right)^T + \left(1 - \sum_{j=1}^{k} \frac{\hat{x}_j^2}{\hat{X}_{jj}}\right)(e^n + e^{n+1})(e^n + e^{n+1})^T,$$

    where $\bar{X}^j$ is the j-th column of matrix $\bar{X}$. Notice that the solution matrix on the left-hand side is a convex combination of $k+1$ many feasible solutions. Therefore, at least one of these solutions must have an objective function value at least as good as $(\hat{X}, \hat{x})$, which is a contradiction.

□

We note that results similar to those of Cases 1 and 2 in Theorem 3 have also been shown in [15] utilizing conic duality whereas our approach uses only the primal problem.

Motivated by the insights derived from Theorem 3 and our extensive computational experiments reported in Section 4.2 (see, in particular, Figure 5 with $\kappa = 2$), we have come up with the following conjecture:

Conjecture 1. *Let $\mathcal{G} = \Delta^n$ and $\kappa = 2$. Then, $\mathcal{O}_{\mathcal{S}^\kappa, \mathcal{S}^\kappa_{P,R,S}}(\alpha, \beta) = 0$ for all $\alpha, \beta \in \mathbb{R}^n \times \mathbb{R}^n$, implying that* $\mathrm{conv}(\mathcal{S}^\kappa) = \mathcal{S}^\kappa_{P,R,S}$.

3.3 Approximation results for general κ

We now switch our attention from exactness results to approximation guarantees for general κ.

3.3.1 Distance-based approximation results

Our first distance-based approximation result below gives an upper bound on the worst-case distance between the nonconvex set $\mathcal{S}^\kappa$ and the set $\mathcal{S}_P^\kappa$ (i.e., the feasible region of the **P** relaxation), which increases with both the exponent κ and the dimension n.

Proposition 8. *Let $n \geq 2$, $\mathcal{G} = \Delta^n$ and $\kappa > 1$. Then, we have*

$$D_{\mathcal{S}^\kappa, \mathcal{S}_P^\kappa} := \max_{(\hat{x},\hat{y}) \in \mathcal{S}_P^\kappa} \min_{(x,y) \in \mathcal{S}^\kappa} \|(x - \hat{x}, y - \hat{y})\|_1 \leq 1 - n^{1-\kappa}.$$

Proof. Since we want to establish an upper bound to $D_{\mathcal{S}^\kappa, \mathcal{S}_P^\kappa}$, we can upper bound the inner minimization problem. This can be done in the inner optimization problem by just substituting a feasible solution. In particular, we will substitute, $x_i := \hat{x}_i$ and $y_i := \hat{x}_i^\kappa$ for all $i = 1, \ldots, n$, where $\sum_{i=1}^n \hat{x}_i = 1$. Then, we have

$$\min_{(x,y) \in \mathcal{S}^\kappa} \|x - \hat{x}\|_1 + \|y - \hat{y}\|_1 \leq \sum_{i=1}^n |\hat{y}_i - \hat{x}_i^\kappa|.$$

Therefore, we are left with solving:

$$\max\nolimits_{(\hat{x},\hat{y}) \in \Delta^n \times \mathbb{R}^n} \left\{ \sum\nolimits_{i=1}^n |\hat{y}_i - \hat{x}_i^\kappa| : \hat{x}_i^\kappa \leq \hat{y}_i \leq \hat{x}_i,\ i = 1, \ldots, n \right\}.$$

Since $\hat{y}_i \geq \hat{x}_i^\kappa$ for points in $\mathcal{S}_P^\kappa$, we obtain

$$\begin{aligned} \max_{(\hat{x},\hat{y}) \in \mathcal{S}_P^\kappa} \sum_{i=1}^n |\hat{y}_i - \hat{x}_i^\kappa| &= \max_{(\hat{x},\hat{y}) \in \mathcal{S}_P^\kappa} \sum_{i=1}^n (\hat{y}_i - \hat{x}_i^\kappa) = \max_{\hat{x} \in \Delta^n} \sum_{i=1}^n (\hat{x}_i - \hat{x}_i^\kappa) \\ &= n \left(\frac{1}{n} - \frac{1}{n^\kappa} \right) = 1 - n^{1-\kappa}, \end{aligned}$$

which completes the proof. $\square$

Notice that $1 - n^{1-\kappa}$ converges to 0 and 1 as $\kappa \to 1^+$ and $\kappa \to \infty$, respectively, and it is equal to $1 - \frac{1}{n}$ for $\kappa = 2$.

How good is the upper bound presented in the above result? Our second distance-based approximation result in Proposition 9 gives a matching lower bound on the worst-case distance between the nonconvex set $\mathcal{S}^\kappa$ and its convex hull, which increases with the exponent κ and dimension n. This lower bound shows that **P** is indeed a "reasonable" convex relaxation, since even $\mathcal{S}^\kappa$ and its convex hull have distances of similar order between them. To prove Proposition 9, we need a technical lemma:

Lemma 3. *Consider the function $f(x) = \left|x - \frac{1}{n}\right| + \left|x^\kappa - \frac{1}{n}\right| + x$. If $n \geq 2$ and $\kappa > 1$, then the minimizer of this function is achieved at $x = \frac{1}{n}$.*

Proof. We may write the function as:

$$f(x) = \begin{cases} \frac{2}{n} - x^\kappa, & x \in [0, \frac{1}{n}] \\ 2x - x^\kappa, & x \in \left[\frac{1}{n}, \left(\frac{1}{n}\right)^{\frac{1}{\kappa}}\right] \\ -\frac{2}{n} + 2x + x^\kappa, & x \in \left[\left(\frac{1}{n}\right)^{\frac{1}{\kappa}}, \infty\right) \end{cases}$$

We see that the function is continuous, is decreasing in the interval $[0, \frac{1}{n}]$, concave in the interval $\left[\frac{1}{n}, \left(\frac{1}{n}\right)^{\frac{1}{\kappa}}\right]$ and increasing in the interval $\left[\left(\frac{1}{n}\right)^{\frac{1}{\kappa}}, \infty\right)$. Therefore, the only possible optimal solutions are $\left\{\frac{1}{n}, \left(\frac{1}{n}\right)^{\frac{1}{\kappa}}\right\}$.

To complete the proof, we need to show that $f(\frac{1}{n}) \leq f((\frac{1}{n})^{\frac{1}{\kappa}})$ or equivalently:

$$\frac{2}{n} - \left(\frac{1}{n}\right)^{\kappa} \leq 2\left(\frac{1}{n}\right)^{\frac{1}{\kappa}} - \frac{1}{n} \;\Leftrightarrow\; \frac{1}{n} \leq \frac{2}{3}\left(\frac{1}{n}\right)^{\frac{1}{\kappa}} + \frac{1}{3}\left(\frac{1}{n}\right)^{\kappa}. \tag{8}$$

Note that for $\kappa = 1$, (8) holds. The proof will be complete by showing that the function $g(\kappa) := \frac{2}{3}\left(\frac{1}{n}\right)^{\frac{1}{\kappa}} + \frac{1}{3}\left(\frac{1}{n}\right)^{\kappa}$ is non-decreasing in κ for $\kappa \geq 1$ and $n \geq 2$.

Observe that $g'(\kappa) = \frac{2}{3}\, n^{-1/\kappa} \frac{\ln(n)}{\kappa^2} - \frac{1}{3}(\ln(n))\, n^{-\kappa} = \frac{\ln(n)}{3}\left(\frac{2\, n^{-1/\kappa}}{\kappa^2} - n^{-\kappa}\right)$. Since $n \geq 2$ implies $\ln(n) > 0$, it suffices to show $\frac{2\, n^{-1/\kappa}}{\kappa^2} \geq n^{-\kappa}$. Taking logarithms, this inequality is equivalent to showing $\ln(2) - 2\ln(\kappa) + (\ln(n))\left(\kappa - \frac{1}{\kappa}\right) \geq 0$. Since $n \geq 2$ and $\kappa - \frac{1}{\kappa} > 0$ for $\kappa > 1$, it is sufficient to prove that

$$h(\kappa) := \ln(2) - 2\ln\kappa + (\ln(2))\left(\kappa - \frac{1}{\kappa}\right) \geq 0, \quad \text{for all } \kappa \geq 1.$$

Observe that $h'(\kappa) = -\frac{2}{\kappa} + (\ln(2))\left(1 + \frac{1}{\kappa^2}\right)$ and $h''(\kappa) = \frac{2}{\kappa^3}\left(\kappa - \ln(2)\right)$. Since $\kappa \geq 1 > \ln(2)$, we have that $h''(\kappa) > 0$ and thus $h(\kappa)$ is convex in the domain $[1, \infty)$. Therefore, h achieves its global minimum at κ^* satisfying, $h'(\kappa^*) = 0$, i.e., κ^* satisfies $\kappa^* + \frac{1}{\kappa^*} = \frac{2}{\ln(2)}$. The only value of κ^* greater than 1 is $\frac{1}{\ln(2)} + \sqrt{(\frac{1}{\ln(2)})^2 - 1}$. Plugging this value into h gives a value of $\ln(2) - 2\ln\left(\frac{1}{\ln(2)} + \sqrt{(\frac{1}{\ln(2)})^2 - 1}\right) + \ln(2)(\sqrt{(\frac{2}{\ln(2)})^2 - 4}) \approx 0.3161 > 0$. Thus, $h(\kappa) > 0$ for all $\kappa > 1$, completing the proof. □

Proposition 9. *Let $n \geq 2$, $\mathcal{G} = \Delta^n$ and $\kappa > 1$. Then, we have*

$$D_{\mathcal{S}^\kappa, \mathrm{conv}(\mathcal{S}^\kappa)} := \max_{(\hat{x},\hat{y}) \in \mathrm{conv}(\mathcal{S}^\kappa)} \min_{(x,y) \in \mathcal{S}^\kappa} \|(x - \hat{x}, y - \hat{y})\|_1 \geq 1 - n^{1-\kappa}.$$

Proof. We just fix $\hat{x}_i = \hat{y}_i = \frac{1}{n}$ for the outer optimization problem. We note that the following is a lower bound:

$$\begin{aligned}
\min_{(x,y)\in\mathcal{S}^\kappa} \sum_{i=1}^{n} \left|x_i - \frac{1}{n}\right| + \sum_{i=1}^{m} \left|y_i - \frac{1}{n}\right| \quad &= \quad \min_{x \in \Delta^n} \sum_{i=1}^{n} \left|x_i - \frac{1}{n}\right| + \sum_{i=1}^{n} \left|x_i^\kappa - \frac{1}{n}\right| \\
&\geq \quad \min_{x \in \mathbb{R}^n_+} \sum_{i=1}^{n} \left|x_i - \frac{1}{n}\right| + \sum_{i=1}^{n} \left|x_i^\kappa - \frac{1}{n}\right| + \sum_{i=1}^{n} x_i - 1,
\end{aligned}$$

where the last inequality follows by taking a particular Lagrangian relaxation. Due to Lemma 3, we know that the minimizer of the single variable function

$$f(x_i) = \sum_{i=1}^{n} \left|x_i - \frac{1}{n}\right| + \sum_{i=1}^{n} \left|x_i^\kappa - \frac{1}{n}\right| + x_i$$

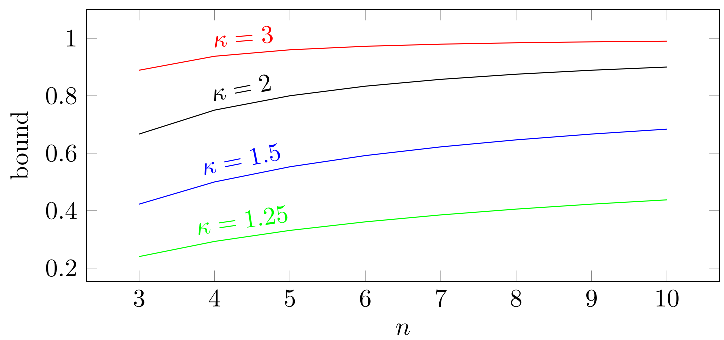
is achieved at $x_i = \frac{1}{n}$. Thus, we obtain

$$\min_{x\in\mathbb{R}^n_+} \sum_{i=1}^{n} \left|x_i - \frac{1}{n}\right| + \sum_{i=1}^{n} \left|x_i^\kappa - \frac{1}{n}\right| + \sum_{i=1}^{n} x_i - 1 = n\left(\frac{1}{n} - \left(\frac{1}{n}\right)^\kappa\right) = 1 - n^{1-\kappa},$$

completing the proof. □

Figure 2 illustrates the comparison of bounds derived in Propositions 8 and 9. As expected, bounds converge to 1 as n increases and the convergence is faster with larger κ.

Figure 2: Bounds derived in Propositions 8 and 9 for different κ and n values.



3.3.2 Objective function-based approximation results

Our first objective function-based approximation result below characterizes the cases under which the **P** relaxation is exact.

Proposition 10. *Let $\mathcal{G} = \Delta^n$ and $\kappa > 1$. Then, we have $O_{\mathcal{S}^\kappa, \mathcal{S}_P^\kappa}(\alpha, \beta) = 0$ under any of the following cases:*

- *Case 1: $J^+ := \{j : \beta_j > 0\} = \{1, \dots, n\}$.*
- *Case 2: $J^- := \{j : \beta_j \leq 0\} = \{1, \dots, n\}$.*
- *Case 3: $J^+ \neq \emptyset$, $J^- \neq \emptyset$ and either of the following holds, where $\hat{J} := \{j \in J^+ : \alpha_j - \mu < 0\}$ with $\mu := \min\{(\alpha_j + \beta_j) : j \in J^-\}$.*
  - *Subcase 3a: $\hat{J} = \emptyset$.*
  - *Subcase 3b: $\hat{J} \neq \emptyset$ and $\sum_{j \in \hat{J}} \left(\frac{\mu - \alpha_j}{\kappa \beta_j}\right)^{\frac{1}{\kappa - 1}} \geq 1$.*

Proof. Consider problems $z(\alpha, \beta) := \min \left\{\alpha^T x + \beta^T y : (x, y) \in \mathcal{S}^\kappa\right\}$ and

$$z_P(\alpha, \beta) := \min \left\{\alpha^T x + \beta^T y : (x, y) \in \mathcal{S}_P^\kappa\right\}. \tag{9}$$

Let us split the indices as $J^+ := \{j : \beta_j > 0\}$ and $J^- := \{j : \beta_j \leq 0\}$. Observe that when optimizing over $\mathcal{S}_P^\kappa$, there exists an optimal solution with the following structure (see Proposition 4): $y_j = x_j^\kappa$ for $j \in J^+$, and $y_j = x_j$ for $j \in J^-$. Therefore, the value of $z_P(\alpha, \beta)$ defined in (9) is equivalent to the following:

$$\min \left\{ \sum_{j \in J^+} (\alpha_j x_j + \beta_j x_j^\kappa) + \sum_{j \in J^-} (\alpha_j + \beta_j) x_j : \sum_{j \in J^+} x_j + \sum_{j \in J^-} x_j = 1, x \geq 0 \right\}.$$

- Case 1: $J^+ = \{1, \dots, n\}$. In this case, problem (9) is exact since $y_j = x_j^\kappa$ for each j, that is, we have $z_P(\alpha, \beta) = z(\alpha, \beta)$.
- Case 2: $J^- = \{1, \dots, n\}$. In this case, problem (9) can be solved as a linear program and the optimal solution is at an extreme point of the simplex. Therefore, we have $x_{j'} = 1$ for $j' = \arg\min\{(\alpha_j + \beta_j)\}$, and $x_j = 0$ for $j \neq j'$. Hence, $y_j = x_j^\kappa$ for each j and the relaxation is exact, that is, we have $z_P(\alpha, \beta) = z(\alpha, \beta)$.
- Case 3: $J^+ \neq \emptyset$ and $J^- \neq \emptyset$. Let $j' = \arg\min\{(\alpha_j + \beta_j) : j \in J^-\}$ and set $\mu = \min\{(\alpha_j + \beta_j) : j \in J^-\}$. Notice that we must have $x_{j'} = 1 - \sum_{j \in J^+}$ and $x_j = 0$ for $j \in J^- \setminus \{j'\}$ in an optimal solution.

With this simplification, we obtain the following equivalent problem for the value of $z_P(\alpha,\beta)$:

$$\begin{aligned}
&\min\left\{\sum_{j\in J^+}(\alpha_j x_j+\beta_j x_j^\kappa)+\mu(1-\sum_{j\in J^+}x_j):\sum_{j\in J^+}x_j\le 1, x_j\ge 0, j\in J^+\right\}\\
=&\mu+\min\left\{\sum_{j\in J^+}[(\alpha_j-\mu)x_j+\beta_j x_j^\kappa]:\sum_{j\in J^+}x_j\le 1, x_j\ge 0, j\in J^+\right\}\\
=&\mu+\min\left\{\sum_{j\in\hat{J}}[(\alpha_j-\mu)x_j+\beta_j x_j^\kappa]:\sum_{j\in\hat{J}}x_j\le 1, x_j\ge 0, j\in \hat{J}\right\},
\end{aligned}$$

where $\hat{J}:=\{j\in J^+:\alpha_j-\mu<0\}$ (note that, in an optimal solution, $x_j=0$ for $j\in J^+\setminus\hat{J}$). In this case, the relaxation is exact if $\sum_{j\in\hat{J}}x_j\in\{0,1\}$. Let us look at subcases:

- Subcase 3a: $\hat{J}=\emptyset$. In this case, the relaxation is exact as $x_{j'}=y_{j'}=1$ and $x_j=y_j=0$ for $j\neq j'$, that is, we have $z_P(\alpha,\beta)=z(\alpha,\beta)$.
- Subcase 3b: $\hat{J}\neq\emptyset$ and $\sum_{j\in\hat{J}}\left(\frac{\mu-\alpha_j}{\kappa\beta_j}\right)^{\frac{1}{\kappa-1}}\ge 1$. In this case, the relaxation is exact as $\sum_{j\in\hat{J}}x_j=1$. In this case, although we do not have a closed form expression of $z_P(\alpha,\beta)$, we know that $z_P(\alpha,\beta)=z(\alpha,\beta)$.

$\square$

Notice that when $J^+\neq\emptyset$ and $J^-\neq\emptyset$, $\hat{J}\neq\emptyset$ and $\sum_{j\in\hat{J}}\left(\frac{\mu-\alpha_j}{\kappa\beta_j}\right)^{\frac{1}{\kappa-1}}<1$, the **P** relaxation is inexact as $\sum_{j\in\hat{J}}x_j\in(0,1)$, that is, we have $z_P(\alpha,\beta)<z(\alpha,\beta)$. However, note that in this case we know the value of $z_P(\alpha,\beta)=\mu+(1-\kappa)\sum_{j\in\hat{J}}\beta_j\left(\frac{\mu-\alpha_j}{\kappa\beta_j}\right)^{\frac{\kappa}{\kappa-1}}$, but not $z(\alpha,\beta)$.

Proposition 10 allows us to estimate the probability that the **P** relaxation is exact when $\mathcal{G}=\Delta^n$. For this purpose, we generate 10^6 random (α,β) vectors and count the number of times the condition $O_{\mathcal{S}^\kappa,\mathcal{S}_P^\kappa}(\alpha,\beta)=0$ is satisfied. We report the results of this simulation in Tables 1 and 2 for two different distributions.

$\kappa\backslash n$	2	4	8	16	32	64
1.25	0.9154	0.8696	0.8937	0.9550	0.9940	0.9999
1.5	0.9064	0.8607	0.8896	0.9540	0.9940	0.9999
1.75	0.8990	0.8551	0.8877	0.9542	0.9939	0.9999
2	0.8933	0.8518	0.8882	0.9540	0.9940	0.9999
2.5	0.8857	0.8490	0.8894	0.9554	0.9942	0.9999
3	0.8804	0.8508	0.8925	0.9571	0.9943	0.9999

Table 1: The estimated probability of the **P** relaxation being exact via simulation when $\mathcal{G}=\Delta^n$ and the objective function coefficients (α,β) are iid sampled from the Unif$(-1,1)$ distribution.

$\kappa \backslash n$	2	4	8	16	32	64	128	256	512
1.25	0.9212	0.8792	0.8788	0.9021	0.9306	0.9551	0.9726	0.9843	0.9913
1.5	0.9123	0.8698	0.8712	0.8960	0.9268	0.9532	0.9721	0.9840	0.9912
1.75	0.9055	0.8630	0.8662	0.8929	0.9254	0.9522	0.9712	0.9835	0.9907
2	0.9001	0.8577	0.8634	0.8910	0.9242	0.9512	0.9708	0.9833	0.9907
2.5	0.8911	0.8539	0.8611	0.8890	0.9220	0.9504	0.9703	0.9829	0.9906
3	0.8848	0.8519	0.8604	0.8896	0.9214	0.9497	0.9697	0.9826	0.9905

Table 2: The estimated probability of the **P** relaxation being exact via simulation when $\mathcal{G} = \Delta^n$ and the objective function coefficients (α, β) are iid sampled from the standard normal distribution.

Interestingly, the probability of Case 3a, which is independent of the value κ, is quite large in the simulation results and it seems to converge to 1 as n increases. This has motivated us to explore whether this probability can be computed analytically and shown to converge to 1. Below, we carry out this analysis for the uniform distribution.

Proposition 11. *Suppose that (α, β) are iid Unif$(-1,1)$ distributed. Then,*

$$\Pr(\text{Case 3a}) = 2^{-n} \sum_{m=1}^{n-1} \binom{n}{m} \left(1 - m2^{-m} \left(I_m + \frac{2^{2(m-n)}}{2n-m}\right)\right),$$

where $I_m := \int_{-1}^{0} \left[\frac{1}{4} - \frac{1}{2}z\right]^{n-m} (1-z)^{m-1} dz$ for $m = 1, \dots, n-1$.

Proof. Let $\mathcal{E}_m$ be the event in which $|J^+| = m$ and $|J^-| = n - m$, $m = 0, \dots, n$. Note that $\Pr(\mathcal{E}_m) = \binom{n}{m} 2^{-n}$. Then, Case 1 and Case 2 are respectively the events $\mathcal{E}_n$ and $\mathcal{E}_0$, which happen with probability 2^{-n} each. Now, let us understand the event $\mathcal{E}_m$, $m = 1, \dots, n-1$.

Let $F(z) := \Pr(\alpha_j + \beta_j \geq z | \beta_j \leq 0)$. Using Bayes's rule and the fact that the density of the Unif$(-1,1)$ distribution is $\frac{1}{2}$ over $[-1,1]$, we deduce that

$$F(z) = \begin{cases} 1 - \frac{1}{4}(z+2)^2 & z \in [-2,-1] \\ \frac{1}{4} - \frac{1}{2}z & z \in [-1,0] \\ \frac{1}{4}(z-1)^2 & z \in [0,1] \end{cases}.$$

Now, let us derive the right-tail probability of $\mu = \min\{\alpha_j + \beta_j : j \in J^-\}$ given the event $\mathcal{E}_m$. Due to the independence of the random variables involved, we have

$$G_m(z) := \Pr(\mu \geq z | \mathcal{E}_m) = [F(z)]^{n-m},$$

Next, we derive the CDF of $\nu := \min\{\alpha_j : j \in J^+\}$ given the event $\mathcal{E}_m$. In fact,

$$H_m(z) := \Pr(\nu \leq z | \mathcal{E}_m) = 1 - \Pr(\nu \geq z | \mathcal{E}_m) = 1 - \left(\frac{1-z}{2}\right)^m = 1 - 2^{-m}(1-z)^m,$$

from which we obtain the PDF of ν given the event $\mathcal{E}_m$ as $h_m(z) = m2^{-m}(1-z)^{m-1}$.

We are now ready to compute the probability of Case 3a given the event $\mathcal{E}_m$ as

$$\begin{aligned} \Pr(\text{Case 3a} | \mathcal{E}_m) &= \Pr(\nu \geq \mu | \mathcal{E}_m) \\ &= 1 - \Pr(\mu \geq \nu | \mathcal{E}_m) = 1 - \int_{-1}^{1} G_m(z) h_m(z) dz \\ &= 1 - \int_{-1}^{0} G_m(z) h_m(z) dz - \int_{0}^{1} G_m(z) h_m(z) dz. \end{aligned} \tag{10}$$

Note that the second integral in the last line of equation (10) is easy to compute as

$$\begin{aligned}\int_0^1 G_m(z)h_m(z)dz &= \int_0^1 \left[\frac{1}{4}(z-1)^2\right]^{n-m} m2^{-m}(1-z)^{m-1}dz \\ &= m2^{m-2n}\int_0^1 (1-z)^{2n-m-1}dz \\ &= m2^{m-2n}\left[-\frac{(1-z)^{2n-m}}{2n-m}\right]_{z=0}^1 = \frac{m2^{m-2n}}{2n-m}.\end{aligned}$$

The computation of the first integral in the last line of equation (10) is more convoluted. Observe that

$$\begin{aligned}\int_{-1}^0 G_m(z)h_m(z)dz &= \int_{-1}^0 \left[\frac{1}{4}-\frac{1}{2}z\right]^{n-m} m2^{-m}(1-z)^{m-1}dz \\ &= m2^{-m}\int_{-1}^0 \left[\frac{1}{4}-\frac{1}{2}z\right]^{n-m}(1-z)^{m-1}dz = m2^{-m}I_m.\end{aligned} \tag{11}$$

It is straightforward to obtain that $I_1 = \int_{-1}^0 \left[\frac{1}{4}-\frac{1}{2}z\right]^{n-1} dz = \left[(-2)\frac{(\frac{1}{4}-\frac{1}{2}z)^n}{n}\right]_{z=-1}^0 = (-2)\frac{(\frac{1}{4})^n-(\frac{3}{4})^n}{n}$. Now, if $m \geq 2$, using integration by parts, we obtain the relation

$$\begin{aligned}I_m &= \left[(-2)(1-z)^{m-1}\frac{(\frac{1}{4}-\frac{1}{2}z)^{n-m+1}}{n-m+1}\right]_{z=-1}^0 \\ &\quad - \frac{2(m-1)}{n-m+1}\int_{-1}^0 \left[\frac{1}{4}-\frac{1}{2}z\right]^{n-m+1}(1-z)^{m-2}dz \\ &= (-2)\frac{(\frac{1}{4})^{n-m+1}-2^{m-1}(\frac{3}{4})^{n-m+1}}{n-m+1} - \frac{2(m-1)}{n-m+1}I_{m-1}.\end{aligned} \tag{12}$$

Since we have already computed I_1, any I_m with $m = 2,\ldots,n-1$ can be obtained recursively using the relation above.

Finally, we conclude that the probability of Case 3a is computed as

$$\begin{aligned}\Pr(\text{Case 3a}) &= \sum_{m=1}^{n-1}\Pr(\text{Case 3a}|\mathcal{E}_m)\Pr(\mathcal{E}_m) = 2^{-n}\sum_{m=1}^{n-1}\binom{n}{m}\Pr(\text{Case 3a}|\mathcal{E}_m) \\ &= 2^{-n}\sum_{m=1}^{n-1}\binom{n}{m}\left(1-m2^{-m}\left(I_m+\frac{2^{2(m-n)}}{2n-m}\right)\right).\end{aligned}$$

□

The above analysis enables us to compute the *exact* probability of the **P** relaxation being exact when $\mathcal{G} = \Delta^n$ and the objective function coefficients (α, β) are iid Unif$(-1, 1)$ distributed. These probabilities are reported in Table 3.

n	2	4	8	16	32	64
Pr(Case 3a)	0.3541	0.6434	0.8242	0.9409	0.9930	0.9999

Table 3: The probability of the **P** relaxation being exact when $\mathcal{G} = \Delta^n$ and the objective function coefficients (α, β) are iid Unif$(-1, 1)$ distributed.

Now, we are ready to prove that the **P** relaxation is exact with high probability for large n when the objective function coefficients are uniformly distributed.

Theorem 4. *Suppose that* (α, β) *are iid* Unif$(-1, 1)$ *distributed. Then,*

$$\Pr(\text{Case 3a}) \to 1 \text{ as } n \to \infty.$$

Hence, $\Pr(O_{\mathcal{S}^\kappa, \mathcal{S}_P^\kappa}(\alpha, \beta) = 0) \to 1$ *as* $n \to \infty$.

Proof. In virtue of Proposition 11, it suffices to show the following:

(i) $2^{-n} \sum_{m=1}^{n-1} \binom{n}{m} \to 1$ as $n \to \infty$.

(ii) $2^{-n} \sum_{m=1}^{n-1} \binom{n}{m} m 2^{-m} I_m \to 0$ as $n \to \infty$.

(iii) $2^{-n} \sum_{m=1}^{n-1} \binom{n}{m} m 2^{-m} \frac{2^{2(m-n)}}{2n-m} \to 0$ as $n \to \infty$.

In the proof, we use the binomial identity $(x + y)^n = \sum_{j=0}^{n} \binom{n}{j} x^j y^{n-j}$ repeatedly.

To prove (i), observe that

$$2^{-n} \sum_{m=1}^{n-1} \binom{n}{m} = 2^{-n} \left(\sum_{m=0}^{n} \binom{n}{m} - \binom{n}{0} - \binom{n}{n} \right) = 1 - 2 \times 2^{-n} \to 1$$

as $n \to \infty$.

To prove (ii), we first note that $I_m \geq 0$ as $m 2^{-m} I_m$ is a probability (see equation (11)). From equation (12), we obtain $I_m \leq \frac{2^m (\frac{3}{4})^{n-m+1}}{n-m+1}$. Then, we have

$$\begin{aligned} 2^{-n} \sum_{m=1}^{n-1} \binom{n}{m} m 2^{-m} I_m &\leq 2^{-n} \sum_{m=1}^{n-1} \frac{n!}{(n-m)!m!} m 2^{-m} \frac{2^m (\frac{3}{4})^{n-m+1}}{n-m+1} \\ &= \left(\frac{3}{8}\right)^n \sum_{m=1}^{n-1} \frac{n!}{(n-m+1)!(m-1)!} \left(\frac{4}{3}\right)^{m-1} \\ &= \left(\frac{3}{8}\right)^n \sum_{m=1}^{n-1} \binom{n}{m-1} \left(\frac{4}{3}\right)^{m-1} 1^{n-m+1} \\ &\leq \left(\frac{3}{8}\right)^n \left(\frac{4}{3} + 1\right)^n = \left(\frac{3}{8}\right)^n \left(\frac{7}{3}\right)^n = \left(\frac{7}{8}\right)^n \to 0 \end{aligned}$$

as $n \to \infty$ (here, the second inequality follows from the binomial identity).

To prove (iii), observe that

$$\begin{aligned} &2^{-n} \sum_{m=1}^{n-1} \binom{n}{m} m 2^{-m} \frac{2^{2(m-n)}}{2n-m} = 2^{-3n} \sum_{m=1}^{n-1} \binom{n}{m} 2^m \frac{m}{2n-m} \\ &\leq 2^{-3n} \sum_{m=1}^{n-1} \binom{n}{m} 2^m = 2^{-3n} \sum_{m=1}^{n-1} \binom{n}{m} 2^m 1^{n-m} \leq 2^{-3n} 3^n = \left(\frac{3}{8}\right)^n \to 0 \end{aligned}$$

as $n \to \infty$ (here, the first inequality follows since $\frac{m}{2n-m} \leq 1$ and the second inequality follows from the binomial identity). $\square$

Our last objective function-based approximation result below gives an upper bound on the largest difference between the optimal objective function values when the same linear function is minimized over the nonconvex set $\mathcal{S}^\kappa$ and the set $\mathcal{S}_P^\kappa$.

Proposition 12. *Let* $\mathcal{G} = \Delta^n$ *and* $\kappa > 1$. *Then, we have*

$$O_{\mathcal{S}^\kappa, \mathcal{S}_P^\kappa}(\alpha, \beta) \leq (-\beta_{j'}) \left(\kappa^{\frac{1}{1-\kappa}} - \kappa^{\frac{\kappa}{1-\kappa}} \right),$$

where $j' = \arg\min\{(\alpha_j + \beta_j) : \beta_j \leq 0\}$.

Proof. Due to Proposition 4, problem $\min_{(x,y)\in\mathcal{S}_P^\kappa}\{\alpha^T x+\beta^T y\}$ is equivalent to

$$\min_{x\in\mathbb{R}^n_+}\left\{\sum_{j:\beta_j>0}(\alpha_j x_j+\beta_j x_j^\kappa)+\sum_{j:\beta_j\le 0}(\alpha_j+\beta_j)x_j:\ \sum_{j:\beta_j>0}x_j+\sum_{j:\beta_j\le 0}x_j=1\right\}.$$

Since the objective function is linear in x_j variables with $\beta_j\le 0$ and they have identical coefficients in the constraint, at most one of them, in particular, the one with the smallest $\alpha_j+\beta_j$ value, can take nonzero value in an optimal solution. Note that the largest possible error is calculated as $\max_{x\in[0,1]}\{x-x^\kappa\}=\kappa^{\frac{1}{1-\kappa}}-\kappa^{\frac{\kappa}{1-\kappa}}$. Hence, the result follows. □

Notice that $\kappa^{\frac{1}{1-\kappa}}-\kappa^{\frac{\kappa}{1-\kappa}}$ converges to 0 and 1 as $\kappa\to 1^+$ and $\kappa\to\infty$, respectively, and it is equal to $\frac{1}{4}$ for $\kappa=2$.

3.4 Valid inequalities obtained via RPT

In this section, we use RPT [35, 5, 11] to derive power cone representable valid inequalities for $\mathcal{S}^\kappa$. Adapting the convention $\frac{0}{0}=0$, the specific application of this technique to our setting yields

$$1\ =\ \sum_{j=1}^n x_j\Rightarrow x_i^\kappa=\sum_{j=1}^n x_j x_i^\kappa\Rightarrow x_i^\kappa=\sum_{j=1}^n\frac{x_j^\kappa x_i^\kappa}{x_j^{\kappa-1}}\Rightarrow y_i=\sum_{j=1}^n\frac{X_{ij}^\kappa}{x_j^{\kappa-1}}\Rightarrow y_i\ge\sum_{j=1}^n\frac{X_{ij}^\kappa}{x_j^{\kappa-1}},$$

where we claim that the last inequality is power cone representable for $i=1,\dots,n$.

To formalize the above derivation, let us first define the RPT set $\mathcal{S}_V^\kappa=\{(x,y)\in\Delta^n\times\mathbb{R}^n:\exists X\in\mathbb{S}^n, w\in\mathbb{R}^{m\times n}:(13)\}$, where

$$X\ge 0, w\ge 0,\ y_i\ge\sum_{j=1}^n w_{ij}, i=1,\dots,n,\ (w_{ij},x_j,X_{ij})\in\mathcal{C}^{1/\kappa}, i,j=1,\dots,n. \tag{13}$$

We now show that the RPT set is an outer-approximation of the nonconvex set $\mathcal{S}^\kappa$, which is given for completeness below.

Proposition 13. *For $\mathcal{G}=\Delta^n$, we have $\mathcal{S}^\kappa\subseteq\mathcal{S}_V^\kappa$.*

Proof. Let $(x,y)\in\mathcal{S}^\kappa$. For $i,j=1,\dots,n$, define

$$X_{ij}=x_i x_j \text{ and } w_{ij}=\begin{cases}\frac{X_{ij}^\kappa}{x_j^{\kappa-1}} & \text{if } x_j>0\\ 0 & \text{if } x_j=0\end{cases}.$$

We will now show that constraints in the definition of the set $\mathcal{S}_V^\kappa$ are satisfied. Firstly, the constraint $y_i\ge\sum_{j=1}^n w_{ij}$ is satisfied since we have

$$y_i=x_i^\kappa=x_i^\kappa\sum_{j=1}^n x_j=x_i^\kappa\sum_{j=1:x_j>0}^n\frac{x_j^\kappa}{x_j^{\kappa-1}}=\sum_{j=1:x_j>0}^n\frac{X_{ij}^\kappa}{x_j^{\kappa-1}}=\sum_{j=1:x_j>0}^n w_{ij}=\sum_{j=1}^n w_{ij}.$$

Secondly, we consider the constraint $(w_{ij},x_j,X_{ij})\in\mathcal{C}^{1/\kappa}$, which is trivially satisfied if $x_i=0$ or $x_j=0$. Now, let us assume that $x_i>0$ and $x_j>0$, and observe that

$$w_{ij}^{1/\kappa}x_j^{1-1/\kappa}=\left(\frac{X_{ij}^\kappa}{x_j^{\kappa-1}}\right)^{1/\kappa}x_j^{1-1/\kappa}=\frac{X_{ij}}{x_j^{1-1/\kappa}}x_j^{1-1/\kappa}=X_{ij}.$$

Hence, we prove that $(x,y)\in\mathcal{S}_V^\kappa$ and the result follows. □

Remark 3. *Although Proposition 13 is given for $\mathcal{G} = \Delta^n$, it can easily be extended to a more general ground set $\mathcal{G} = \{x \in [0,1]^n : Ax = b\}$ with $A \in \mathbb{R}_+^{m\times n}$ and $b \in \mathbb{R}_+^m$. In this case, the RPT set is obtained as $\mathcal{S}_V^\kappa = \{(x,y) \in \mathcal{G} \times \mathbb{R}^n : \exists X \in \mathbb{S}^n, w \in \mathbb{R}^{m\times n} : (14)\}$, where*

$$\begin{gathered} X \geq 0, w \geq 0,\, b_\iota y_i \geq \sum_{j=1}^n A_{\iota j} w_{ij}, \iota = 1,\ldots,m, i = 1,\ldots,n, \\ (w_{ij}, x_j, X_{ij}) \in \mathcal{C}^{1/\kappa}, i,j = 1,\ldots,n. \end{gathered} \tag{14}$$

4 Computational Experiments

In this section, we present the results of our computational experiments, where the ground set $\mathcal{G}$ is chosen as the standard simplex Δ^n.

4.1 Computational setting

We use a 64-bit workstation with two Intel(R) Xeon(R) Gold 6248R CPU (3.00GHz) processors (256 GB RAM) and the Python programming language in our computational study. We utilize BARON 25.7.29 [34] to solve the nonconvex problems via Pyomo 6.4.0, and MOSEK 10.0.40 to solve the conic programming relaxations via CVXPY 1.5.2 with the default settings.

We compare the strength of the following nine relaxations against the nonconvex program defined over the set $\mathcal{S}^\kappa$, which we will refer to as the **NON** model:

- The **P** relaxation, defined over $\mathcal{S}_P^\kappa$
- The **PR** relaxation, defined over $\mathcal{S}_{P,R}^\kappa$
- The **PRs** relaxation, defined over $\mathcal{S}_{P,R,s}^\kappa$
- The **PRs**^{3} relaxation, defined over $\mathcal{S}_{P,s^3}^\kappa := \{(x,y) \in \mathcal{S}_P^\kappa : \exists X \in \mathbb{S}^n : (6), (15), (5)\}$, where

$$X \geq 0, \quad \begin{bmatrix} X_{ii} & X_{ij} & x_i \\ X_{ij} & X_{jj} & x_j \\ x_i & x_j & 1 \end{bmatrix} \succeq 0, 1 \leq i < j \leq n. \tag{15}$$

  Note that this relaxation only requires a subset of all 3×3 principal minors to be positive semidefinite.

- The **PRS** relaxation, defined over $\mathcal{S}_{P,R,S}^\kappa$
- The **PRV** relaxation, obtained by adding RPT constraints (13) to the **PR** relaxation
- The **PRsV** relaxation, obtained by adding RPT constraints (13) to the **PRs** relaxation
- The **PRs**^{3}**V** relaxation, obtained by adding RPT constraints (13) to the **PRs**^{3} relaxation
- The **PRSV** relaxation, obtained by adding RPT constraints (13) to the **PRS** relaxation

We run an extensive set of experiments parametrized by three key components:

- The distribution of objective function coefficients $\mathcal{D}$: We choose the objective function coefficients randomly with respect to two different distributions: i) Unif$(-1, 1)$, ii) standard normal.
- The value of exponent κ: We choose six different κ values given by the following set: $\{1.25, 1.5, 1.75, 2, 2.5, 3\}$.
- The value of dimension n: We choose nine different n values given by the following set: $\{2, 3, \ldots, 10\}$.

We will call a triplet of $(\mathcal{D}, \kappa, n)$ a *setting* (notice that we have $2 \times 6 \times 9 = 108$ settings in total). For each setting given by the triplet $(\mathcal{D}, \kappa, n)$, we repeat the experiment 1000 times, and solve the **NON** model by BARON and the nine relaxations listed above by MOSEK. Therefore, in total, 108000 instances are created and each instance is solved ten times. MOSEK has given an `UNKNOWN` status for four instances for at least one relaxation, which are excluded from the analysis below.

We record the objective function value of each optimization problem solved as $z^{\mathbf{M}}_{\mathcal{D},\kappa,n,r}$ and the times it takes to solve it as $t^{\mathbf{M}}_{\mathcal{D},\kappa,n,r}$, where r stands for the replication index in the setting $(\mathcal{D}, \kappa, n)$ and $\mathbf{M}$ is the model solved (either the **NON** model or any of the nine relaxations listed above). Then, for each instance and model, we compute the *absolute dual gap* as $z^{\mathbf{NON}}_{\mathcal{D},\kappa,n,r} - z^{\mathbf{M}}_{\mathcal{D},\kappa,n,r}$, and define the cumulative absolute gap of a setting $(\mathcal{D}, \kappa, n)$ for model $\mathbf{M}$ as

$$\text{Cumulative Gap}^{\mathbf{M}}_{\mathcal{D},\kappa,n} = \sum_{r=1}^{1000} \left(z^{\mathbf{NON}}_{\mathcal{D},\kappa,n,r} - z^{\mathbf{M}}_{\mathcal{D},\kappa,n,r} \right).$$

Note that, by construction, Cumulative $\text{Gap}^{\mathbf{NON}}_{\mathcal{D},\kappa,n} = 0$ for every setting. The cumulative time of a setting $(\mathcal{D}, \kappa, n)$ for model $\mathbf{M}$ is computed as

$$\text{Cumulative Time}^{\mathbf{M}}_{\mathcal{D},\kappa,n} = \sum_{r=1}^{1000} t^{\mathbf{M}}_{\mathcal{D},\kappa,n,r}.$$

Finally, we will say that a relaxation is *exact* for an instance if the gap is less than 10^{-4} and we record the number of times a model $\mathbf{M}$ is exact in the setting $(\mathcal{D}, \kappa, n)$ as

$$\#\text{Exact}^{\mathbf{M}}_{\mathcal{D},\kappa,n} = \sum_{r=1}^{1000} \mathbf{1} \left(z^{\mathbf{NON}}_{\mathcal{D},\kappa,n,r} - z^{\mathbf{M}}_{\mathcal{D},\kappa,n,r} \leq 10^{-4} \right),$$

where $\mathbf{1}(\cdot)$ is the indicator function. These three metrics, Cumulative $\text{Gap}^{\mathbf{M}}_{\mathcal{D},\kappa,n}$, Cumulative $\text{Time}^{\mathbf{M}}_{\mathcal{D},\kappa,n}$ and $\#\text{Exact}^{\mathbf{M}}_{\mathcal{D},\kappa,n}$, will be our main performance criteria in the following discussion.

4.2 Results

We now provide the details of our computational study. We note that each statistic given below is averaged over the settings considered.

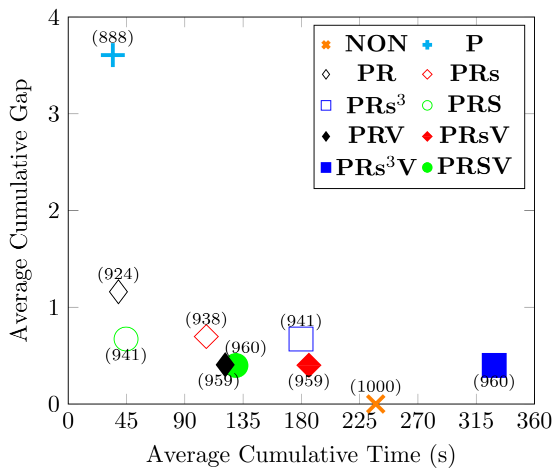
4.2.1 Aggregate Results

We first provide the most aggregate results when all the 108 settings are considered together in Figure 3. We observe that the **P** relaxation, which is the weakest one, can solve 888 instances on the average and the average cumulative gap is 3.607 while it only takes 34.429 seconds to solve. The addition of the RLT constraints has the largest marginal effect, increasing the average number of exact instances to 924 and reducing the average cumulative gap to 1.160 in the **PR** relaxation by the increase of a mere 4 seconds. The **PRS** relaxation further increases the average number of exact instances to 941 and reduces the average cumulative gap to 0.672 while only taking 6 seconds longer than the **PR** relaxation. Interestingly, the average cumulative gaps reported by the **PRs** and **PRs**^{3} relaxations are very close to the **PRS** relaxation, albeit the computational cost of these weaker relaxations are relatively high. The addition of the RPT constraints (13) further improves the strength of the relaxations. Relaxations **PRV** and **PRSV** are almost indistinguishable as the number of exact instances (959 vs. 960), the gaps (0.405 vs. 0.398) and the times (121.177 vs. 129.492) are all very close to each other. We again observe that weaker relaxations **PRsV** and **PRs**^{3}**V** take more time than the **PRSV** relaxation.

The comparison of **PRS** and **PRV** relaxations leads to interesting observations. As proven in Theorem 1, both relaxations are exact when $n = 2$ and $\kappa = 2$. The **PRS** relaxation is also exact for $n \geq 3$ and $\kappa = 2$ in our experiments whereas the average cumulative gap of the **PRV** relaxation is 0.023 and the average number of exact instances is 992 across these 16 settings. In the remaining 90 settings where $\kappa \neq 2$, the **PRV** relaxation is more successful with an average cumulative gap of 0.482 and average number of exact instances of 952 whereas the **PRS** relaxation has an average cumulative gap of 0.806 and average number of exact instances of 929. We note that these relaxations are not comparable in general as there exist instances in which the bound given by the **PRS** relaxation is stronger than that of the **PRV** relaxation and vice versa.

To summarize, we can say that in circumstances in which the time budget is limited, **PR** and **PRS** relaxations offer reasonable alternatives to the **NON** model solved by BARON. When more computational budget is available, the addition of RPT constraints (13) can further improve the relaxation quality, with the **PRV** relaxation in particular being an effective option.

Figure 3: Average Cumulative Gap vs. Average Cumulative Time with respect to all settings considered. The average number of exact instances is given in parenthesis.



We again note that the strengths of **PRs** and **PRs**^{3} relaxations are very close to the **PRS** relaxation, and this observation remains valid with the addition of the RPT constraints (13). Related to the **PRs**^{3} relaxation, we have the following conjecture:

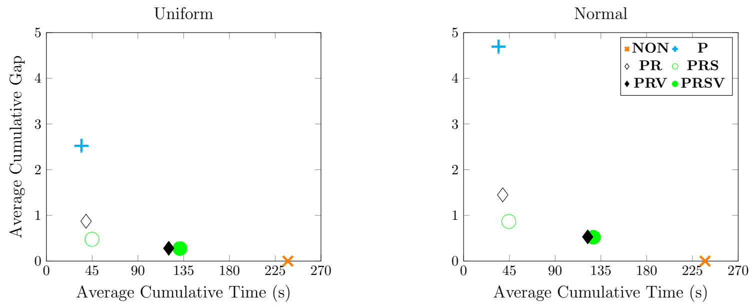
Conjecture 2. *We have* $\mathcal{S}^{\kappa}_{P,R,s^3} = \mathcal{S}^{\kappa}_{P,R,S}$.

Although it might be interesting to pursue exploring this observation in the future theoretically, from a computational standpoint, **PRs**, **PRs**^{3}, **PRsV** and **PRs**^{3}**V** relaxations are simply too expensive alternatives and dominated by stronger relaxations in terms of computational effort. Therefore, we omit them from the discussion below.

4.2.2 The effect of distribution $\mathcal{D}$

We now analyze the effect of the distribution of objective function coefficients $\mathcal{D}$ on the results, which are reported in Figure 4. Although Unif$(-1, 1)$ and standard normal both have mean zero, their variances are 1 and 1/6, respectively. This is reflected on the outcomes as the average cumulative gap values for the standard normal are higher than those of the uniform distribution. We also note that these observations are consistent with the simulation results that we have conducted in Section 3.3.2 (see Tables 1 and 2).

Figure 4: Average Cumulative Gap vs. Average Cumulative Time with respect to different distributions for the objective function coefficients.

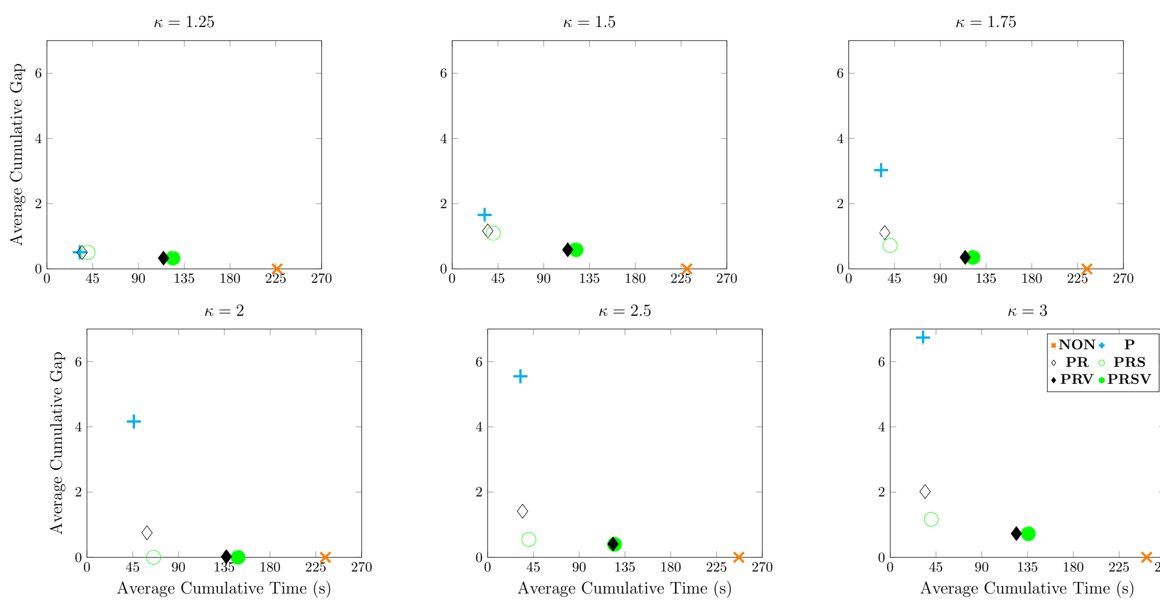


In terms of computational cost, the choice of the distribution does not seem to make a significant difference.

4.2.3 The effect of exponent κ

We next analyze the effect of exponent κ on the results, which are reported in Figure 5. We observe that the average cumulative gap of the **P** relaxation increases with κ, which is an expected outcome due to Proposition 12. Interestingly, the average cumulative gap of other relaxations first increases, then decreases with respect to κ. As stated in Conjecture 1, the **PRS** relaxation is exact when $\kappa = 2$ in all our experiments. We also note that all relaxations are quite strong around $\kappa = 2$, which suggests that the effect of the RLT constraints is most influential around this value.

Figure 5: Average Cumulative Gap vs. Average Cumulative Time with respect to different κ values.



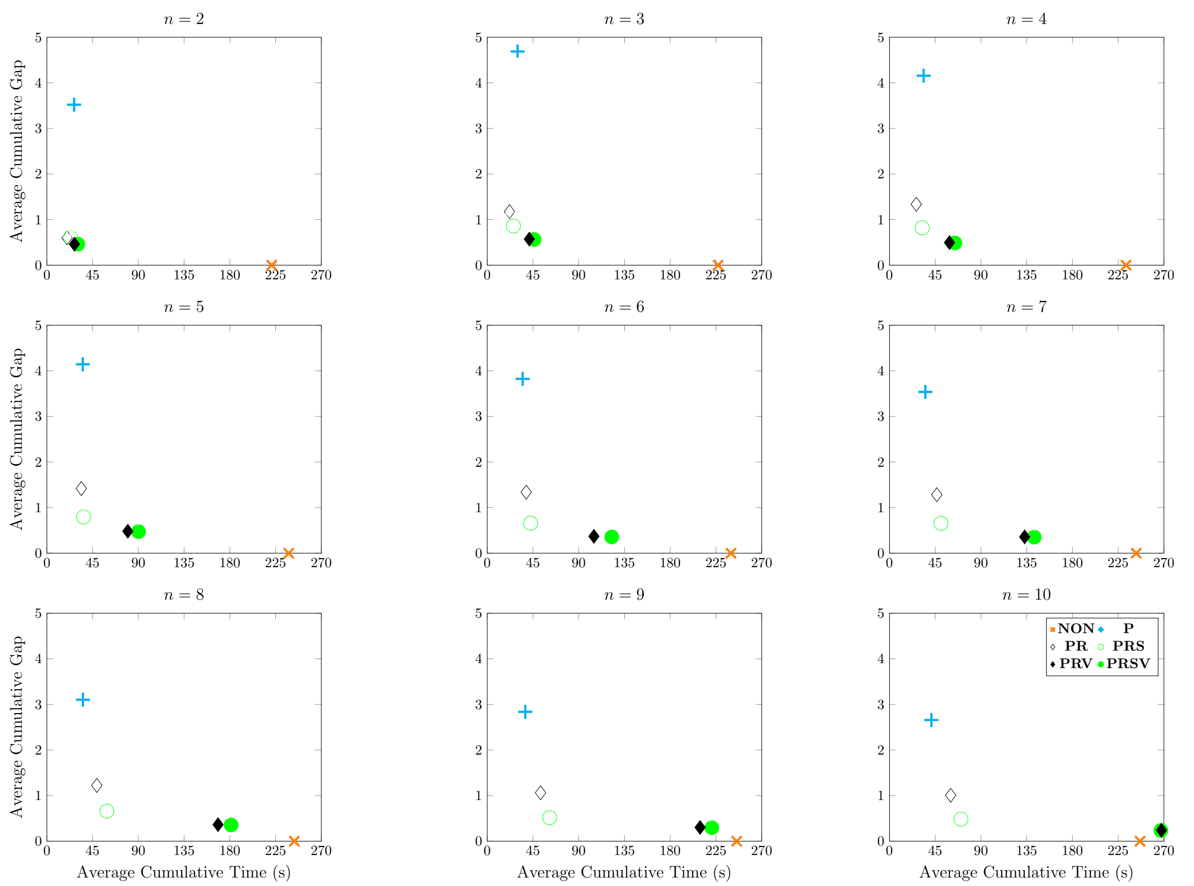
In terms of computational cost, we do not observe a significant difference or a consistent trend for different values of κ.

4.2.4 The effect of dimension n

Finally, we analyze the effect of dimension n on the results, which are reported in Figure 6. The most interesting observation here is that the average cumulative gap of the **P** relaxation is highest when $n = 3$.

We also note that the average cumulative gap gradually decreases with higher values of n, which is in agreement with the simulation results that we have reported in Section 3.3.2 (see Tables 1 and 2) and Theorem 4. The other relaxations give similar average cumulative gaps for small values of n whereas their gaps are getting more different with larger values of n.

Figure 6: Average Cumulative Gap vs. Average Cumulative Time with respect to different n values.



In terms of computational cost, we observe a slight increase in the solution of the **NON** model with n whereas the times of **P**, **PR** and **PRS** relaxations increase moderately. We note that the times of **PRV** and **PRSV** relaxations are most sensitive to the dimension n and they become even more expensive to solve than the **NON** model for $n = 10$. Combined with the fact that the effect of the RPT constraints (13) diminishes with dimension, it is not advisable to use them in higher dimensions.

5 Conclusion

In this paper, we studied the nonconvex set $\mathcal{S}^{\kappa}$, which is directly related to separable standard quadratic programming and appears as a substructure in potential-based flow networks. We proposed several conic relaxations for this nonconvex set, and compared these relaxations both theoretically and via extensive computational experiments. Our main conclusions are as follows. The **P** relaxation, which is the weakest relaxation among those considered, provides quite strong bounds when a linear function is optimized over it. In particular, we derived worst-case bounds on the distance between the **P** relaxation and $\mathcal{S}^{\kappa}$, and performed a probabilistic analysis which demonstrates that this relaxation is exact with high probability in higher dimensions. Beyond **P**, our results highlight that the RLT relaxation is very important. In particular, the exact convex hull in low-dimensions for $\kappa = 2$ is obtained using the **PR** relaxation. Computationally, we see that **PR** gives significantly better bounds in comparison to **P**. Interestingly, we proved that addition of PSD constraints do not add any value on top on **P**, while when these are added to **PR** can produce non-trivial improvement to bounds. Finally, **PRS** and **PRV** are incomparable to each other, both produce improvements over **PR**, with **PRS** being more efficient to solve, but giving slightly worse bounds than **PRV** on average.

We would like to note that although the separable function $\sum_{j=1}^{n}(\alpha_j x_j + \beta_j x_j^\kappa)$ we consider in this study may seem special at first glance, most of our results extend to the more general case where we replace each function $\alpha_j x_j + \beta_j x_j^\kappa$ with a general convex function of x_j that is non-decreasing on the ground set.

There are several future research avenues we would like to explore. For instance, resolving Conjecture 1, which states that the **PRS** relaxation is exact for $\kappa = 2$, and Conjecture 2, which states that $\mathbf{PRS}^3$ and **PRS** relaxations are equivalent, are two immediate directions for further theoretical analysis. Moreover, analyzing the cases with a separable function with multiple exponents, exponents where $\kappa < 1$ and different ground sets from both theoretical and empirical aspects are also promising.

Acknowledgments

We gratefully acknowledge several helpful discussions with Nick Sahinidis on this topic.

References

[1] Warren Adams, Akshay Gupte, and Yibo Xu. Error bounds for monomial convexification in polynomial optimization. *Mathematical Programming*, 175(1):355–393, 2019.

[2] Xiaowei Bao, Aida Khajavirad, Nikolaos V. Sahinidis, and Mohit Tawarmalani. Global optimization of nonconvex problems with multilinear intermediates. *Mathematical Programming Computation*, 7(1):1–37, 2015.

[3] Pietro Belotti, Andrew J. Miller, and Mahdi Namazifar. Valid inequalities and convex hulls for multilinear functions. *Electronic Notes in Discrete Mathematics*, 36:805–812, 2010.

[4] Harold P. Benson. Concave envelopes of monomial functions over rectangles. *Naval Research Logistics (NRL)*, 51(4):467–476, 2004.

[5] D. Bertsimas, D. de Moor, D. den Hertog, T. Koukouvinos, and J. Zhen. Cone product reformulation for global optimization. *Optimization Online, 2023b. URL https://optimization-online. org, b*, 2023.

[6] I. M. Bomze and E. De Klerk. Solving standard quadratic optimization problems via linear, semidefinite and copositive programming. *Journal of Global Optimization*, 24(2):163–185, 2002.

[7] I. M. Bomze and M. Locatelli. Separable standard quadratic optimization problems. *Optimization Letters*, 6(5):857–866, 2012.

[8] I. M. Bomze, W. Schachinger, and R. Ullrich. The complexity of simple models—a study of worst and typical hard cases for the standard quadratic optimization problem. *Mathematics of Operations Research*, 43(2):651–674, 2018.

[9] P. Börner, M. Klimm, A. Lutz, M. E. Pfetsch, M. Skutella, and L. Strubberg. Valid cuts for the design of potential-based flow networks. In *International Conference on Integer Programming and Combinatorial Optimization*, pages 142–156. Springer, 2025.

[10] S. Burer. On the copositive representation of binary and continuous nonconvex quadratic programs. *Mathematical Programming*, 120(2):479–495, 2009.

[11] S. S. Dey and A. Khajavirad. A second-order cone representable class of nonconvex quadratic programs. *arXiv preprint arXiv:2508.18435*, 2025.

[12] S. S. Dey, B. Kocuk, and N. Redder. Node-based valid inequalities for the optimal transmission switching problem. *Discrete Optimization*, 43:100683, 2022.

[13] Claudia D'Ambrosio, Andrea Lodi, Sven Wiese, and Cristiana Bragalli. Mathematical programming techniques in water network optimization. *European Journal of Operational Research*, 243(3):774–788, 2015.

[14] M. R. Garey and D. S. Johnson. *Computers and intractability*, volume 29. wh freeman New York, 2002.

[15] Y. G. Gökmen and E. A. Yıldırım. On standard quadratic programs with exact and inexact doubly nonnegative relaxations. *Mathematical Programming*, 193(1):365–403, 2022.

[16] J. Gondzio and E. A. Yıldırım. Global solutions of nonconvex standard quadratic programs via mixed integer linear programming reformulations. *Journal of Global Optimization*, 81(2):293–321, 2021.

[17] M. Gross, M. E. Pfetsch, L. Schewe, M. Schmidt, and M. Skutella. Algorithmic results for potential-based flows: Easy and hard cases. *Networks*, 73(3):306–324, 2019.

[18] Taotao He, Siyue Liu, and Mohit Tawarmalani. Convexification techniques for fractional programs. *Mathematical Programming*, pages 1–43, 2024.

[19] J. Júdice, V. Sessa, and M. Fukushima. A two-phase sequential algorithm for global optimization of the standard quadratic programming problem. *Journal of Global Optimization*, pages 1–51, 2024.

[20] Y. Li, S. S. Dey, and N. V. Sahinidis. A reformulation-enumeration minlp algorithm for gas network design. *Journal of Global Optimization*, 90(4):931–963, 2024.

[21] Y. Li, S. S. Dey, N. V. Sahinidis, N. Susarla, M. A. Zamarripa, and M. G. Drouven. Optimizing the design and operation of water networks: Two decomposition approaches. *Computers & Chemical Engineering*, 192:108897, 2025.

[22] G. Liuzzi, M. Locatelli, and V. Piccialli. A new branch-and-bound algorithm for standard quadratic programming problems. *Optimization Methods and Software*, 34(1):79–97, 2019.

[23] John E Maxfield and Henryk Minc. On the matrix equation x' x= a. *Proceedings of the Edinburgh Mathematical Society*, 13(2):125–129, 1962.

[24] Clifford A. Meyer and Christodoulos A. Floudas. Convex envelopes for edge-concave functions. *Mathematical programming*, 103(2):207–224, 2005.

[25] B. C. Okumusoglu and B. Kocuk. Global optimization of gas transportation and storage: Convex hull characterizations and relaxations. *arXiv preprint arXiv:2503.15143*, 2025.

[26] Anatoliy D. Rikun. A convex envelope formula for multilinear functions. *Journal of Global Optimization*, 10(4):425–437, Jun 1997.

[27] R. Z. Ríos-Mercado and C. Borraz-Sánchez. Optimization problems in natural gas transportation systems: A state-of-the-art review. *Applied Energy*, 147:536–555, 2015.

[28] Hong Seo Ryoo and Nikolaos V. Sahinidis. Analysis of bounds for multilinear functions. *Journal of Global Optimization*, 19(4):403–424, 2001.

[29] H. D. Sherali and W. P. Adams. *A reformulation-linearization technique for solving discrete and continuous nonconvex problems*, volume 31. Springer Science & Business Media, 2013.

[30] Hanif D. Sherali. Convex envelopes of multilinear functions over a unit hypercube and over special discrete sets. *Acta mathematica vietnamica*, 22(1):245–270, 1997.

[31] Mohit Tawarmalani and Nikolaos V. Sahinidis. *Convexification and global optimization in continuous and mixed-integer nonlinear programming: theory, algorithms, software, and applications*, volume 65. Springer Science & Business Media, 2002.

[32] Hoang Tuy. *Convex analysis and global optimization*, volume 110. Springer, 2016.

[33] Luze Xu and Jon Lee. Gaining or losing perspective for convex multivariate functions on box domains. *Mathematical Programming*, 211(1):319–339, 2025.

[34] Y. Zhang and N. V. Sahinidis. Solving continuous and discrete nonlinear programs with baron. *Computational Optimization and Applications*, pages 1–39, 2024.

[35] J. Zhen, D. de Moor, and D. den Hertog. An extension of the reformulation-linearization technique to nonlinear optimization. *Available at Optimization Online*, 2021.