Noname manuscript No.
(will be inserted by the editor)

Convexification techniques for fractional programs

Taotao He · Siyue Liu · Mohit Tawarmalani

Received: date / Accepted: date

Abstract This paper develops a correspondence relating convex hulls of fractional functions with those of polynomial functions over the same domain. Using this result, we develop a number of new reformulations and relaxations for fractional programming problems. First, we relate $0-1$ problems involving a ratio of affine functions with the boolean quadric polytope, and use inequalities for the latter to develop tighter formulations for the former. Second, we derive a new formulation to optimize a ratio of quadratic functions over a polytope using copositive programming. Third, we show that univariate fractional functions can be convexified using moment hulls. Fourth, we develop a new hierarchy of relaxations that converges finitely to the simultaneous convex hull of a collection of ratios of affine functions of $0-1$ variables. Finally, we demonstrate theoretically and computationally that our techniques close a significant gap relative to state-of-the-art relaxations, require much less computational effort, and can solve larger problem instances.

Keywords fractional programming; boolean quadric polytopes; quadratic optimization; copositive optimization; moment hull; assortment optimization; distillation

1 Introduction

Fractional terms arise in many combinatorial and nonconvex optimization problems [37,71,17] to model choice behavior, financial and performance ratios, variational principles, engineering models, and geometric characteristics. For example,

Taotao He
Antai College of Economics and Management, Shanghai Jiao Tong University
hetaotao@sjtu.edu.cn, https://taotaohe.github.io/

Siyue Liu
Tepper School of Business, Carnegie Mellon University
siyueliu@andrew.cmu.edu

Mohit Tawarmalani
Mitchell E. Daniels, Jr. School of Business, Purdue University
mtawarma@purdue.edu, http://web.ics.purdue.edu/∼mtawarma/

Luce-type choice models [53] are used in feature selection [61], assortment optimization [59], and facility location problems [6]; financial and performance ratios are used to measure return on investment [49,26], fairness [80], quality of regression models [39], efficiency of queuing systems [9], performance of restricted policies [3], quality of set-covers [2], consistent biclustering [22], and other graph properties such as density [50] or connectedness; engineering models capture physical processes such as Underwood equations in distillation configuration design [79] and exergy models in chemical processes [41]; geometrical measures are used to measure flatness of shapes and condition numbers of matrices [24].

We are interested in optimization problems of the following form:

$$\max_{x \in \mathcal{X}} \quad \sum_{i=1}^{m} \frac{f_i(x)}{g_i(x)},$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$ and $g_i : \mathbb{R}^n \to \mathbb{R}$ are polynomial functions, and $\mathcal{X}$ is a set described using polynomial inequalities. We will particularly be interested in the following special cases: (i) $\mathcal{X} = \bigl\{x \in \mathbb{R}^n \bigm| Cx \leq d,\ x_i \in \{0,1\} \text{ for } i \leq p\bigr\}$, where $C \in \mathbb{R}^{r \times n}$ and $d \in \mathbb{R}^r$, and (ii) functions $f_i(\cdot)$ and $g_i(\cdot)$ are affine and/or quadratic functions.

The literature on fractional programming can be divided into two categories. The first category consists of problems that can be solved in polynomial time. Two prominent examples include problems treated in [23] and [56]. These two seminal works focused on fractional programming problems where the objective function is a ratio of affine functions and the feasible region $\mathcal{X}$ is described, respectively, by linear inequalities or a combinatorial set over which linear functions can be optimized easily. The second category of problems do not admit polynomial time algorithms. A simple example is where the objective is to maximize a sum of two ratios of affine functions defined over a polytope [54]. The resulting problem is NP-hard. Many applied problems belong to the second category, and are solved using branch & bound algorithms that rely on convexification or reformulation strategies [34,74,8,59,17,69,68,57]. However, an intriguing gap exists in the literature. Specifically, if problem instances of the first category are solved using techniques adopted for the second category, the resulting relaxations are not tight and the guarantee of polynomial time convergence is lost. This is arguably because the results that enable polynomial-time solution algorithms do not directly yield convexification results. In this paper, we discover a projective transformation that permits these results to be viewed from a convexification lens. Leveraging this insight, we improve the quality of relaxations for a variety of fractional programming problems that belong to the second category.

Below, we highlight a few of our contributions:

1. We reduce the task of convexifying fractional programming problems to that of convexifying semi-algebraic sets relying on a projective transformation that is of independent interest. The reduction unifies approaches to solve single-ratio fractional optimization problems with those to create convex relaxations for harder problems.
2. We establish a precise connection between 0-1 linear fractional programming, where the objective is a ratio of affine functions, and the convexification of the boolean quadric polytope (BQP) and/or cut polytope. This allows us to

leverage valid inequalities for BQP [5,62] to improve relaxations for $0-1$ fractional programming problems. Our computational results demonstrate that the developed relaxations are significantly tighter on a collection of assortment optimization problems.
3. We show that the problem of optimizing a ratio of quadratic functions over a polytope, possibly with binary constraints, can be reformulated into a copostive programming problem. This result extends prior results in copositive optimization literature [18].
4. We discover a one-to-one correspondence between moment hulls [48,66] and the convex hull of inverses of univariate linear forms. These inverses occur in practical design problems involving chemical separations [79,40,41] and we show that the resulting relaxations are significantly tighter than those typically used in the literature.
5. When the objective is a sum-of-ratios of binary variables, we establish a finite hierarchy of reformulations that converges to the convex hull of the problem. The result reveals an intriguing decomposition. We show that convexifying each ratio individually, in a specific way, leads to the simultaneous hull of multiple ratios.
6. We provide explicit convex hulls for small-dimensional sets extending previous results [76,19] that are used in solvers to develop relaxations of factorable functions.
7. We demonstrate theoretically and computationally that our reformulations and relaxations dominate state-of-the-art schemes in the literature [64,77,57]. For example, on a collection of assortment planning problems [58] and sets arising in chemical process design, we show that the new relaxations close approximately 50% of the gap. Our reformulations solve faster and enable an exact solution of larger problem instances.

The paper is organized as follows. In Section 2, we establish a precise correspondence between fractional optimization and polynomial optimization. This allows us to unify the techniques of [23] and [56] within a common framework. In Section 3, we exploit inequalities for BQP to tighten relaxations for $0-1$ fractional programs. In Section 4, we consider problems involving optimization of a ratios of quadratic functions and, when the domain is a polytope, provide a reformulation using copositive programming. In Section 5.1, we establish a correspondence between simultaneously convexifying inverses of univariate affine functions and the moment hull. In Section 5.2, we develop a hierarchy of relaxations that converges to the convex hull of a sum-of-ratios optimization problem over binary variables. In Section 6, we propose a relaxation for linear fractional programming and prove that our relaxation dominates existing relaxations in the literature. We also present computational experiments with $0-1$ fractional programs and a set that appears in the design of distillation trains. Concluding remarks are provided in Section 7. All missing proofs are included in Appendix A.

Notation For a positive integer n, let $[n]$ denote $\{1,\dots,n\}$. We use $\mathbb{R}$ to denote the set of real numbers, and $\mathbb{R}_+$ (resp. $\mathbb{R}_{++}$) to denote the set of nonnegative (resp. positive) real numbers. For a set $S \subseteq \mathbb{R}^n$, $\mathrm{conv}(S)$ denotes the convex hull of S, $\mathrm{rec}(S)$ denotes the recession cone of S, $\mathrm{ri}(S)$ denote the relative interior of S, and $\mathrm{cl}\, S$ denote the closure of S. For $\lambda \geq 0$, $\lambda S := \{\lambda x \mid x \in S\}$ if $\lambda > 0$ and $\lambda S = \mathrm{rec}(S)$ if $\lambda = 0$. We will write $\frac{x}{\alpha^\intercal x}$, where $x \in \mathbb{R}^n$, to denote that each element of x is divided by $\alpha^\intercal x$. For any d, we use $\mathbb{S}^d$ to denote the set of symmetric $d \times d$ matrices, $\mathbb{S}^d_+$ to denote the set of symmetric positive semidefinite matrices. We will use $X \succeq 0$ to denote $X \in \mathbb{S}^d_+$. For two $m \times n$ matrices X and Y, its inner product is given by $\langle X, Y\rangle = \mathrm{Tr}(XY^\intercal) = \sum_{i\in[m]}\sum_{j\in[n]} X_{ij}Y_{ij}$.

2 Simultaneous convexification of fractional terms

Fractional programming has been studied extensively and has a diverse set of applications. A seminal result of [23] shows how to optimize a linear fractional function over a polytope as long as the denominator is positive over this region. Formally, the result solves a class of linear fractional programs in the following form:

$$\min\left\{ \frac{b_0 + b^\intercal x}{a_0 + a^\intercal x} \;\middle|\; Cx \le d,\ x \ge 0 \right\}, \tag{1}$$

where $C \in \mathbb{R}^{r\times n}$, d is an r dimensional vector, (a_0, a) is a vector in $\mathbb{R} \times \mathbb{R}^n$ such that $a_0 + a^\intercal x$ is positive over the feasible region, and (b_0, b) is a vector in $\mathbb{R} \times \mathbb{R}^n$. The key step reduces (1) to the following linear program which can be solved in polynomial time

$$\min\left\{ b_0\rho + b^\intercal y \;\middle|\; Cy \le d\rho,\ a_0\rho + a^\intercal y = 1,\ y \ge 0,\ \rho \ge 0 \right\}.$$

Another fundamental result in [56] states that if for any $c \in \mathbb{Z}^n$ the problem

$$\min\left\{ c^\intercal x \;\middle|\; x \in \mathcal{X} \subseteq \{0,1\}^n \right\} \tag{2}$$

can be solved in $O(p(n))$ comparisons and $O(q(n))$ additions then, for $(a_0, a) \in \mathbb{Z} \times \mathbb{Z}^n$ and $(b_0, b) \in \mathbb{Z} \times \mathbb{Z}^n$ such that $a_0 + a^\intercal x$ is positive over $\mathcal{X}$, the problem

$$\min\left\{ \frac{b_0 + b^\intercal x}{a_0 + a^\intercal x} \;\middle|\; x \in \mathcal{X} \subseteq \{0,1\}^n \right\} \tag{3}$$

can be solved in $O\left(p(n)\left(p(n) + q(n)\right)\right)$ time. These results enable polynomial-time solution procedures for a variety of fractional programming problems. Regardless, if either (1) or (3) is provided to a global optimization solver, the resulting relaxation does not exploit these tractability results, raising a question is whether the above results provide any insights for relaxation construction.

We remark that the polynomial-time algorithms do not imply that the convex hull of a fractional function over $\mathcal{X}$ can be separated in polynomial time. To see this, for any $c \in \mathbb{R}^n$, $\min\{c^\intercal x \mid x \in \mathcal{X} = \{0,1\}^n\}$ is easy to solve, as it attains optimum at $x_i = 1$ if $c_i < 0$ and $x_i = 0$ otherwise. Yet, it is hard to separate the convex hull of the graph a fractional function over $\{0,1\}^n$, given as follows:

$$\left\{ \left(x, \frac{b_0 + b^\intercal x}{a_0 + a^\intercal x}\right) \;\middle|\; x \in \{0,1\}^n \right\}.$$

This is because, if the separation problem could be solved in polynomial time, we could use the ellipsoid algorithm to solve the following optimization problem in polynomial time, see Corollary 2.9 in [75],

$$\min\left\{ \frac{b_0 + b^\intercal x}{a_0 + a^\intercal x} + c^\intercal x \;\middle|\; x \in \{0,1\}^n \right\}.$$

However, the latter problem is NP-hard as shown in [44] (also, see the later discussion in Example 3). Therefore, unless P = NP, the separation problem for the graph of a linear fractional function over $\{0,1\}^n$ is not solvable in polynomial time. Moreover, as [54] showed, it is also NP-hard to solve the following optimization problem:

$$\min_x \Big\{ x_1 - \frac{1}{x_2} \;\Big|\; Cx \le d,\ x \ge 0 \Big\},$$

which implies that it is NP-hard to find a hyperplane separating an arbitrary point from the convex hull of a set involving a single fractional term, that is

$$\left\{ \left(x, \frac{1}{x_2} \right) \;\middle|\; Cx \le d,\ x \ge 0 \right\}.$$

It follows that Charnes-Cooper reformulation and Megiddo's algorithm do not convexify the fractional function, but only allow us to optimize this function, over the corresponding domain. We will derive a convexification result, in a transformed space, that recovers polynomial-time solvability of (1) and (3) under conditions described above. More importantly, this result yields insights into relaxations of fractional functions for general mixed-integer nonlinear programming problems.

Our relaxation scheme introduces following variables to represent the fractional terms

$$\rho = \frac{1}{a_0 + a^\intercal x} \quad \text{and} \quad y_i = \frac{x_i}{a_0 + a^\intercal x} \quad \text{for } i \in [n].$$

Using these new variables, the problems (1) and (3) are rewritten as:

$$\min \Big\{ b_0 \rho + b^\intercal y \;\Big|\; x \in \mathcal{X},\ (\rho, y) = \frac{(1,x)}{a_0 + a^\intercal x} \Big\},$$

where, (1) is modeled with $\mathcal{X} = \{x \mid Cx \le d, x \ge 0\}$, whereas, for (3), $\mathcal{X}$ is a subset of $\{0,1\}^n$ such that (2) is easily solvable. Even though

$$\left\{ \left(x, \frac{b_0 + b^\intercal x}{a_0 + a^\intercal x} \right) \;\middle|\; x \in \mathcal{X} \right\}$$

cannot be easily separated over these domains, we will argue that convex hull of the following set can be separated in polynomial time:

$$\Big\{ (\rho, y) \;\Big|\; x \in \mathcal{X},\ (\rho, y) = \frac{(1,x)}{a_0 + a^\intercal x} \Big\}. \tag{4}$$

We highlight the main differences between these sets. Specifically, the x variables are projected out in (4) and one fractional function is replaced with many fractional terms.

The convex hull of (4) will follow from a more general fact that relates the convex hull of a set to that of its projective transform. Consider a transformation given as follows

$$\Phi(\rho, x) = \left(\frac{1}{\rho}, \frac{x}{\rho} \right) \qquad \text{for every } (\rho, x) \in \mathbb{R}_{++} \times \mathbb{R}^n . \tag{5}$$

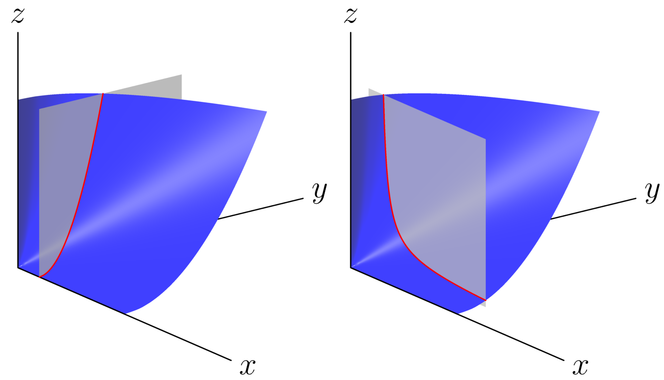
To fix ideas, the transformation Φ maps the set (4) to $\big\{(a_0 + a^\intercal x, x) \mid x \in \mathcal{X}\big\}$. Theorem 1, which we show below, allows us to use the convex hull of the transformed set to derive the convex hull of (4). For a subset $\mathcal{S}$ of $\mathbb{R}_{++} \times \mathbb{R}^n$, the image of $\mathcal{S}$ under Φ can be visualized as follows. First, we homogenize the set $\mathcal{S}$ with a scaling variable σ, and obtain a cone

$$\{(\sigma, \sigma\rho, \sigma x) \mid (\rho, x) \in \mathcal{S},\ \rho > 0,\ \sigma > 0\}.$$

$\mathcal{S}$ can be viewed as the intersection of this cone with $\{\sigma = 1\}$ (after projecting out the first coordinate). Then, we consider a different cross-section of this cone,

$$\{(\sigma, \sigma\rho, \sigma x) \mid (\rho, x) \in \mathcal{S},\ \rho > 0,\ \sigma > 0,\ \sigma\rho = 1\}.$$

Now, the image of $\mathcal{S}$ under Φ is obtained by projecting the new cross-section onto the space of the first and last coordinates. In Figure 1, we exemplify this transformation by showing that quadratic x^2 and the reciprocal $\frac{1}{x}$ as different sections of the same cone.



(a) Slice: $\{(y,z) \mid z \geq y^2, y \geq 0\}$ (b) Slice: $\left\{(x,z) \,\middle|\, z \geq \frac{1}{x}\right\}$

Fig. 1: Two views of the cone $\{(x,y,z) \in \mathbb{R}^3_+ \mid y \leq \sqrt{xz}\}$. When x-coordinate (resp. y-coordinate) is fixed, we obtain the slice shown in Figure(1a) (resp. Figure(1b)). For $\lambda, \gamma, \theta \geq 0$, consider rays $(25\lambda, 5\lambda, \lambda)$ and $(25\gamma, 20\gamma, 16\gamma)$ in the cone and the ray $(25\theta, 8\theta, 4\theta)$ that is in the convex cone generated by these rays. This observation leads to two convex representations of interest. Specifically, when y-coordinate (resp. x-coordinate) is scaled to 1, we have $\left(\frac{25}{8}, 1, \frac{1}{2}\right) = \frac{1}{2}\left(5, 1, \frac{1}{5}\right) + \frac{1}{2}\left(\frac{5}{4}, 1, \frac{4}{5}\right)$ (resp. $\left(1, \frac{8}{25}, \frac{4}{25}\right) = \frac{4}{5}\left(1, \frac{1}{5}, \frac{1}{25}\right) + \frac{1}{5}\left(1, \frac{4}{5}, \frac{16}{25}\right)$). The former is a combination of points on the curve $z = \frac{1}{x}$, while the latter is a convex combination of points on $z = y^2$.

Theorem 1 *For a subset $\mathcal{S}$ of $\mathbb{R}_{++} \times \mathbb{R}^n$, we have*

$$\begin{aligned} \operatorname{conv}(\mathcal{S}) &= \{(\rho, x) \mid (1, x) \in \rho \operatorname{conv}(\Phi(\mathcal{S})),\ \rho > 0\} \\ \operatorname{conv}(\Phi(\mathcal{S})) &= \{(\sigma, y) \mid (1, y) \in \sigma \operatorname{conv}(\mathcal{S}),\ \sigma > 0\}. \end{aligned}$$

Proof First, notice that

$$\Phi\big(\Phi(\mathcal{S})\big) = \Phi\left(\left\{\left(\frac{1}{\rho}, \frac{x}{\rho}\right) \;\middle|\; (\rho, x) \in \mathcal{S}\right\}\right) = \mathcal{S}.$$

Thus, it suffices to prove the first equation. Let R denote the right-hand-side of the first equation. First, we show that $R \subseteq \operatorname{conv}(\mathcal{S})$. Let $(\rho, x) \in R$. Then, $(\frac{1}{\rho}, \frac{x}{\rho}) \in \operatorname{conv}\big(\Phi(\mathcal{S})\big)$, and thus there exist convex multipliers λ and a set of points (ρ^j, x^j) of $\mathcal{S}$ such that

$$\left(\frac{1}{\rho}, \frac{x}{\rho}\right) = \sum_j \lambda_j \left(\frac{1}{\rho^j}, \frac{x^j}{\rho^j}\right).$$

Thus,

$$(\rho, x, 1) = \rho\left(1, \frac{x}{\rho}, \frac{1}{\rho}\right) = \rho \sum_j \lambda_j \left(1, \frac{x^j}{\rho^j}, \frac{1}{\rho^j}\right) = \sum_j \frac{\rho\lambda_j}{\rho^j}(\rho^j, x^j, 1).$$

This shows that $(\rho, x) \in \operatorname{conv}(\mathcal{S})$ since $(\rho^j, x^j) \in \mathcal{S}$, $\sum_j \frac{\rho\lambda_j}{\rho^j} = 1$ and $\frac{\rho\lambda_j}{\rho^j} \geq 0$. Therefore, $R \subseteq \operatorname{conv}(\mathcal{S})$. To show the reverse containment $\operatorname{conv}(\mathcal{S}) \subseteq R$, we consider a point $(\rho, x) \in \mathcal{S}$. It follows readily that

$$\left(\frac{1}{\rho}, \frac{x}{\rho}\right) = \Phi(\rho, x) \in \Phi(\mathcal{S}).$$

This shows that $(1, x) \in \rho\Phi(\mathcal{S})$. Therefore, $(\rho, x) \in R$, and thus $\mathcal{S} \subseteq R$. In addition, the relaxation R is a convex set since it is the projection onto the space of (ρ, x) variables of the following intersection

$$\left\{(\rho, \sigma', x) \;\middle|\; (\sigma', x) \in \rho \operatorname{conv}\big(\Phi(\mathcal{S})\big),\ \rho > 0\right\} \cap \big\{(\rho, \sigma', x) \;\big|\; \sigma' = 1\big\},$$

where the former set is the smallest convex cone that includes $\operatorname{conv}\big(\phi(\mathcal{S})\big)$, see Corollary 2.6.3 in [65]. Hence, $\operatorname{conv}(\mathcal{S}) \subseteq R$ as R is convex. □

Remark 1 In this remark, we show a polynomial-time equivalence of separating $\operatorname{conv}(\mathcal{S})$ and $\operatorname{conv}\big(\Phi(\mathcal{S})\big)$ from arbitrary points. This, together with the polynomial-time equivalence of separation and optimization [42], implies a polynomial time equivalence of optimization over the two convex hulls. Let $(\bar{\rho}, \bar{x}) \in \mathbb{R}_{++} \times \mathbb{R}^n$. We will use the separation oracle of $\operatorname{conv}\big(\Phi(\mathcal{S})\big)$ to separate $(\bar{\rho}, \bar{x})$ from $\operatorname{conv}(\mathcal{S})$. If $(\bar{\sigma}, \bar{y}) := \left(\frac{1}{\bar{\rho}}, \frac{\bar{x}}{\bar{\rho}}\right) \in \operatorname{conv}\big(\Phi(\mathcal{S})\big)$, then, by Theorem 1, we conclude $(\bar{\rho}, \bar{x}) \in \operatorname{conv}(\mathcal{S})$. Otherwise, the separation oracle returns a linear inequality $\big\langle \alpha, (\sigma, y)\big\rangle + \beta \leq 0$ to separate $(\bar{\sigma}, \bar{y})$ from $\operatorname{conv}(\Phi(\mathcal{S}))$. Then, by Theorem 1, the inequality $\big\langle \alpha, (1, x)\big\rangle + \beta \cdot \rho \leq 0$ separates $(\bar{\rho}, \bar{x})$ from $\operatorname{conv}(\mathcal{S})$. Conversely, using Theorem 1 and a separation oracle of $\operatorname{conv}(\mathcal{S})$, we can devise a separation algorithm for $\operatorname{conv}\big(\Phi(\mathcal{S})\big)$. □

Next, we present an application of Theorem 1 that, under certain conditions, relates the convex hull of fractional functions sharing a denominator to that of a set involving just the numerators. We show that convexifying these two sets, denoted $\mathcal{G}$ and $\mathcal{F}$ respectively, is polynomially equivalent, so that a formulation or separation procedure for one yields that for the other. Consider a vector of base functions $f : \mathcal{X} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ denoted as $f(x) = \big(f_1(x), \ldots, f_m(x)\big)$ and another vector of functions obtained from the base functions by dividing each of them with a linear form of f, *i.e.*, for $\alpha \in \mathbb{R}^m$, we consider

$$\mathcal{F} = \left\{ f(x) \;\middle|\; x \in \mathcal{X} \right\} \quad \text{and} \quad \mathcal{G} = \left\{ \frac{f(x)}{\sum_{i \in [m]} \alpha_i f_i(x)} \;\middle|\; x \in \mathcal{X} \right\}. \tag{6}$$

In Theorem 2, we will show that, under certain conditions, the convex hull of one set can be described using that of the other as follows:

$$\operatorname{conv}(\mathcal{G}) = \left\{ g \in \mathbb{R}^m \;\middle|\; \exists \rho \geq 0 \,\text{s.t.}\, g \in \rho \operatorname{conv}(\mathcal{F}),\ \alpha^\intercal g = 1 \right\} \tag{7a}$$

$$\operatorname{conv}(\mathcal{F}) = \left\{ f \in \mathbb{R}^m \;\middle|\; \exists \sigma \geq 0 \,\text{s.t.}\, f \in \sigma \operatorname{conv}(\mathcal{G}),\ f_1 = 1 \right\}. \tag{7b}$$

One of the conditions for (7a) to hold is the boundedness of $\mathcal{F}$. The next example shows that such condition is required.

Example 1 Consider a discrete set defined as

$$\mathcal{G} = \left\{ \frac{(1, x)}{1 + x} \;\middle|\; x \geq 0,\ x \in \mathbb{Z} \right\}.$$

Let $\mathcal{F} = \{(1, x) \mid x \geq 0,\ x \in \mathbb{Z}\}$ and notice that the convex hull of $\mathcal{F}$ is its continuous relaxation. For this setting, Theorem 2 is not applicable since $\mathcal{F}$ is not bounded. In fact, the equality in (7a) does not hold since the right hand side of (7a) reduces to $\{g \mid \exists \rho \geq 0 \,\text{s.t.}\, 1 = g_1 + g_2,\ g_1 = \rho, g_2 \geq 0\}$, and consists of $(0, 1)$, which does not belong to $\operatorname{conv}(\mathcal{G})$ because $\frac{1}{1+x} > 0$ for all $x \in \mathcal{G}$. □

Example 1 shows that the right hand side of (7a) can include points not in $\operatorname{conv}(\mathcal{G})$. Assume that $\mathcal{F}$ is a polyhedron. These extra points arise because, when $\mathcal{F}$ is homogenized, with the homogenizing variable set to 0, we obtain the recession cone of the polyhedron instead of a set containing just the origin. However, we can show that these additional points are typically included in $\operatorname{cl}\operatorname{conv}(\mathcal{G})$.

Theorem 2 *Let $\mathcal{F}$ and $\mathcal{G}$ be nonempty sets defined as in (6), and assume that there exists an $\epsilon > 0$ such that $\sum_{i \in [m]} \alpha_i f_i(x) > \epsilon$ for all $x \in \mathcal{X} \subseteq \mathbb{R}^n$. If $\mathcal{F}$ is bounded, (7a) holds and if, in addition, $f_1(x) = 1$ then (7b) holds. Moreover,*

$$\operatorname{cl}\operatorname{conv}(\mathcal{G}) = \left\{ g \in \mathbb{R}^m \;\middle|\; \exists \rho \geq 0 \,\text{s.t.}\, g \in \rho \operatorname{cl}\operatorname{conv}(\mathcal{F}),\ \alpha^\intercal g = 1 \right\}, \tag{8}$$

where $0 \operatorname{cl}\operatorname{conv}(\mathcal{F})$ denote the recession cone of $\operatorname{cl}\operatorname{conv}(\mathcal{F})$.

Proof Here, we focus on proving the correctness of (7). The correctness of (8) is shown in Appendix A.1. Consider a lift of $\mathcal{F}$ and $\mathcal{G}$, defined as follows, respectively,

$$\mathcal{S} := \left\{ (\alpha^\intercal f, f) \;\middle|\; f \in \mathcal{F} \right\} \quad \text{and} \quad \mathcal{T} := \Phi(\mathcal{S}) = \left\{ (\rho, g) := \frac{(1, f)}{\alpha^\intercal f} \;\middle|\; f \in \mathcal{F} \right\}. \tag{9}$$

We will invoke Theorem 1 to obtain a convex hull expression of $\mathcal{S}$ (resp. $\mathcal{T}$), which, after projection, yields the equality in (7a) (resp. (7b)).

First, we establish the equality in (7a). Since $\mathcal{G}$ is the projection of $\mathcal{T}$ onto the space of g variable and convexification commutes with projection, it suffices to derive the convex hull of $\mathcal{T}$, which is obtained as follows:

$$\begin{aligned} \operatorname{conv}(\mathcal{T}) &= \left\{ (\rho, g) \;\middle|\; (1, g) \in \rho \operatorname{conv}(\mathcal{S}),\ \rho > 0 \right\} \\ &= \left\{ (\rho, g) \;\middle|\; g \in \rho \operatorname{conv}(\mathcal{F}),\ \alpha^\intercal g = 1,\ \rho > 0 \right\} \\ &= \left\{ (\rho, g) \;\middle|\; g \in \rho \operatorname{conv}(\mathcal{F}),\ \alpha^\intercal g = 1,\ \rho \geq 0 \right\}, \end{aligned}$$

where the first equality holds due to Theorem 1 and $\mathcal{S} = \Phi(\mathcal{T})$. The second equality is derived using $\operatorname{conv}(\mathcal{S}) = \{(\alpha^\intercal f, f) \mid f \in \operatorname{conv}(\mathcal{F})\}$, where the equality hods since $\mathcal{S}$ is obtained as a linear transformation of $\mathcal{F}$. The last equality holds since if the rhs contains a point (ρ, g) with $\rho = 0$, i.e. $g \in 0 \operatorname{conv}(\mathcal{F})$ then the boundedness of $\mathcal{F}$ implies $g = 0$, contradicting to the equality constraint $\alpha^\intercal g = 1$.

Then, we establish the equality in (7b). Similarly, since $\mathcal{F}$ is the projection of $\mathcal{S}$, it suffices to derive the convex hull of $\mathcal{S}$,

$$\begin{aligned}\operatorname{conv}(\mathcal{S}) &= \{(\sigma, f) \mid (1, f) \in \sigma \operatorname{conv}(\mathcal{T}),\ \sigma > 0\} \\ &= \{(\sigma, f) \mid f \in \sigma \operatorname{conv}(\mathcal{G}),\ f_1 = 1,\ \sigma > 0\} \\ &= \{(\sigma, f) \mid f \in \sigma \operatorname{conv}(\mathcal{G}),\ f_1 = 1,\ \sigma \geq 0\},\end{aligned}$$

where the first equality holds due to Theorem 1 and $\mathcal{T} = \Phi(\mathcal{S})$. The second equality holds since, under the assumption that $f_1(x) = 1$, $\operatorname{conv}(\mathcal{T}) = \{(\rho, g) \mid \rho = g_1,\ g \in \operatorname{conv}(\mathcal{G})\}$. To see the last equality, we observe that $\operatorname{conv}(\mathcal{G})$ is bounded since $\sum_i \alpha_i f_i(x) > \epsilon$ and the boundedness of $\mathcal{F}$ implies that $\mathcal{G}$ is bounded. □

It is useful to interpret (7a) as arising from a two-step procedure where we first homgenize $\operatorname{conv}(\mathcal{F})$ and then intersect the resulting cone with $\alpha^\intercal g = 1$. To illustrate the relevance of Theorem 2 in convexifying fractional programs, we revisit the result of [56].

Example 2 We describe a polynomial time algorithm to solve (3) using a polynomial-time algorithm to optimize linear functions over $\mathcal{X}$. Using the equivalence of separation and optimization, (3) can be solved in polynomial time if $\operatorname{conv}(\mathcal{G})$ can be separated in polynomial time, where

$$\mathcal{G} = \left\{ \frac{(1, x)}{a_0 + a^\intercal x} \;\middle|\; x \in \mathcal{X} \right\}.$$

As we showed in Theorem 2, the convex hull of $\mathcal{G}$ can be derived from that of $\mathcal{X}$. We use (ρ, y) to denote g. The use of ρ is reasonable because the first coordinate is 1 in $\operatorname{conv}(\mathcal{F})$ which, by $g \in \rho \operatorname{conv}(\mathcal{F})$, implies that $g_1 = \rho$. Specifically, the separation problem of $\operatorname{conv}(\mathcal{G})$ is solved as follows. If a point $(\bar{\rho}, \bar{y}) \notin \operatorname{conv}(\mathcal{G})$ then $\bar{\rho} > 0$ and, therefore, $\frac{\bar{y}}{\bar{\rho}} \notin \operatorname{conv}(\mathcal{X})$. Assume $\alpha^\intercal x \leq b$ is a valid inequality for $\operatorname{conv}(\mathcal{X})$ that $\frac{\bar{y}}{\bar{\rho}}$ does not satisfy. Then, $\alpha^\intercal \bar{y} - b\bar{\rho} > 0$ while $\alpha^\intercal y - b\rho \leq 0$ for each $(\rho, y) \in \operatorname{conv}(\mathcal{G})$. □

As shown in [44], minimizing the sum of a linear fractional function and a linear function over $\{0,1\}^n$ is NP-hard. We relate this hardness result to Theorem 2 in the next example.

Example 3 Consider a linear fractional problem defined as

$$\min \left\{ \frac{b_0 + b^\intercal x}{a_0 + a^\intercal x} + c^\intercal x \;\middle|\; x \in \{0,1\}^n \right\}, \tag{10}$$

where $(a_0, a) \in \mathbb{Z} \times \mathbb{Z}^n$, $(b_0, b) \in \mathbb{Z} \times \mathbb{Z}^n$, and $c \in \mathbb{Z}^n$. We show that (10) is NP-hard under Turing reductions. By the equivalence of separation and optimization, (10) is polynomial time solvable for all (b_0, b) if and only if the separation of the convex hull of $\mathcal{G}$ is polynomial time solvable, where

$$\mathcal{G} = \left\{ \left(\frac{(1, x)}{1 + a^\intercal x}, x \right) \;\middle|\; x \in \{0,1\}^n \right\}.$$

By Remark 1 and Theorem 2, if $\mathrm{conv}(\mathcal{G})$ has a polynomial time separation algorithm, so does $\mathrm{conv}(\mathcal{F})$, where

$$\mathcal{F} = \left\{\big(1, x, x(1 + a^\intercal x)\big) \;\middle|\; x \in \{0,1\}^n\right\}.$$

By the ellipsoid algorithm, the following problem that minimizes a linear function with integer coefficients over $\mathcal{F}$ is polynomial time solvable:

$$\min\left\{\alpha + \beta^\intercal x + (\gamma^\intercal x)(1 + a^\intercal x) \;\middle|\; x \in \{0,1\}^n\right\}, \tag{11}$$

where $\alpha \in \mathbb{Z}$, and β and γ are vectors in $\mathbb{Z}^n$. However, (11) is NP-hard since the subset-sum problem can be polynomially reduced to it as follows. Consider $w \in \mathbb{Z}^n_+$ and $K \in \mathbb{Z}_+$. The subset sum is to decide if there exists an $x \in \{0,1\}^n$ such that $w^\intercal x = K$. Let $a = w$, $\alpha = 0$, $\beta = (-2K-1)\cdot w$ and $\gamma = w$. Now, observe that

$$\alpha + \beta^\intercal x + \gamma^\intercal x(1 + a^\intercal x) = (w^\intercal x)\cdot(w^\intercal x) - 2K \cdot w^\intercal x = (w^\intercal x - K)^2 - K^2,$$

which is $-K^2$ if and only if there exists an $x \in \{0,1\}^n$ satisfying $w^\intercal x = K$. $\square$

3 Connections between zero-one fractional and quadratic programming

Let $G = (V, E)$ be a graph, with V and E as the set of nodes and edges respectively, that has no self-loops. The boolean quadric polytope (BQP) associated with the graph G, denoted as QP_G, is defined as the convex hull of

$$\left\{(x, z) \in \mathbb{R}^{|V|+|E|} \;\middle|\; z_{ij} = x_i x_j \text{ for } (i,j) \in E,\ x \in \{0,1\}^{|V|}\right\}.$$

In particular, we denote by QP_G as QP if G is a complete graph. This polytope was introduced in [62], and its polyhedral structure and related algorithmic problems have been studied extensively [15,16,33,14,47]. BQP is related to the cut polytope [5] via a bijective linear transformation [27]. We show that valid inequalities and convex hull characterizations in special cases of BQP and/or cut-polytope lead to improved relaxations for 0-1 fractional programs.

We present several BQP inequalities in a homogenized form using an additional variable ρ to make them easier to use later [62]. The McCormick inequalities [55] for $i, j \in [n]$ with $i \neq j$ are as follows:

$$-z_{ij} \le 0,\ -z_{ij} + x_i + x_j \le \rho,\ z_{ij} - x_i \le 0,\ z_{ij} - x_j \le 0, \tag{McCormick}$$

the triangle inequalities for $i, j, k \in [n]$ are:

$$\begin{aligned} x_i + x_j + x_k - z_{ij} - z_{jk} - z_{ik} &\le \rho \\ -x_i + z_{ij} + z_{ik} - z_{jk} &\le 0 \\ -x_j + z_{ij} - z_{ik} + z_{jk} &\le 0 \\ -x_k - z_{ij} + z_{ik} + z_{jk} &\le 0, \end{aligned} \tag{Triangle}$$

and the odd-cycle inequalities are described next. Consider a cycle $C \subseteq E$ and a subset $D \subseteq C$ that has an odd number of elements. Let $S_0 = \{u \in V \mid \exists e, f \in D, e \neq f,\ \text{s.t. } e \cap f = u\}$ and $S_1 = \{u \in V \mid \exists e, f \in C \setminus D, e \neq f,\ \text{s.t. } e \cap f = u\}$, *i.e.*, S_0 (resp. S_1) is the set of nodes with both (resp. none) of the adjacent edges in D. Then, the odd-cycle inequality is

$$\sum_{i \in S_0} x_i - \sum_{i \in S_1} x_i + \sum_{(i,j) \in C \setminus D} z_{ij} - \sum_{(i,j) \in D} z_{ij} \leq \frac{|D|-1}{2}\rho. \qquad (\textsc{Odd-Cycle})$$

Odd-cycle inequalities can be separated in polynomial-time [27,5].

Besides the explicit inequalities in the space of BQP variables, we will also exploit the reformulation-linearization technique (RLT) to relax BQP [70] by introducing new variables that represent monomials of higher degrees. Specifically, given a positive integer $2 \leq k \leq |V|$, for each subset e of V of cardinality from 2 to k, we represent the monomial $\prod_{i \in e} x_i$ with a variable z_e. Then, the k^{th} RLT relaxation of BQP, denoted as RLT_k, is the system of linear inequalities obtained from the following system of polynomial inequalities by using relation $x_i = x_i^2$ and replacing each monomial with its linearization[1]:

$$\prod_{i \in S} x_i \prod_{j \in T} (1 - x_j) \geq 0 \qquad \text{for all } S \cap T = \emptyset \text{ and } |S \cup T| = k + 1.$$

Lemma 1 ([70]) *For a given $d \in \mathbb{Z}_+$ with $d \geq 2$, the convex hull of $\bigl\{(x, z) \mid x \in \{0,1\}^d,\ z_e = \prod_{i \in e} x_i\ \forall e \subseteq [d]$ s.t. $|e| \geq 2\bigr\}$ is given by RLT_{d-1}.*

3.1 Zero-one linear fractional programs and BQP

In this subsection, we focus on the problem of minimizing the sum of a linear fractional function and a linear function with binary variables

$$\min\left\{ \frac{b_0 + b^\intercal x}{a_0 + a^\intercal x} + c^\intercal x \;\middle|\; x \in \{0,1\}^n \right\}, \qquad (12)$$

where (a_0, a) and (b_0, b) are vectors in $\mathbb{R} \times \mathbb{R}^n$, c is a vector in $\mathbb{R}^n$, and the $a_0 + a^\intercal x$ is assumed to be positive on $\{0,1\}^n$. For a subset I of $[n]$, consider the linear fractional zero-one set defined as:

$$\mathcal{G}_I := \left\{ \left(\frac{(1, x)}{a_0 + a^\intercal x}, (x_i)_{i \in I} \right) \;\middle|\; x \in \{0,1\}^n \right\},$$

where $a_0 + a^\intercal x$ is assumed to be positive on $\{0,1\}^n$. We will use valid inequalities of BQP to study the convex hull of $\mathcal{G}_{[n]}$, and will refer to this convex hull as the *linear fractional polytope* (LFP). Clearly, the fractional program (12) is equivalent to maximizing a linear function over LFP. The linear fractional set $\mathcal{G}_I$ has also appears in reformulations of linear fractional zero-one programs [74,17,57], where the objective is to optimize the sum of ratios of affine functions subject to a set of linear constraints. This class of problems has many applications including assortment planning, facility location, and mean dispersion, see [17].

A tractable (polynomial-size) description of LFP is not possible, unless P=NP, since such a description would yield a polynomial time algorithm to solve (12),

[1] Observe that unlike the standard notation, our usage of RLT_k linearizes degree $k + 1$ monomials.

which is shown to be NP-hard in Example 3. Instead, we propose a hierarchy of polyhedral relaxations of LFP. For $k \in [n]$, the *k-term relaxation* of LFP is defined as follows:

$$\bigcap_{I \subseteq [n]:\ |I|=k} \operatorname{conv}(\mathcal{G}_I),$$

where $\operatorname{conv}(\mathcal{G}_I)$ is extended to the space of $\mathcal{G}_{[n]}$ by appending the remaining variables so that the intersection is well-defined. In particular, we will use valid inequalities of BQP to describe the k-term relaxation of LFP. To make this connection apparent, we begin with a remark on a slightly general setting.

Remark 2 Let $f(\cdot)$ be a vector of functions mapping from $\{0,1\}^n$ to $\mathbb{R}^m$, and consider two sets

$$\mathcal{G} := \left\{ \left(\frac{1}{a_0 + a^\intercal x}, \frac{x}{a_0 + a^\intercal x}, \frac{f(x)(a_0 + a^\intercal x)}{a_0 + a^\intercal x} \right) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \;\middle|\; x \in \{0,1\}^n \right\}$$
$$\mathcal{F} := \left\{ (1, x, f) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m \;\middle|\; f = f(x)(a_0 + a^\intercal x),\ x \in \{0,1\}^n \right\}.$$

Assume that $a_0 + a^\intercal x$ is positive on $\{0,1\}^n$. By setting $\alpha = (a_0, a^\intercal, 0) \in \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^m$ and invoking Theorem 2, the convex hull of $\mathcal{G}$ is given as follows:

$$\operatorname{conv}(\mathcal{G}) = \left\{ (\rho, y, g) \;\middle|\; (\rho, y, g) \in \rho \operatorname{conv}(\mathcal{F}),\ a_0\rho + a^\intercal y = 1,\ \rho \geq 0 \right\}. \square$$

Now, we use Remark 2 to relate the convex hull of $\mathcal{G}_I$ and BQP. For $I \subseteq [n]$, let $G(I)$ be the graph with $V = [n]$ and $E = \{(i,j) \mid i \in I,\ j \neq i\}$. Letting $f(x) = (x_i)_{i \in I}$, Remark 2 yields

$$\begin{aligned} \operatorname{conv}(\mathcal{G}_I) = \Big\{ (\rho, y, (x_i)_{i\in I}) \;\Big|\; & (y, w) \in \rho\, \mathrm{QP}_{G(I)},\ \rho \geq 0,\ a_0\rho + a^\intercal y = 1, \\ & x_i = \ell_i(y, w) \text{ for } i \in I \Big\}, \end{aligned} \tag{13}$$

where $\ell_i(y, w) := (a_0 + a_i)y_i + \sum_{j \neq i} a_j w_{ij}$. Observe that we relax $x_i \times (a_0 + a^\intercal x)$ instead of $x_i \times \frac{1}{a_0 + a^\intercal x}$ over $\{0,1\}^n$ using techniques in [64,76]. Since we can distribute the product over the summation and allows us to leverage BQP literature to derive inequalities, our approach performs significantly better (see Section 6).

Next, we use valid inequalities of BQP and (13) to derive, for various settings, explicit descriptions of the k-term relaxations. For $I \subseteq V$ with $|I| = 1$, the graph $G(I)$ is acyclic. Proposition 8 in [62] shows that, in this case, $\mathrm{QP}_{G(I)}$ is described by McCormick inequalities. This, together with (13), yields that the following explicit description of the 1-term relaxation of LFP:

$$\begin{aligned} \Big\{ (\rho, y, x) \;\Big|\; & (\rho, y, w) \text{ satisfies (McCormick)},\ a_0\rho + a^\intercal y = 1,\ \rho \geq 0, \\ & x_i = \ell_i(y, w) \text{ for } i \in [n] \Big\}. \end{aligned} \tag{1-Term}$$

We next characterize the k-term relaxation for $k \geq 2$ using a classical decomposition result (see [67] and Theorem 1 in [30]).

Lemma 2 *For $i \in [m]$, let D_i be a subset in the space of variable (x_i, y), and let $D := \{(x, y) \mid (x_i, y) \in D_i \ \text{ for } i \in [m]\}$. If $\operatorname{proj}_y(D_1) = \cdots = \operatorname{proj}_y(D_m) = V$ and V is a set of finite affinely independent points, $\operatorname{conv}(D) = \{(x, y) \mid (x_i, y) \in \operatorname{conv}(D_i) \text{ for } i \in [m]\}$.*

Corollary 1 *The 2-term relaxation is given by (1-Term) and (Triangle).*

The k-term relaxation requires a characterization of $\mathrm{QP}_{G(I)}$, for $I \subseteq V$ with $|I| = k$. Unlike the case for $k = 2$, when $k \geq 3$, Lemma 2 does not decompose $\mathrm{QP}_{G(I)}$ into lower dimensional sets. However, k^{th}-level RLT can be used to describe the k-term relaxation. This result relies on the observation that k^{th}-level RLT introduces variables so that the decomposition lemma can be invoked.

Proposition 1 *For $2 \leq k \leq n-1$, the k-term relaxation is given as follows*

$$\Big\{(\rho, y, x) \;\Big|\; (y, w) \in \rho\,\mathrm{RLT}_k,\ a_0\rho + a^\intercal y = 1,\ 0 \leq y \leq \rho,\\ x_i = (a_0 + a_i)y_i + \sum_{j \neq i} a_{ij} w_{ij} \text{ for } i \in [n]\Big\},$$

where, for each $S \subseteq [n]$ with $2 \leq |S| \leq k+1$, the variable w_S represents $\frac{\prod_{j \in S} x_j}{a_0 + a^\intercal x}$.

3.2 Bilinear fractional programs with box constraints and BQP

In this subsection, we consider the bilinear fractional programming problem with box constraints

什么是box constraints

$$\min\left\{\frac{x^\intercal B x + d^\intercal x + d_0}{x^\intercal A x + c^\intercal x + c_0} \;\middle|\; x \in [0,1]^n\right\}, \tag{14}$$

where A and B are two $n \times n$ upper triangular matrices with *zero-diagonal*, and $c, d \in \mathbb{R}^n$. We assume that the denominator $x^\intercal A x + c^\intercal x + c_0$ is positive over the box. The main result of this subsection, Proposition 2, implies that the continuous optimization problem (14) is equivalent to its discrete counterpart, that is, there exists an optimal solution x^* such that $x^* \in \{0,1\}^n$. This class of optimization problems appears in many application areas, *e.g.*, bond portfolio optimization [49], maximum mean dispersion problem [63], and feature selection [58].

We will show that (14) is closely related to BQP. Towards this end, we linearize the objective function by introducing variables. Let $G = (V, E)$ be a graph with $V = [n]$ and

$$E = \{(i,j) \mid a_{ij} \neq 0\} \cup \{(i,j) \mid b_{ij} \neq 0\},$$

which consists of the indices of the non-zero entries in A and B. With the graph G, we denote the denominator as

$$q_G(x) = \sum_{(i,j) \in E} a_{ij} x_i x_j + \sum_{i \in V} c_i x_i + c_0. \tag{15}$$

Now, we introduce a variable $\rho \in \mathbb{R}$, a variable y_i for each node $i \in V$, and w_{ij} for each edge $(i,j) \in E$, and require that $(\rho, y, w) \in \mathcal{G}_G$, where

$$\mathcal{G}_G := \left\{ \left(\frac{(1, x, z)}{q_G(x)}\right) \;\middle|\; x \in [0,1]^{|V|},\ z_{ij} = x_i x_j \text{ for } (i,j) \in E \right\}.$$

It follows readily that the bilinear fractional program (14) is equivalent to

$$\min\left\{ \sum_{(i,j)\in E} b_{ij} w_{ij} + \sum_{i\in V} d_i y_i + d_0 \rho \;\middle|\; (\rho, y, w) \in \mathcal{G}_G \right\}.$$

The set $\mathcal{G}_G$ above can be replaced with $\operatorname{conv}(\mathcal{G}_G)$ without affecting the optimal value since the objective function is linear. In other words, it suffices to study $\operatorname{conv}(\mathcal{G}_G)$.

Now, we show that the convexification of $\mathcal{G}_G$ and of the boolean quadric polytope QP_G are equivalent, and, thus polyhedrality of QP_G implies that of $\operatorname{conv}(\mathcal{G}_G)$.

Proposition 2 *Assume that $q_G(\cdot)$ is positive on $[0,1]^{|V|}$. Then,*

$$\operatorname{conv}(\mathcal{G}_G) = \left\{ (\rho, y, w) \;\middle|\; (y, w) \in \rho\, \mathrm{QP}_G,\ \rho \geq 0,\ \sum_{(i,j)\in E} a_{ij} w_{ij} + \sum_{i\in V} c_i y_i + c_0 \rho = 1 \right\}.$$

Moreover, $\mathrm{QP}_G = \big\{ (x, z) \bigm| (1, x, z) \in \sigma \operatorname{conv}(\mathcal{G}_G),\ \sigma \geq 0 \big\}$.

Next, we discuss some consequences of this equivalence result. It is shown in [36] that, for a complete graph G, there does not exist a polynomial-sized extended formulation for $\operatorname{conv}(\mathcal{G}_G)$. Neither does a polynomial-time separation algorithm exist unless P = NP. Nevertheless, we exploit BQP to construct relaxations of $\mathcal{G}_G$. Let the *odd-cycle* relaxation of $\mathcal{G}_G$ be as follows:

什么是odd-circle

$$\left\{ (\rho, y, w) \;\middle|\; (\rho, y, w) \text{ satisfies (}\textsc{Odd-Cycle}\text{) for all possible } C \text{ and } D,\ \rho \geq 0,\ \sum_{(i,j)\in E} a_{ij} w_{ij} + \sum_{i\in V} c_i y_i + c_0 \rho = 1 \right\}.$$

什么是series-parallel graph

Using standard results obtained for QP_G, this relaxation describes the convex hull of $\mathcal{G}_G$ if G is a series-parallel graph. A graph is *series-parallel* if it arises from a forest by repeatedly replacing edges by parallel edges or by edges in series.

Corollary 2 *If G is series-parallel graph then (14) is polynomial-time solvable using the odd-cycle relaxation of $\mathcal{G}_G$.*

Remark 3 Proposition 2 generalizes to multilinear fractional programming with box constraints. This problem is equivalent to minimizing a linear function over the *multilinear polytope*, a polytope that has been studied extensively, see [28, 29, 30, 31, 32] and references therein. We elaborate on how this research can be leveraged for multilinear fractional programming. Let $G = (V, E)$ be a hypergraph, where V is the set of nodes of G, and E is a set of subsets of V of cardinality at least two, named the hyperedges of G. With a hypergraph G, the multilinear polytope MP_G is defined as the convex hull of

$$\left\{ (x, z) \;\middle|\; x \in \{0,1\}^{|V|},\ z_e = \prod_{i\in e} x_i \text{ for } e \in E \right\}.$$

On the other hand, minimizing the ratio of two multilinear functions over the box constraints is equivalent to minimizing a linear function over a set $\mathcal{M}_G$, which is defined as follows. With a hypergraph G, we associate a multilinear function $q_G(x) = \sum_{e \in E} a_e \prod_{i \in e} x_i + c^\intercal x + c_0$ and a set

$$\mathcal{M}_G = \Bigg\{ (\rho, y, w) \ \Bigg|\ x \in [0,1]^{|V|},\ \rho = \frac{1}{q_G(x)}, \\ y_i = x_i \rho \text{ for } i \in V,\ w_e = \rho \prod_{i \in e} x_i \text{ for } e \in E \Bigg\}.$$

Then, by Theorem 2, it follows that

$$\operatorname{conv}(\mathcal{M}_G) = \Bigg\{ (\rho, y, w) \ \Bigg|\ (y, w) \in \rho \, \mathrm{MP}_G,\ \rho \geq 0,\ \sum_{e \in E} a_e w_e + \sum_{i \in V} c_i x_i + c_0 \rho = 1 \Bigg\}.$$

Here, although the original problem is stated over $[0,1]^n$, the restriction to $\{0,1\}^{|V|}$ in MP_G does not change the optimal value because simultaneous convex hull of multilinear functions over $[0,1]^n$ is the same as that over $\{0,1\}^n$, see Corollary 2.7 in [73]. □

4 Ratio of quadratics via copositive optimization

In this section, we consider a class of quadratic fractional optimization problems defined as follows

$$\min \Bigg\{ \frac{x^\intercal B x + b^\intercal x + b_0}{x^\intercal A x + a^\intercal x + a_0} \ \Bigg|\ x \in \mathcal{X} \cap \mathcal{L} \Bigg\}, \tag{16}$$

where $\mathcal{X}$ is a subset of $\mathbb{R}^n$, and $\mathcal{L}$ is an affine set of $\mathbb{R}^n$. Assume that $\mathcal{X} \cap \mathcal{L}$ is non-empty and bounded, and the denominator $q(x) := x^\intercal A x + a^\intercal x + a_0$ is positive over $\mathcal{X} \cap \mathcal{L}$. Let

$$\mathcal{G} = \Bigg\{ \frac{(1, x, xx^\intercal)}{q(x)} \ \Bigg|\ x \in \mathcal{X} \cap \mathcal{L} \Bigg\}.$$

Then, the quadratic fractional program (16) is equivalent to minimizing the linear function $b_0 \rho + \langle b, y \rangle + \langle B, Y \rangle$ over $\mathcal{G}$. Thus, optimzing the linear function over the convex hull of $\mathcal{G}$ would yield a relaxation with no gap. In the following, we will establish an equivalence between the convexification of $\mathcal{G}$ and $\mathcal{F}$, where

$$\mathcal{F} = \big\{ (x, X) \ \big|\ x \in \mathcal{X},\ X = xx^\intercal \big\}.$$

The set $\mathcal{F}$ has been studied extensively in various settings; see [72, 18, 1, 21, 45, 81, 46] for various recent results and [19] for a survey of earlier results.

To describe the convex hull characterizations of $\mathcal{F}$, we introduce the following notation that we will use throughout the rest of the section. A matrix $M \in \mathbb{S}^d$ is called completely positive if there exists k nonnegative vectors $h_1, h_2, \ldots, h_k$ in $\mathbb{R}^d_+$ such that $M = \sum_{i \in [k]} h_i h_i^\intercal$. We will use $\mathcal{CP}_d$ to denote the set of $d \times d$ completely positive matrices. Given $(\rho, x, X) \in \mathbb{R} \times \mathbb{R}^d \times \mathbb{S}^d$, let

$$Z(\rho, x, X) := \begin{pmatrix} \rho & x^\intercal \\ x & X \end{pmatrix} \in \mathbb{S}^{d+1}.$$

Now, we prove the equivalence of convexification between $\mathcal{G}$ and $\mathcal{F}$. Besides the presence of fractions in $\mathcal{G}$, the sets $\mathcal{G}$ and $\mathcal{F}$ also differ since the domain in $\mathcal{G}$ is intersected with an affine set. Thus, the equivalence result requires a facial decomposition besides Theorem 2. The techniques are inspired by the proof of Theorem 2.6 in [18].

Proposition 3 *Assume that $q(x) > 0$ over a non-empty bounded set $\mathcal{X} \cap \mathcal{L}$. Suppose that $\mathcal{L} = \{x \mid Cx = d\}$, where $C \in \mathbb{R}^{m\times n}$ and d is a vector in $\mathbb{R}^m$. Then,*

$$\begin{aligned}\operatorname{conv}(\mathcal{G}) = \Bigl\{(\rho, y, Y) \Bigm| (y, Y) \in \rho \operatorname{conv}(\mathcal{F}),\ \langle A, Y\rangle + \langle a, y\rangle + a_0\rho = 1,\ \rho \geq 0,\\ \operatorname{Tr}\bigl(CYC^\intercal - Cyd^\intercal - dy^\intercal C^\intercal + \rho dd^\intercal\bigr) = 0\Bigr\}.\end{aligned}$$

If the affine set $\mathcal{L}$ is $\mathbb{R}^n$ then $\operatorname{conv}(\mathcal{F}) = \bigl\{(x, X) \bigm| (1, x, X) \in \sigma \operatorname{conv}(\mathcal{G}),\ \sigma \geq 0\bigr\}$.

Next, we specialize the results in Proposition 3 to three cases. First, we consider the case when $\mathcal{X}$ is the non-negative orthant. It follows readily that the convex hull of $\mathcal{F}$ can be described by the cone of completely positive matrices, that is

$$\operatorname{conv}(\mathcal{F}) = \bigl\{(x, X) \bigm| Z(1, x, X) \in \mathcal{CP}_{n+1}\bigr\}.$$

Thus, by Proposition 3, we obtain a copositive programming reformulation for minimizing a quadratic fractional function over a polytope. This result extends the copositive programming formulation for minimizing a quadratic function over a polytope shown in [18,19].

Corollary 3 *If $\mathcal{X}$ is the non-negative orthant then (16) can be formulated as*

$$\begin{aligned}\min \quad & b_0\rho + b^\intercal y + \langle B, Y\rangle \\ \text{s.t.} \quad & Z(\rho, y, Y) \in \mathcal{CP}_{n+1} \\ & \langle A, Y\rangle + \langle a, y\rangle + a_0\rho = 1 \\ & \operatorname{Tr}\bigl(CYC^\intercal - Cyd^\intercal - dy^\intercal C^\intercal + \rho dd^\intercal\bigr) = 0.\end{aligned}$$

Remark 4 Binary variables can be handled in this setup easily. We will rewrite $0 \leq x_i \leq 1$ as $x_i + s_i = 1$ with $x_i, s_i \geq 0$. Assume that in the matrix X the row (column) index of variable x_i is p and that of s_i is q. All entries in a completely positive matrix and, therefore, X_{pq} is non-negative. In other words, if we require $X_{pq} = 0$, we are restricted to a face of the set of completely positive matrices. Since each extreme ray of the completely positive matrices is of the form $hh^\intercal$ for some $h \geq 0$, it follows that extreme rays of the cone that belong to this face satisfy that either h_p or h_q is zero. Moreover, since all extreme rays satisfy $h_p + h_q = 1$, because $x_i + s_i = 1$ is imposed as $\mathcal{L}$ is enforced in Proposition 3, it follows that each extreme ray, h, satisfies the binary conditions. □

Second, we consider the case when $\mathcal{X}$ is a ball of radius 1 centered at origin, that is $\{x \mid \|x\|_2 \leq 1\}$ where $\|\cdot\|_2$ is the Euclidean norm. In this case, it follows from [72,19] that the convex hull of $\mathcal{F}$ can be described using the Shor's semidefinite programming (SDP) relaxation and reformulation-linearization constraints,

$$\operatorname{conv}(\mathcal{F}) = \bigl\{(x, X) \bigm| Z = Z(1, x, X),\ Z \succeq 0,\ \langle L, Z\rangle \geq 0\bigr\},$$

where $L := \text{Diag}(1, -1, \ldots, -1)$. This, together with Proposition 3, yields a polynomial time solvable SDP formulation for (16). When we specialize this formulation to the case where the affine space $\mathcal{L}$ is $\mathbb{R}^n$, we obtain a polynomial time solvable SDP formulation for optimizing the ratio of two quadratics over an ellipsoid, thus recovering the result of [7]. On the other hand, this result can be generalized by using results that convexify $\mathcal{F}$ when $\mathcal{X}$ is jointly defined by a ball and additional non-intersecting constraints [21,45,81,46].

Last, we use Proposition 3 to derive tractable envelopes of bivariate quadratic fractional functions. More specifically, we consider the graph of the ratio of bivariate quadratic and linear functions over a convex quadrilateral

$$\mathcal{S} = \left\{ (t, x) \in \mathbb{R} \times \mathbb{R}^2 \;\middle|\; t = \frac{x^\intercal B x + b^\intercal x + b_0}{a^\intercal x + a_0},\ x \in \mathcal{X} \right\},$$

where $\mathcal{X}$ is a convex quadrilateral in $\mathbb{R}^2$, and $a^\intercal x + a_0$ is assumed to be positive over the domain $\mathcal{X}$. For the special case where the domain is a box in the positive orthant, the numerator is a non-negative linear function, and the denominator is a positive linear function, [76] relies on the linearity of the numerator to limit the domain of the function, and then derives the convex and concave envelope of the fraction. Special cases of these relaxations are used in mixed-integer nonlinear programming solvers [78,60,10] to relax factorable functions. We will now present an SDP formulation for the convex hull of $\mathcal{S}$. This generalizes the class of fractional functions to allow for a nonlinear numerator, a setting in which the convex extension argument does not apply.

Corollary 4 *Assume that $\mathcal{X} = \{x \mid Cx \leq d\}$ is a convex quadrilateral in the plane, and $a^\intercal x + a_0$ is positive on $\mathcal{X}$. Then, an SDP formulation for* $\text{conv}(\mathcal{S})$ *is given as follows*

$$\begin{aligned}
&t = b_0\rho + b^\intercal y + \langle B, Y \rangle \\
&x_i = a_1 Y_{i1} + a_2 Y_{i2} + a_0 y_i,\ i = 1, 2 \\
&a_0\rho + a^\intercal y = 1 \\
&Z = Z(\rho, y, Y),\ Z \succeq 0 \\
&CYC^\intercal - Cyd^\intercal - dy^\intercal C^\intercal + \rho dd^\intercal \geq 0.
\end{aligned}$$

5 Multiple denominators

In this section, we treat the case where there are multiple denominator expressions. In Section 5.1, we consider the univariate case in two settings. First, we treat the case where the variable is raised to various powers in the denominator, and second, we consider the case where expressions in the denominator involve shifts of this variable by different constants. In each case, we derive convex hull formulation using the moment-hull characterization. In Section 5.2 we consider the multivariate case for $0-1$ variables, where multiple linear expressions form denominators. In this case, we develop a hierarchy of relaxations converging to the convex hull.

5.1 Univariate case: convex hull formulation

In this section, we are interested in developing convex hulls for:

$$\mathcal{G}_{p,q} = \left\{ \left(\frac{1}{x^p}, \frac{1}{x^{p-1}}, \cdots, \frac{1}{x}, 1, x, \cdots, x^q \right) \;\middle|\; x \in \mathcal{X} \right\},$$

and

$$\mathcal{G}_r = \left\{ \left(1, \frac{1}{x - r_1}, \ldots, \frac{1}{x - r_n}, x - r_0 \;\middle|\; x \in \mathcal{X} \right\},$$

where $p, q \in \mathbb{Z}_+$, $r \in \mathbb{R}^n$ is a vector such that $r_1 < \ldots < r_n$, $\mathcal{X}$ is a closed set such that none of the denominators attain zero. The first result generalizes the moment hull to allow fractional terms. This is useful to simultaneously relax fractions and powers instead of the current practice of relaxing them separately. The set $\mathcal{G}_r$ occurs in chemical process design problems where Underwood equations are modeled using functions in $\mathcal{G}_r$ and r_i are relative volatilities of components in a chemical mixture [79].

We will relate the convex hulls of $\mathcal{G}_r$ and $\mathcal{G}_{p,q}$ to the convex hull of the moment curve which is described as follows. Let $d \in \mathbb{Z}_+$ and $\mathcal{X}$ be a subset of $\mathbb{R}$. Then, the moment curve in $\mathbb{R}^{d+1}$ with support $\mathcal{X}$ is defined as follows

$$\mathcal{M}_d = \left\{ (1, x, x^2, \ldots, x^d) \;\middle|\; x \in \mathcal{X} \right\}.$$

When $\mathcal{X}$ is a discrete set (resp. an interval) the convex hull of $\mathcal{M}_d$ is referred to as a *cyclic polytope* (resp. *moment hull*). The Gale evenness condition in [38] provides a characterization of facets of a cyclic polytope, [13] provides an extended formulation for a cyclic polytope, and [35] provides a polynomial-sized SDP formulation for a certain class of cyclic polytopes. The moment hull is SDP representable as follows, see Section 3.5.4 in [12]. Suppose that $\mathcal{X} = [a, b]$, and let $H : \mathbb{R}^{2d+1} \to \mathbb{S}^{d+1}$ be a linear transformation that maps $\mu = (\mu_0, \mu_1, \ldots, \mu_{2d})$ to a Hankel matrix $H(\mu)$, that is

$$H(\mu) = \begin{pmatrix} \mu_0 & \mu_1 & \mu_2 & \cdots & \mu_d \\ \mu_1 & \mu_2 & \mu_3 & \cdots & \mu_{d+1} \\ \mu_2 & \mu_3 & \mu_4 & \cdots & \mu_{d+1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \mu_d & \mu_{d+1} & \mu_{d+2} & \cdots & \mu_{2d} \end{pmatrix}.$$

If d is odd, $(\mu_0, \mu_1, \ldots, \mu_d) \in \operatorname{conv}(\mathcal{M}_d)$ if and only if

$$\begin{aligned} H(\mu_1, \ldots, \mu_d) - aH(\mu_0, \ldots, \mu_{d-1}) &\succeq 0 \\ bH(\mu_0, \ldots, \mu_{d-1}) - H(\mu_1, \ldots, \mu_d) &\succeq 0 \\ \mu_0 &= 1. \end{aligned}$$

If d is even, $(\mu_0, \mu_1, \ldots, \mu_d) \in \operatorname{conv}(\mathcal{M}_d)$ if and only if

$$\begin{aligned} H(\mu_0, \ldots, \mu_d) &\succeq 0 \\ -H(\mu_2, \ldots, \mu_d) + (a + b)H(\mu_1, \ldots, \mu_{d-1}) - abH(\mu_0, \ldots, \mu_{d-2}) &\succeq 0 \\ \mu_0 &= 1. \end{aligned}$$

In our study, we will use the homogenization of $\mathrm{conv}(\mathcal{M}_d)$, that is, the conic hull of $\mathcal{M}_d$, denoted as $\mathrm{cone}(\mathcal{M}_d)$. An explicit description of $\mathrm{cone}(\mathcal{M}_d)$ is obtained by dropping the constraint $\mu_0 = 1$ in a polyhedral/SDP representation of $\mathrm{conv}(\mathcal{M}_d)$.

We start by establishing that convexifying the moment curve and the curve $\mathcal{G}_{p,q}$ are equivalent. The proof of this result follows directly from Theorem 2.

Corollary 5 *Let $p, q \in \mathbb{Z}_+$, and assume that $x^p > 0$ for every $x \in \mathcal{X}$. Then,*

$$\begin{aligned}\mathrm{conv}(\mathcal{G}_{p,q}) &= \bigl\{(\nu_0, \nu_1, \dots, \nu_{p+q}) \bigm| \nu \in \mathrm{cone}(\mathcal{M}_{p+q}),\ \nu_p = 1\bigr\}\\ \mathrm{conv}(\mathcal{M}_{p+q}) &= \bigl\{(\mu_0, \nu_1, \dots, \mu_{p+q}) \bigm| \mu \in \mathrm{cone}(\mathcal{G}_{p,q}),\ \mu_0 = 1\bigr\}.\end{aligned}$$

Next, we consider $\mathcal{G}_r$ defined in the beginning of this section, which we refer to as $\mathcal{G}$ by dropping the subscript for the remainder of this section. In the following, we show that the convex hull of $\mathcal{G}$ is isomorphic to the convex hull of $\mathcal{M}_{n+1}$. Our proof consists of two steps. Using partial fraction decomposition and Theorem 1, we will establish an isomorphism between the convex hull of $\mathcal{G}$ and that of the following curve given by polynomials,

$$\mathcal{F} = \Bigl\{\bigl(f_0(x), f_1(x), \dots, f_{n+1}(x)\bigr) \Bigm| x \in \mathcal{X}\Bigr\},$$

where $f_0(x) = \prod_{j=1}^n (x - r_j)$, for $i \in [n]$, $f_i(x) = \prod_{j \neq i,0}(x - r_j)$, and $f_{n+1}(x) = \prod_{j=0}^n (x - r_j)$. To use this result, we first establish an isomorphism between $\mathcal{F}$ and the moment curve $\mathcal{M}_{n+1}$ in the following lemma.

Lemma 3 *Let r_{n+1} be a real number not in $\{r_0, \dots, r_n\}$, and consider an $(n+2) \times (n+2)$ matrix T with entries $(\beta_{ij})_{i,j=0}^{n+1}$, where for each row $i \in \{0\} \cup [n+1]$*

$$\beta_{ij} = \begin{cases} \dfrac{(r_j)^i}{f_j(r_j)} & j \in [n] \\[2ex] \dfrac{(r_0)^i - \sum_{k=1}^n \beta_{ik} f_k(r_0)}{f_0(r_0)} & j = 0 \\[2ex] \dfrac{(r_{n+1})^i - \sum_{k=0}^n \beta_{ik} f_k(r_{n+1})}{f_{n+1}(r_{n+1})} & j = n+1. \end{cases}$$

Then, $\mathcal{M}_{n+1} = T\mathcal{F}$, and $\mathcal{F} = T^{-1}\mathcal{M}_{n+1}$.

We first assume that $\prod_{i=1}^n (x - r_i) > 0$ for all $x \in \mathcal{X}$. The case with $\prod_{i=1}^n (x - r_i) < 0$ can be handled similarly by considering $-f_i(x)$ for some i instead. The applications described above satisfy that $\prod_{i=1}^n (x - r_i)$ is sign-invariant over $\mathcal{X}$.

Proposition 4 *Assume $r_0, r_1, \dots, r_n$ are distinct reals such that $\prod_{i=1}^n (x - r_i) > 0$. Then,*

$$\begin{aligned}\mathrm{conv}(\mathcal{G}) &= \bigl\{\nu = (\nu_0, \nu_1, \dots, \nu_{n+1}) \bigm| T\nu \in \mathrm{cone}(\mathcal{M}_{n+1}),\ \nu_0 = 1\bigr\}\\ \mathrm{conv}(\mathcal{M}_{n+1}) &= \bigl\{\mu = (\mu_0, \mu_1, \dots, \mu_{n+1}) \bigm| T^{-1}\mu \in \mathrm{cone}(\mathcal{G}),\ \mu_0 = 1\bigr\}.\end{aligned}$$

The assumption that $\prod_{i=1}^n (x - r_i)$ does not change sign can be relaxed as follows. We can split $\mathcal{X}$ into two subsets, one where the number of r_i values larger than x is even and another where this number is odd. Then, since $\prod_{i=1}^n (x - r_i)$ is sign-invariant over these subsets, the convex hull of their union can be obtained using disjunctive programming [4].

5.2 Multivariate case: a hierarchy of relaxations

Consider the following problem:

$$\max\left\{ \sum_{i\in[m]} \frac{b_{i0} + b_i^\intercal x}{a_{i0} + a_i^\intercal x} \;\middle|\; x \in \mathcal{X} \subseteq \{0,1\}^n \right\}, \tag{17}$$

where $m, n \in \mathbb{Z}_+$, $a_i, b_i \in \mathbb{R}^n$, $a_{i0}, b_{i0} \in \mathbb{R}$, and $a_{i0} + a_i^\intercal x > 0$ for every $x \in \mathcal{X}$. Assume $\mathcal{X}$ can be described using a system of linear inequalities, that is, $\mathcal{X} = \{x \mid Cx \geq d\} \cap \{0,1\}^n$, where C is an $r \times n$ matrix and d is an r-dimensional vector. Clearly, this maximization problem is equivalent to maximizing $\sum_{i\in[m]} b_{i0}\rho^i + b_i^\intercal y^i$ over the following set

$$\mathcal{G} := \left\{ (\rho, y) \;\middle|\; x \in \mathcal{X},\; (\rho^i, y^i) = \frac{(1, x)}{a_{i0} + a_i^\intercal x} \text{ for } i \in [m] \right\},$$

where $\rho = (\rho^1, \ldots, \rho^m)$ and $y = (y^1, \ldots, y^m)$. In this section, we present a hierarchy of convex outer-approximations for $\mathcal{G}$. For $k \in [n]$, the k^{th} level relaxation of $\mathcal{G}$ is related to the k^{th} level RLT of the feasible region $\mathcal{X}$, denoted as $\mathrm{RLT}_k(\mathcal{X})$. The relaxation $\mathrm{RLT}_k(\mathcal{X})$ is defined as the system of linear inequalities obtained by linearizing the following system of polynomial inequalities using the relation $x_i = x_i^2$ and representing the monomial $\prod_{i\in e} x_i$ with an introduced variable z_e after distributing products:

$$(Cx - d) \prod_{i\in S} x_i \prod_{j\in T} (1 - x_j) \geq 0 \text{ for } S \cap T = \emptyset \text{ with } |S \cup T| = k.$$

We begin with a base relaxation using ideas in Section 3.1. Consider the following lifting $\mathcal{G}^1$ of $\mathcal{G}$ in the space of (ρ, y, x) variables, that is

$$\mathcal{G}^1 = \left\{ (\rho, y, x) \;\middle|\; x \in \mathcal{X},\; (\rho^i, y^i) = \frac{(1, x)}{a_{i0} + a_i^\intercal x} \text{ for } i \in [m] \right\}.$$

More generally, $\mathcal{G}^k$ will denote the set used to construct the k^{th} level relaxation in our proposed hierarchy. Then, we project $\mathcal{G}^1$ into m subspaces, where for $i \in [m]$, $\mathcal{G}_i^1$ denotes the projection of $\mathcal{G}^1$ onto the space of (ρ^i, y^i, x) variables. To relax each projection, we introduce variable w_{jr}^i to linearize $\frac{x_j x_r}{a_{i0} + a_i^\intercal x}$, and construct the following relaxation:

$$\begin{aligned} \mathcal{R}_i^1 = \Bigg\{ (\rho^i, y^i, x) \;\Bigg|\; & (y^i, w^i) \in \rho^i \, \mathrm{RLT}_1(\mathcal{X}),\; a_{i0}\rho^i + a_i^\intercal y^i = 1,\; \rho^i \geq 0, \\ & x_j = (a_{i0} + a_{ij}) y_j^i + \sum_{r \neq j} a_{ir} w_{jr}^i \text{ for } j \in [n] \Bigg\}, \end{aligned}$$

where the four linear constraints are obtained by linearizing the following constraints, respectively,

$$
\begin{aligned}
\frac{(x, xx^\intercal)}{a_{i0} + a_i^\intercal x} &\in \frac{1}{a_{i0} + a_i^\intercal x}\,\mathrm{RLT}_1(\mathcal{X})\\
\frac{a_{i0}}{a_{i0} + a_i^\intercal x} + \sum_{r\in[m]} \frac{a_{ir}x_r}{a_{i0} + a_i^\intercal x} &= 1\\
\frac{1}{a_{i0} + a_i^\intercal x} &\geq 0\\
x_j = \frac{(a_{i0} + a_{ij})x_j + \sum_{r\neq j} a_{ir}x_j x_r}{a_{i0} + a_i^\intercal x}&.
\end{aligned}
$$

Now, the base relaxation $\mathcal{R}^1$ is the intersection of relaxations $\mathcal{R}_i^1$, that is

$$\mathcal{R}^1 = \bigl\{(\rho, y, x) \bigm| (\rho^i, y^i, x) \in \mathcal{R}_i^1 \text{ for } i \in [m]\bigr\}.$$

We motivate our relaxation hierarchy by first considering a simple way to tighten $\mathcal{R}^1$, which does not yield the convex hull of $\mathcal{G}$. This relaxation replaces $\mathrm{RLT}_1(\mathcal{X})$ in the definition of $\mathcal{R}_i^1$ with the n^{th}-level RLT relaxation of $\mathcal{X}$.

Example 4 Consider the case where $m = 2$ and $n = 2$. Let $a_{10} = 1$ and $a_1 = (2, 4)$, and $a_{20} = 1$ and $a_2 = (3, 5)$. In this case, the set $\mathcal{G}$ of our interest consists of the following four points listed in the space of variables $(\rho^1, y_1^1, y_2^1, \rho^2, y_1^2, y_2^2)$:

$$\Bigl(1, 0, 0, 1, 0, 0\Bigr) \quad \Bigl(\frac{1}{3}, \frac{1}{3}, 0, \frac{1}{4}, \frac{1}{4}, 0\Bigr) \quad \Bigl(\frac{1}{5}, 0, \frac{1}{5}, \frac{1}{6}, 0, \frac{1}{6}\Bigr) \text{ and } \Bigl(\frac{1}{7}, \frac{1}{7}, \frac{1}{7}, \frac{1}{9}, \frac{1}{9}, \frac{1}{9}\Bigr).$$

It can be verified that the linear inequality $25\rho^1 - 4y_1^1 - 24\rho^2 \leq 1$ is valid for $\mathcal{G}$. However, this inequality is not valid for the relaxation obtained by individually convexifying each linear fractional function. To see this, let $\mathcal{G}_i^1 = \bigl\{(\rho^i, y^i, x) \bigm| (\rho^i, y^i) = \frac{(1,x)}{1+a_i^\intercal x},\ x \in \{0,1\}^2\bigr\}$. Now, maximizing $25\rho^1 - 4y_1^1 - 24\rho^2$ over $\{(\rho, y, x) \mid (\rho^i, y^i, x) \in \mathrm{conv}(\mathcal{G}_i^1) \text{ for } i = 1, 2\}$ yields 9, showing that intersecting $\mathrm{conv}(\mathcal{G}_i^1)$ does not yield $\mathrm{conv}(\mathcal{G})$. □

This discrepancy occurs because we have not exploited that $(y^i, w^i) \in \rho^i\,\mathrm{RLT}_n(\mathcal{X})$ and $(y^j, w^j) \in \rho^j\,\mathrm{RLT}_n(\mathcal{X})$ are scalar multiples of one another. To exploit this, first observe that the convex hull of U is a simplex, where

$$U = \Bigl\{u \Bigm| x \in \mathcal{X},\ u_S = \prod_{j\in S} x_j \text{ for } \emptyset \neq S \subseteq [n]\Bigr\}.$$

Therefore, for any point in the convex hull, there is a unique set of barycentric coordinates (convex multipliers associated with the extreme points of $\mathrm{conv}(U)$.) Since (y^i, w^i) scales to (y^j, w^j), for all $i \neq j$, we write:

$$u_S = \frac{\bigl(a_{i0} + \sum_{j\in S} a_{ij}\bigr) \prod_{t\in S} x_t + \sum_{r\notin S} a_{ir} \prod_{t\in S\cup\{r\}} x_t}{a_{i0} + a_i^\intercal x},$$

and introduce variable w_S^i to represent $\frac{\prod_{j\in S} x_j}{a_{i0}+a_i^\intercal x}$. For each $i \in [m]$ and $S \subseteq [n]$, the above equality describes a linear relation between u_S and w_S^i. Moreover, every function of 0−1 variables, and, therefore, w_S^i, can be expressed as a linear combination of $(u_S)_{S\subseteq[n]}$.

Formally, for each $2 \leq k \leq n$, we consider a lift of $\mathcal{G}$ given as follows,

$$\mathcal{G}^k = \left\{ (\rho, y, u) \;\middle|\; (\rho^i, y^i) = \frac{(1, x)}{a_{i0} + a_i^\intercal x} \text{ for } i \in [m], x \in \mathcal{X}, \right. \\ \left. u_S = \prod_{j \in S} x_j \text{ for } S \subseteq [n] \text{ with } 1 \leq |S| \leq k \right\}.$$

For each $i \in [m]$, let $\mathcal{G}_i^k$ denote the projection of $\mathcal{G}^k$ onto the space of (ρ^i, y^i, u) variables. A relaxation of $\mathcal{G}_i^k$ is given as follows

$$\mathcal{R}_i^k = \left\{ (\rho^i, y^i, u) \;\middle|\; w^i \in \rho^i \operatorname{RLT}_k(\mathcal{X}),\; y_j^i = w_{\{j\}}^i \text{ for } j \in [n],\; a_{i0}\rho^i + a_i^\intercal y^i = 1, \right. \\ \left. u_S = a_{i0} + \sum_{j \in S} \left(a_{ij} w^i_{S \cup \{j\}} \right) \text{ for } S \subseteq [n] \text{ s.t. } 1 \leq |S| \leq k \right\},$$

where $w_T^i = y_T^i$ if $T = \{j\}$ for some $j \in [n]$. Let $(\mathcal{R}^k)_{k=1}^n$ be a hierarchy of relaxations for $\mathcal{G}$, where

$$\mathcal{R}^k = \left\{ (\rho, y, u) \;\middle|\; (\rho^i, y^i, u) \in \mathcal{R}_i^k \text{ for } i \in [m] \right\} \quad \text{for } k \in [n].$$

Clearly, $\mathcal{R}^1 \supseteq \cdots \supseteq \mathcal{R}^n$. In fact, $\operatorname{proj}_{(\rho,y)}(\mathcal{R}^n) = \operatorname{conv}(\mathcal{G})$.

Theorem 3 *The following holds:* $\operatorname{proj}_{(\rho,y)}(\mathcal{R}^1) \supseteq \cdots \supseteq \operatorname{proj}_{(\rho,y)}(\mathcal{R}^n) = \operatorname{conv}(\mathcal{G})$.

6 Theoretical comparisons and computational experiments

In this section, we consider the constrained mixed-binary fractional programs, propose a relaxation for them using quadratic programming, and show that the resulting relaxation is tighter than relaxations used in the fractional programming literature. Then, we perform a computational study of relaxations for binary fractional programs occurring in assortment optimization using our reformulation based on (1-Term). Then, we consider a set arising from chemical process design and demonstrate that our relaxation based on the construction in Section 5.1 closes significant gap relative to standard relaxation techniques. Our computational experiments are performed on a MacBook Pro with Apple M1 Pro with 10-cores CPU and 16 GB of memory. The code is written in Julia v1.6 [11], and our formulations are modeled using JuMP v1.4.0 [52] and solved using Gurobi v10.0.2 [43] as an MIP solver and SCIP v8.0 [10] as an MINLP solver.

6.1 A polynomial-sized relaxation for fractional programs and theoretical comparisons

In this section, we leverage Theorem 2 and simple outer-approximations for quadratic programs to propose a polynomial-sized relaxation (see (20) and (22)/(23) below) for constrained fractional programs, and theoretically compare its tightness relative to existing relaxations. Our results show that this approach yields relaxations that are tighter than those in the literature. Specifically, we will focus our attention on:

$$\max_{x\in\mathcal{X}} \quad \sum_{i=1}^{m} \frac{b_{i0}+b_i^\intercal x}{a_{i0}+a_i^\intercal x} + c^\intercal x, \tag{FP}$$

where we assume that $\mathcal{X} = \bigl\{x \in \{0,1\}^p \times \mathbb{R}^{n-p} \bigm| Cx \le d\bigr\}$ for $C \in \mathbb{R}^{r\times n}$, $d \in \mathbb{R}^r$, $\mathcal{X}$ is bounded, and $a_{i0} + a_i^\intercal x > 0$ for all x that satisfy $Cx \le d$ and $x_j \in [0,1]$ for $j \in [p]$. The special case with $p = n$ will be referred to as *binary fractional programming* (BFP).

An approach used to reformulate (FP) is to introduce $\rho^i = \frac{1}{a_{i0}+a_i^\intercal x}$ and $y_j^i = \frac{x_j}{a_{i0}+a_i^\intercal x}$ for all $i \in [m]$ and $j \in [n]$. Then, (FP) is reformulated as

$$\begin{aligned}
\max \quad & \sum_{i=1}^{m} \bigl(b_{i0}\rho^i + b_i^\intercal y^i\bigr) + c^\intercal x & & \\
\text{s.t.} \quad & x_i \in \{0,1\} & i \in [p] & \qquad (18a)\\
& y_j^i = \rho^i x_j & i \in [m],\ j \in [n] & \qquad (18b)\\
& a_{i0}\rho^i + a_i^\intercal y^i = 1 & i \in [m] & \qquad (18c)\\
& \overline{C}x \le \bar{d}, & & \qquad (18d)
\end{aligned}$$

where $\overline{C}x \le \bar{d}$ consists of the inequalities $Cx \le d$, and $x_j \in [0,1]$ for $j \in [p]$. This reformulation leads to one of the most popular polyhedral relaxations for (FP) [51,57]. The relaxation replaces (18b) with McCormick inequalities. In particular, letting $\rho^i(L)$ and $\rho^i(U)$ (resp. $x_j(L)$ and $x_j(U)$) be lower and upper bounds on variable ρ^i (resp. x_j), a polyhedral relaxation for the feasible region of (18) is given as follows:

$$\mathcal{R}_{\mathsf{LEF}} := \left\{ (\rho, y, x) \;\middle|\; \begin{array}{ll} y_j^i \le x_j(U)\rho^i + \rho^i(L)x_j - x_j(U)\rho^i(L) & \text{for } i \in [m], j \in [n] \\ y_j^i \le x_j(L)\rho^i + \rho^i(U)x_j - x_j(L)\rho^i(U) & \text{for } i \in [m], j \in [n] \\ y_j^i \ge x_j(U)\rho^i + \rho^i(U)x_j - x_j(U)\rho^i(U) & \text{for } i \in [m], j \in [n] \\ y_j^i \ge x_j(L)\rho^i + \rho^i(L)x_j - x_j(L)\rho^i(L) & \text{for } i \in [m], j \in [n] \\ \text{(18c), (18d)} & \end{array} \right\}, \tag{19}$$

where $\rho := (\rho^i)_{i=1}^m$ and $y := (y^i)_{i=1}^m$. The strength of McCormick inequalities relies heavily on the bounds of the associated variables. For unconstrained BFP, we have that $x_j(L) = 0$, $x_j(U) = 1$, $\rho^i(L) = \frac{1}{\sum_{j=0}^n a_{ij}}$, and $\rho^i(U) = \frac{1}{a_{i0}}$. For the general case, we assume that these bounds are obtained by minimizing/maximizing their defining functions over the linear constraints using linear programming.

We now consider an alternate relaxation, based on the techniques developed in this paper. To solve (FP), it suffices to convexify $\mathcal{G} := \cap_{i=1}^m \mathcal{G}^i$, where

$$\mathcal{G}^i := \left\{ \bigl(\rho^i, y^i, x\bigr) \;\middle|\; \begin{array}{l} x \in \{0,1\}^p \times \mathbb{R}^{n-p}, \overline{C}x \le \bar{d}, \\ \rho^i = \dfrac{1}{a_{i0}+a_i^\intercal x},\ y_j^i = \dfrac{x_j}{a_{i0}+a_i^\intercal x} \text{ for } j \in [n] \end{array} \right\}.$$

Although, tighter relaxations can be derived using copositive programming as in Section 4 or the hierarchy of Section 5.2, we focus, here, on the polynomial-sized relaxations from Section 3 and RLT constraints obtained at the first level. Applying Theorem 2, we get

$$\operatorname{conv}(\mathcal{G}^i) = \Big\{(\rho^i, y^i, x) \;\Big|\; (y^i, W^i) \in \rho^i \operatorname{conv}(\mathcal{F}),\ \rho^i \geq 0,\ a_{i0}\rho^i + a_i^\intercal y^i = 1, \\ x_j = (a_{i0} + a_{ij})y_j^i + \sum_{k \neq j} a_{ik} W_{jk}^i,\ \text{for } j \in [n]\Big\},$$

where $\mathcal{F} = \{(x, xx^\intercal) \mid \overline{C}x \leq \bar{d},\ x \in \{0,1\}^p \times \mathbb{R}^{n-p}\}$. As mentioned before, since it is NP-hard to optimize linear functions over $\mathcal{F}$, we will replace $\operatorname{conv}(\mathcal{F})$ with its 1st-level RLT relaxation, $\widehat{F}$. In particular, $\widehat{\mathcal{F}}$ will be defined as:

$$\widehat{F} = \Big\{(x, X) \;\Big|\; \overline{C}X\overline{C}^\intercal - \bar{d}x^\intercal\overline{C}^\intercal - \overline{C}x\bar{d}^\intercal + \bar{d}\bar{d}^\intercal \geq 0,\ X_{jj} = x_j \text{ for } j \in [p]\Big\}.$$

This construction yields a polyhedral relaxation for $\operatorname{conv}(\mathcal{G})$,

$$\mathcal{R}_{\mathrm{QP}} := \left\{ (\rho, y, x) \;\middle|\; \begin{array}{l} \overline{C}W^i\overline{C}^\intercal - \bar{d}(y^i)^\intercal\overline{C}^\intercal - \overline{C}y^i\bar{d}^\intercal + \rho^i\bar{d}\bar{d}^\intercal \geq 0 \text{ for } i \in [m] \\ a_{i0}\rho^i + a_i^\intercal y^i = 1,\ \rho^i \geq 0 \quad \text{for } i \in [m] \\ x_j = a_{i0}y_j^i + \sum_k a_{ik}W_{jk}^i \quad \text{for } i \in [m],\ j \in [n] \\ \overline{C}y^i \leq \rho^i\bar{d} \quad \text{for } i \in [m] \\ W_{jj}^i = y_j^i \quad \text{for } j \in [p] \\ \overline{C}x \leq \bar{d} \end{array} \right\}. \tag{20}$$

It can be shown that this relaxation is tighter than the relaxation $\mathcal{R}_{\mathsf{LEF}}$.

Proposition 5 $\mathcal{R}_{\mathrm{QP}} \subseteq \mathcal{R}_{\mathsf{LEF}}$.

The containment in Proposition 5 is often strict as our computations and the example in Appendix A.12 demonstrate. In mixed-integer nonlinear programming (MINLP) literature, another form of McCormick relaxation is used to solve (FP), see [64] for example. This relaxation writes $z_i = \frac{b_{i0}+b_i^\intercal x}{a_{i0}+a_i^\intercal x}$ as $z_i\big(a_{i0}+a_i^\intercal x\big) = b_{i0}+b_i^\intercal x$ and relaxes the left hand side using McCormick inequalities. Since McCormick inequalities require bounds, we minimize/maximize $\frac{b_{i0}+b_i^\intercal x}{a_{i0}+a_i^\intercal x}$ (resp. $a_{i0} + a_i^\intercal x$) over $\overline{C}x \leq \bar{d}$ to derive bounds for z_i (resp. $a_{i0} + a_i^\intercal x$). We show that the McCormick inequalities, so derived, are also implied by $\mathcal{R}_{\mathrm{QP}}$.

Proposition 6 *Assume $z_i = \frac{b_{i0}+b_i^\intercal x}{a_{i0}+a_i^\intercal x}$ (resp. $d_i = a_{i0} + a_i^\intercal x$) is bounded between z_i^L and z_i^U (resp. $d_i^L > 0$ and d_i^U) for all x that satisfy $\overline{C}x \leq \bar{d}$. Then, $\mathcal{R}_{\mathrm{QP}}$ implies the inequalities obtained by relaxing the left-hand-side of $z_i d_i = b_{i0} + b_i^\intercal x$ using McCormick inequalities.*

The relaxation described in Proposition 6 can be tightened when the bounds on $a_{i0} + a_i^\intercal x$ are implied by bounds on x [77]. The idea is to disaggregate the product $z_i(a_{i0} + a_i^\intercal x)$ by distributing the product over summation and relaxing $z_i x_j$ using McCormick envelopes. We show in Appendix A.14 that the inequalities so obtained are also implied in $\mathcal{R}_{\mathrm{QP}}$.

The following conic inequalities were proposed in [68] to tighten $\mathcal{R}_{\mathsf{LEF}}$:

$$y^i_j(a_{i0} + a_i^\intercal x) \geq x_j^2 \quad \text{for } i \in [m],\ j \in [p] \tag{21a}$$
$$\rho^i(a_{i0} + a_i^\intercal x) \geq 1 \quad \text{for } i \in [m]. \tag{21b}$$

Inequality (21b) is commonly used in MINLP literature as $\rho^i \geq \frac{1}{a_{i0}+a_i^\intercal x}$, exploiting that the defining function of ρ^i is convex. It is also used to derive nonlinear underestimators for $x_i \times \frac{1}{a_{i0}+a_i^\intercal x}$ by viewing it as a product of two convex functions [55,64,76]. A conic relaxation for the feasible region of (18) that utilizes (21) is then given as follows:

$$\mathcal{R}_{\mathsf{CEF}} = \Big\{\big((\rho^i)_{i=1}^m, (y^i)_{i=1}^m, x\big) \;\Big|\; \big((\rho^i)_{i=1}^m, (y^i)_{i=1}^m, x\big) \in \mathcal{R}_{\mathsf{LEF}} \text{ and satisfies } (21)\Big\}.$$

In [57], formulations based on $\mathcal{R}_{\mathsf{CEF}}$ are found to be the best performing formulations. Here, we show that (21) arises naturally from standard relaxations used in quadratic programming literature. In particular, it is standard to relax $\begin{pmatrix}1\\x\end{pmatrix}\begin{pmatrix}1 & x^\intercal\end{pmatrix} \succeq 0$ into $\begin{pmatrix}1 & x^\intercal\\ x & X\end{pmatrix} \succeq 0$. Following Theorem 2, we homogenize this matrix inequality as follows:

$$\begin{pmatrix}\rho^i & (y^i)^\intercal\\ y^i & W^i\end{pmatrix} \succeq 0 \quad \text{for } i \in [m], \tag{22}$$

and augment $\mathcal{R}_{\mathrm{QP}}$ with this inequality. We show this relaxation implies (21). First, we rewrite (22) in an equivalent form:

$$\begin{pmatrix}a_{i0} + x^\intercal a_i & 1 & x^\intercal\\ 1 & \rho^i & (y^i)^\intercal\\ x & y^i & W^i\end{pmatrix} \succeq 0 \quad \text{for } i \in [m]. \tag{23}$$

Proposition 7 *Adding (23) to $\mathcal{R}_{\mathrm{QP}}$ is equivalent to adding (22) to $\mathcal{R}_{\mathrm{QP}}$.*

Since the matrix in (23) is positive-semidefinite, the determinants of its 2×2 principal minors are non-negative. Therefore, Proposition 7 shows that, adding (23) or, equivalently, (22) to $\mathcal{R}_{\mathrm{QP}}$ implies the inequalities in (21). In particular, (21a) (resp. (21b)) follows by selecting rows and columns indexed 1 and $j+2$ (resp. 1 and 2). Moreover, (23) implies the following conic inequalities:

$$y^i_j\rho^i \geq \big(y^i_j\big)^2 \quad \text{for } i \in [m] \text{ and } j \in [n] \tag{24a}$$
$$y^i_j y^i_k \geq \big(W^i_{jk}\big)^2 \quad \text{for } i \in [m] \text{ and } j, k \in [n]. \tag{24b}$$

using other 2×2 principal minors of the associated matrix. Clearly, (24a) is redundant since $0 \leq y^i_j \leq \rho^i$. However, (24b) tightens the linear relaxation $\mathcal{R}_{\mathrm{QP}}$ above.

6.2 Computational comparison on binary fractional programming

In this subsection, we present numerical experiments on synthetically generated instances of binary fractional programming (BFP). First, we present our formulation 1Term-Conic for BFP. Recall that (18) reformulates (FP). Here, we use a relaxation of $\mathcal{R}_{\mathrm{QP}}$ along with the conic constraint (21b) to relax the feasible region of (18) and obtain a mixed-integer conic formulation for BFP, referred to as 1Term-Conic:

$$
\begin{aligned}
\max\quad & \sum_{i=1}^{m} \left(b_{i0}\rho^i + b_i^\intercal y^i\right) + c^\intercal x && && \\
\text{s.t.}\quad & x \in \mathcal{X} \subseteq \{0,1\}^n,\ (19) && && \\
& \max\{0,\ y_j^i + y_k^i - \rho^i\} \le W_{jk}^i \le \min\{y_j^i,\ y_k^i\} && \text{for } i \in [m], j,k \in [n] && (25a)\\
& a_{i0}\rho^i + a_i^\intercal y^i = 1,\ \rho^i \ge 0 && \text{for } i \in [m] && (25b)\\
& x_j = a_{i0} y_j^i + \sum_k a_{ik} W_{jk}^i && \text{for } i \in [m], j \in [n] && (25c)\\
& y_j^i \le \rho^i && \text{for } i \in [m] && (25d)\\
& W_{jj}^i = y_j^i && \text{for } j \in [n] && (25e)\\
& \rho^i (a_{i0} + a_i^\intercal x) \ge 1 && \text{for } i \in [m]. && (25f)
\end{aligned}
$$

Constraints (25a) - (25e) are derived by specializing $\mathcal{R}_{\mathrm{QP}}$ to unconstrained BFP. Namely, for unconstrained BFP, $\mathcal{R}_{\mathrm{QP}}$ is equivalent to constructing the McCormick envelopes of $x_j x_k$ for $j,k \in [n]$ before homogenizing these inequalities using ρ^i. This was referred to as the (1-Term) relaxation in Section 3. Although (19) is redundant in this formulation, we include these constraints because we will relax (25a) and (25c) partially to allow the formulation to scale to larger problem sizes. The last constraint (25f) is the conic constraint (21b). We remark that the formulation 1Term-Conic is related to the formulation CH proposed in [25] for assortment planning in the context of quick-commerce. There are two types of constraints in formulation CH of [25], with one being the constraint (21b) and another that relaxes certain telescoping sums.

We will compare our formulation 1Term-Conic with two formulations, LEF and CEF, studied in [57]. The formulation LEF and CEF are built on the relaxation $\mathcal{R}_{\mathsf{LEF}}$ and $\mathcal{R}_{\mathsf{CEF}}$, respectively. More specifically, the formulation LEF (resp. CEF) maximize the objective function of 1Term-Conic over $\mathcal{R}_{\mathsf{LEF}}$ (resp. $\mathcal{R}_{\mathsf{CEF}}$) and the binary constraint $x \in \{0,1\}^n$.

In our first experiment, we compare the three formulations, LEF, CEF and 1Term-Conic, in terms of their continuous relaxations on instances of unconstrained BFP, generated using the procedure described in [57]. Specifically, for each $(n,m) \in \{(30,3),(50,5),(70,7)\}$, we randomly generated 30 instances of unconstrained BFP as follows. The coefficient a_{ij} is sampled from a $U[0,20]$ distribution expect for a_{i0} which is sampled from $U[1,20]$, the coefficient b_{ij} is sampled from $U[-20,0]$, and the coefficient $c_i = 0$ for all $i \in [n]$. We solve the continuous relaxation of each of the three formulations for each instance, and then compute the closed LEF gap as

$$
100\% \times \frac{v_{\mathsf{LEF}} - v_{\mathsf{model}}}{v_{\mathsf{LEF}} - \hat{v}} \qquad \text{for } \mathsf{model} \in \{\mathsf{LEF}, \mathsf{CEF}, \mathsf{1Term\text{-}Conic}\},
$$

where $\hat{v}$ is the objective value of the best integer solution obtained from LEF after 600 seconds of computation, and v_{model} is the optimal value of the continuous relaxation of the formulation model. Table 1 shows the continuous relaxation of 1Term-Conic closes up to 60% of the relaxation gap while that of CEF closes up to 30%.

(n, m)	model	closed LEF gap (%)			
		avg.	min.	max.	std.
$(30, 3)$	1Term-Conic	64.2	44.7	100	14.7
	CEF	33.1	19.7	60.1	8.9
$(50, 5)$	1Term-Conic	41.8	33.5	50.7	4.2
	CEF	20.3	15.6	26.9	2.9
$(70, 7)$	1Term-Conic	33.5	26.4	40.7	3.6
	CEF	15.4	11.5	18.7	1.8

Table 1: Comparison of continuous relaxations of LEF, CEF and 1Term-Conic

Our second numerical experiment compares the three formulations by solving assortment optimization problems to global optimality. For 1Term-Conic, we do not introduce W^i_{jk} variables. Instead, we begin with 1Term-Conic without Constraints (25a) and (25c). Then, for the root relaxation solution, we use the active bound from Constraints (25a) to replace W^i_{jk} in (25c) and introduce the resulting constraint. The formulation so obtained will be referred to as 1Term-Conic-R. All three formulations are solved using Gurobi with a time limit of 1 hour. The problems are generated using the procedure described in [57]. Specifically, for $(n, m) \in \{(50, 5), (100, 10), (200, 20), (400, 40), (800, 80)\}$, we randomly generate 10 instances as follows: $a_{ij} \sim U[0, 1]$ and $b_{ij} = a_{ij} r_i$ with $r_i \sim U[1, 3]$ for all $i \in [m]$ and $j \in [n]$, $a_{i0} = 0.1n$ and $b_{i0} = 0$ for all $i \in [m]$, $c_j = 0$ for all $j \in [n]$, and $\mathcal{X} = \{x \in \{0, 1\}^n \mid \sum_{i \in [n]} x_i \leq \kappa\}$ with $\kappa = 20\% \cdot n$. Table 2 shows that 1Term-Conic-R outperforms the other formulations. More specifically, for most problems, 1Term-Conic-R a few branch-and-bound nodes to discover the optimal solution while LEF and CEF require larger branch-and-bound trees. In particular, LEF is unable to solve problems with $(n, m) \in \{(200, 20), (400, 40), (800, 80)\}$ and even a few problems of smaller size. The formulation CEF performs significantly better than LEF, but cannot solve instances with $(n, m) \in \{(400, 40), (800, 80)\}$. For smaller problems, the time taken by LEF and CEF is often an order of magnitude or more higher than that by 1Term-Conic-R.

6.3 Univariate fractional terms in chemical process design

Consider

$$
\begin{aligned}
\max \quad & -ax - b^\intercal y + c^\intercal z \\
\text{s.t.} \quad & z_i = \frac{y_i}{x - r_i} \quad \text{for } i \in [m] \\
& x^L \leq x \leq x^U,\ y_i^L \leq y_i \leq y_i^U \quad \text{for } i \in [m],
\end{aligned} \tag{PD}
$$

where $a \in \mathbb{R}^m$, $b \in \mathbb{R}$, $c \in \mathbb{R}^m$, $r \in \mathbb{R}^m$, and x^L, x^U, y_i^L and y_i^U are constants. The motivation for this problem comes from chemical process design. In particular, to

(n,m)	formulation	time (seconds) avg.	min.	max.	std.	nodes	solved	rgap
(50, 5)	1Term-Conic-R	0.10	0.07	0.15	0.03	1	10	0.0%
	LEF	0.56	0.30	1.58	0.40	2173.4	10	0.0%
	CEF	0.80	0.20	2.33	0.60	1	10	0.0%
(100, 10)	1Term-Conic-R	2.4	0.4	14.0	4.2	2	10	0.0%
	LEF	1125.2	26.5	3600	1432.6	5189583	8	1.5%
	CEF	9.6	2.0	17.0	4.4	14.6	10	0.0%
(200, 20)	1Term-Conic-R	5.8	2.9	13.0	3.1	5.4	10	0.0%
	LEF	3600	3600	3600	0	1749502	0	4.0%
	CEF	50.0	26.1	137.4	36.1	263.4	10	0%
(400, 40)	1Term-Conic-R	65.2	38.3	118.0	23.1	1	10	0.0%
	LEF	3600	3600	3600	0	969919	0	5.7%
	CEF	3600	3600	3600	0	3	0	†
(800, 80)	1Term-Conic-R	626.4	433.9	805.3	115.9	1	10	0.0%
	LEF	3600	3600	3600	0	325.1	0	5.9%
	CEF	3600	3600	3600	0	1	0	†

Table 2: Comparison of LEF, CEF and 1Term-Conic-R on instances of assortment planning. Here, (n,m) specify problem dimensions. Each row consists of data accumulated over 10 instances. The "nodes" column reports the number of branch & bound nodes required to solve the problem, "solved" is the number of instances solved to optimality, and "rgap" is the average of relative gaps, $\frac{z_{\text{best}}-z_{\text{opt}}}{z_{\text{LEF}}-z_{\text{opt}}}$, remaining at the end of the solution process. The "†" symbol denotes that Gurobi was unable to fully process the root node of the branch-and-bound tree within the time limit of 1 hour.

perform distillation, a column requires vapor flow, and generating this vapor flow requires energy. As such, research has focused on identifying the configuration for a given separation task with the minimum vapor flow. Under standard technical assumptions, the minimum vapor flow can be computed using Underwood's method, see [40] for details. Similarly, [41] shows that the liquid phase mole fraction in the condenser/reboiler that is in equilibrium with the vapor can be modeled using a univeriate fraction such as z_i in (PD).

A typical approach to relax the feasible region of (PD) is to reformulate $z_i = \frac{y_i}{x-r_i}$ as $z_i(x-r_i) = y_i$, and then relax the product using McCormick inequalities. To derive a lower (resp. an upper) bound on z_i, we minimize (resp. maximize) $\frac{y_i}{x-r_i}$ over $x \in [x^L, x^U]$ and $y_i \in [y_i^L, y_i^U]$. We refer to this relaxation as Uni-MC. Instead, we relax (PD) using the techniques developed in Section 5.1. Let

$$\mathcal{G} = \Bigl\{(1,\nu,x) \Bigm| 0 \le x \le 1,\ \nu_i = \frac{1}{x-r_i} \text{ for } i \in [m]\Bigr\}.$$

We derive the convex hull of $\mathcal{G}$ using Proposition 4, and relax $z_i = y_i\nu_i$ using McCormick inequalities. Then, we obtain the following relaxation, referred to as Uni-MH,

$$\begin{aligned}
&(1,\nu,x) \in \operatorname{conv}(\mathcal{G}) \\
&z_i \ge \max\Bigl\{\frac{1}{1-r_i}y_i + \nu_i - \frac{1}{1-r_i},\ -\frac{1}{r_i}y_i + 2\nu_i + 2\frac{1}{r_i}\Bigr\} \qquad \text{for } i \in [m] \\
&z_i \le \min\Bigl\{\frac{1}{1-r_i}y_i + 2\nu_i - 2\frac{1}{1-r_i},\ -\frac{1}{r_i}y_i + \nu_i + \frac{1}{r_i}\Bigr\} \qquad \text{for } i \in [m].
\end{aligned}$$

We numerically test our relaxation Uni-MH on synthetically generated problem instances. For each $m \in \{5, 9, 13\}$, we fix $x^L = 0$, $x^U = 1$, $y_i^L = 1$ and $y_i^U = 2$, and randomly generate 100 instances of (PD) as follows: $c_i \sim U[-1, 1]$, $r_i \sim U[-1, -0.1]$ for $i \in [\frac{m+1}{2}]$ and $r_i \sim U[1.1, 2]$ otherwise, and (a, b) is the gradient of $\sum_{i\in[m]} c_i \frac{y_i}{x-r_i}$ at a point uniformly sampled from $[0, 1] \times [1, 2]^m$. We solve Uni-MC and Uni-MH for each instance, and then compute the remaining gap as

$$\text{Relative remaining gap} = 100\% \times \frac{v_{\text{Uni-MH}} - v_{\text{SCIP}}}{v_{\text{Uni-MC}} - v_{\text{SCIP}}},$$

where v_{SCIP} is the optimal objective value of (PD) returned by SCIP [10], and $v_{\text{Uni-MH}}$ (resp. $v_{\text{Uni-MC}}$) is the optimal objective value of Uni-MH (resp. Uni-MC). In Figure 2, we show the empirical distribution function of the relative remaining gap for the 100 instances. Observe that for 50% of the instances, the relative remaining gap is below 5%, and for 80% of the instances, the relative remaining gap is below 25%.

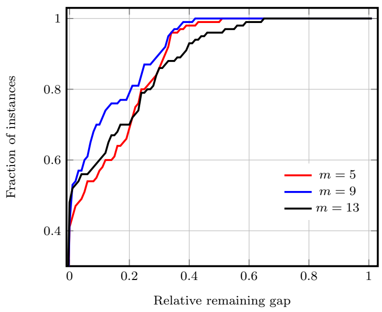


Fig. 2: Cumulative distribution plot of the relative remaining gap for instances of (PD).

7 Conclusions

In this paper, we show that the study of fractional programming problems can leverage results in polynomial optimization literature. This result hinges upon a projective transformation that involves homogenization and renormalization of the convex hull of polynomial functions. We exemplify the usage of this result by connecting with existing studies on a number of convexification techniques. Numerical results indicate that the reformulations and relaxations developed in this paper outperform existing literature in solving assortment optimization problems and close significant gap on a set appearing in chemical process design literature.

It is necessary to investigate the practical applications of our convexification techniques. This includes using our formulations on non-synthetic instances, such as assortment planning and distillation configuration design, and incorporating recent developments towards making copositive and semidefinite programming based relaxations tractable from a computational standpoint. It is also interesting to study whether the projective transformation, introduced in Theorem 1, can be applied in designing approximation algorithms for fractional programming problems.

A Missing Proofs and Discussions

A.1 Proof of Theorem 2

Let CG denote the right hand side set of (8). Since $0\,\mathrm{cl}\,\mathrm{conv}(\mathcal{F})$ represents the recession cone of $\mathrm{cl}\,\mathrm{conv}(\mathcal{F})$, by Theorem 9.6 in [65], it follows that CG is closed. By using Theorem 1 and lifts of $\mathcal{F}$ and $\mathcal{G}$ defined as in (9), we obtain that $\mathrm{conv}(\mathcal{G}) \subseteq \mathrm{CG}$. Therefore, it follows that $\mathrm{cl}\,\mathrm{conv}(\mathcal{G}) \subseteq \mathrm{CG}$. Now, we show the reverse inclusion. Let $g \in \mathrm{CG}$. First, assume $\rho > 0$. Let $f = \frac{g}{\rho}$ and consider a sequence $(f^k)_k \subseteq \mathrm{conv}(\mathcal{F})$ such that $f^k \to f$. Since $\alpha^\intercal g = 1$, it follows that $\alpha^\intercal f = \frac{1}{\rho} > 0$. Then, there exists a K such that for all $k \geq K$, $\alpha^\intercal f^k > \frac{1}{2\rho}$. We limit the sequence to $(f^k)_{k \geq K}$. Observe that

$$0 \leq \left\| \frac{f^k}{\alpha^\intercal f^k} - \frac{f}{\alpha^\intercal f} \right\| \leq \left\| \frac{f^k}{\alpha^\intercal f^k} - \frac{f}{\alpha^\intercal f^k} \right\| + \left\| \frac{f}{\alpha^\intercal f^k} - \frac{f}{\alpha^\intercal f} \right\| \to 0,$$

where the convergence holds since $\|f^k - f\| \to 0$, $\alpha^\intercal f^k > \frac{1}{2\rho}$, and $\alpha^\intercal f = \frac{1}{\rho} > 0$. Thus, $g^k := \frac{f^k}{\alpha^\intercal f^k}$ converges to $g := \frac{f}{\alpha^\intercal f}$. Moreover, $g^k \in \mathrm{conv}(\mathcal{G})$, implying that $g \in \mathrm{cl}\,\mathrm{conv}(\mathcal{G})$. To see that $g^k \in \mathrm{conv}(\mathcal{G})$, we invoke Theorem 1 and lifts defined as in (9) as follows. As $f^k \in \mathrm{conv}(\mathcal{F})$, its lift $\big(\alpha^\intercal f^k, f^k\big)$ belongs to $\mathrm{conv}(\mathcal{S})$, where $\mathcal{S}$ is defined as in (9). It follows from Theorem 1 that $\big(\frac{1}{\alpha^\intercal f^k}, \frac{f^k}{\alpha^\intercal f^k}\big) \in \mathrm{conv}\big(\Phi(\mathcal{S})\big)$, and we have $g^k = \frac{f^k}{\alpha^\intercal f^k} \in \mathrm{conv}(\mathcal{G})$ by projecting out the first coordinate. To complete the proof, we treat the case where $\rho = 0$. In this case, g belongs to the recession cone of $\mathrm{cl}\,\mathrm{conv}(\mathcal{F})$ and $\alpha^\intercal g = 1$. Since $\mathcal{F} \neq \emptyset$, we may choose $f \in \mathrm{ri}\big(\mathrm{conv}(\mathcal{F})\big)$. Consequently, $f + \beta g \in \mathrm{conv}(\mathcal{F})$ for all $\beta > 0$ (see Corollary 8.3.1 in [65]). As above, $\frac{(f+\beta g)}{\alpha^\intercal f + \beta} \in \mathrm{conv}(\mathcal{G})$, where we have used that $\alpha^\intercal g = 1$. Taking $\beta \to \infty$, we have that $g \in \mathrm{cl}\,\mathrm{conv}(\mathcal{G})$. □

A.2 Proof of Corollary 1

Let $I = \{i,j\}$ where $i,j \in [n]$ with $i \neq j$. Due to (13), it suffices to describe $\mathrm{QP}_{G(I)}$. For $k \in V \setminus \{i,j\}$, let C_k denote the triangle subgraph of $G(I)$ whose vertex set is V and edge set is $\big\{(i,j),(j,k),(i,k)\big\}$, and define $\mathcal{F}_{C_k} = \big\{(x_i,x_j,x_k,z_{ij},z_{jk},z_{ik}) \in \{0,1\}^6 \;\big|\; z_{ij} = x_ix_j,\ z_{jk} = x_jx_k,\ z_{ik} = x_ix_k\big\}$. Then, it follows that $\mathrm{QP}_{G(I)}$ is the convex hull of $\big\{(x,z) \;\big|\; (x_i,x_j,x_k,z_{ij},z_{jk},z_{ik}) \in \mathcal{F}_{C_k} \forall k \in V \setminus \{i,j\}\big\}$. It follows from [62] that $\mathrm{conv}(\mathcal{F}_{C_k})$ can be described using McCormick and triangle inequalities. Moreover, for each $k \neq i,j$, the projection of $\mathcal{F}_{C_k}$ onto the space of (x_i,x_j,z_{ij}) variables is $\{(x_i,x_j,x_ix_j) \mid x_i \in \{0,1\},\ x_j \in \{0,1\}\}$, a set of affinely independent points. Thus, the proof is complete by invoking Lemma 2 to obtain $\mathrm{QP}_{G(I)}$ as the intersection of $\mathrm{conv}(\mathcal{F}_{C_k})$. □

A.3 Proof of Proposition 1

Let $I \subseteq V$ with $|I| = k$. Due to (13), it suffices to show that RLT_k yields a description of $\mathrm{QP}_{G(I)}$. Consider a set defined by all multilinear monomials in variable $(x_i)_{i \in I}$, that is, $\mathcal{M}_0 = \big\{(x,z^0) \;\big|\; x \in \{0,1\}^n,\ z^0_e = \prod_{t \in e} x_t\ \forall e \subseteq I \text{ s.t. } |e| \geq 2\big\}$. For each $j \notin I$, consider a set defined by all multilinear monomials in variable $(x_i)_{i \in I \cup \{j\}}$,

$$\mathcal{M}_j := \left\{ (x,z^0,z^j) \;\middle|\; (x,z^0) \in \mathcal{M}_0, z^j_e = \prod_{t \in e} x_t \forall e \subseteq I \cup \{j\} \text{ s.t. } |e| \geq 2 \text{ and } j \in e \right\}.$$

Let $\mathcal{M} := \{(x, z^0, (z^j)_{j\notin I}) \mid (x, z^0, z^j) \in \mathcal{M}_j \ \forall j \notin I\}$. Clearly, the convex hull of $\mathcal{M}$ is an extended formulation of $\mathrm{QP}_{G(I)}$. The proof is complete if we argue that RLT_k describes the convex hull of $\mathcal{M}$. It follows readily from Lemma 1 that RLT_k describes the convex hull of $\mathcal{M}_j$ for every $j \notin I$. Now, observe that for each $j \notin I$, $\mathrm{proj}_{(x,z^0)}(\mathcal{M}_j) = \mathcal{M}_0$, and $\mathcal{M}_0$ is a set of affinely independent points. Thus, we can invoke the Lemma 2 to obtain $\mathrm{conv}(\mathcal{M}) = \{(x, z^0, (z^j)_{j\notin I}) \mid (x, z^0, z^j) \in \mathrm{conv}(\mathcal{M}_j) \ \forall j \notin I\}$. □

A.4 Proof of Proposition 2

Let $f_G(x) = (1, x, z)$, where $z_{ij} = x_i x_j$ for $(i,j) \in E$, and consider $\mathcal{F}_G = \{f_G(x) \mid x \in [0,1]^V\}$. Then, $\mathrm{conv}(\mathcal{F}_G) = \mathrm{conv}\{f_G(x) \mid x \in \{0,1\}^V\} = \{(1, x, z) \mid (x,z) \in \mathrm{QP}_G\}$, where the first equality follows from Theorem 1 in [20] and the second equality is by the definition of QP_G. Moreover, by (15) and boundedness of $\mathcal{G}_G$, Theorem 2 yields the expression of $\mathrm{conv}(\mathcal{G}_G)$ and QP_G. □

A.5 Proof of Corollary 2

If G is series-parallel graph, by Theorem 10 of [62], the $\rho\,\mathrm{BQP}_G$ is described using the odd-cycle inequalities, and thus, by Proposition 2, the convex hull of $\mathcal{G}_G$ is given by the odd-cycle relaxation. Moreover, by [5], the separation problem of BQP_G is polynomial time solvable and, thus, by Remark 1, so is the separation problem of the odd cycle relaxation. Therefore, due to the polynomial equivalence between separation and optimization [42], minimizing a linear function over the odd-cycle relaxation is polynomial-time solvable. Hence, (14) is polynomial-time solvable. □

A.6 Proof of Proposition 3

Lemma 4 (Lemma 5.1 in [4]) *Consider a subset $\mathcal{S}$ of $\mathbb{R}^n$, and a hyperplane $H := \{x \mid \ell(x) = \beta\}$ so that $\ell(x) \leq \beta$ is a valid linear inequality of $\mathcal{S}$. Then,* $\mathrm{conv}(\mathcal{S} \cap H) = \mathrm{conv}(\mathcal{S}) \cap H$.

If $\mathcal{L} = \mathbb{R}^n$, that is $C = 0$ and $d = 0$, the result follows from Theorem 2 by letting $f(x) = (1, x, xx^\intercal)$ and $\alpha = (1, a, A)$. Now, assume that $\mathcal{L}$ is not $\mathbb{R}^n$. By Theorem 2, to obtain the convex hull of $\mathcal{G}$, it suffices to derive the convex hull of $\mathcal{F}'$, where $\mathcal{F}' := \{(x, X) \mid x \in \mathcal{X} \cap \mathcal{L},\ X = xx^\intercal\}$. We argue that $\mathcal{F}'$ can be obtained as the intersection of $\mathcal{F}$ and a supporting hyperplane. Let c^i be the i^{th} row of C, and d_i be the i^{th} entry of d. Then, x satisfies $c^i x = d_i$ for $i \in [m]$ if and only if $\mathrm{Tr}(Cxx^\intercal C^\intercal - Cxd^\intercal - dx^\intercal C^\intercal + dd^\intercal) = \sum_{i\in[m]}(c^i x - d_i)^2 = 0$. Thus, $\mathcal{F}' = \{(x, X) \in \mathcal{F} \mid \mathrm{Tr}(CXC^\intercal - Cxd^\intercal - dx^\intercal C^\intercal + dd^\intercal) = 0\}$. Since $\mathrm{Tr}(CXC^\intercal - Cxd^\intercal - dx^\intercal C^\intercal + dd^\intercal) \geq 0$ is a valid linear inequality for $\mathcal{F}$, it follows from Lemma 4 that

$$\mathrm{conv}(\mathcal{F}') = \Big\{(x, X) \ \Big|\ \mathrm{Tr}\big(CXC^\intercal - Cxd^\intercal - dx^\intercal C^\intercal + dd^\intercal\big) = 0, (x, X) \in \mathrm{conv}(\mathcal{F})\Big\}.$$

This, together with Theorem 2, gives the description of $\mathrm{conv}(\mathcal{G})$. □

A.7 Proof of Corollary 4

Lift $\mathcal{S}$ to $\{(t, x) \mid t = b_0\rho + b^\intercal y + \langle B, Y\rangle,\ x_i = a_1 Y_{i1} + a_2 Y_{i2} + y_i\ i = 1, 2,\ (\rho, y, Y) \in \mathcal{G}\}$, where

$$\mathcal{G} = \left\{ \frac{(1, x, xx^\intercal)}{a^\intercal x + a_0} \ \middle|\ Cx \leq d \right\}.$$

Since $\mathcal{S}$ is obtained as a linear transformation of $\mathcal{G}$, it suffices to convexify $\mathcal{G}$. Let $\mathcal{F} = \{(x, X) \mid Cx \leq d,\ X = xx^\intercal\}$. By Theorem 2 in [19], we obtain $\mathrm{conv}(\mathcal{F}) = \{(x, X) \mid Z(1, x, X) \succeq 0,\ CXC^\intercal - Cxd^\intercal - dx^\intercal C^\intercal + dd^\intercal \geq 0\}$. Thus, using Proposition 3, we obtain a convex hull description of $\mathcal{G}$. □

A.8 Proof of Lemma 3

Let $\mathcal{P}^{n+1}$ be the vector space of univariate polynomials with degree at most $n+1$. The set of monomials $\{x^0, x^1, \dots x^{n+1}\}$ is a basis of $\mathcal{P}^{n+1}$. Now, we show that $\{f_0(\cdot), f_1(\cdot), \dots, f_{n+1}(\cdot)\}$ is also a basis of $\mathcal{P}^{n+1}$, and T is an invertible map between two bases. For each $i \in \{0\} \cup [n+1]$, the linear combination of $f_0(\cdot), \dots, f_{n+1}(\cdot)$ with weights $\beta_{i0}, \dots, \beta_{i(n+1)}$ interpolates the monomial x^i over the given set of reals $\{r_0, \dots, r_{n+1}\}$, that is, $x^i = \beta_{i0} f_0(x) + \cdots + \beta_{i(n+1)} f_{n+1}(x)$ for $x \in \{r_0, \dots, r_{n+1}\}$. This can be easily verified because $f_i(x)$ is 0 when $x = r_j$ for $j \neq i$. Since $r_0, \dots, r_{n+1}$ are $n+2$ distinct reals, the interpolating function is unique. In other words, $(x^0, x, \dots, x^{n+1}) = T(f_0(x), \dots, f_{n+1}(x))$ for all x. Thus, by a counting argument, we conclude that $\{f_i(\cdot)\}_{i=0}^{n+1}$ is also a basis and T is invertible. In particular, $\mathcal{M}_{n+1} = T\mathcal{F}$ and $\mathcal{F} = T^{-1}\mathcal{M}_{n+1}$, completing the proof. □

A.9 Proof of Proposition 4

First, we obtain the convex hull characterization of $\mathcal{G}$. Observe that

$$\begin{aligned}\operatorname{conv}(\mathcal{G}) &= \bigl\{\nu \bigm| \nu \in \rho \operatorname{conv}(\mathcal{F}),\ \nu_0 = 1,\ \rho \geq 0\bigr\} = \bigl\{\nu \bigm| \nu \in \rho \operatorname{conv}(T^{-1}\mathcal{M}_{n+1}),\ \nu_0 = 1,\ \rho \geq 0\bigr\} \\ &= \bigl\{\nu \bigm| T\nu \in \rho \operatorname{conv}(\mathcal{M}_{n+1}),\ \nu_0 = 1,\ \rho \geq 0\bigr\} = \bigl\{\nu \bigm| T\nu \in \operatorname{cone}(\mathcal{M}_{n+1}),\ \nu_0 = 1\bigr\}.\end{aligned}$$

where the first equality follows from Theorem 2, the second equality holds due to $\mathcal{F} = T^{-1}\mathcal{M}_{n+1}$ in Lemma 3, the third equality holds since convexification commutes with linear transformation, and the last equality holds by the definition of $\operatorname{cone}(\mathcal{M}_{n+1})$.

Next, we derive the convex hull of $\mathcal{M}_{n+1}$ in terms of $\operatorname{conv}(\mathcal{G})$. By partial fraction decomposition, there exists real numbers $\alpha_1, \dots, \alpha_n$ such that

$$\frac{1}{f_0(x)} = \alpha_1 \frac{1}{x - r_1} + \ldots + \alpha_n \frac{1}{x - r_n}.$$

This implies that $1 = \sum_{i \in [n]} \alpha_i f_i(x)$. As defined in the statement of Lemma 3 and shown in its proof, $(\beta_{0i})_{i=0}^{n+1}$ is the unique weight vector such that $1 = \sum_{i=0}^{n+1} \beta_{0i} f_i(x)$. Thus, $\alpha_i = \beta_{0i}$ for $i \in [n]$ and $\beta_{00} = \beta_{0(n+1)} = 0$. We will use Theorem 1 to obtain the convex hull of $\mathcal{F}$. To do so, we derive the convex hull of $\Phi(\mathcal{F})$ (see (5)), which is as follows

$$\begin{aligned}\operatorname{conv}\bigl(\Phi(\mathcal{F})\bigr) &= \operatorname{conv}\Bigl\{\Bigl(\frac{1}{f_0(x)}, \frac{1}{x - r_1}, \dots, \frac{1}{x - r_n}, x - r_0\Bigr) \Bigm| x \in \mathcal{X}\Bigr\} \\ &= \operatorname{conv}\Bigl\{\Bigl(\sum_{i \in [n]} \alpha_i \nu_i, \nu_1, \dots, \nu_{n+1}\Bigr) \Bigm| (1, \nu_1, \dots, \nu_{n+1}) \in \mathcal{G}\Bigr\} \\ &= \Bigl\{\Bigl(\sum_{i \in [n]} \alpha_i \nu_i, \nu_1, \dots, \nu_{n+1}\Bigr) \Bigm| (1, \nu_1, \dots, \nu_{n+1}) \in \operatorname{conv}(\mathcal{G})\Bigr\},\end{aligned}$$

where the first equality holds by the definition of Φ, the second equality follows from the partial fraction decomposition, and the last equality follows because convexification commutes with linear transformation. Using Theorem 1, we obtain

$$\begin{aligned}\operatorname{conv}(\mathcal{F}) &= \Bigl\{(f_0, f_1, \dots, f_{n+1}) \Bigm| (1, f_1, \dots, f_{n+1}) \in f_0 \operatorname{conv}\bigl(\Phi(\mathcal{F})\bigr),\ f_0 \geq 0\Bigr\} \\ &= \Bigl\{(f_0, f_1, \dots, f_{n+1}) \Bigm| \sum_{i \in [n]} \alpha_i f_i = 1,\ (f_0, f_1, \dots, f_{n+1}) \in \operatorname{cone}(\mathcal{G})\Bigr\},\end{aligned}$$

where the first equality holds from Theorem 1, and the second equality follows from the convex hull description of $\Phi(\mathcal{F})$. Let β_0 be the 0^{th} row of T, that is $\beta_0 := (\beta_{00}, \dots, \dots, \beta_{0(n+1)})$. Now,

$$\begin{aligned}\operatorname{conv}(\mathcal{M}_{n+1}) &= \operatorname{conv}(T\mathcal{F}) = T \operatorname{conv}(\mathcal{F}) \\ &= \bigl\{\mu \bigm| \langle \beta_0, T^{-1}\mu \rangle = 1,\ T^{-1}\mu \in \operatorname{cone}(\mathcal{G})\bigr\} \\ &= \bigl\{\mu \bigm| \mu_0 = 1,\ T^{-1}\mu \in \operatorname{cone}(\mathcal{G})\bigr\},\end{aligned}$$

where the first equality follows from Lemma 3, the third equality is obtained by using the fact that $\beta_{0i} = \alpha_i$ for $i \in [n]$ and $\beta_{00} = \beta_{0(n+1)} = 0$. □

A.10 Proof of Theorem 3

We will show that $\mathrm{conv}(\mathcal{G}^n) = \mathcal{R}^n$. This implies that

$$\mathrm{proj}_{(\rho,y)}(\mathcal{R}^n) = \mathrm{proj}_{(\rho,y)}\big(\mathrm{conv}(\mathcal{G}^n)\big) = \mathrm{conv}\big(\mathrm{proj}_{(\rho,y)}(\mathcal{G}^n)\big) = \mathrm{conv}(\mathcal{G}),$$

where the second equality holds since convexification commutes with projection, and the last equality holds since $\mathrm{proj}_{(\rho,y)}(\mathcal{G}^n) = \mathcal{G}$. For each $i \in [m]$, we obtain that $\mathrm{proj}_u(\mathcal{G}_i^n) = U$. Since U is a set of affinely independent points, it follows from Lemma 2 that the convex hull of $\mathcal{G}^n$ is $\big\{(\rho, y, u) \bigm| (\rho^i, y^i, u) \in \mathrm{conv}(\mathcal{G}_i^n)\ \forall i \in [m]\big\}$. The proof is complete if we show that the convex hull of $\mathcal{G}_i^n$ is $\mathcal{R}_i^n$. We prove this by invoking Theorem 2. Let $f^i(x) = (1, x, v^i)$, where $v_S^i = \prod_{j\in S} x_j(a_{i0} + a_i^\intercal x)$ for nonempty $S \subseteq [n]$, and let $\mathcal{F}^i = \big\{f^i(x) \mid x \in \{0,1\}^n\big\}$. By Theorem 3 in [70], the n^{th} level RLT of $\mathcal{X}$ describes the convex hull of U. Thus, we obtain

$$\begin{aligned}\mathrm{conv}(\mathcal{F}^i) =\Big\{ \big(1, x^i, v^i\big) \Bigm| & z^i \in \mathrm{RLT}_n(\mathcal{X}),\ x_j^i = z_{\{j\}}^i \text{ for } j \in [n],\\ & v_S^i = a_{i0} + \sum_{j\in S}\Big(a_{ij} z^i_{S\cup\{j\}}\Big) \text{ for } S \subseteq [n] \text{ s.t. } 1 \le |S| \le k\Big\}.\end{aligned}$$

Let $\alpha_i = (a_{i0}, a_i, 0)$. Then, $a_{i0} + a_i^\intercal x = \alpha_i^\intercal f^i(x)$. Defining $y^i = \rho^i x^i$, $u = \rho^i v^i$ and $w^i = \rho^i z^i$, Theorem 2 shows that $\mathrm{conv}(\mathcal{G}_i^n)$ is $\mathcal{R}_i^n$. □

A.11 Proof of Proposition 5

Since every point in $\mathcal{R}_{\mathrm{QP}}$ satisfies (18c) and (18d), it suffices to show that the McCormick constraints used to relax (18b) are implied in $\mathcal{R}_{\mathrm{QP}}$. Towards this end, we will show a more general result. Assume $\alpha^\intercal x \le \bar{\alpha}$ and $\beta^\intercal x \le \bar{\beta}$ are implied by $\overline{C}x \le \bar{d}$. We show that $\mathcal{R}_{\mathrm{QP}}$ implies the inequality $\beta^\intercal W^i \alpha - \bar{\alpha}\beta^\intercal y^i - \bar{\beta}(y^i)^\intercal\alpha + \rho^i\bar{\alpha}\bar{\beta} \ge 0$, which is obtained by linearizing $\big(\alpha^\intercal x - \bar{\alpha}\big)\big(\beta^\intercal x - \bar{\beta}\big) \ge 0$ and then homogenizing the resulting inequality using ρ^i. Since the two inequalities in the product are implied by the linear inequalities, there exists an $\omega \ge 0$ and a $\nu \ge 0$ such that $\overline{C}^\intercal\omega = \alpha$, $\bar{d}^\intercal\omega \le \bar{\alpha}$, $\nu^\intercal\overline{C} = \beta^\intercal$, and $\nu^\intercal\bar{d} \le \bar{\beta}$. Therefore, it follows that every point in $\mathcal{R}_{\mathrm{QP}}$ satisfies:

$$\begin{aligned}&\nu^\intercal\overline{C}W^i\overline{C}^\intercal\omega - \nu^\intercal\bar{d}\Big((y^i)^\intercal\overline{C}^\intercal - \rho^i\bar{d}^\intercal\Big)\omega - \nu^\intercal\overline{C}y^i\bar{d}^\intercal\omega \ge 0\\ &\Rightarrow \beta^\intercal W^i\alpha - \bar{\beta}\Big((y^i)^\intercal\overline{C}^\intercal - \rho^i\bar{d}^\intercal\Big)\omega - \beta^\intercal y^i\bar{d}^\intercal\omega \ge 0\\ &\Rightarrow \beta^\intercal W^i\alpha - \bar{\beta}(y^i)^\intercal\alpha - \big(\beta^\intercal y^i - \rho^i\bar{\beta}\big)\bar{d}^\intercal\omega \ge 0\\ &\Rightarrow \beta^\intercal W^i\alpha - \bar{\beta}(y^i)^\intercal\alpha - \big(\beta^\intercal y^i - \rho^i\bar{\beta}\big)\bar{\alpha} \ge 0,\end{aligned}\tag{26}$$

where the first implication is because $\big(\nu^\intercal\bar{d} - \bar{\beta}\big)\big((y^i)^\intercal\overline{C}^\intercal - \rho^i\bar{d}^\intercal\big)\omega \ge 0$, and the third implication is because $\big(\beta^\intercal y^i - \rho^i\bar{\beta}\big)\big(\bar{d}^\intercal\omega - \bar{\alpha}\big) \ge 0$. Now, this shows that the homogenization of the linearization of $\big(x_j(L) - x_j\big)\Big(a_{i0} + a_i^\intercal x - \frac{1}{\rho^i(L)}\Big) \ge 0$ is implied in $\mathcal{R}_{\mathrm{QP}}$, which in turn implies that $\mathcal{R}_{\mathrm{QP}}$ satisfies

$$\begin{aligned}&-e_j^\intercal W^i a_i + x_j(L)(y^i)^\intercal a_i + \big(y_j^i - \rho^i x_j(L)\big)\Big(\frac{1}{\rho^i(L)} - a_{i0}\Big) \ge 0\\ &\Rightarrow -e_j^\intercal W^i a_i + x_j(L) + \frac{y_j^i}{\rho^i(L)} - \frac{\rho^i x_j(L)}{\rho^i(L)} - a_{i0}y_j^i \ge 0\\ &\Rightarrow -x_j + x_j(L) + \frac{y_j^i}{\rho^i(L)} - \frac{\rho^i x_j(L)}{\rho^i(L)} \ge 0\\ &\Rightarrow -\rho^i(L)x_j + \rho^i(L)x_j(L) + y_j^i - \rho^i x_j(L) \ge 0,\end{aligned}$$

which is one of the McCormick inequalities. The first implication above is because $\rho^i a_{i0} + (y^i)^\intercal a_i = 1$, and the second implication is because $x_j = a_{i0}y_j^i + e_j^\intercal W^i a_i$. The remaining McCormick inequalities follow using a similar argument with other bounds on x_j and ρ^i.

□

A.12 Example illustrating $\mathcal{R}_{\mathrm{QP}} \subsetneq \mathcal{R}_{\mathsf{LEF}}$

We give an example to illustrate that the containment may be strict. Let $m = 1$, $n = 2$, $a_0 = a_1 = a_2 = 1$ and consider the unconstrained BFP, in which case $\mathcal{R}_{\mathrm{QP}} = \mathrm{conv}(\mathcal{G}^1)$ because when $n = 2$ the 1st-level RLT relaxation $\widehat{F}$ coincides with the convex hull of $\mathcal{F}$. We verify that $p = (\rho, y_1, y_2, x_1, x_2) = (\frac{1}{2}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}) \in \mathcal{R}_{\mathsf{LEF}} \setminus \mathrm{conv}(\mathcal{G}^1)$ when taking $x_j(L) = 0$, $x_j(U) = 1$, for $j = 1, 2$ and $\rho(L) = \frac{1}{3}$, $\rho(U) = 1$. To verify that $p \notin \mathrm{conv}(\mathcal{G}^1)$, observe that $x_1 = \frac{x_1(1+x_1+x_2)}{1+x_1+x_2} \geq \frac{2x_1}{1+x_1+x_2} = 2y_1$. To see that $p \in \mathcal{R}_{\mathsf{LEF}}$, it follows easily that p satisfies (18c). Moreover,

$$\begin{pmatrix} \rho \\ x_j \\ y_j \end{pmatrix} = \frac{3}{4}\begin{pmatrix} \rho(L) \\ x_j(L) \\ x_j(L)\rho(L) \end{pmatrix} + \frac{1}{4}\begin{pmatrix} \rho(U) \\ x_j(U) \\ x_j(U)\rho(U) \end{pmatrix}$$

which implies that y_j satisfies the McCormick inequalities in (19) which relax $y_j = x_j \rho$.

A.13 Proof of Proposition 6

All such inequalities are obtained as linearizations of products of certain inequalities. Therefore, we will first show a more general result. Assume that the inequality $\gamma_0 \rho^i + \gamma^\intercal y^i - \bar{\gamma} \leq 0$ is implied by $\overline{C}\bar{y} \leq \rho^i \bar{d}$ and $a_{i0}\rho^i + a_i^\intercal y^i = 1$, and $\beta^\intercal x - \bar{\beta} \leq 0$ is implied by $\overline{C}x \leq \bar{d}$. Then, we will show that the homogenization of

$$\big(\gamma_0 \rho^i + \gamma^\intercal y^i - \bar{\gamma}\big)\big(\beta^\intercal x - \bar{\beta}\big) \geq 0 \tag{27}$$

$$\text{expressed as } \gamma_0 (y^i)^\intercal \beta + \gamma^\intercal W^i \beta - \bar{\gamma} x^\intercal \beta - \bar{\beta}\gamma_0 \rho^i - \bar{\beta}\gamma^\intercal y^i + \bar{\gamma}\bar{\beta} \geq 0, \tag{28}$$

is implied by the constraints in $\mathcal{R}_{\mathrm{QP}}$. Since the two inequalities whose product is homogenized are implied by linear inequalities in $\mathcal{X}$, there exists an $\omega \geq 0$, $\theta \in \mathbb{R}$, and a $\nu \geq 0$ such that $\overline{C}^\intercal \omega + \theta a_i = \gamma$, $-\bar{d}^\intercal \omega + \theta a_{i0} = \gamma_0$, $\theta \leq \bar{\gamma}$, $\nu^\intercal \overline{C} = \beta$, and $\nu^\intercal \bar{d} \leq \bar{\beta}$. It follows that $0 \geq x^\intercal \overline{C}^\intercal \omega - \bar{d}^\intercal \omega = x^\intercal(\gamma - \theta a_i) + (\gamma_0 - \theta a_{i0})$. Now, points in $\mathcal{R}_{\mathrm{QP}}$ satisfy:

$$\begin{aligned}
&(\gamma - \theta a_i)^\intercal W^i \beta - (\theta a_{i0} - \gamma_0)(y^i)^\intercal \beta - (\gamma - \theta a_i)^\intercal y^i \bar{\beta} + \rho^i(\theta a_{i0} - \gamma_0)\bar{\beta} \\
&\quad + \theta\Big(\big(a_{i0} y^i + a_i^\intercal W^i - x^\intercal\big)\beta - \big(a_{i0}\rho^i + a_i^\intercal y^i - 1\big)\bar{\beta}\Big) \geq 0 \\
&\Rightarrow \gamma^\intercal W^i \beta + \gamma_0 (y^i)^\intercal \beta - \gamma^\intercal y^i \bar{\beta} - \rho^i \gamma_0 \bar{\beta} - \theta(x^\intercal \beta - \bar{\beta}) \geq 0 \\
&\Rightarrow \gamma^\intercal W^i \beta + \gamma_0 (y^i)^\intercal \beta - \gamma^\intercal y^i \bar{\beta} - \rho^i \gamma_0 \bar{\beta} - \bar{\gamma}(x^\intercal \beta - \bar{\beta}) \geq 0.
\end{aligned}$$

The first inequality follows since $\mathcal{R}_{\mathrm{QP}}$ satisfies $a_{i0} y^i + a_i^\intercal W^i = x^\intercal$, $a_{i0}\rho^i + a_i^\intercal y^i = 1$, and (26), where $\alpha = \gamma - \theta a_i$ and $\bar{\alpha} = \theta a_{i0} - \gamma_0$. The modified form of (26) holds because $x^\intercal(\gamma - \theta a_i) + (\gamma_0 - \theta a_{i0}) \leq 0$ and $\beta^\intercal x - \bar{\beta} \leq 0$ are a consequence of $\overline{C}x \leq \bar{d}$. The second implication follows since $\big(\theta - \bar{\gamma}\big)\big(x^\intercal \beta - \bar{\beta}\big) \geq 0$. Replacing z_i with $b_{i0}\rho^i + b^\intercal y^i$ and d_i with $a_{i0} + a_i^\intercal x$, we reduce $\big(z_i - z_i^L\big)\big(d_i - d_i^L\big) \geq 0$ into an inequality of the form (27). Then, our earlier argument shows that its linearization, which yields one of the McCormick inequalities, is implied by the inequality corresponding to (28), after $b_{i0}\rho^i + b^\intercal y^i$ (resp. $b_{i0} y^\intercal + b^\intercal W^i$) is replaced by z_i (resp. variables representing $z_i x^\intercal$). A similar argument shows that the remaining McCormick inequalities are also implied in $\mathcal{R}_{\mathrm{QP}}$. □

A.14 Disaggregation of products

Proposition 8 *Consider the equality $z_i(a_{i0} + a_i^\intercal x) = b_{i0} + b_i^\intercal x$, where z_i represents $\frac{b_{i0} + b_i^\intercal x}{a_{i0} + a_i^\intercal x}$. Assume h_{ij} represents $z_i x_j$. Then, $\mathcal{R}_{\mathrm{QP}}$ implies the following inequalities:*

1. *McCormick inequalities that relax $h_{ij} = z_i x_j$ based on bounds of z_i and x_j obtained over $\overline{C}x \leq \bar{d}$.*
2. *The equality $a_{i0} z_i + a_i^\intercal h_i = b_{i0} + b_i^\intercal x$.*

Proof Proof. We have that:

$$\begin{aligned} b_i^\intercal W_{\cdot j}^i + b_{i0} y_j^i - (b_{i0}\rho^i + b_i^\intercal y^i) x_j(L) - z_i^L (x_j - x_j(L)) \geq 0 \\ \Rightarrow h_{ij} - z_i x_j(L) - z_i^L (x_j - x_j(L)) \geq 0. \end{aligned}$$

The first inequality follows since (28) is implied in $\mathcal{R}_{\mathrm{QP}}$ with β as the j^{th} standard vector, $\bar{\beta} = x_j(L)$, $\gamma = b_i$, $\gamma_0 = b_{i0}$, and $\bar{\gamma} = z_i^L$. This is because $x_j - x_j(L) \geq 0$ and $b_{i0}\rho^i + b_i^\intercal y^i - z_i^L \geq 0$ are implied by $\overline{C}x \leq \bar{d}$. The implication follows since we will represent $b_i^\intercal W_{\cdot j}^i + b_{i0} y_j^i$ as h_{ij} and $b_{i0}\rho^i + b_i^\intercal y^i$ as z_i. This shows that one of the McCormick inequalities is implied in $\mathcal{R}_{\mathrm{QP}}$. The proof for the remaining McCormick inequalities is similar. It also follows that:

$$a_{i0} z_i + h_i^\intercal a_i = b_{i0} a_{i0} \rho^i + a_{i0} b_i^\intercal y^i + b_i^\intercal W^i a_i + b_{i0} (y^i)^\intercal a_i = b_{i0} + b_i^\intercal x,$$

where the first equality is by the definitions of z_i and h_i given above and the second equality is because $W^i a_i + a_{i0} y^i = x$ and $a_{i0}\rho^i + (y^i)^\intercal a_i = 1$. $\square$

A.15 Proof of Proposition 7

Clearly, (23) implies (22) since the matrix in the latter constraint is a principal submatrix of the matrix in (23). The following shows the reverse implication:

$$\begin{aligned} (22) &\Rightarrow \begin{pmatrix} a_{i0} & a_i^\intercal \\ 1 & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} \rho^i & (y^i)^\intercal \\ y^i & W^i \end{pmatrix} \begin{pmatrix} a_{i0} & 1 & 0 \\ a_i & 0 & I \end{pmatrix} \succeq 0, \quad \forall i \in [m] \\ &\Rightarrow \begin{pmatrix} a_{i0}^2 \rho^i + a_{i0} a_i^\intercal y^i + (a_{i0}(y^i)^\intercal + a_i^\intercal W^i) a_i & a_{i0}\rho^i + a_i^\intercal y^i & a_{i0}(y^i)^\intercal + a_i^\intercal W^i \\ a_{i0}\rho^i + a_i^\intercal y^i & \rho^i & (y^i)^\intercal \\ a_{i0} y^i + W^i a_i & y^i & W^i \end{pmatrix} \succeq 0, \forall i \in [m] \\ &\Rightarrow (23). \end{aligned}$$

The last implication follows because $a_{i0}\rho^i + a_i^\intercal y^i = 1$ and $a_{i0}(y^i)^\intercal + a_i^\intercal W^i = x^\intercal$. $\square$

References

1. Anstreicher, K.M., Burer, S.: Computable representations for convex hulls of low-dimensional quadratic forms. Mathematical Programming **124**(1-2), 33–43 (2010)
2. Arora, S., MC, P., Swarup, K.: The set covering problem with linear fractional functional. Indian Journal of Pure and Applied Mathematics **8**, 578–588 (1977)
3. Azar, Y., Cohen, E., Fiat, A., Kaplan, H., Racke, H.: Optimal oblivious routing in polynomial time. In: Proceedings of the thirty-fifth annual ACM symposium on Theory of computing, pp. 383–388 (2003)
4. Balas, E.: Disjunctive programming: Properties of the convex hull of feasible points. Discrete Applied Mathematics **89**(1-3), 3–44 (1998)
5. Barahona, F., Mahjoub, A.R.: On the cut polytope. Mathematical Programming **36**, 157–173 (1986)
6. Barros, A.I.: Discrete and fractional programming techniques for location models, vol. 3. Springer Science & Business Media (1998)
7. Beck, A., Teboulle, M.: A convex optimization approach for minimizing the ratio of indefinite quadratic functions over an ellipsoid. Mathematical Programming **118**(1), 13–35 (2009)
8. Benson, H.P.: Using concave envelopes to globally solve the nonlinear sum of ratios problem. Journal of Global Optimization **22**(1-4), 343–364 (2002)
9. Berman, O., Larson, R.C., Chiu, S.S.: Optimal server location on a network operating as an M/G/1 queue. Operations Research **33**(4), 746–771 (1985)
10. Bestuzheva, K., Chmiela, A., Müller, B., Serrano, F., Vigerske, S., Wegscheider, F.: Global optimization of mixed-integer nonlinear programs with SCIP 8. arXiv preprint arXiv:2301.00587 (2023)

11. Bezanson, J., Edelman, A., Karpinski, S., Shah, V.B.: Julia: A fresh approach to numerical computing. SIAM Review **59**(1), 65–98 (2017)
12. Blekherman, G., Parrilo, P.A., Thomas, R.R.: Semidefinite optimization and convex algebraic geometry. SIAM (2012)
13. Bogomolov, Y., Fiorini, S., Maksimenko, A., Pashkovich, K.: Small extended formulations for cyclic polytopes. Discrete & Computational Geometry **53**, 809–816 (2015)
14. Bonami, P., Günlük, O., Linderoth, J.: Globally solving nonconvex quadratic programming problems with box constraints via integer programming methods. Mathematical Programming Computation **10**(3), 333–382 (2018)
15. Boros, E., Crama, Y., Hammer, P.L.: Chvátal cuts and odd cycle inequalities in quadratic 0–1 optimization. SIAM Journal on Discrete Mathematics **5**(2), 163–177 (1992)
16. Boros, E., Hammer, P.L.: Cut-polytopes, boolean quadric polytopes and nonnegative quadratic pseudo-boolean functions. Mathematics of Operations Research **18**(1), 245–253 (1993)
17. Borrero, J.S., Gillen, C., Prokopyev, O.A.: Fractional 0–1 programming: applications and algorithms. Journal of Global Optimization **69**, 255–282 (2017)
18. Burer, S.: On the copositive representation of binary and continuous nonconvex quadratic programs. Mathematical Programming **120**(2), 479–495 (2009)
19. Burer, S.: A gentle, geometric introduction to copositive optimization. Mathematical Programming **151**(1), 89–116 (2015)
20. Burer, S., Letchford, A.N.: On nonconvex quadratic programming with box constraints. SIAM Journal on Optimization **20**(2), 1073–1089 (2009)
21. Burer, S., Yang, B.: The trust region subproblem with non-intersecting linear constraints. Mathematical Programming **149**(1-2), 253–264 (2015)
22. Busygin, S., Prokopyev, O.A., Pardalos, P.M.: Feature selection for consistent biclustering via fractional 0–1 programming. Journal of Combinatorial Optimization **10**, 7–21 (2005)
23. Charnes, A., Cooper, W.W.: Programming with linear fractional functionals. Naval Research Logistics Quarterly **9**(3-4), 181–186 (1962)
24. Chen, X., Womersley, R.S., Ye, J.J.: Minimizing the condition number of a Gram matrix. SIAM Journal on Optimization **21**(1), 127–148 (2011)
25. Chen, Y., He, T., Rong, Y., Wang, Y.: An integer programming approach for quick-commerce assortment planning. arXiv preprint arXiv:2405.02553 (2024)
26. Cheridito, P., Kromer, E.: Reward-risk ratios. Journal of Investment Strategies **3**(1) (2013)
27. De Simone, C.: The cut polytope and the boolean quadric polytope. Discrete Mathematics **79**(1), 71–75 (1990)
28. Del Pia, A., Khajavirad, A.: A polyhedral study of binary polynomial programs. Mathematics of Operations Research **42**(2), 389–410 (2017)
29. Del Pia, A., Khajavirad, A.: The multilinear polytope for acyclic hypergraphs. SIAM Journal on Optimization **28**(2), 1049–1076 (2018)
30. Del Pia, A., Khajavirad, A.: On decomposability of multilinear sets. Mathematical Programming **170**(2), 387–415 (2018)
31. Del Pia, A., Khajavirad, A.: The running intersection relaxation of the multilinear polytope. Mathematics of Operations Research **46**(3), 1008–1037 (2021)
32. Del Pia, A., Khajavirad, A.: The multilinear polytope of beta-acyclic hypergraphs has polynomial extension complexity. arXiv preprint arXiv:2212.11239 (2022)
33. Deza, M.M., Laurent, M., Weismantel, R.: Geometry of cuts and metrics, vol. 2. Springer (1997)
34. Falk, J.E., Palocsay, S.W.: Image space analysis of generalized fractional programs. Journal of Global Optimization **4**, 63–88 (1994)
35. Fawzi, H., Saunderson, J., Parrilo, P.A.: Sparse sums of squares on finite abelian groups and improved semidefinite lifts. Mathematical Programming **160**, 149–191 (2016)
36. Fiorini, S., Massar, S., Pokutta, S., Tiwary, H.R., De Wolf, R.: Linear vs. semidefinite extended formulations: exponential separation and strong lower bounds. In: Proceedings of the forty-fourth annual ACM symposium on Theory of computing, pp. 95–106 (2012)
37. Frenk, J.B., Schaible, S.: Fractional programming. Springer (2005)
38. Gale, D.: Neighborly and cyclic polytopes. In: Symposia in Pure Mathematics, vol. 7, pp. 225–232 (1963)
39. Gómez, A., Prokopyev, O.A.: A mixed-integer fractional optimization approach to best subset selection. INFORMS Journal on Computing **33**(2), 551–565 (2021)
40. Gooty, R.T., Agrawal, R., Tawarmalani, M.: Advances in MINLP to identify energy-efficient distillation configurations. Operations Research (2022)

41. Gooty, R.T., Mathew, T.J., Tawarmalani, M., Agrawal, R.: An MINLP formulation to identify thermodynamically-efficient distillation configurations. Computers & Chemical Engineering **178**, 108369 (2023)
42. Grötschel, M., Lovász, L., Schrijver, A.: Geometric algorithms and combinatorial optimization, vol. 2. Springer Science & Business Media (2012)
43. Gurobi Optimization, LLC: Gurobi Optimizer Reference Manual (2022). URL `https://www.gurobi.com`
44. Hansen, P., Poggi de Aragão, M.V., Ribeiro, C.C.: Boolean query optimization and the 0-1 hyperbolic sum problem. Annals of Mathematics and Artificial Intelligence **1**, 97–109 (1990)
45. Ho-Nguyen, N., Kilinc-Karzan, F.: A second-order cone based approach for solving the trust-region subproblem and its variants. SIAM Journal on Optimization **27**(3), 1485–1512 (2017)
46. Joyce, A., Yang, B.: Convex hull results on quadratic programs with non-intersecting constraints. Mathematical Programming pp. 1–20 (2023)
47. Jünger, M., Mallach, S.: Exact facetial odd-cycle separation for maximum cut and binary quadratic optimization. INFORMS Journal on Computing **33**(4), 1419–1430 (2021)
48. Karlin, S., Shapley, L.S., CA, R.C.S.M.: Geometry of moment spaces. Rand Corporation (1951)
49. Konno, H., Inori, M.: Bond portfolio optimization by bilinear fractional programming. Journal of the Operations Research Society of Japan **32**(2), 143–158 (1989)
50. Lanciano, T., Miyauchi, A., Fazzone, A., Bonchi, F.: A survey on the densest subgraph problem and its variants. ACM Computing Surveys (2023)
51. Li, H.L.: A global approach for general 0–1 fractional programming. European Journal of Operational Research **73**(3), 590–596 (1994)
52. Lubin, M., Dowson, O., Garcia, J.D., Huchette, J., Legat, B., Vielma, J.P.: JuMP 1.0: recent improvements to a modeling language for mathematical optimization. Mathematical Programming Computation pp. 1–9 (2023)
53. Luce, R.D.: Individual choice behavior: A theoretical analysis. Courier Corporation (2012)
54. Matsui, T.: NP-hardness of linear multiplicative programming and related problems. Journal of Global Optimization **9**, 113–119 (1996)
55. McCormick, G.P.: Computability of global solutions to factorable nonconvex programs: Part i - Convex underestimating problems. Mathematical Programming **10**(1), 147–175 (1976)
56. Megiddo, N.: Combinatorial optimization with rational objective functions. Mathematics of Operations Research **4**(4), 414–424 (1979)
57. Mehmanchi, E., Gómez, A., Prokopyev, O.A.: Fractional 0–1 programs: links between mixed-integer linear and conic quadratic formulations. Journal of Global Optimization **75**, 273–339 (2019)
58. Mehmanchi, E., Gómez, A., Prokopyev, O.A.: Solving a class of feature selection problems via fractional 0–1 programming. Annals of Operations Research **303**(1-2), 265–295 (2021)
59. Méndez-Díaz, I., Miranda-Bront, J.J., Vulcano, G., Zabala, P.: A branch-and-cut algorithm for the latent-class logit assortment problem. Discrete Applied Mathematics **164**, 246–263 (2014)
60. Misener, R., Floudas, C.A.: ANTIGONE: algorithms for continuous/integer global optimization of nonlinear equations. Journal of Global Optimization **59**(2-3), 503–526 (2014)
61. Nguyen, H., Franke, K., Petrovic, S.: Optimizing a class of feature selection measures. In: NIPS 2009 Workshop on Discrete Optimization in Machine Learning: Submodularity, Sparsity & Polyhedra (DISCML), Vancouver, Canada (2009)
62. Padberg, M.: The boolean quadric polytope: some characteristics, facets and relatives. Mathematical programming **45**(1), 139–172 (1989)
63. Prokopyev, O.A., Kong, N., Martinez-Torres, D.L.: The equitable dispersion problem. European Journal of Operational Research **197**(1), 59–67 (2009)
64. Quesada, I., Grossmann, I.E.: A global optimization algorithm for linear fractional and bilinear programs. Journal of Global Optimization **6**, 39–76 (1995)
65. Rockafellar, R.T.: Convex analysis. Princeton university press (2015)
66. Schmüdgen, K.: The moment problem, vol. 277. Springer (2017)
67. Schrijver, A.: Short proofs on the matching polyhedron. Journal of Combinatorial Theory, Series B **34**(1), 104–108 (1983)
68. Sen, A., Atamtürk, A., Kaminsky, P.: A conic integer optimization approach to the constrained assortment problem under the mixed multinomial logit model. Operations Research **66**(4), 994–1003 (2018)

69. Shen, K., Yu, W.: Fractional programming for communication systems—part i: Power control and beamforming. IEEE Transactions on Signal Processing **66**(10), 2616–2630 (2018)
70. Sherali, H.D., Adams, W.P.: 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)
71. Stancu-Minisian, I.M.: Fractional programming: theory, methods and applications, vol. 409. Springer Science & Business Media (2012)
72. Sturm, J.F., Zhang, S.: On cones of nonnegative quadratic functions. Mathematics of Operations research **28**(2), 246–267 (2003)
73. Tawarmalani, M.: Inclusion certificates and simultaneous convexification of functions (2010). Working paper
74. Tawarmalani, M., Ahmed, S., Sahinidis, N.V.: Global optimization of 0-1 hyperbolic programs. Journal of Global Optimization **24**(4), 385–416 (2002)
75. Tawarmalani, M., Richard, J.P.P., Xiong, C.: Explicit convex and concave envelopes through polyhedral subdivisions. Mathematical Programming **138**(1-2), 531–577 (2013)
76. Tawarmalani, M., Sahinidis, N.V.: Semidefinite relaxations of fractional programs via novel convexification techniques. Journal of Global Optimization **20**(2), 133–154 (2001)
77. Tawarmalani, M., Sahinidis, N.V.: Global optimization of mixed-integer nonlinear programs: A theoretical and computational study. Mathematical programming **99**(3), 563–591 (2004)
78. Tawarmalani, M., Sahinidis, N.V.: A polyhedral branch-and-cut approach to global optimization. Mathematical Programming **103**(2), 225–249 (2005)
79. Underwood, A.: Fractional distillation of multicomponent mixtures. Industrial & Engineering Chemistry **41**(12), 2844–2847 (1949)
80. Xinying Chen, V., Hooker, J.N.: A guide to formulating fairness in an optimization model. Annals of Operations Research **326**(1), 581–619 (2023)
81. Yang, B., Anstreicher, K., Burer, S.: Quadratic programs with hollows. Mathematical Programming **170**, 541–553 (2018)