Mathematics of Operations Research

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Exploiting Sign Symmetries in Minimizing Sums of Rational Functions

Feng Guo, Jie Wang, Jianhao Zheng

https://pubsonline.informs.org/journal/moor

MATHEMATICS OF OPERATIONS RESEARCH
Articles in Advance, pp. 1–20
ISSN 0364-765X (print), ISSN 1526-5471 (online)

Exploiting Sign Symmetries in Minimizing Sums of Rational Functions

Feng Guo,[a] Jie Wang,[b,*] Jianhao Zheng[a]

[a] School of Mathematical Sciences, Dalian University of Technology, Dalian, Liaoning Province 116024, China; [b] State Key Laboratory of Mathematical Sciences, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

*Corresponding author

Contact: fguo@dlut.edu.cn, https://orcid.org/0000-0003-3106-2653 (FG); wangjie212@amss.ac.cn, https://orcid.org/0000-0002-9681-1451 (JW); jhzheng2025@163.com (JZ)

Received: May 25, 2024
Revised: April 5, 2025
Accepted: October 5, 2025
Published Online in Articles in Advance: December 1, 2025

MSC2020 Subject Classifications: Primary: 90C23, 90C32; Secondary: 90C22, 90C26

https://doi.org/10.1287/moor.2024.0541

Abstract. This paper is devoted to the problem of minimizing a sum of rational functions over a basic semialgebraic set. We provide a hierarchy of sum-of-squares (SOS) relaxations that is dual to the generalized moment problem approach proposed by Bugarin, Henrion, and Lasserre. The investigation of the dual SOS aspect offers two benefits: (1) it allows us to conduct a convergence rate analysis for the hierarchy; (2) it leads to a sign symmetry–adapted hierarchy consisting of block-diagonal semidefinite relaxations. When the problem possesses correlative sparsity as well as sign symmetries, we propose sparse semidefinite relaxations by exploiting both structures. Various numerical experiments are performed to demonstrate the efficiency of our approach. Finally, an application to maximizing sums of generalized Rayleigh quotients is presented.

Funding: F. Guo was supported by the Chinese National Natural Science Foundation [Grant 12471478]. J. Wang was funded by the National Key R&D Program of China [Grant 2023YFA1009401] and the Natural Science Foundation of China [Grants 12201618 and 12171324].

Keywords: sum of rational functions • sign symmetry • semidefinite relaxation • correlative sparsity • generalized Rayleigh quotient

1. Introduction

In this paper, we consider the optimization problem of minimizing a sum of rational functions:

$$\rho := \inf_{\mathbf{x}\in\mathbf{K}} \sum_{i=1}^{N} \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}, \tag{SRFO}$$

where

$$\mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n \,|\, g_j(\mathbf{x}) \ge 0,\ j = 1, \ldots, m\}, \tag{1}$$

and p_i, q_i, g_j are polynomials in variables $\mathbf{x} = (x_1, \ldots, x_n)$. Throughout the paper, we make the following assumption on (SRFO):

Assumption 1. (i) $\mathbf{K}$ *is compact;* (ii) $q_i > 0$ *on* $\mathbf{K}$ *for* $i = 1, \ldots, N$.

The Problem (SRFO) has applications in various fields, including computer vision (Hartley et al. [10], Kahl et al. [12]), multiuser multiple input, multiple output (MIMO) systems (Primolevo et al. [24]), and sparse Fisher discriminant analysis in pattern recognition (Dundar et al. [7], Fung and Ng [8], Wu et al. [38]).

Given the potentially large number of terms N and the absence of convexity (respectively, concavity) assumptions on p_i (respectively, q_i), globally solving (SRFO) or achieving a close approximation of the optimum ρ presents significant challenges. In literature, (SRFO) is often called a *sum-of-ratios program*, which stands as one of the most complex fractional programs encountered to date. For strategies of tackling sum-of-ratios programs under specific assumptions regarding concavity and linearity of ratios, the reader may refer to the survey (Schaible and Shi [27] and Wu et al. [37, table 1]).

In the scenario where $N = 1$, by utilizing the polynomial structure in (SRFO) and Positivstellensätze from real algebraic geometry, some hierarchies of semidefinite programming (SDP) relaxations were proposed in Jibetean and de Klerk [11] and Nie et al. [22]. The basic idea is to maximize a number $\gamma \in \mathbb{R}$ subject to the nonnegativity of $p_1 - \gamma q_1$ on $\mathbf{K}$, which can be ensured by certain sum-of-squares (SOS) representations. In Guo et al. [9], the problem of minimizing a rational function was reformulated as a polynomial optimization problem and solved by the exact Jacobian SDP relaxation method proposed by Nie [20].

For the case that $N > 1$, one could attempt to reduce all rational functions p_i/q_i to the same denominator and apply the hierarchies of SDP relaxations mentioned above. However, because of the potentially large value of N, the resulting unified denominator may have a significantly high degree, which makes it impractical to even solve the first order relaxations of the hierarchies. To overcome this difficulty, Bugarin et al. [6] reformulated (SRFO) as an equivalent infinite-dimensional linear program, which is a particular instance of the generalized moment problem (GMP) with N unknown measures. The GMP can be relaxed into a hierarchy of SDPs, which provides increasingly tight lower bounds on ρ. When a correlative sparsity pattern is present in the polynomial data of (SRFO), a sparse GMP reformulation for (SRFO) and a corresponding hierarchy of sparse SDP relaxations were also proposed in Bugarin et al. [6]. By employing pushforward measures, Lasserre et al. [15] gave an approach yielding convergent upper bounds on ρ.

1.1. Contributions

Our main contributions are summarized as follows:

1. By studying the Lagrange dual problem, we unveil that the core strategy of the GMP approach is to replace the terms p_i/q_i, $i = 2, \ldots, N$ by polynomial approximations from below on $\mathbf{K}$, leading to a problem with a single denominator. As a result, we provide a hierarchy of SOS relaxations that is dual to the GMP approach proposed by Bugarin, Henrion, and Lasserre. Moreover, we are able to conduct a convergence rate analysis for the hierarchy of SDP relaxations for (SRFO).

2. We present a sign symmetry–adapted hierarchy consisting of block-diagonal SDP relaxations for (SRFO). Furthermore, when both correlative sparsity and sign symmetries are present in the polynomial data of (SRFO), we show that the exploitation of sign symmetries can be naturally incorporated into the sparse hierarchy of SDP relaxations for (SRFO), thereby further reducing the computational cost. Various numerical experiments demonstrate the efficiency of our approach.

The rest of the paper is organized as follows. We first recall some preliminaries and the GMP approach for (SRFO) in Section 2. In Section 3, we investigate the dual aspect of the GMP approach and conduct convergence rate analysis for the SDP hierarchy. In Section 4, we develop sparse SDP relaxations for (SRFO) by exploiting sign symmetries as well as correlative sparsity present in the problem data. Numerical experiments are presented in Section 5, and an application to maximizing sums of generalized Rayleigh quotients is provided in Section 6.

2. Notation and Preliminaries

Let $\mathbb{N}$ denote the set of nonnegative integers. For $n \in \mathbb{N} \setminus \{0\}$, let $[n] := \{1, 2, \ldots, n\}$. For $k \in \mathbb{N}$, let $\mathbb{N}^n_k := \{\boldsymbol{\alpha} = (\alpha_i) \in \mathbb{N}^n \mid \sum_{i=1}^n \alpha_i \leq k\}$. We use $|\cdot|$ to denote the cardinality of a set. Let $\mathbb{R}[\mathbf{x}] := \mathbb{R}[x_1, \ldots, x_n]$ be the ring of multivariate polynomials in n variables $\mathbf{x}$ and $\mathbb{R}[\mathbf{x}]_k$ denote the subset of polynomials of degree no greater than k. A polynomial $f \in \mathbb{R}[\mathbf{x}]$ can be written as $f = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$ with $f_{\boldsymbol{\alpha}} \in \mathbb{R}$ and $\mathbf{x}^{\boldsymbol{\alpha}} := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. The support of f is defined by $\operatorname{supp}(f) := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid f_{\boldsymbol{\alpha}} \neq 0\}$. For $\mathcal{A} \subseteq \mathbb{N}^n$, $\mathbb{R}[\mathcal{A}]$ denotes the set of polynomials with supports contained in $\mathcal{A}$; that is, $\mathbb{R}[\mathcal{A}] := \{f \in \mathbb{R}[\mathbf{x}] \mid \operatorname{supp}(f) \subseteq \mathcal{A}\}$. For $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$, $\|\mathbf{u}\|$ denotes the standard Euclidean norm of $\mathbf{u}$. For $t \in \mathbb{R}$, we use $\lceil t \rceil$ to denote the smallest integer that is not smaller than t. We use $A \succeq 0$ to indicate that the matrix A is positive semidefinite. For two matrices A, B of the same size, let $A \circ B$ denote the Hadamard product, defined by $[A \circ B]_{ij} = A_{ij} B_{ij}$.

2.1. Sums of Squares and Moments

We recall some background about SOS polynomials and the dual theory of *moment matrices*. A polynomial $f(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ is said to be a *sum of squares* if it can be written as $f(\mathbf{x}) = \sum_{i=1}^t f_i(\mathbf{x})^2$ for some $f_1(\mathbf{x}), \ldots, f_t(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$. Let $\Sigma[\mathbf{x}]$ denote the set of SOS polynomials in $\mathbb{R}[\mathbf{x}]$.

Let $\mathbf{g} := \{g_1, \ldots, g_m\}$ be the set of polynomials that defines the semialgebraic set $\mathbf{K}$ in (1). We denote by

$$Q(\mathbf{g}) := \left\{ \sigma_0 + \sum_{j=1}^m \sigma_j g_j \;\middle|\; \sigma_j \in \Sigma[\mathbf{x}], j \in \{0\} \cup [m] \right\}$$

the *quadratic module* generated by $\mathbf{g}$ and denote by

$$Q_k(\mathbf{g}) := \left\{ \sigma_0 + \sum_{j=1}^m \sigma_j g_j \;\middle|\; \sigma_0, \sigma_j \in \Sigma[\mathbf{x}], \deg(\sigma_0), \deg(\sigma_j g_j) \leq 2k, j \in [m] \right\}$$

the k-th *truncated quadratic module*. It is clear that if $f \in Q(\mathbf{g})$, then $f(\mathbf{x}) \geq 0$ for any $\mathbf{x} \in \mathbf{K}$, though the converse is not necessarily true.

Given a (pseudomoment) sequence of real numbers $\mathbf{y} := (y_{\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^n}$, the k-th order *moment matrix* is the matrix $\mathbf{M}_k(\mathbf{y})$ indexed by $\mathbb{N}^n_k$, with the $(\boldsymbol{\alpha}, \boldsymbol{\beta})$-th entry being $y_{\boldsymbol{\alpha}+\boldsymbol{\beta}}$. Given a polynomial $f(\mathbf{x}) = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}$, the k-th order *localizing matrix* $\mathbf{M}_k(f\mathbf{y})$ indexed by $\mathbb{N}^n_k$ is defined by $[\mathbf{M}_k(f\mathbf{y})]_{\boldsymbol{\beta},\boldsymbol{\gamma}} = \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}+\boldsymbol{\beta}+\boldsymbol{\gamma}}$. The Riesz functional $L_{\mathbf{y}}$ on $\mathbb{R}[\mathbf{x}]$ is defined by $L_{\mathbf{y}}(\sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}}) := \sum_{\boldsymbol{\alpha}} f_{\boldsymbol{\alpha}} y_{\boldsymbol{\alpha}}$.

Definition 1 (Archimedean Condition). We say that $Q(\mathbf{g})$ is *Archimedean* if there exists $M > 0$ such that $M - x_1^2 - \cdots - x_n^2 \in Q(\mathbf{g})$.

Note that the Archimedean condition implies that $\mathbf{K}$ is compact and the converse is not necessarily true. However, for any compact set $\mathbf{K}$, we could always force the associated quadratic module to be Archimedean by adding a redundant constraint $M - x_1^2 - \cdots - x_n^2 \geq 0$ in the description of $\mathbf{K}$ for sufficiently large M.

Theorem 1. (Putinar [25, Putinar's Positivstellensatz]). *Suppose that $Q(\mathbf{g})$ is Archimedean. If a polynomial $f \in \mathbb{R}[\mathbf{x}]$ is positive on $\mathbf{K}$, then $f \in Q_k(\mathbf{g})$ for some $k \in \mathbb{N}$.*

2.2. The GMP Reformulation and SDP Relaxations

In Bugarin et al. [6], they reformulated (SRFO) as the following GMP:

$$\begin{cases} \inf\limits_{\mu_i \in \mathcal{M}(\mathbf{K})_+} & \sum_{i=1}^N \int_{\mathbf{K}} p_i \mathrm{d}\mu_i \\ \text{s.t.} & \int_{\mathbf{K}} q_1 \mathrm{d}\mu_1 = 1, \\ & \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_i \mathrm{d}\mu_i = \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_1 \mathrm{d}\mu_1, \quad \forall \boldsymbol{\alpha} \in \mathbb{N}^n, i \in [N] \setminus \{1\}, \end{cases} \tag{P}$$

where $\mathcal{M}(\mathbf{K})_+$ denotes the set of finite positive Borel measures supported on $\mathbf{K}$.

Let

$$d_j := \lceil \deg(g_j)/2 \rceil, \quad j \in [m],$$
$$d_{\min} := \max\{\lceil \deg(p_i)/2 \rceil, \lceil \deg(q_i)/2 \rceil, i \in [N]; d_j, j \in [m]\}.$$

Based on the Reformulation (P), Bugarin et al. further proposed the following hierarchy of SDP relaxations for (SRFO) ($k \geq d_{\min}$):

$$\begin{cases} \inf\limits_{\mathbf{y}_i} & \sum_{i=1}^N L_{\mathbf{y}_i}(p_i) \\ \text{s.t.} & \mathbf{M}_k(\mathbf{y}_i) \succeq 0, \quad i \in [N], \\ & \mathbf{M}_{k-d_j}(g_j \mathbf{y}_i) \succeq 0, \quad i \in [N], j \in [m], \\ & L_{\mathbf{y}_1}(q_1) = 1, \\ & L_{\mathbf{y}_i}(\mathbf{x}^{\boldsymbol{\alpha}} q_i) = L_{\mathbf{y}_1}(\mathbf{x}^{\boldsymbol{\alpha}} q_1), \quad \forall \boldsymbol{\alpha} \in \mathbb{N}^n_{2k - \max\{\deg(q_1), \deg(q_i)\}}, i \in [N] \setminus \{1\}. \end{cases} \tag{Pk}$$

It was shown in Bugarin et al. [6] that under Assumption 1 and the Archimedean condition, the sequence of optima of (Pk) converges to ρ as $k \to \infty$.

3. The Dual Perspective and Convergence Rate Analysis

In this section, we will unveil the underlying principle of the GMP Reformulation (P) for (SRFO) from the dual perspective, which enables us to achieve a convergence rate analysis of the hierarchy of SDP relaxations for (SRFO).

3.1. The Dual Perspective

Let us derive the Lagrange dual problem of the GMP Reformulation (P) for (SRFO). Note that there are infinitely many constraints involved in (P). To formulate the dual problem, we need to embed these constraints into an appropriate functional space paired with a dual space (Shapiro [29]). Let $\mathcal{Y}$ be the space of all functions $\omega : \mathbb{N}^n \to \mathbb{R}$

equipped with natural algebraic operations of addition and multiplication by a scalar. We associate this space with the dual space $\mathcal{Y}^*$ consisting of functions $\omega^*: \mathbb{N}^n \to \mathbb{R}$ such that only a finite number of values $\omega^*(\boldsymbol{\alpha}), \boldsymbol{\alpha} \in \mathbb{N}^n$ are nonzero. For $\omega \in \mathcal{Y}$ and $\omega^* \in \mathcal{Y}^*$, define the scalar product

$$\langle \omega^*, \omega \rangle := \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} \omega^*(\boldsymbol{\alpha})\omega(\boldsymbol{\alpha}),$$

where the summation is performed over $\boldsymbol{\alpha}$ in the finite support set of ω^*. Equivalently, we can take $\mathcal{Y}^*$ to be the polynomial ring $\mathbb{R}[\mathbf{x}]$. For any $h(\mathbf{x}) = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} h_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]$ and $\omega \in \mathcal{Y}$, define the scalar product

$$\langle h, \omega \rangle := \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} h_{\boldsymbol{\alpha}} \omega(\boldsymbol{\alpha}).$$

Clearly, for each $i \in [N] \setminus \{1\}$, the mapping $\omega_i : \mathbb{N}^n \to \mathbb{R}$ defined by

$$\omega_i(\boldsymbol{\alpha}) := \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_i \mathrm{d}\mu_i - \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_1 \mathrm{d}\mu_1, \quad \forall \boldsymbol{\alpha} \in \mathbb{N}^n,$$

belongs to the space $\mathcal{Y}$, and then the second part of constraints of (P) can be expressed as $\omega_i = 0, i \in [N] \setminus \{1\}$. Let $h_i = \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} h^i_{\boldsymbol{\alpha}} \mathbf{x}^{\boldsymbol{\alpha}} \in \mathbb{R}[\mathbf{x}]$ be the dual variable associated with the constraint $\omega_i = 0, i \in [N] \setminus \{1\}$, and let c be the dual variable associated with the constraint $\int_{\mathbf{K}} q_1 \mathrm{d}\mu_1 = 1$. The Lagrangian of (P) now can be written as

$$\begin{aligned}
\mathcal{L}(\{\mu_i\}, \{h_i\}, c) &:= \sum_{i=1}^{N} \int_{\mathbf{K}} p_i \mathrm{d}\mu_i - c\left(\int_{\mathbf{K}} q_1 \mathrm{d}\mu_1 - 1\right) - \sum_{i=2}^{N} \langle h_i, \omega_i \rangle \\
&= c + \sum_{i=1}^{N} \int_{\mathbf{K}} p_i \mathrm{d}\mu_i - \int_{\mathbf{K}} c q_1 \mathrm{d}\mu_1 - \sum_{i=2}^{N} \sum_{\boldsymbol{\alpha} \in \mathbb{N}^n} h^i_{\boldsymbol{\alpha}} \left(\int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_i \mathrm{d}\mu_i - \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_1 \mathrm{d}\mu_1\right) \\
&= c + \int_{\mathbf{K}} \left(p_1 - c q_1 + q_1 \sum_{i=2}^{N} h_i\right) \mathrm{d}\mu_1 + \sum_{i=2}^{N} \int_{\mathbf{K}} (p_i - q_i h_i) \mathrm{d}\mu_i.
\end{aligned}$$

The Lagrange dual function of (P) is

$$\inf_{\mu_i \in \mathcal{M}(\mathbf{K})_+} \mathcal{L}(\{\mu_i\}, \{h_i\}, c) = \begin{cases} c, & \text{if } p_1 - q_1(c - \sum_{i=2}^{N} h_i) \geq 0, \ p_i - q_i h_i \geq 0 \text{ on } \mathbf{K}, \\ -\infty, & \text{otherwise.} \end{cases}$$

Hence, the Lagrange dual problem of (P) is

$$\begin{cases} \sup\limits_{c, h_i} & c \\ \text{s.t.} & p_1(\mathbf{x}) + \left(\sum\limits_{i=2}^{N} h_i(\mathbf{x}) - c\right) q_1(\mathbf{x}) \geq 0, \quad \forall \mathbf{x} \in \mathbf{K}, \\ & p_i(\mathbf{x}) - h_i(\mathbf{x}) q_i(\mathbf{x}) \geq 0, \quad \forall \mathbf{x} \in \mathbf{K}, i \in [N] \setminus \{1\}, \\ & h_i \in \mathbb{R}[\mathbf{x}], \quad i \in [N] \setminus \{1\}. \end{cases} \tag{D}$$

For each $i \in [N] \setminus \{1\}$, define

$$\mathcal{H}_i := \left\{ h \in \mathbb{R}[\mathbf{x}] \,\middle|\, \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} \geq h(\mathbf{x}), \ \forall \mathbf{x} \in \mathbf{K} \right\}.$$

Because each $q_i > 0$ on $\mathbf{K}$, the Lagrange dual problem of (P) can be also expressed as

$$\sup_{h_i \in \mathcal{H}_i} \inf_{\mathbf{x} \in \mathbf{K}} \frac{p_1(\mathbf{x})}{q_1(\mathbf{x})} + \sum_{i=2}^{N} h_i(\mathbf{x}).$$

Consequently, the underlying principle of the GMP reformulation (P) for (SRFO), derived from the dual aspect, can be unfolded as follows: (1) replacing the terms $\frac{p_i}{q_i}$, $i \in [N] \setminus \{1\}$ by polynomial approximations $h_i \in \mathbb{R}[\mathbf{x}]$ from below on $\mathbf{K}$; (2) computing the infimum of the resulting function, which contains only a single denominator; and (3) letting the approximations h_i's vary and taking the supremum.

Theorem 2. *Under Assumption 1, the optimum of (D) equals ρ.*

Proof. Denote the optimum of (D) by τ. Clearly, we have $\tau \le \rho$. Fix an arbitrary $\varepsilon > 0$. By the Stone–Weierstrass theorem, for each $i \in [N] \setminus \{1\}$, there exists $\hat{h}_i \in \mathbb{R}[\mathbf{x}]$ such that

$$\sup_{\mathbf{x}\in\mathbf{K}} \left|\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \hat{h}_i(\mathbf{x})\right| \le \frac{\varepsilon}{2(N-1)}.$$

Letting $h_i := \hat{h}_i - \varepsilon/2(N-1)$, we have

$$\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \frac{\varepsilon}{N-1} \le h_i(\mathbf{x}) \le \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}, \quad \forall \mathbf{x} \in \mathbf{K}.$$

Hence, $h_i \in \mathcal{H}_i$, and

$$\tau \ge \inf_{\mathbf{x}\in\mathbf{K}} \frac{p_1(\mathbf{x})}{q_1(\mathbf{x})} + \sum_{i=2}^{N} h_i(\mathbf{x}) \ge \inf_{\mathbf{x}\in\mathbf{K}} \sum_{i=1}^{N} \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \varepsilon \ge \rho - \varepsilon.$$

Thus, $(\rho - \varepsilon, h_2, \ldots, h_N)$ is feasible to (D). As ε is arbitrary, we obtain $\tau = \rho$. □

Remark 1. Theorem 2 can also be derived from strong duality between (P) and (D). In fact, as $\mathbf{K}$ is compact, there exist *constant* polynomials $\hat{h}_i$'s and $\hat{c} \in \mathbb{R}$ such that $p_i/q_i > \hat{h}_i$, $i \in [N] \setminus \{1\}$, and $p_1/q_1 + \sum_{i=2}^{N} \hat{h}_i > \hat{c}$ on $\mathbf{K}$. That is, $(\hat{c}, \hat{h}_2, \ldots, \hat{h}_N)$ is a strictly feasible solution of (D), and hence, there is no dual gap between (P) and (D).

By truncating the polynomial degree, we obtain the following hierarchy of dual SOS relaxations for (SRFO) ($k \ge d_{\min}$):

$$\rho_k := \begin{cases} \sup\limits_{c, h_i} & c \\ \text{s.t.} & p_1(\mathbf{x}) + \left(\sum\limits_{i=2}^{N} h_i(\mathbf{x}) - c\right) q_1(\mathbf{x}) \in Q_k(\mathbf{g}), \\ & p_i(\mathbf{x}) - h_i(\mathbf{x}) q_i(\mathbf{x}) \in Q_k(\mathbf{g}), \quad i \in [N] \setminus \{1\}, \\ & h_i \in \mathbb{R}[\mathbf{x}]_{2k - \max\{\deg(q_1), \deg(q_i)\}}, \quad i \in [N] \setminus \{1\}. \end{cases} \tag{Dk}$$

Theorem 3. *Under Assumption 1 and the Archimedean condition, it holds $\rho_k \nearrow \rho$ as $k \to \infty$.*

Proof. It is obvious that $\rho_k \le \rho_{k+1} \le \rho$ for all $k \ge d_{\min}$. Fix an arbitrary $\varepsilon > 0$, and we will prove that there exists $k_\varepsilon \in \mathbb{N}$ such that $\rho_{k_{\bar{\varepsilon}}} \ge \rho - \varepsilon$. For $\bar{\varepsilon} := \varepsilon/3$, let $(\rho - \bar{\varepsilon}, \bar{h}_2, \ldots, \bar{h}_N)$ be the feasible solution to (D) provided in the proof of Theorem 2. For each $i = 2, \ldots, N$, let

$$h_i := \bar{h}_i - \frac{\bar{\varepsilon}}{N-1}.$$

Then, for $\mathbf{x} \in \mathbf{K}$, it holds that

$$p_i(\mathbf{x}) - h_i(\mathbf{x}) q_i(\mathbf{x}) = p_i(\mathbf{x}) - \bar{h}_i(\mathbf{x}) q_i(\mathbf{x}) + \frac{\bar{\varepsilon}}{N-1} q_i(\mathbf{x}) \ge \frac{\bar{\varepsilon}}{N-1} q_i(\mathbf{x}) > 0. \tag{2}$$

By Theorem 1, there exists $k_i \in \mathbb{N}$ such that $h_i \in \mathbb{R}[\mathbf{x}]_{2k_i - \deg(q_i)}$ and $p_i - h_i q_i \in Q_{k_i}(\mathbf{g})$. Moreover, for $\mathbf{x} \in \mathbf{K}$, we have

$$\begin{aligned} & p_1(\mathbf{x}) + \left(\sum_{i=2}^{N} h_i(\mathbf{x}) - (\rho - \varepsilon)\right) q_1(\mathbf{x}) \\ &= p_1(\mathbf{x}) + q_1(\mathbf{x}) \sum_{i=2}^{N} \bar{h}_i(\mathbf{x}) - \bar{\varepsilon} q_1(\mathbf{x}) - \rho q_1(\mathbf{x}) + 3\bar{\varepsilon} q_1(\mathbf{x}) \\ &\ge (\rho - \bar{\varepsilon}) q_1(\mathbf{x}) - \rho q_1(\mathbf{x}) + 2\bar{\varepsilon} q_1(\mathbf{x}) \ge \bar{\varepsilon} q_1(\mathbf{x}) > 0. \end{aligned}$$

By Theorem 1 again, there exists $k_1 \in \mathbb{N}$ such that

$$p_1(\mathbf{x}) + \left(\sum_{i=2}^{N} h_i(\mathbf{x}) - (\rho - \varepsilon)\right) q_1(\mathbf{x}) \in Q_{k_1}(\mathbf{g}).$$

Let $k_\varepsilon := \max_{1 \le i \le N} k_i$. Then, $(\rho - \varepsilon, h_2, \ldots, h_N)$ is feasible to (Dk) with $k = k_\varepsilon$, which implies $\rho_{k_\varepsilon} \ge \rho - \varepsilon$. □

3.2. Convergence Rate Analysis

The exploration of the dual aspect of the GMP Reformulation (P) allows us to perform a convergence rate analysis for the hierarchy of SDP relaxations for (SRFO) by utilizing some existing results from the literature.

Lemma 1. *For each $i \in [N] \setminus \{1\}$, there exists a constant c_i depending on p_i, q_i, and $\mathbf{K}$ such that for any $k \in \mathbb{N}$, there exists $h \in \mathbb{R}[\mathbf{x}]_k$ satisfying*

$$\sup_{\mathbf{x}\in\mathbf{K}} \left| \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - h(\mathbf{x}) \right| \le \frac{c_i}{k}.$$

Proof. Denote by

$$\omega_{\mathbf{K}}\left(\frac{p_i}{q_i}, t\right) := \sup\left\{ \left| \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \frac{p_i(\mathbf{y})}{q_i(\mathbf{y})} \right| : \mathbf{x}, \mathbf{y} \in \mathbf{K}, \|\mathbf{x}-\mathbf{y}\| \le t \right\}$$

the standard modulus of continuity of p_i/q_i on $\mathbf{K}$. By the multivariate version of Jackson's theorem (see Timan [32]), there exists a constant $\hat{c}_i$ depending on p_i, q_i, and $\mathbf{K}$ such that for any $k \in \mathbb{N}$, there exists $h \in \mathbb{R}[\mathbf{x}]_k$ satisfying

$$\sup_{\mathbf{x}\in\mathbf{K}} \left| \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - h(\mathbf{x}) \right| \le \hat{c}_i \omega_{\mathbf{K}}\left(\frac{p_i}{q_i}, \frac{1}{k}\right).$$

As $q_i(\mathbf{x}) > 0$ on $\mathbf{K}$, p_i/q_i is Lipschitz on $\mathbf{K}$. So there is a constant L_i such that $\omega_{\mathbf{K}}(p_i/q_i, t) \le L_i t$. Then, the conclusion follows by letting $c_i := \hat{c}_i L_i$. □

For a polynomial $\phi \in \mathbb{R}[\mathbf{x}]$, denote by $\|\phi\|$ the max norm of ϕ on $[-1,1]^n$; that is, $\|\phi\| = \max_{\mathbf{x}\in[-1,1]^n} |\phi(\mathbf{x})|$.

Assumption 2. (i) *Normalized Archimedean condition: $1 - x_1^2 - \cdots - x_n^2 \in Q(\mathbf{g})$;* (ii) *for all $j \in [m]$, $\|g_j\| \le \frac{1}{2}$; and* (iii) *the origin belongs to the interior of* $\mathbf{K}$.

As we always assume that $\mathbf{K}$ is compact, Assumption 2(i) can be satisfied by first rescaling $\mathbf{K}$ (if necessary) and then adding the redundant inequality constraint $1 - x_1^2 - \cdots - x_n^2 \ge 0$ to the description of $\mathbf{K}$. Similarly, Assumption 2(ii) can be fulfilled by appropriately rescaling each g_j, $j \in [m]$.

Recently, significant progress has been made in understanding the complexity of Putinar's Positivstellensatz and analyzing the convergence rates of various SOS hierarchies; see Baldi [3], Baldi and Mourrain [4], Korda and Henrion [13], Korda et al. [14], Nie and Schweighofer [21], Slot [31], and the references therein. To establish the convergence rate analysis of the hierarchy of SDP relaxations for (SRFO), we leverage the effective version of Putinar's Positivstellensatz presented in Baldi and Mourrain [4]. We first recall the constant and exponent of Łojasiewicz inequalities associated with the semialgebraic set $\mathbf{K}$.

Theorem 4. (Baldi and Mourrain [4], Bochnak et al. [5]). *For $\mathbf{x} \in [-1,1]^n$, let*

$$G(\mathbf{x}) := |\min\{g_1(\mathbf{x}), \ldots, g_m(\mathbf{x}), 0\}|, \quad \operatorname{dist}(\mathbf{x}, \mathbf{K}) := \min\{\|\mathbf{x} - \mathbf{x}'\| : \mathbf{x}' \in \mathbf{K}\}.$$

There exist positive real numbers $\mathfrak{c}$ and $Ł$ such that for $\mathbf{x} \in [-1,1]^n$,

$$\operatorname{dist}(\mathbf{x}, \mathbf{K})^{L} \le \mathfrak{c}\, G(\mathbf{x}).$$

Theorem 5 (Baldi and Mourrain [4, Theorem 1.7]). *Assume that $n \ge 2$ and Assumption 2,* (i)–(ii) *hold. Let $\mathfrak{c}$ and L be the Łojasiewicz coefficient and exponent given in Theorem 4. If $\phi(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ is strictly positive on $\mathbf{K}$, then $\phi \in Q_k(\mathbf{g})$ whenever*

$$k \ge \gamma(n, \mathbf{g}) \deg(\phi)^{3.5nL} \left(\frac{\|\phi\|}{\min_{\mathbf{x}\in\mathbf{K}} \phi(\mathbf{x})} \right)^{2.5nL},$$

for some constant $\gamma(n, \mathbf{g})$ depending only on n and g_j's.

To bound the norm $\|\phi\|$, we need the following result.

Proposition 1 (Schlosser et al. [28, Lemma A.1]). *Assume that $\mathbf{K} \subset [-1,1]^n$ and Assumption 2*(iii) *holds. If $\phi(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ is nonnegative on $\mathbf{K}$, then*

$$\|\phi\| \le \left(1 + \frac{\deg(\phi)^2}{4} \left(\frac{2}{t}\right)^{\deg(\phi)+1} \right) \max_{\mathbf{x}\in\mathbf{K}} \phi(\mathbf{x}),$$

where $t \in (0,1)$ is such that $[-t,t]^n \subset \mathbf{K}$.

Let

$$\kappa := \frac{1}{\sup\{t > 0 : [-t, t]^n \subseteq \mathbf{K}\}},$$

and

$$q^{\max} := \max_{2\le i\le N, \mathbf{x}\in\mathbf{K}} q_i(\mathbf{x}),\ q^{\min} := \min_{2\le i\le N, \mathbf{x}\in\mathbf{K}} q_i(\mathbf{x}),\ \rho^{\max} := \max_{\mathbf{x}\in\mathbf{K}} \sum_{i=1}^{N} \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}.$$

Using Theorem 5, we can establish the following convergence rate of the hierarchy (Dk).

Theorem 6. *Assume that $N\ge 2, n\ge 2$; Assumptions 1 and 2 hold. Then, there exist constants C_1 and C_2 depending on n, p_i's, q_i's, g_j's, and $\mathbf{K}$ such that for any $\varepsilon>0$, we have $\rho_k \ge \rho-\varepsilon$ whenever*

$$\begin{aligned} k &\ge C_2 D(N,\varepsilon)^{3.5nL}\left(\left(1+\frac{D(N,\varepsilon)^2}{4}(2\omega)^{D(N,\varepsilon)+1}\right)\left(\frac{(\rho^{\max}-\rho+\frac{3\varepsilon}{4})q^{\max}}{\frac{\varepsilon}{4(N-1)}q^{\min}}\right)\right)^{2.5nL} \\ &= O\left(\left(\frac{N-1}{\varepsilon}\right)^{11nL}(2\omega)^{\frac{10C_1(N-1)nL}{\varepsilon}}\right), \end{aligned}$$

where $\omega=\max\{1,\kappa\}$, and

$$D(N,\varepsilon) := \max\left\{\deg(p_i),\ \left\lceil\frac{4C_1(N-1)}{\varepsilon}\right\rceil + \deg(q_i) : i\in[N]\right\}.$$

Proof. Fix an arbitrary $\varepsilon>0$. For each $i\in[N]\setminus\{1\}$, let c_i be the constant in Lemma 1 and $C_1 := \max_{2\le i\le N} c_i$. Then, there exists $h_i\in\mathbb{R}[\mathbf{x}]$ of degree $\lceil 4c_i(N-1)/\varepsilon\rceil$ satisfying

$$\sup_{\mathbf{x}\in\mathbf{K}}\left|\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}-h_i(\mathbf{x})\right| \le \frac{c_i}{\lceil 4c_i(N-1)/\varepsilon\rceil} \le \frac{\varepsilon}{4(N-1)}.$$

For $i\in[N]\setminus\{1\}$, let $\hat{h}_i := h_i - \varepsilon/2(N-1)$. Then, for any $\mathbf{x}\in\mathbf{K}$, it holds that

$$\begin{aligned} \phi_i(\mathbf{x}) &:= p_i(\mathbf{x}) - \hat{h}_i(\mathbf{x})q_i(\mathbf{x}) = q_i(\mathbf{x})\left(\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}-\hat{h}_i(\mathbf{x})\right) \\ &= q_i(\mathbf{x})\left(\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}-h_i(\mathbf{x})+\frac{\varepsilon}{2(N-1)}\right) > 0,\ \forall i\in[N]\setminus\{1\}, \end{aligned}$$

and

$$\begin{aligned} \phi_1(\mathbf{x}) &:= p_1(\mathbf{x}) + \left(\sum_{i=2}^{N}\hat{h}_i(\mathbf{x}) - (\rho-\varepsilon)\right)q_1(\mathbf{x}) \\ &= q_1(\mathbf{x})\left(\frac{p_1(\mathbf{x})}{q_1(\mathbf{x})} + \sum_{i=2}^{N} h_i(\mathbf{x}) - \rho + \frac{\varepsilon}{2}\right) \\ &\ge q_1(\mathbf{x})\left(\sum_{i=1}^{N}\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \frac{\varepsilon}{4} - \rho + \frac{\varepsilon}{2}\right) > 0. \end{aligned}$$

By Theorem 5, there exists a constant C_2 depending on n and g_j's such that for each $i\in[N]$, $\phi_i\in Q_k(\mathbf{g})$ whenever

$$k \ge C_2 \deg(\phi_i)^{3.5nL}\left(\frac{\|\phi_i\|}{\min_{\mathbf{x}\in\mathbf{K}}\phi_i(\mathbf{x})}\right)^{2.5nL} =: k_i.$$

It remains to find an upper bound of $k_i, i \in [N]$. Clearly, we have

$$\begin{aligned}
\min_{\mathbf{x}\in\mathbf{K}} \phi_i(\mathbf{x}) &= \min_{\mathbf{x}\in\mathbf{K}}(p_i(\mathbf{x}) - \hat{h}_i(\mathbf{x})q_i(\mathbf{x})) \geq \frac{\varepsilon}{4(N-1)}q^{\min}, i \in [N]\setminus\{1\},\\
\min_{\mathbf{x}\in\mathbf{K}} \phi_1(\mathbf{x}) &= \min_{\mathbf{x}\in\mathbf{K}}\left(p_1(\mathbf{x}) + \left(\sum_{i=2}^{N}\hat{h}_i(\mathbf{x}) - (\rho-\varepsilon)\right)q_1(\mathbf{x})\right)\\
&\geq \min_{\mathbf{x}\in\mathbf{K}}\left(q_1(\mathbf{x})\left(\sum_{i=1}^{N}\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \frac{\varepsilon}{4} - \rho + \frac{\varepsilon}{2}\right)\right) \geq \frac{\varepsilon}{4}q^{\min},
\end{aligned}$$

and

$$\begin{aligned}
\max_{\mathbf{x}\in\mathbf{K}} \phi_i(\mathbf{x}) &= \max_{\mathbf{x}\in\mathbf{K}}(p_i(\mathbf{x}) - \hat{h}_i(\mathbf{x})q_i(\mathbf{x})) \leq \frac{3\varepsilon}{4(N-1)}q^{\max}, i \in [N]\setminus\{1\},\\
\max_{\mathbf{x}\in\mathbf{K}} \phi_1(\mathbf{x}) &= \max_{\mathbf{x}\in\mathbf{K}}\left(p_1(\mathbf{x}) + \left(\sum_{i=2}^{N}\hat{h}_i(\mathbf{x}) - (\rho-\varepsilon)\right)q_1(\mathbf{x})\right)\\
&\leq \max_{\mathbf{x}\in\mathbf{K}}\left(q_1(\mathbf{x})\left(\sum_{i=1}^{N}\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} + \frac{\varepsilon}{4} - \rho + \frac{\varepsilon}{2}\right)\right)\\
&\leq \left(\rho^{\max} - \rho + \frac{3\varepsilon}{4}\right)q^{\max}.
\end{aligned}$$

Note that

$$\begin{aligned}
&\frac{\varepsilon}{4(N-1)}q^{\min} \leq \frac{\varepsilon}{4}q^{\min}, \quad \frac{3\varepsilon}{4(N-1)}q^{\max} \leq \left(\rho^{\max} - \rho + \frac{3\varepsilon}{4}\right)q^{\max},\\
&\deg(\phi_i) = \deg(p_i - \hat{h}_i(\mathbf{x})q_i), \deg(\phi_1) = \deg\left(p_1 + \left(\sum_{i=2}^{N}\hat{h}_i - (\rho-\varepsilon)\right)q_1\right) \leq D(N,\varepsilon).
\end{aligned}$$

Therefore, by Proposition 1, it holds that

$$k_i \leq C_2 D(N,\varepsilon)^{3.5nL}\left(\left(1 + \frac{D(N,\varepsilon)^2}{4}(2\omega)^{D(N,\varepsilon)+1}\right)\left(\frac{(\rho^{\max} - \rho + \frac{3\varepsilon}{4})q^{\max}}{\frac{\varepsilon}{4(N-1)}q^{\min}}\right)\right)^{2.5nL},$$

for all $i \in [N]$. Hence, whenever

$$\begin{aligned}
k &\geq C_2 D(N,\varepsilon)^{3.5nL}\left(\left(1 + \frac{D(N,\varepsilon)^2}{4}(2\omega)^{D(N,\varepsilon)+1}\right)\left(\frac{(\rho^{\max} - \rho + \frac{3\varepsilon}{4})q^{\max}}{\frac{\varepsilon}{4(N-1)}q^{\min}}\right)\right)^{2.5nL}\\
&= O\left(\left(\frac{N-1}{\varepsilon}\right)^{11nL}(2\omega)^{\frac{10C_1(N-1)nL}{\varepsilon}}\right),
\end{aligned}$$

$(\rho - \varepsilon, \hat{h}_2, \ldots, \hat{h}_N)$ is feasible to (Dk), and thus, $\rho_k \geq \rho - \varepsilon$. □

4. Sign Symmetry–Adapted SDP Relaxations

In this section, we propose block-diagonal SDP relaxations for (SRFO) by exploiting sign symmetries as well as correlative sparsity, which would significantly reduce the computational burden for structured problems.

4.1. Sparse SDP Relaxations by Exploiting Sign Symmetries

Let us first define the sign symmetries of a subset $\mathcal{A} \subseteq \mathbb{N}^n$.

Definition 2. Given a finite set $\mathcal{A} \subseteq \mathbb{N}^n$, the sign symmetries $\mathcal{R}(\mathcal{A})$ of $\mathcal{A}$ consist of all vectors $\mathbf{r} \in \mathbb{Z}_2^n := \{0,1\}^n$ satisfying $\mathbf{r}^\top\boldsymbol{\alpha} \equiv 0 \,(\text{mod } 2)$ for all $\boldsymbol{\alpha} \in \mathcal{A}$. Moreover, we define the associated set $\overline{\mathcal{A}} := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid \mathbf{r}^\top\boldsymbol{\alpha} \equiv 0 \,(\text{mod } 2),\ \ \forall \mathbf{r} \in \mathcal{R}(\mathcal{A})\}$.

Note that $\mathcal{R}(\mathcal{A})$ is a linear subspace of $\mathbb{Z}_2^n$ for any $\mathcal{A} \subseteq \mathbb{N}^n$. The notion of sign symmetries stems from the invariance of polynomials under sign flips on variables (Löfberg [16], Wang et al. [34]). For any $\mathbf{r} \in \mathbb{Z}_2^n$, we define the map $\theta_{\mathbf{r}} : \mathbb{R}[\mathbf{x}] \to \mathbb{R}[\mathbf{x}]$ by

$$\theta_{\mathbf{r}}(f)(x_1, \ldots, x_n) = f((-1)^{r_1} x_1, \ldots, (-1)^{r_n} x_n). \tag{3}$$

A polynomial f is said to have the sign symmetry $\mathbf{r}$ if $\theta_{\mathbf{r}}(f) = f$. Note that a polynomial f has the set of sign symmetries $\mathcal{R} \subseteq \mathbb{Z}_2^n$ if and only if $\text{supp}(f) \subseteq \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid \mathbf{r}^\top \boldsymbol{\alpha} \equiv 0 \ (\text{mod } 2), \ \ \forall \mathbf{r} \in \mathcal{R}\}$.

For $i \in [N] \setminus \{1\}$, let

$$\mathcal{A}_i := \text{supp}(p_i) \cup \text{supp}(q_i) \cup \bigcup_{j=1}^{m} \text{supp}(g_j), \tag{4}$$

and let

$$\mathcal{A}_1 := \text{supp}(p_1) \cup \text{supp}(q_1) \cup \bigcup_{i=2}^{N} \mathcal{A}_i. \tag{5}$$

For the remainder of this section, we denote the set of sign symmetries of $\mathcal{A}_i$ by $\mathcal{R}_i$, $i \in [N]$.

Given $\mathbf{r} \in \mathbb{Z}_2^n$, let $\mathbf{r}(\mathbf{a}) := ((-1)^{r_1} a_1, \ldots, (-1)^{r_n} a_n)$ for $\mathbf{a} \in \mathbb{R}^n$, and for a set $S \subseteq \mathbb{R}^n$, let $\mathbf{r}(S) := \{\mathbf{r}(\mathbf{a}) \in \mathbb{R}^n \mid \mathbf{a} \in S\}$.

Lemma 2. *One has* $\mathbf{r}(\mathbf{K}) = \mathbf{K}$, $\theta_{\mathbf{r}}(p_i) = p_i$, *and* $\theta_{\mathbf{r}}(q_i) = q_i$ *for all* $\mathbf{r} \in \mathcal{R}_i$, $i \in [N]$.

Proof. It is straightforward to verify from the definitions. □

In order to take the inherent sign symmetries of (SRFO) into account, let us consider the following sign symmetry–adapted version of (D):

$$\rho^{\mathrm{s}} := \begin{cases} \sup\limits_{c, h_i} & c \\ \text{s.t.} & \dfrac{p_1(\mathbf{x})}{q_1(\mathbf{x})} + \displaystyle\sum_{i=2}^{N} h_i(\mathbf{x}) \geq c, \quad \forall \mathbf{x} \in \mathbf{K}, \\ & \dfrac{p_i(\mathbf{x})}{q_i(\mathbf{x})} \geq h_i(\mathbf{x}), \quad \forall \mathbf{x} \in \mathbf{K}, \\ & c \in \mathbb{R}, h_i \in \mathbb{R}[\overline{\mathcal{A}_i}], \quad i \in [N] \setminus \{1\}, \end{cases} \tag{SD}$$

where each $\overline{\mathcal{A}_i}$ is defined as in Definition 2.

Theorem 7. *Under Assumption 1, it holds that* $\rho^{\mathrm{s}} = \rho$.

Proof. Clearly, we have $\rho^{\mathrm{s}} \leq \rho$ by Theorem 2. To show the converse, fix an arbitrary $\varepsilon > 0$, and let $(\rho - \varepsilon, h_2, \ldots, h_N)$ be the feasible solution to (D) provided in the proof of Theorem 2 so that

$$\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \frac{\varepsilon}{N-1} \leq h_i(\mathbf{x}) \leq \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}, \quad \forall \mathbf{x} \in \mathbf{K}. \tag{6}$$

For each $i \in [N] \setminus \{1\}$, let $\tilde{h}_i := \frac{1}{|\mathcal{R}_i|} \sum_{\mathbf{r} \in \mathcal{R}_i} \theta_{\mathbf{r}}(h_i)$. We have $\sum_{\mathbf{r} \in \mathcal{R}_i} \theta_{\mathbf{r}}(\mathbf{x}^{\boldsymbol{\alpha}}) = 0$ for each $\boldsymbol{\alpha} \in \mathbb{N}^n \setminus \overline{\mathcal{A}_i}$. In fact, as $\boldsymbol{\alpha} \in \mathbb{N}^n \setminus \overline{\mathcal{A}_i}$, there exists $\tilde{\mathbf{r}} \in \mathcal{R}_i$ such that

$$-\sum_{\mathbf{r} \in \mathcal{R}_i} \theta_{\mathbf{r}}(\mathbf{x}^{\boldsymbol{\alpha}}) = \theta_{\tilde{\mathbf{r}}}\left(\sum_{\mathbf{r} \in \mathcal{R}_i} \theta_{\mathbf{r}}(\mathbf{x}^{\boldsymbol{\alpha}})\right) = \sum_{\mathbf{r} \in \mathcal{R}_i} \theta_{\tilde{\mathbf{r}}+\mathbf{r}}(\mathbf{x}^{\boldsymbol{\alpha}}) = \sum_{\mathbf{r} \in \tilde{\mathbf{r}}+\mathcal{R}_i} \theta_{\mathbf{r}}(\mathbf{x}^{\boldsymbol{\alpha}}) = \sum_{\mathbf{r} \in \mathcal{R}_i} \theta_{\mathbf{r}}(\mathbf{x}^{\boldsymbol{\alpha}}).$$

Therefore, we have $\tilde{h}_i \in \mathbb{R}[\overline{\mathcal{A}_i}]$. Moreover, it follows from Lemma 2 and the inequalities in (6) that

$$\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \frac{\varepsilon}{N-1} \leq \tilde{h}_i(\mathbf{x}) \leq \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}, \quad \forall \mathbf{x} \in \mathbf{K}. \tag{7}$$

Thus, we have

$$\inf_{\mathbf{x} \in \mathbf{K}} \frac{p_1(\mathbf{x})}{q_1(\mathbf{x})} + \sum_{i=2}^{N} \tilde{h}_i(\mathbf{x}) \geq \inf_{\mathbf{x} \in \mathbf{K}} \sum_{i=1}^{N} \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \varepsilon = \rho - \varepsilon.$$

Therefore, $(\rho - \varepsilon, \tilde{h}_2, \ldots, \tilde{h}_N)$ is feasible to (SD). As $\varepsilon > 0$ is arbitrary, we obtain $\rho^{\mathrm{s}} \geq \rho$. □

Given $\mathbf{r} \in \mathbb{Z}_2^n$ with $\mathbf{r}(\mathbf{K}) = \mathbf{K}$, for a measure $\mu \in \mathcal{M}(\mathbf{K})_+$, we define a new measure $\mu^{\mathbf{r}}$ by $\mu^{\mathbf{r}}(S) = \mu(\mathbf{r}(S))$ for any Borel set $S \subseteq \mathbf{K}$. A measure is said to be *invariant* with respect to the sign symmetries $\mathcal{R}$ if $\mu^{\mathbf{r}} = \mu$ for all $\mathbf{r} \in \mathcal{R}$.

Lemma 3. *Let* $\mathbf{r} \in \mathbb{Z}_2^n$, $\mu \in \mathcal{M}(\mathbf{K})_+$, *and* $f \in \mathbb{R}[\mathbf{x}]$.

1. *For* $\boldsymbol{\alpha} \in \mathbb{N}^n$, $\int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} \mathrm{d}\mu^{\mathbf{r}} = (-1)^{\mathbf{r}^\top \boldsymbol{\alpha}} \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} \mathrm{d}\mu$.
2. *If* $\mathbf{r} \in \mathcal{R}(\mathrm{supp}(f))$, *then* $\int_{\mathbf{K}} f \mathrm{d}\mu^{\mathbf{r}} = \int_{\mathbf{K}} f \mathrm{d}\mu$.
3. *If* $\mathbf{r} \in \mathcal{R}(\mathrm{supp}(f))$, *then for* $\boldsymbol{\alpha} \in \mathbb{N}^n$, $\int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} f \mathrm{d}\mu^{\mathbf{r}} = (-1)^{\mathbf{r}^\top \boldsymbol{\alpha}} \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} f \mathrm{d}\mu$.

Proof. It is straightforward to verify from the definitions. □

Let $\mathcal{M}(\mathbf{K})_+^{\mathcal{R}_i}$ denote the set of finite positive Borel measures that are invariant with respect to the sign symmetries $\mathcal{R}_i$. Then, the dual of (SD) reads as

$$\begin{cases} \inf\limits_{\mu_i \in \mathcal{M}(\mathbf{K})_+^{\mathcal{R}_i}} & \sum\limits_{i=1}^{N} \int_{\mathbf{K}} p_i \mathrm{d}\mu_i \\ \text{s.t.} & \int_{\mathbf{K}} q_1 \mathrm{d}\mu_1 = 1, \\ & \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_i \mathrm{d}\mu_i = \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_1 \mathrm{d}\mu_1, \quad \forall \boldsymbol{\alpha} \in \overline{\mathcal{A}_i},\ i \in [N] \setminus \{1\}. \end{cases} \tag{SP}$$

We now point out that (SP) cannot be viewed as a consequence of (P) equipped with invariant measures when $\overline{\mathcal{A}_1} \setminus \cap_{i=2}^N \overline{\mathcal{A}_i} \neq \emptyset$. Indeed, the latter would involve the constraints

$$\int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_i \mathrm{d}\mu_i = \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_1 \mathrm{d}\mu_1, \quad \forall \boldsymbol{\alpha} \in \overline{\mathcal{A}_1},\ i \in [N] \setminus \{1\}, \tag{8}$$

where each $\mu_i \in \mathcal{M}(\mathbf{K})_+^{\mathcal{R}_i}$. Now, suppose $\boldsymbol{\alpha} \in \bar{\mathcal{A}}_1 \setminus \cap_{i=2}^N \overline{\mathcal{A}_i}$ such that $\boldsymbol{\alpha} \notin \bar{\mathcal{A}}_j$ for some $j \in [N] \setminus \{1\}$. Then, there exists $\tilde{\mathbf{r}} \in \mathcal{R}_j$ such that $\tilde{\mathbf{r}}^\top \boldsymbol{\alpha} \equiv 1 \,(\mathrm{mod}\ 2)$, and by Lemma 3, it follows that

$$\int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_j \mathrm{d}\mu_j = \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_j \mathrm{d}\mu_j^{\tilde{\mathbf{r}}} = (-1)^{\tilde{\mathbf{r}}^\top \boldsymbol{\alpha}} \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_j \mathrm{d}\mu_j = -\int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_j \mathrm{d}\mu_j,$$

which implies $\int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_j \mathrm{d}\mu_j = \int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_1 \mathrm{d}\mu_1 = 0$. Therefore, (8) harbors implicit constraints

$$\int_{\mathbf{K}} \mathbf{x}^{\boldsymbol{\alpha}} q_1 \mathrm{d}\mu_1 = 0, \quad \forall \boldsymbol{\alpha} \in \overline{\mathcal{A}_1} \setminus \bigcap_{i=2}^{N} \overline{\mathcal{A}_i},$$

which are not involved in (SP). This (somewhat surprising) fact particularly highlights the significance of considering the SOS Problem (D) when one exploits sign symmetries for (SRFO).

The Sign Symmetry–Adapted Reformulation (SD) allows us to consider block-diagonal SDP relaxations for (SRFO). For $\mathcal{A} \subseteq \mathbb{N}^n$ and $k \in \mathbb{N}$, let us define a binary matrix $B_k^{\mathcal{A}}$ indexed by $\mathbb{N}_k^n$ through

$$[B_k^{\mathcal{A}}]_{\boldsymbol{\alpha},\boldsymbol{\beta}} = \begin{cases} 1, & \text{if } \boldsymbol{\alpha} + \boldsymbol{\beta} \in \overline{\mathcal{A}}, \\ 0, & \text{otherwise.} \end{cases}$$

It could be easily seen that $B_k^{\mathcal{A}}$ is block diagonal up to appropriate row/column permutations (Wang et al. [34]). We define the sparse quadratic module $Q_k(\mathbf{g}, \mathcal{A})$ associated with $\mathcal{A}$ by

$$Q_k(\mathbf{g}, \mathcal{A}) := \left\{ \sigma_0 + \sum_{j=1}^{m} \sigma_j g_j \,\middle|\, \sigma_j = [\mathbf{x}]_{k-d_j}^\top G_j [\mathbf{x}]_{k-d_j}, G_j \in \mathbb{S}^+(B_{k-d_j}^{\mathcal{A}}), j \in \{0\} \cup [m] \right\},$$

where $d_0 := 0$; for $s \in \mathbb{N}$, $[\mathbf{x}]_s := [1, x_1, x_2, \ldots, x_n^s]$ denotes the canonical vector of monomials up to degree s; and

$\mathbb{S}^+(B_s^{\mathcal{A}})$ denotes the set of positive semidefinite matrices with sparsity pattern being specified by $B_s^{\mathcal{A}}$. The sign symmetry–adapted version of (Dk) is given by

$$\rho_k^{\mathrm{s}} := \begin{cases} \sup\limits_{c,h_i} & c \\ \text{s.t.} & p_1(\mathbf{x}) + \left(\sum\limits_{i=2}^N h_i(\mathbf{x}) - c\right) q_1(\mathbf{x}) \in Q_k(\mathbf{g}, \mathcal{A}_1), \\ & p_i(\mathbf{x}) - h_i(\mathbf{x}) q_i(\mathbf{x}) \in Q_k(\mathbf{g}, \mathcal{A}_i), \quad i \in [N] \setminus \{1\}, \\ & c \in \mathbb{R},\, h_i(\mathbf{x}) \in \mathbb{R}[\overline{\mathcal{A}_i}] \cap \mathbb{R}[\mathbf{x}]_{2k - \max\{\deg(q_1), \deg(q_i)\}}, \quad i \in [N] \setminus \{1\}. \end{cases} \tag{SDk}$$

The following theorem is a sign symmetry–adapted version of Theorem 1.

Theorem 8 (Wang et al. [34, Theorem 6.11]). *Let $f \in \mathbb{R}[\mathbf{x}]$ and $\mathcal{A} = \mathrm{supp}(f) \cup \cup_{j=1}^m \mathrm{supp}(g_j)$. Assume that the quadratic module $Q(\mathbf{g})$ is Archimedean and f is positive on $\mathbf{K}$. Then, $f \in Q_k(\mathbf{g}, \mathcal{A})$ for some $k \in \mathbb{N}$.*

Remark 2. By a similar argument as for Theorem 8 (see Wang et al. [34]), one can actually show that if $f \in Q_k(\mathbf{g})$, then $f \in Q_k(\mathbf{g}, \mathcal{A})$ with $\mathcal{A} = \mathrm{supp}(f) \cup \bigcup_{j=1}^m \mathrm{supp}(g_j)$.

Theorem 9. *The following statements hold true:*

i. *For $k \geq d_{\min}$, $\rho_k^{\mathrm{s}} \leq \rho_k$. Moreover, if $\mathcal{R}_1 = \mathcal{R}_2 = \cdots = \mathcal{R}_N$, then $\rho_k^{\mathrm{s}} = \rho_k$.*
ii. *Under Assumption 1 and the Archimedean condition, $\rho_k^{\mathrm{s}} \nearrow \rho$ as $k \to \infty$.*

Proof. (i) It follows from the fact that any feasible solution $(c, h_2, \dots, h_N)$ to (SDk) is also feasible to (Dk). Now, assume $\mathcal{R}_1 = \mathcal{R}_2 = \cdots = \mathcal{R}_N$. To show $\rho_k^{\mathrm{s}} \geq \rho_k$, let $(c, h_2, \dots, h_N)$ be any feasible solution to (Dk). For each $i \in [N] \setminus \{1\}$, let $\tilde{h}_i := \frac{1}{|\mathcal{R}_1|} \sum_{\mathbf{r} \in \mathcal{R}_1} \theta_{\mathbf{r}}(h_i) \in \mathbb{R}[\bar{\mathcal{A}}_1] \cap \mathbb{R}[\mathbf{x}]_{2k - \max\{\deg(q_1), \deg(q_i)\}}$. From $p_i(\mathbf{x}) - h_i(\mathbf{x}) q_i(\mathbf{x}) \in Q_k(\mathbf{g})$ and Remark 2, we deduce that

$$p_i(\mathbf{x}) - \tilde{h}_i(\mathbf{x}) q_i(\mathbf{x}) \in Q_k(\mathbf{g}, \mathcal{A}_1). \tag{9}$$

Moreover, from $p_1(\mathbf{x}) + (\sum_{i=2}^N h_i(\mathbf{x}) - c) q_1(\mathbf{x}) \in Q_k(\mathbf{g})$ and Remark 2, it follows that

$$p_1(\mathbf{x}) + \left(\sum_{i=2}^N \tilde{h}_i(\mathbf{x}) - c\right) q_1(\mathbf{x}) \in Q_k(\mathbf{g}, \mathcal{A}_1).$$

Therefore, $(c, \tilde{h}_2, \dots, \tilde{h}_N)$ is feasible to (SDk), which implies $\rho_k^{\mathrm{s}} \geq \rho_k$.

(ii) It is clear that $\rho_k^{\mathrm{s}} \leq \rho_{k+1}^{\mathrm{s}}$ for any $k \geq d_{\min}$. To show the convergence, fix an arbitrary $\varepsilon > 0$. For $\bar{\varepsilon} := \varepsilon/3$, let $(\rho - \bar{\varepsilon}, \bar{h}_2, \dots, \bar{h}_N)$ be the feasible solution to (SD) provided in the proof of Theorem 7. For each $i \in [N] \setminus \{1\}$, let

$$h_i := \bar{h}_i - \frac{\bar{\varepsilon}}{N-1} \in \mathbb{R}[\overline{\mathcal{A}_i}].$$

Then, for $\mathbf{x} \in \mathbf{K}$, it holds that

$$p_i(\mathbf{x}) - h_i(\mathbf{x}) q_i(\mathbf{x}) = p_i(\mathbf{x}) - \bar{h}_i(\mathbf{x}) q_i(\mathbf{x}) + \frac{\bar{\varepsilon}}{N-1} q_i(\mathbf{x}) \geq \frac{\bar{\varepsilon}}{N-1} q_i(\mathbf{x}) > 0. \tag{10}$$

By Theorem 8, there exists $k_i \in \mathbb{N}$ such that $h_i \in \mathbb{R}[\overline{\mathcal{A}_i}] \cap \mathbb{R}[\mathbf{x}]_{2k_i - \deg(q_i)}$ and $p_i - h_i q_i \in Q_{k_i}(\mathbf{g}, \mathcal{A}_i)$. Moreover, for $\mathbf{x} \in \mathbf{K}$, we have

$$\begin{aligned} & p_1(\mathbf{x}) + \left(\sum_{i=2}^N h_i(\mathbf{x}) - (\rho - \varepsilon)\right) q_1(\mathbf{x}) \\ &= p_1(\mathbf{x}) + q_1(\mathbf{x}) \sum_{i=2}^N \bar{h}_i(\mathbf{x}) - \bar{\varepsilon} q_1(\mathbf{x}) - \rho q_1(\mathbf{x}) + 3\bar{\varepsilon} q_1(\mathbf{x}) \\ &\geq (\rho - \bar{\varepsilon}) q_1(\mathbf{x}) - \rho q_1(\mathbf{x}) + 2\bar{\varepsilon} q_1(\mathbf{x}) \geq \bar{\varepsilon} q_1(\mathbf{x}) > 0. \end{aligned}$$

By Theorem 8 again, there exists $k_1 \in \mathbb{N}$ such that

$$p_1(\mathbf{x}) + \left(\sum_{i=2}^N h_i(\mathbf{x}) - (\rho - \varepsilon)\right) q_1(\mathbf{x}) \in Q_{k_1}(\mathbf{g}, \mathcal{A}_1).$$

Table 1. Computational results for Example 1.

k	sup (Dk)	sup (SDk)		
		Case (1)	Case (2)	Case (3)
2	−0.3563	−0.4275	−0.4513	−0.4738
3	−0.3465	−0.3469	−0.3546	−0.3550
4		−0.3465	−0.3465	−0.3465

Let $k_\varepsilon := \max_{1\le i\le N} k_i$. Then, $(\rho-\varepsilon, h_2,\dots,h_N)$ is feasible to (SDk) with $k=k_\varepsilon$, which implies $\rho^{\mathrm{s}}_{k_\varepsilon} \ge \rho-\varepsilon$. As ε is arbitrary, we prove the convergence of $\{\rho^{\mathrm{s}}_k\}$ to ρ. □

Note that the optimum ρ^{s}_k of (SDk) may depend on which ratio being chosen as p_1/q_1. We give an example to illustrate this phenomenon.

Example 1. Let

$$f := \frac{x^2+y^2-yz}{1+2x^2+y^2+z^2} + \frac{y^2+x^2z}{1+x^2+2y^2+z^2} + \frac{z^2-x+y}{1+x^2+y^2+2z^2}.$$

Consider the minimization of f over the unit ball $\{(x,y,z)\in\mathbb{R}^3 \mid 1-x^2-y^2-z^2\ge 0\}$. We consider three cases: (1) $p_1=x^2+y^2-yz, q_1=1+2x^2+y^2+z^2$; (2) $p_1=y^2+x^2z, q_1=1+x^2+2y^2+z^2$; and (3) $p_1=z^2-x+y, q_1=1+x^2+y^2+2z^2$. We present the computational results in Table 1. For this problem, −0.3465 can be certified to be globally optimal.

Remark 3. To guarantee $\rho^{\mathrm{s}}_k=\rho_k$, one may simply letting $\mathcal{A}_1=\cdots=\mathcal{A}_N$ by Theorem 9.

Remark 4. Other types of symmetries (e.g., variable permutations) could be exploited in a similar but more complicated way (where group theory will come into play) as we do in this section. In particular, Theorem 7 holds true for other types of symmetries as well. We leave the full exploration of more general symmetries in sum-of-rational-functions optimization for future work.

4.2. Sparse SDP Relaxations by Exploiting Both Correlative Sparsity and Sign Symmetries

In this subsection, we present sparse SDP relaxations for (SRFO) by exploiting both correlative sparsity and sign symmetries. Let us begin by recalling a correlative sparsity–adapted GMP reformulation and a corresponding hierarchy of sparse SDP relaxations for (SRFO) proposed in Bugarin et al. [6]. We first describe the correlative sparsity pattern in (SRFO). For $I\subseteq[n]$, $\mathbf{x}(I)$ denotes the set of variables x_i with $i\in I$.

Assumption 3. *The index sets $[n]$ and $[m]$ can be decomposed as $[n]=\cup_{i=1}^N I_i$ and $[m]=\cup_{i=1}^N J_i$ such that*

i. *For every $i\in[N]$, $p_i,q_i\in\mathbb{R}[\mathbf{x}(I_i)]$;*

ii. *For every $j\in J_i$, $g_j\in\mathbb{R}[\mathbf{x}(I_i)]$;*

iii. *For every $i\in[N]$, there exists $k\in J_i$ such that $g_k=M_i-\sum_{\ell\in I_i}x_\ell^2$ for some $M_i>0$;*

iv. *The subsets $\{I_i\}_{i=1}^N$ satisfy the running intersection property (RIP); that is, for every $i\in[N]\setminus\{1\}$, $I_i\cap(\cup_{j=1}^{i-1}I_j)\subseteq I_k$ for some $k\in[i-1]$.*

Remark 5. If $\mathbf{K}$ is compact and one knows some $M>0$ such that $M-x_1^2-\cdots-x_n^2\ge 0$ for all $\mathbf{x}\in\mathbf{K}$, then we can always add redundant constraints $M-\sum_{j\in I_i}x_j^2\ge 0, i\in[N]$ to the description of $\mathbf{K}$ so that Assumption 3(iii) is satisfied.

For every $i\in[N]$, let

$$\mathbf{K}_i := \left\{\mathbf{x}(I_i)\in\mathbb{R}^{|I_i|} \mid g_j(\mathbf{x}(I_i))\ge 0,\ j\in J_i\right\},$$

and $\pi_i:\mathcal{M}(\mathbf{K})_+\to\mathcal{M}(\mathbf{K}_i)_+$ be the projection on $\mathbf{K}_i$; that is, for any $\mu\in\mathcal{M}(\mathbf{K})_+$,

$$\pi_i(\mu(B)) := \mu(\{\mathbf{x}:\mathbf{x}\in\mathbf{K}, \mathbf{x}(I_i)\in B\}), \quad \forall B\in\mathcal{B}(\mathbf{K}_i),$$

where $\mathcal{B}(\mathbf{K}_i)$ is the usual Borel σ-algebra associated with $\mathbf{K}_i$. For every pair $(i,j)\in[N]\times[N]$ with $i\ne j$ and $I_i\cap I_j\ne\emptyset$, let

$$\mathbf{K}_{ij}=\mathbf{K}_{ji} := \{\mathbf{x}(I_i\cap I_j):\mathbf{x}(I_i)\in\mathbf{K}_i,\ \mathbf{x}(I_j)\in\mathbf{K}_j\},$$

and the projection $\pi_{ij}:\mathcal{M}(\mathbf{K}_i)_+\to\mathcal{M}(\mathbf{K}_{ij})_+$ is defined in an obvious similar manner. Moreover, let us define the sets

$$U_i := \{j \in \{i+1,\ldots,N\} \mid I_i \cap I_j \neq \emptyset\},\ i \in [N-1], \quad U_N := \emptyset,$$

and

$$V_1 := \emptyset, \quad V_i := \{j \in \{1,\ldots,i-1\} \mid I_i \cap I_j \neq \emptyset\},\ i \in [N] \setminus \{1\}.$$

Then, the correlative sparsity–adapted GMP reformulation for (SRFO) is given by

$$\rho^{\mathrm{c}} := \begin{cases} \inf\limits_{\mu_i \in \mathcal{M}(\mathbf{K}_i)_+} & \sum\limits_{i=1}^{N} \int_{\mathbf{K}_i} p_i \mathrm{d}\mu_i \\ \text{s.t.} & \int_{\mathbf{K}_i} q_i \mathrm{d}\mu_i = 1, \quad i \in [N], \\ & \pi_{ij}(q_i \mathrm{d}\mu_i) = \pi_{ji}(q_j \mathrm{d}\mu_j), \quad j \in U_i,\ i \in [N-1]. \end{cases} \tag{CP}$$

For $I \subseteq [n]$ and $f \in \mathbb{R}[\mathbf{x}(I)]$, let $\mathbf{M}_k(\mathbf{y}, I)$ (respectively, $\mathbf{M}_k(f\mathbf{y}, I)$) be the moment (respectively, localizing) submatrix obtained by retaining only those rows and columns of $\mathbf{M}_k(\mathbf{y})$ (respectively, $\mathbf{M}_k(f\mathbf{y})$) indexed by $\boldsymbol{\alpha} \in \mathbb{N}^n$ with $\alpha_i = 0$ whenever $i \notin I$. For $1 \le i < j \le N$ with $I_i \cap I_j \neq \emptyset$, let $\mathbb{N}^{(ij)} := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid \alpha_k = 0,\ \forall k \notin I_i \cap I_j\}$. Then, the hierarchy of correlative sparsity–adapted SDP relaxations for (SRFO) is given by

$$\rho_k^{\mathrm{c}} := \begin{cases} \inf\limits_{\mathbf{y}_i} & \sum\limits_{i=1}^{N} L_{\mathbf{y}_i}(p_i) \\ \text{s.t.} & \mathbf{M}_k(\mathbf{y}_i, I_i) \succeq 0, \quad i \in [N], \\ & \mathbf{M}_{k-d_j}(g_j \mathbf{y}_i, I_i) \succeq 0, \quad j \in J_i, i \in [N], \\ & L_{\mathbf{y}_i}(q_i) = 1, \quad i \in [N], \\ & L_{\mathbf{y}_i}(\mathbf{x}^{\boldsymbol{\alpha}} q_i) = L_{\mathbf{y}_j}(\mathbf{x}^{\boldsymbol{\alpha}} q_j), \\ & \forall \boldsymbol{\alpha} \in \mathbb{N}^{(ij)}_{2k-\max\{\deg(q_i),\deg(q_j)\}}, j \in U_i, i \in [N-1]. \end{cases} \tag{CPk}$$

Theorem 10 (Bugarin et al. [6, Theorems 3.1 and 3.2]). *Let Assumption 3 hold, and assume that $q_i > 0$ on $\mathbf{K}_i$ for $i \in [N]$. It holds that $\rho^{\mathrm{c}} = \rho$ and $\rho_k^{\mathrm{c}} \nearrow \rho$ as $k \to \infty$.*

We can derive the Lagrange dual of (CPk), which reads as

$$\begin{cases} \sup\limits_{c_i, h_{i,j}} & \sum\limits_{i=1}^{N} c_i \\ \text{s.t.} & p_i - \left(c_i + \sum\limits_{j \in U_i} h_{i,j} - \sum\limits_{j \in V_i} h_{j,i} \right) q_i \in Q_k(\{g_j\}_{j \in J_i}),\ c_i \in \mathbb{R},\ i \in [N], \\ & h_{i,j} \in \mathbb{R}[\mathbf{x}(I_i \cap I_j)]_{2k-\max\{\deg(q_i),\deg(q_j)\}}, \quad j \in U_i,\ i \in [N-1]. \end{cases} \tag{CDk}$$

Proposition 2. *Under Assumption 3(iii), the strong duality holds between (CPk) and (CDk) for all $k \ge d_{\min}$.*

Proof. Let $\{\mu_i\}$ be a feasible solution of (CP) (see the proof of Bugarin et al. [6, theorem 3.1] for the existence). For each $i \in [N]$, let $\mathbf{y}_i = (y_{i\boldsymbol{\alpha}})_{\boldsymbol{\alpha} \in \mathbb{N}^{|I_i|}_{2k}}$ be such that $y_{i\boldsymbol{\alpha}} = \int_{\mathbf{K}_i} \mathbf{x}^{\boldsymbol{\alpha}} \mathrm{d}\mu_i$ for $\boldsymbol{\alpha} \in \mathbb{N}^{|I_i|}_{2k}$. Then, $\{\mathbf{y}_i\}$ is a feasible solution of (CPk). As Assumption 3(iii) holds, according to the proof of Bugarin et al. [6, theorem 2.2], the feasible set of (CPk) is compact, which implies that the optimal solution set of (CPk) is nonempty and bounded. Therefore, (CDk) is strictly feasible (Trnovská [33]), and the strong duality holds (Shapiro and Scheinber [30, theorem 4.1.3]). □

Let

$$\mathcal{A} := \bigcup_{i=1}^{N} (\mathrm{supp}(p_i) \cup \mathrm{supp}(q_i)) \cup \bigcup_{j=1}^{m} \mathrm{supp}(g_j), \tag{11}$$

and $\mathcal{R}$ be the set of sign symmetries of $\mathcal{A}$. Now, we consider the following sign symmetry–adapted version of (CP):

$$\rho^{\mathrm{cs}} := \begin{cases} \inf\limits_{\mu_i \in \mathcal{M}(\mathbf{K}_i)_+^{\mathcal{R}}} & \sum\limits_{i=1}^{N} \int_{\mathbf{K}_i} p_i \mathrm{d}\mu_i \\ \text{s.t.} & \int_{\mathbf{K}_i} q_i \mathrm{d}\mu_i = 1, \quad i \in [N], \\ & \pi_{ij}(q_i \mathrm{d}\mu_i) = \pi_{ji}(q_j \mathrm{d}\mu_j), \quad j \in U_i, i \in [N-1]. \end{cases} \tag{CSP}$$

Theorem 11. *Let Assumption 3 hold, and assume that $q_i > 0$ on $\mathbf{K}_i$ for $i \in [N]$. It holds that $\rho^{\mathrm{cs}} = \rho$.*

Proof. Because $\mathcal{M}(\mathbf{K}_i)_+^{\mathcal{R}}$ is a subset of $\mathcal{M}(\mathbf{K}_i)_+$ for each $i \in [N]$, we immediately have $\rho^{\mathrm{cs}} \geq \rho$. For the converse, fix an arbitrary $\varepsilon > 0$, and suppose that $\{\mu_i\}$ is any feasible solution of (CP) with $\sum_{i=1}^N \int_{\mathbf{K}} p_i \mathrm{d}\mu_i < \rho + \varepsilon$. For each $i \in [N]$, let us define a new measure $\tilde{\mu}_i \in \mathcal{M}(\mathbf{K}_i)_+^{\mathcal{R}}$ by $\tilde{\mu}_i := \frac{1}{|\mathcal{R}|}\sum_{\mathbf{r}\in\mathcal{R}} \mu_i^{\mathbf{r}}$. For any $j \in U_i, i \in [N-1]$, by Lemma 3, one can check that $\pi_{ij}(q_i \mathrm{d}\tilde{\mu}_i) = \pi_{ji}(q_j \mathrm{d}\tilde{\mu}_j)$. Moreover, by Lemma 3 again, $\sum_{i=1}^N \int_{\mathbf{K}} p_i \mathrm{d}\tilde{\mu}_i = \sum_{i=1}^N \int_{\mathbf{K}} p_i \mathrm{d}\mu_i < \rho + \varepsilon$. As $\varepsilon > 0$ is arbitrary, $\rho^{\mathrm{cs}} \leq \rho$. □

For each $i \in [N]$ and each $k \in \mathbb{N}$, define a binary matrix $B_{i,k}^{\mathcal{A}}$ indexed by $\mathbb{N}_k^{|I_i|}$ (we embed $\mathbb{N}_k^{|I_i|}$ into $\mathbb{N}_k^n$ in the natural way) such that

$$[B_{i,k}^{\mathcal{A}}]_{\boldsymbol{\alpha},\boldsymbol{\beta}} := \begin{cases} 1, & \text{if } \boldsymbol{\alpha} + \boldsymbol{\beta} \in \overline{\mathcal{A}}, \\ 0, & \text{otherwise.} \end{cases}$$

The sign symmetry–adapted version of (CPk) is given by

$$\rho_k^{\mathrm{cs}} := \begin{cases} \inf\limits_{\mathbf{y}_i} & \sum\limits_{i=1}^{N} L_{\mathbf{y}_i}(p_i) \\ \text{s.t.} & B_{i,k}^{\mathcal{A}} \circ \mathbf{M}_k(\mathbf{y}_i, I_i) \succeq 0, \quad i \in [N], \\ & B_{i,k-d_j}^{\mathcal{A}} \circ \mathbf{M}_{k-d_j}(g_j \mathbf{y}_i, I_i) \succeq 0, \quad j \in J_i, i \in [N], \\ & L_{\mathbf{y}_i}(q_i) = 1, \quad i \in [N], \\ & L_{\mathbf{y}_i}(\mathbf{x}^{\boldsymbol{\alpha}} q_i) = L_{\mathbf{y}_j}(\mathbf{x}^{\boldsymbol{\alpha}} q_j), \\ & \forall \boldsymbol{\alpha} \in \mathbb{N}^{(ij)}_{2k-\max\{\deg(q_i),\deg(q_j)\}} \cap \overline{\mathcal{A}}, j \in U_i, i \in [N-1]. \end{cases} \tag{CSPk}$$

Theorem 12. *Let Assumption 3 hold, and assume that $q_i > 0$ on $\mathbf{K}_i$ for $i \in [N]$. It holds that $\rho_k^{\mathrm{cs}} = \rho_k^{\mathrm{c}}$ for all $k \geq d_{\min}$. Consequently, $\rho_k^{\mathrm{cs}} \nearrow \rho$ as $k \to \infty$.*

Proof. Let $\{\mathbf{y}_i\}$ be a feasible solution to (CPk). Note that $B_{i,k}^{\mathcal{A}} \circ \mathbf{M}_k(\mathbf{y}_i, I_i)$ (respectively, $B_{i,k-d_j}^{\mathcal{A}} \circ \mathbf{M}_{k-d_j}(g_j \mathbf{y}_i, I_i)$, $j \in J_i$) consists of diagonal blocks of $\mathbf{M}_k(\mathbf{y}_i, I_i)$ (respectively, $\mathbf{M}_{k-d_j}(g_j \mathbf{y}_i, I_i)$, $j \in J_i$) for all $i \in [N]$. Thus, $\{\mathbf{y}_i\}$ is also feasible to (CSPk). So $\rho_k^{\mathrm{cs}} \leq \rho_k^{\mathrm{c}}$.

On the other hand, let $\{\mathbf{y}_i\}$ be any feasible solution to (CSPk). For every $i \in [N]$, we define a pseudo–moment sequence $\mathbf{y}_i' = (y_{i\boldsymbol{\alpha}}')_{\boldsymbol{\alpha} \in \mathbb{N}_k^{|I_i|}}$ as follows:

$$y_{i\boldsymbol{\alpha}}' = \begin{cases} y_{i\boldsymbol{\alpha}}, & \text{if } \boldsymbol{\alpha} \in \overline{\mathcal{A}}, \\ 0, & \text{otherwise.} \end{cases}$$

By the definition of $\mathcal{A}$, one can easily check that $\{\mathbf{y}_i'\}$ is a feasible solution to (CPk) and $\sum_{i=1}^N L_{\mathbf{y}_i'}(p_i) = \sum_{i=1}^N L_{\mathbf{y}_i}(p_i)$. So $\rho_k^{\mathrm{cs}} \geq \rho_k^{\mathrm{c}}$, and it follows that $\rho_k^{\mathrm{cs}} = \rho_k^{\mathrm{c}}$ as desired. □

The dual of (CSPk) reads as

$$\begin{cases} \sup\limits_{c_i, h_{i,j}} & \sum\limits_{i=1}^{N} c_i \\ \text{s.t.} & p_i - \left(c_i + \sum\limits_{j \in U_i} h_{i,j} - \sum\limits_{j \in V_i} h_{j,i}\right) q_i \in Q_k(\{g_j\}_{j \in J_i}, \mathcal{A}),\ c_i \in \mathbb{R},\ i \in [N], \\ & h_{i,j} \in \mathbb{R}[\mathbf{x}(I_i \cap I_j)]_{2k - \max\{\deg(q_i), \deg(q_j)\}} \cap \mathbb{R}[\overline{\mathcal{A}}], \quad j \in U_i, i \in [N-1]. \end{cases} \tag{CSDk}$$

Proposition 3. *Under Assumption 3*(iii), *the strong duality holds between* (CSPk) *and* (CSDk) *for all* $k \geq d_{\min}$.

Proof. Note that for any feasible solution $\{\mathbf{y}_i\}$ of (CSPk), only the entries $y_{i\boldsymbol{\alpha}}, \boldsymbol{\alpha} \in \overline{\mathcal{A}}$ are involved in (CSPk). By the proof of Theorem 12, the corresponding point $\{\mathbf{y}_i'\}$ is feasible to (CPk). Then, the proof of Proposition 2 indicates that the feasible set of (CSPk) is compact. Hence, the strong duality holds between (CSPk) and (CSDk), as proved in Proposition 2. □

5. Numerical Experiments

In this section, we conduct numerical experiments to test the performance of the sign symmetry–adapted approaches against the approaches without exploiting sign symmetries and the approaches in Bugarin et al. [6]. We use the Julia package `TSSOS`[1] (Magron and Wang [17]) to build the SDP relaxations and rely on `MOSEK` (Andersen and Andersen [2]) to solve them. Throughout this section, k stands for the relaxation order, the symbol "-" indicates that `MOSEK` runs out of memory, and the symbol "–" indicates that computational time exceeds 3,600 seconds. All numerical experiments were performed on a desktop computer with Intel(R) Core(TM) i9-14900 CPU of 2.00 GHz and 128 GB RAM.

5.1. Comparison of (SDk) with (Pk) and (Dk)

We begin by presenting three examples without correlative sparsity, where we compare the Sign Symmetry–Adapted Relaxation (SDk) with the Moment Relaxation (Pk) and the SOS Relaxation (Dk).

Example 2. Consider the problem

$$\min_{\mathbf{x} \in \mathbb{R}^3} \sum_{a=1/M}^{1-1/M} \frac{p_a(\mathbf{x})}{q_a(\mathbf{x})} \quad \text{s.t. } x_1^2 + x_2^2 + x_3^2 = 3, \tag{12}$$

where

$$\begin{aligned} p_a(\mathbf{x}) &= a^4(x_1^{6d} + x_2^{6d} + x_3^{6d}) + (x_1^{4d}x_2^{2d} + x_2^{4d}x_3^{2d} + x_3^{4d}x_1^{2d}) + a^8(x_1^{2d}x_2^{4d} + x_2^{2d}x_3^{4d} + x_3^{2d}x_1^{4d}), \\ q_a(\mathbf{x}) &= 2a^6(x_1^{4d}x_2^{2d} + x_2^{4d}x_3^{2d} + x_3^{4d}x_1^{2d}) + 2a^2(x_1^{2d}x_2^{4d} + x_2^{2d}x_3^{4d} + x_3^{2d}x_1^{4d}) + 3(1 - 2a^2 + a^4 - 2a^6 + a^8)x_1^{2d}x_2^{2d}x_3^{2d}, \end{aligned}$$

and M, d are positive integers. Replacing x_i^d's by y_i's in $p_a(\mathbf{x}), q_a(\mathbf{x})$ and denoting the resulting polynomials by $\tilde{p}_a(\mathbf{y}), \tilde{q}_a(\mathbf{y})$ with $\mathbf{y} = (y_1, y_2, y_3)$, one can show that for all $0 < a < 1$, $\tilde{p}_a(\mathbf{y}) - \tilde{q}_a(\mathbf{y}) \geq 0$ on $\mathbb{R}^3$ with 10 zeros: $\{(1, \pm 1, \pm 1), (\pm a, 1, 0), (0, \pm a, 1), (1, 0, \pm a)\}$ (Reznick [26]). Then, it is obvious that the optimum of (12) is $M - 1$. Moreover, we can rewrite $q_a(\mathbf{x})$ as

$$q_a(\mathbf{x}) = 2a^6(x_1^{4d}x_2^{2d} + x_2^{4d}x_3^{2d} + x_3^{4d}x_1^{2d} - 3x_1^{2d}x_2^{2d}x_3^{2d}) + 2a^2(x_1^{2d}x_2^{4d} + x_2^{2d}x_3^{4d} + x_3^{2d}x_1^{4d} - 3x_1^{2d}x_2^{2d}x_3^{2d}) + 3(1 + a^4 + a^8)x_1^{2d}x_2^{2d}x_3^{2d}.$$

Then, by the arithmetic-geometric inequality, we see that $q_a(\mathbf{x}) > 0$ for all feasible points $\mathbf{x}$. Hence, (12) satisfies Assumption 1 for all positive integers M and d. For $M \in \{6, 8, 10\}$ and $d \in \{2, 3, 4\}$, we solve (12) using the minimum relaxation order $k = 3d$ and present the computational results in Table 2. From the table, we see that the

Table 2. Computational results for (12).

M	inf (Pk)/time (seconds)			sup (Dk)/time (seconds)			sup (SDk)/time (seconds)		
	$d = 2$	$d = 3$	$d = 4$	$d = 2$	$d = 3$	$d = 4$	$d = 2$	$d = 3$	$d = 4$
6	5.00/8.83	5.00/3,072	—	5.00/0.23	5.00/2.77	5.00/30.2	5.00/0.04	5.00/0.33	5.00/0.77
8	7.00/14.7	—	—	7.00/0.34	7.00/3.37	7.00/42.8	7.00/0.06	7.00/0.43	7.00/1.27
10	9.00/18.1	—	—	9.00/0.54	9.00/4.03	9.00/78.3	9.00/0.09	9.00/0.47	9.00/1.95

Table 3. Computational results for (13).

n	inf (Pk)/time (seconds)			sup (Dk)/time (seconds)			sup (SDk)/time (seconds)		
5	$k=3$ $-49.45/16.8$	$k=4$ $-12.47/1{,}251$	$k=5$ —	$k=3$ $-49.45/2.40$	$k=4$ $-12.47/27.5$	$k=5$ $-10.41/205$	$k=3$ $-49.45/0.15$	$k=4$ $-12.47/1.35$	$k=5$ $-10.41/3.35$
10	$k=2$ $-25.24/21.1$	$k=3$ —	$k=4$ —	$k=2$ $-25.24/3.58$	$k=3$ $-19.24/1{,}307$	$k=4$ —	$k=2$ $-25.24/0.08$	$k=3$ $-19.24/1.97$	$k=4$ $-10.21/29.3$

primal approach is significantly slower than our dual approach, and by exploiting sign symmetries, we gain further speedup by approximately one to two magnitudes.

Example 3. Consider the following problem adapted from Bugarin et al. [6]:

$$\min_{\mathbf{x}\in\mathbb{R}^n} -\sum_{i=1}^{N} \frac{1}{\sum_{j=1}^{n}(x_j^2 - a_{ij})^2 + c_i} \quad \text{s.t. } 60 - \sum_{i=1}^{n}(x_i^2 - 5)^2 \geq 0. \tag{13}$$

The data $\{a_{ij}\}, \{c_i\}$ are given in Ali et al. [1, table 16] with $N = 30, n \in \{5, 10\}$. For $n = 5$, the optimum of (13) is -10.41, and for $n = 10$, the optimum is -10.21. We present the computational results in Table 3. From the table, we see that (1) the global optimum is achieved at $k = 5$ for $n = 5$ and is achieved at $k = 4$ for $n = 10$, (2) the primal approach is significantly slower than our dual approach, and (3) by exploiting sign symmetries, we gain further speedup by approximately one to three magnitudes.

Example 4. Now, we carry out randomly generated problems of the form

$$\min_{\mathbf{x}\in\mathbb{R}^n} -\sum_{i=1}^{N} \frac{1}{f_i^2 + 1} \quad \text{s.t. } x_1^2 + \cdots + x_n^2 \leq 1, \tag{14}$$

which are constructed as follows: (1) randomly choose a subset of nonconstant monomials $\mathscr{M}$ from $[\mathbf{x}]_d$ with prescribed probability ξ, and (2) randomly assign three elements in $\mathscr{M}$ to each f_i coupled with random coefficients in [0, 1]. It is clear that the optimal value of (14) is $-N$. We denote the set of (14) generated in such a way by RandSRFO (N, n, d, ξ) and construct the following 12 instances:[2]

$$\begin{aligned} P_1, P_2, P_3 &\in \text{RandSRFO}(10, 6, 4, 0.05), \\ P_4, P_5, P_6 &\in \text{RandSRFO}(8, 5, 5, 0.10), \\ P_7, P_8, P_9 &\in \text{RandSRFO}(6, 8, 4, 0.03), \\ P_{10}, P_{11}, P_{12} &\in \text{RandSRFO}(12, 7, 5, 0.05). \end{aligned}$$

We solve those instances using the minimum relaxation order $k = d$ and present the computational results in Table 4. From the table, again, we see that the primal approach is significantly slower than our dual approach, and by exploiting sign symmetries, we gain further speedup by approximately one to three magnitudes.

5.2. Comparison of (CSDk) with (CPk) and (CDk)

In this subsection, we compare the sign symmetry–adapted sparse relaxation (CSDk) with the sparse moment relaxation (CPk) and the sparse SOS relaxation (CDk).

Table 4. Computational results for (14).

	P_1	P_2	P_3	P_4	P_5	P_6
inf (Pk)/time (seconds)	$-10.00/1{,}812$	$-10.00/2{,}897$	$-10.00/2{,}251$	—	—	—
sup (Dk)/time (seconds)	$-10.00/19.4$	$-10.00/16.9$	$-10.00/19.2$	$-8.000/19.7$	$-8.000/19.3$	$-8.000/19.3$
sup (SDk)/time (seconds)	$-10.00/0.93$	$-10.00/0.70$	$-10.00/0.71$	$-8.000/1.57$	$-8.000/1.41$	$-8.000/1.61$
	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}
inf (Pk)/time (seconds)	—	—	—	—	—	—
sup (Dk)/time (seconds)	$-6.000/3{,}078$	$-6.000/3{,}294$	$-6.000/3{,}375$	—	—	—
sup (SDk)/time (seconds)	$-6.000/1.84$	$-6.000/1.79$	$-6.000/2.50$	$-12.00/18.5$	$-12.00/19.6$	$-12.00/20.3$

Table 5. Computational results for (15).

N	inf (CPk)/time (seconds)		sup (CDk)/time (seconds)		sup (CSDk)/time (seconds)	
	$k=4$	$k=5$	$k=4$	$k=5$	$k=4$	$k=5$
20	9.350*/15.8	80.00/547	16.61*/1.10	80.00/6.80	16.24*/0.16	80.00/0.70
40	17.92*/38.2	160.0/871	36.01*/3.53	160.0/14.5	34.26*/0.48	160.0/1.10
60	31.54*/50.5	240.0/1,094	52.59*/3.46	240.0/18.1	52.78*/0.53	240.0/1.61

Example 5. Consider

$$\begin{aligned} &\min_{\mathbf{x}\in\mathbb{R}^{2N+2}} \quad \sum_{i=1}^{N} \frac{(x_{2i-1}^2 + x_{2i}^2 + x_{2i+1}^2)x_{2i-1}^2 x_{2i}^2 x_{2i+1}^2 + x_{2i+2}^8}{x_{2i-1}^2 x_{2i}^2 x_{2i+1}^2 x_{2i+2}^2} \\ &\text{s.t.} \quad x_{2i-1}^2 + x_{2i}^2 + x_{2i+1}^2 + x_{2i+2}^2 = 4, \quad i \in [N]. \end{aligned} \tag{15}$$

For each i, the polynomial

$$(x_{2i-1}^2 + x_{2i}^2 + x_{2i+1}^2 - 4x_{2i+2}^2)x_{2i-1}^2 x_{2i}^2 x_{2i+1}^2 + x_{2i+2}^8$$

is nonnegative with zeros $(\pm 1, \pm 1, \pm 1, \pm 1)$ (Motzkin [18]). Therefore, the optimum of (15) equals $4N$. It is clear that Assumption 3 holds with $I_i = \{2i-1, 2i, 2i+1, 2i+2\}$ and $J_i = \{i\}$, $i \in [N]$. For $N \in \{20, 40, 60\}$, we solve (15) and present the computational results in Table 5. With the relaxation order $k = 4$, `MOSEK` returns "UNKNOWN_RESULT_STATUS" (marked by "*"), which is possibly because the SDP relaxations with a lower relaxation order are ill-conditioned.[3] It can be seen from the table that the primal approach is significantly slower than our dual approach, and by exploiting sign symmetries, we gain further speedup by one magnitude.

Example 6. Consider

$$\min_{\mathbf{x}\in\mathbb{R}^{2N+1}} \sum_{i=1}^{N} \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} \quad \text{s.t.} \quad x_{2i-1}^2 + x_{2i}^2 + x_{2i+1}^2 = 3, \; i \in [N], \tag{16}$$

where

$$\begin{aligned} p_i(\mathbf{x}) &= x_{2i-1}^{6d} + x_{2i}^{6d} + x_{2i+1}^{6d} + 3x_{2i-1}^{2d} x_{2i}^{2d} x_{2i+1}^{2d}, \\ q_i(\mathbf{x}) &= x_{2i-1}^{4d} x_{2i}^{2d} + x_{2i-1}^{2d} x_{2i}^{4d} + x_{2i-1}^{4d} x_{2i+1}^{2d} + x_{2i-1}^{2d} x_{2i+1}^{4d} + x_{2i}^{4d} x_{2i+1}^{2d} + x_{2i}^{2d} x_{2i+1}^{4d}, \end{aligned}$$

for some $d \geq 1$. From Example 2, it is easy to see that the optimum of (16) equals N. Moreover, Assumption 3 holds with $I_i = \{2i-1, 2i, 2i+1\}$ and $J_i = \{i\}$, $i \in [N]$. For $N \in \{5, 10, 20\}$ and $d \in \{2, 3, 4\}$, we solve (16) using the minimum relaxation order $k = 3d$ and present the computational results in Table 6. Again, it can be seen from the table that the primal approach is significantly slower than our dual approach, and by exploiting sign symmetries, we gain further speedup by one magnitude.

5.3. Comparison with the Epigraph Approach

In this subsection, we further compare with the epigraph approach, which translates (SRFO) into a polynomial optimization problem:

$$\begin{cases} \inf\limits_{c_i, \mathbf{x}} & \sum\limits_{i=1}^{N} c_i \\ \text{s.t.} & p_i(\mathbf{x}) - c_i q_i(\mathbf{x}) = 0, \quad i \in [N], \\ & g_j(\mathbf{x}) \geq 0, \quad j \in [m]. \end{cases} \tag{SEA}$$

Table 6. Computational results for (16).

N	inf (CPk)/time (seconds)			sup (CDk)/time (seconds)			sup (CSDk)/time (seconds)		
	$d=2$	$d=3$	$d=4$	$d=2$	$d=3$	$d=4$	$d=2$	$d=3$	$d=4$
5	5.000/6.17	5.000/844	—	5.000/0.33	5.000/2.10	5.000/10.7	5.000/0.04	5.000/0.27	5.000/0.71
10	10.00/13.9	10.00/1,638	—	10.00/0.54	10.00/3.28	10.00/25.2	10.00/0.11	10.00/0.37	10.00/1.01
20	20.00/24.7	—	—	20.00/0.74	20.00/6.05	20.00/68.4	20.00/0.15	15.00/0.73	20.00/2.62

Table 7. Computational results for (17). The "EPI" column records the lower bounds obtained by the epigraph approach.

N	EPI/time (seconds) $k=4$	inf (CPk)/time (seconds) $k=2$	sup (CDk)/time (seconds) $k=2$	sup (CSDk)/time (seconds) $k=2$
50	50.0/0.36	50.0/0.03	50.0/0.04	50.0/0.02
100	100.0/1.01	100.0/0.08	100.0/0.10	100.0/0.04
150	150.0/2.47	150.0/0.12	150.0/0.18	150.0/0.08

Note that (SEA) inherits the correlative sparsity and sign symmetries of (SRFO), which can be exploited to derive sparse SDP relaxations for (SEA) (Wang et al. [36]).

Example 7. Consider

$$\max_{\mathbf{x}\in\mathbb{R}^{N+1}} \sum_{i=1}^{N} \frac{1}{100(x_{i+1}^2 - x_i^2)^2 + (x_i^2 - 1)^2 + 1} \quad \text{s.t.} \;\; 16 - x_i^2 \geq 0, \;\; i \in [N+1], \tag{17}$$

which is modified from the well-known Rosenbrock problem. Clearly, the optimum of (17) equals N, and Assumption 3 holds with $I_i = J_i = \{i, i+1\}$, $i \in [N]$. We solve (17) for $N \in \{50, 100, 150\}$. It turns out that the SDP relaxations (CPk), (CDk), and (CSDk) attain global optimality at $k = 2$, whereas the epigraph approach attains global optimality at $k = 4$. We present the computational results in Table 7, from which we can see that the epigraph approach is much slower than the others.

Example 8. Consider

$$\min_{\mathbf{x}\in\mathbb{R}^{N+s}} \sum_{i=1}^{N} \frac{\sum_{j=1}^{s} x_{i+j-1}x_{i+j}}{1 + \sum_{j=1}^{s+1} j x_{i+j-1}^2} \quad \text{s.t.} \;\; 1 - x_i^2 \geq 0, \;\; i \in [N+s], \tag{18}$$

for some $s \geq 1$. It is clear that Assumption 3 holds with $I_i = J_i = \{i, i+1, \ldots, i+s\}$, $i \in [N]$. Let $s = 6$. We solve (18) for $N \in \{20, 40, 60\}$ using the minimum relaxation order $k = 3$ and present the computational results in Table 8. From the table, we see that the epigraph approach yields slightly tighter bounds, whereas the Sign Symmetry–Adapted Sparse Relaxation (CSDk) is the most efficient approach.

Finally, we would like to mention that for the problems presented in the previous two subsections, the epigraph approach either yields very bad bounds or spends a substantially long time.

6. An Application to Maximizing Sums of Generalized Rayleigh Quotients

In this section, we apply the sparse SDP relaxations to the problem of maximizing a sum of generalized Rayleigh quotients, which arises from signal processing.

In the downlink of a multiuser MIMO system, the base station can multiplex signals intended for different users on the same spectral resource. A challenging problem arising in such a scenario is the joint optimization of channel assignment (scheduling) and beamforming aimed at maximizing the sum rate in each time slot. Assuming zero-forcing linear beamforming at the base station, Primolevo et al. [24] addressed this task by investigating a greedy method to approximately maximize the sum rate. At each iteration of their method, a spatial channel needs to be determined to allocate a specific user, which reduces to maximizing a sum of generalized Rayleigh quotients of the form

$$\max_{\mathbf{z}\in\mathbb{C}^n} \sum_{i=1}^{N} \frac{\mathbf{z}^{\mathrm{H}} A_i \mathbf{z}}{\mathbf{z}^{\mathrm{H}} B_i \mathbf{z}} \quad \text{s.t.} \;\; \|\mathbf{z}\|^2 = 1, \tag{19}$$

where $A_i, B_i \in \mathbb{C}^{n\times n}$ are positive semidefinite Hermitian matrices, $\mathbf{z}^{\mathrm{H}}$ denotes the conjugate transpose of $\mathbf{z}$, n is the

Table 8. Computational results for (18). The "EPI" column records the lower bounds obtained by the epigraph approach.

N	EPI/time (seconds) $k=3$	inf (CPk)/time (seconds) $k=3$	sup (CDk)/time (seconds) $k=3$	sup (CSDk)/time (seconds) $k=3$
20	−4.5892/19.6	−4.6173/810	−4.6173/50.5	−4.6173/17.1
40	−8.7429/50.2	−8.8778/1,323	−8.8778/69.2	−8.8778/42.5
60	−12.901/98.9	−13.138/2,171	−13.138/107	−13.138/64.3

Table 9. Computational results for (20).

(n,N)	inf (Pk)/time (seconds)		sup (Dk)/time (seconds)		sup (SDk)/time (seconds)	
	$k=2$	$k=3$	$k=2$	$k=3$	$k=2$	$k=3$
(3, 20)	73.55/0.54	39.20/93.1	73.55/0.39	39.20/9.28	73.55/0.22	39.20/3.80
(4, 10)	157.8/2.30	157.8/927	157.8/0.79	157.8/66.5	157.8/0.53	157.8/28.2
(5, 5)	22.76/4.93	—	22.76/1.72	19.28/564	22.76/1.15	19.28/274

number of assigned spatial channels at the current iteration, and N is the number of available users in the time slot.

When $N=2$, the real counterpart of the Problem (19) also appears in the sparse Fisher discriminant analysis in pattern recognition (see Dundar et al. [7], Fung and Ng [8], Wu et al. [38]). The single generalized Rayleigh quotient optimization problem corresponds to the classical eigenvalue problem and can be solved in polynomial time (Parlett [23]). However, when $N \geq 2$, solving (19) is much more challenging and requires sophisticated techniques (see Nguyen et al. [19], Wang et al. [35], Zhang [39] for the real case of $N=2$). In particular, Primolevo et al. [24] restricted $\mathbf{z}$ in (19) to be columns of an identity matrix to give a suboptimal solution.

Here, we convert (19) into a real problem by specifying the real and imaginary parts of the complex variables and coefficients appearing in (19). Let $\mathbf{z} = \mathbf{x} + \mathbf{i}\mathbf{y}$, $A_i = A_{i,1} + \mathbf{i}A_{i,2}$, and $B_i = B_{i,1} + \mathbf{i}B_{i,2}$ with $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $A_{i,1}, A_{i,2}, B_{i,1}, B_{i,2} \in \mathbb{R}^{n\times n}$. Then, (19) becomes

$$\max_{\mathbf{x},\mathbf{y}\in\mathbb{R}^n} \sum_{i=1}^{N} \frac{\mathbf{x}^\top A_{i,1}\mathbf{x} - 2\mathbf{x}^\top A_{i,2}\mathbf{y} + \mathbf{y}^\top A_{i,1}\mathbf{y}}{\mathbf{x}^\top B_{i,1}\mathbf{x} - 2\mathbf{x}^\top B_{i,2}\mathbf{y} + \mathbf{y}^\top B_{i,1}\mathbf{y}} \quad \text{s.t. } \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = 1. \tag{20}$$

In our numerical experiments, we assume that A_i's are Hermitian and B_i's are positive definite. To satisfy this condition, we generate random matrices $C_{i,1}, C_{i,2}, D_{i,1}, D_{i,2} \in \mathbb{R}^{n\times n}$ with each entry being drawn from the uniform distribution on [0, 1], and let

$$A_i = (C_{i,1} + \mathbf{i}C_{i,2}) + (C_{i,1} + \mathbf{i}C_{i,2})^{\mathrm{H}}, \quad B_i = (D_{i,1} + \mathbf{i}D_{i,2})^{\mathrm{H}}(D_{i,1} + \mathbf{i}D_{i,2}).$$

We solve one instance of (20) with (Pk), (Dk), and (SDk) for $(n,N) \in \{(3,20),(4,10),(5,5)\}$, and present the computational results in Table 9. From the table, we see that the primal approach is slower than our dual approach, and by exploiting sign symmetries, we gain around twice speedup.

Acknowledgments
The authors sincerely thank the associate editor and reviewers for their insightful comments, which have significantly improved the manuscript.

Endnotes
[1] TSSOS is publicly available at https://github.com/wangjie212/TSSOS.

[2] They are available at https://wangjie212.github.io/jiewang/code.html.

[3] This issue typically affects the reliability of the solution but does not substantially affect the solving time.

References
[1] Ali MM, Khompatraporn C, Zabinsky ZB (2005) A numerical evaluation of several stochastic algorithms on selected continuous global optimization test problems. *J. Global Optim.* 31:635–672.
[2] Andersen ED, Andersen KD (2000) The Mosek interior point optimizer for linear programming: An implementation of the homogeneous algorithm. Frenk H, Roos K, Terlaky T, Zhang S, eds. *High Performance Optimization*, Applied Optimization, vol. 33 (Springer, Boston), 197–232.
[3] Baldi L (2022) Effective representations in real algebraic geometry and polynomial optimization. PhD thesis, Inria d'Université Côte d'Azur, Valbonne, France.
[4] Baldi L, Mourrain B (2023) On the effective Putinar's Positivstellensatz and moment approximation. *Math. Programming* 200(1):71–103.
[5] Bochnak J, Coste M, Roy MF (1998) *Real Algebraic Geometry*, A Series of Modern Surveys in Mathematics, vol. 36 (Springer, Berlin, Heidelberg).
[6] Bugarin F, Henrion D, Lasserre JB (2016) Minimizing the sum of many rational functions. *Math. Programming Comput.* 8(1):83–111.
[7] Dundar MM, Fung G, Bi J, Sathyakama S, Rao B (2005) Sparse Fisher discriminant analysis for computer aided detection. Kargupta H, Kamath C, Srivastava J, Goodman A, eds. *Proc. 2005 SIAM Internat. Conf. Data Mining (SDM)* (Society for Industrial and Applied Mathematics, Philadelphia), 476–480.
[8] Fung E, Ng M (2007) On sparse Fisher discriminant method for microarray data analysis. *Bioinformation* 2(5):230–234.

[9] Guo F, Wang L, Zhou G (2014) Minimizing rational functions by exact Jacobian SDP relaxation applicable to finite singularities. *J. Global Optim.* 58(2):261–284.
[10] Hartley R, Kahl F, Olsson C, Seo Y (2013) Verifying global minima for L2 minimization problems in multiple view geometry. *Internat. J. Comput. Vision* 101:288–304.
[11] Jibetean D, de Klerk E (2006) Global optimization of rational functions: A semidefinite programming approach. *Math. Programming* 106(1):93–109.
[12] Kahl F, Agarwal S, Chandraker MK, Kriegman D, Belongie S (2008) Practical global optimization for multiview geometry. *Internat. J. Comput. Vision* 79:271–284.
[13] Korda M, Henrion D (2018) Convergence rates of moment-sum-of-squares hierarchies for volume approximation of semialgebraic sets. *Optim. Lett.* 12(3):435–442.
[14] Korda M, Henrion D, Jones CN (2017) Convergence rates of moment-sum-of-squares hierarchies for optimal control problems. *Systems Control Lett.* 100:1–5.
[15] Lasserre JB, Magron V, Marx S, Zahm O (2021) Minimizing rational functions: A hierarchy of approximations via pushforward measures. *SIAM J. Optim.* 31(3):2285–2306.
[16] Löfberg J (2009) Pre- and post-processing sum-of-squares programs in practice. *IEEE Trans. Automatic Control* 54(5):1007–1011.
[17] Magron V, Wang J (2021) TSSOS: A Julia library to exploit sparsity for large-scale polynomial optimization. *Effective Methods Algebraic Geometry (Tromso, Norway).*
[18] Motzkin T (1967) The arithmetic-geometric inequality. Shisha O, ed. *Inequalities* (Academic Press, New York), 205–224.
[19] Nguyen VB, Sheu RL, Xia Y (2016) Maximizing the sum of a generalized Rayleigh quotient and another Rayleigh quotient on the unit sphere via semidefinite programming. *J. Global Optim.* 64(2):399–416.
[20] Nie J (2013) An exact Jacobian SDP relaxation for polynomial optimization. *Math. Programming* 137:225–255.
[21] Nie J, Schweighofer M (2007) On the complexity of Putinar's Positivstellensatz. *J. Complexity* 23(1):135–150.
[22] Nie J, Demmel J, Gu M (2008) Global minimization of rational functions and the nearest GCDs. *J. Global Optim.* 40(4):697–718.
[23] Parlett BN (1998) *The Symmetric Eigenvalue Problem* (Society for Industrial and Applied Mathematics, Philadelphia, PA).
[24] Primolevo G, Simeone O, Spagnolini U (2006) Towards a joint optimization of scheduling and beamforning for MIMO downlink. *2006 IEEE Ninth Internat. Sympos. Spread Spectrum Tech. Appl.* (IEEE, Pisctaway, NJ), 493–497.
[25] Putinar M (1993) Positive polynomials on compact semi-algebraic sets. *Indiana Univ. Math. J.* 42(3):969–984.
[26] Reznick B (2000) Some concrete aspects of Hilbert's 17th problem. *Contemporary Mathematics*, vol. 253 (American Mathematical Society, Providence, RI), 251–272.
[27] Schaible S, Shi J (2003) Fractional programming: The sum-of-ratios case. *Optim. Methods Software* 18(2):219–229.
[28] Schlosser C, Tacchi-Bénard M, Lazarev A (2026) Convergence rates for the moment-SOS hierarchy. *Numerical Algebra, Control Optim.* 16:105–156.
[29] Shapiro A (2009) Semi-infinite programming, duality, discretization and optimality conditions. *Optimization* 58(2):133–161.
[30] Shapiro A, Scheinber K (2000) Duality, optimality conditions and perturbation analysis. Wolkowicz H, Saigal R, Vandenberghe L, eds. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications* (Kluwer Academic Publisher, Boston), 67–110.
[31] Slot L (2022) Asymptotic analysis of semidefinite bounds for polynomial optimization and independent sets in geometric hypergraphs. PhD thesis, Tilburg University, Tilburg, Netherlands.
[32] Timan A (1963) *Theory of Approximation of Functions of a Real Variable*, International Series of Monographs on Pure and Applied Mathematics (Pergamon Press, Oxford, UK).
[33] Trnovská M (2005) Strong duality conditions in semidefinite programming. *J. Electr. Engrg.* 56(1):1–5.
[34] Wang J, Magron V, Lasserre JB (2021) TSSOS: A moment-SOS hierarchy that exploits term sparsity. *SIAM J. Optim.* 31(1):30–58.
[35] Wang X, Wang L, Xia Y (2018) An efficient global optimization algorithm for maximizing the sum of two generalized Rayleigh quotients. *Comput. Appl. Math.* 37(4):4412–4422.
[36] Wang J, Magron V, Lasserre JB, Mai NHA (2022) CS-TSSOS: Correlative and term sparsity for large-scale polynomial optimization. *ACM Trans. Math. Software* 48(4):1–26.
[37] Wu WY, Sheu RL, Birbil Şİ (2008) Solving the sum-of-ratios problem by a stochastic search algorithm. *J. Global Optim.* 42(1):91–109.
[38] Wu MC, Zhang L, Wang Z, Christiani DC, Lin X (2009) Sparse linear discriminant analysis for simultaneous testing for the significance of a gene set/pathway and gene selection. *Bioinformatics* 25(9):1145–1151.
[39] Zhang LH (2013) On optimizing the sum of the Rayleigh quotient and the generalized Rayleigh quotient on the unit sphere. *Comput. Optim. Appl.* 54(1):111–139.