

Exploiting Sign Symmetries in Minimizing Sums of Rational Functions*

Feng Guo[†], Jie Wang[‡], and Jianhao Zheng[§]

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 due to 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.

Key words. sum of rational functions, sign symmetry, semidefinite relaxation, correlative sparsity, generalized Rayleigh quotient

MSC codes. 90C23, 90C22, 90C26

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

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

where

$$(1.1) \quad \mathbf{K} := \{\mathbf{x} \in \mathbb{R}^n \mid g_j(\mathbf{x}) \geq 0, j = 1, \dots, m\},$$

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

Assumption 1.1. (i) \mathbf{K} is compact; (ii) $q_i > 0$ on \mathbf{K} for $i = 1, \dots, N$.

The problem (SRFO) has applications in various fields, including computer vision [7, 9], multi-user MIMO systems [21], sparse Fisher discriminant analysis in pattern recognition [4, 5, 32].

Given the potentially large number of terms N and the absence of convexity (resp. concavity) assumptions on p_i (resp. 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

*Submitted to the editors DATE.

Funding: This work was funded by National Key R&D Program of China under grant No. 2023YFA1009401, Natural Science Foundation of China under grant No. 12201618 and 12171324.

[†]School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, Liaoning Province, China (fguo@dlut.edu.cn).

[‡]Academy of Mathematics and Systems Science, Chinese academy of sciences, Beijing 100190, China (wangjie212@amss.ac.cn).

[§]School of Mathematical Sciences, Dalian University of Technology, Dalian 116024, Liaoning Province, China.

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 [24] and [33, 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 [8, 18]. 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 [6], 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 [17].

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, due to 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, Henrion, and Lasserre [3] 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 [3]. By employing pushforward measures, Lasserre et al. [12] gave an approach yielding convergent upper bounds on ρ .

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 , $i = 2, \dots, N$ 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 due to 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 $[n] := \{1, 2, \dots, n\}$. For $k \in \mathbb{N}$, let $\mathbb{N}_k^n := \{\boldsymbol{\alpha} = (\alpha_i) \in \mathbb{N}^n \mid \sum_{i=1}^n \alpha_i \leq k\}$. We use

$|\cdot|$ to denote the cardinality of a set. Let $\mathbb{R}[\mathbf{x}] := \mathbb{R}[x_1, \dots, x_n]$ 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 $f = \sum_{\alpha \in \mathbb{N}^n} f_{\alpha} \mathbf{x}^{\alpha}$ with $f_{\alpha} \in \mathbb{R}$ and $\mathbf{x}^{\alpha} := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$. The support of f is defined by $\text{supp}(f) := \{\alpha \in \mathbb{N}^n \mid f_{\alpha} \neq 0\}$. For $\mathcal{A} \subseteq \mathbb{N}^n$, $\mathbb{R}[\mathcal{A}]$ denotes the set of polynomials with supports contained in \mathcal{A} , i.e., $\mathbb{R}[\mathcal{A}] := \{f \in \mathbb{R}[\mathbf{x}] \mid \text{supp}(f) \subseteq \mathcal{A}\}$. For $\mathbf{u} = (u_1, \dots, u_n) \in \mathbb{R}^n$, $\|\mathbf{u}\|$ 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 $f_1(\mathbf{x}), \dots, f_t(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$. Let $\Sigma[\mathbf{x}]$ denote the set of SOS polynomials in $\mathbb{R}[\mathbf{x}]$.

Let $\mathbf{g} := \{g_1, \dots, g_m\}$ be the set of polynomials that defines the semialgebraic set \mathbf{K} in (1.1). We denote by

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

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

$$\mathcal{Q}_k(\mathbf{g}) := \left\{ \sigma_0 + \sum_{j=1}^m \sigma_j g_j \mid \sigma_0, \sigma_j \in \Sigma[\mathbf{x}], \deg(\sigma_0), \deg(\sigma_j g_j) \leq 2k, j \in [m] \right\}$$

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

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

Definition 2.1 (Archimedean condition). We say that $\mathcal{Q}(\mathbf{g})$ is *Archimedean* if there exists $M > 0$ such that $M - x_1^2 - \cdots - x_n^2 \in \mathcal{Q}(\mathbf{g})$.

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 2.2. [22, Putinar's Positivstellensatz] Suppose that $\mathcal{Q}(\mathbf{g})$ is Archimedean. If a polynomial $f \in \mathbb{R}[\mathbf{x}]$ is positive on \mathbf{K} , then $f \in \mathcal{Q}_k(\mathbf{g})$ for some $k \in \mathbb{N}$.

2.2. The GMP reformulation and SDP relaxations. In [3], Bugarin et al. reformulated (SRFO) as the following GMP:

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

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

Let

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

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

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

It was shown in [3] that under Assumption 1.1 and the Archimedean condition, the sequence of optima of (Pk) converges to ρ as $k \rightarrow \infty$.

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 [25]. Let \mathcal{Y} be the space of all functions $\omega : \mathbb{N}^n \rightarrow \mathbb{R}$ 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 $\omega^* : \mathbb{N}^n \rightarrow \mathbb{R}$ such that only a finite number of values $\omega^*(\alpha), \alpha \in \mathbb{N}^n$ are nonzero. For $\omega \in \mathcal{Y}$ and $\omega^* \in \mathcal{Y}^*$, define the scalar product

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

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

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

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

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

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 $h_i = \sum_{\alpha \in \mathbb{N}^n} h_\alpha^i \mathbf{x}^\alpha \in \mathbb{R}[\mathbf{x}]$ 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 $\int_{\mathbf{K}} q_1 d\mu_1 = 1$. The Lagrangian of (P) now can be written as

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

The Lagrange dual function of (P) is

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

Hence, the Lagrange dual problem of (P) is

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

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

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

Since 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; 3) letting the approximations h_i 's vary and taking the supremum.

Theorem 3.1. *Under Assumption 1.1, the optimum of (D) equals ρ .*

Proof. Denote the optimum of (D) by τ . Clearly, we have $\tau \leq \rho$. 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

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

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

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

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

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

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

Remark 3.2. Theorem 3.1 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, $(\hat{c}, \hat{h}_2, \dots, \hat{h}_N)$ 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) ($k \geq d_{\min}$):

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

Theorem 3.3. *Under Assumption 1.1 and the Archimedean condition, it holds $\rho_k \nearrow \rho$ as $k \rightarrow \infty$.*

Proof. It is obvious that $\rho_k \leq \rho_{k+1} \leq \rho$ for all $k \geq d_{\min}$. Fix an arbitrary $\varepsilon > 0$ and we will prove that there exists $k_\varepsilon \in \mathbb{N}$ such that $\rho_{k_\varepsilon} \geq \rho - \varepsilon$. For $\bar{\varepsilon} := \varepsilon/3$, let $(\rho - \bar{\varepsilon}, \bar{h}_2, \dots, \bar{h}_N)$

be the feasible solution to (D) provided in the proof of Theorem 3.1. For each $i = 2, \dots, N$, let

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

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

$$(3.1) \quad p_i(\mathbf{x}) - h_i(\mathbf{x})q_i(\mathbf{x}) = p_i(\mathbf{x}) - \bar{h}_i(\mathbf{x})q_i(\mathbf{x}) + \frac{\bar{\varepsilon}}{N-1}q_i(\mathbf{x}) \geq \frac{\bar{\varepsilon}}{N-1}q_i(\mathbf{x}) > 0.$$

By Theorem 2.2, 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

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

By Theorem 2.2 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 $k_\varepsilon := \max_{1 \leq i \leq N} k_i$. Then $(\rho - \varepsilon, h_2, \dots, h_N)$ is feasible to (Dk) with $k = k_\varepsilon$, which implies $\rho_{k_\varepsilon} \geq \rho - \varepsilon$. ■

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 3.4. *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*

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

Proof. Denote by

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

the standard modulus of continuity of p_i/q_i on \mathbf{K} . By the multivariate version of Jackson's theorem (see [27]), 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

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

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 $\omega_{\mathbf{K}}(p_i/q_i, t) \leq L_i t$. Then, the conclusion follows by letting $c_i := \hat{c}_i L_i$. ■

Assumption 3.5. The origin belongs to the interior of \mathbf{K} .

The following result is a consequence of the fundamental result [19, Theorem 6] and [11, Corollary 1].

Theorem 3.6. [10, Theorem 3] *Let the Archimedean condition and Assumption 3.5 hold. If $\phi(\mathbf{x}) \in \mathbb{R}[\mathbf{x}]$ is strictly positive on \mathbf{K} , then $\phi \in Q_k(\mathbf{g})$ whenever*

$$k \geq C_{\mathbf{K}} \exp \left[\left(3^{\deg(\phi)+1} \kappa^{\deg(\phi)} (\deg(\phi))^2 n^{\deg \phi} \frac{\max_{\mathbf{x} \in \mathbf{K}} \phi(\mathbf{x})}{\min_{\mathbf{x} \in \mathbf{K}} \phi(\mathbf{x})} \right)^{C_{\mathbf{K}}} \right],$$

for some constant $C_{\mathbf{K}}$ depending only on g_j 's and

$$(3.2) \quad \kappa := \frac{1}{\sup \{t > 0: [-t, t]^n \subseteq \mathbf{K}\}}.$$

Let

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

Using Theorem 3.6, we can establish the following convergence rate of the hierarchy (Dk).

Theorem 3.7. *Let the Archimedean condition and Assumptions 1.1, 3.5 hold. Then, there exist constants C_1 and C_2 depending on p_i 's, q_i 's, g_j 's and \mathbf{K} , such that for any $\varepsilon > 0$, we have $\rho_k \geq \rho - \varepsilon$ whenever*

$$\begin{aligned} k &\geq C_2 \exp \left[\left(3(3\omega n)^{D(N,\varepsilon)} D(N,\varepsilon)^2 \frac{(4(\rho^{\max} - \rho) + 3\varepsilon)(N-1) q^{\max}}{\varepsilon} \frac{q^{\max}}{q^{\min}} \right)^{C_2} \right] \\ &= O \left(\exp \left[\frac{1}{\varepsilon^{3C_2}} (3\omega n)^{\frac{4(N-1)C_1 C_2}{\varepsilon}} \right] \right), \end{aligned}$$

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 3.4 and $C_1 := \max_{2 \leq i \leq N} c_i$. Then, there exists $h_i \in \mathbb{R}[\mathbf{x}]$ of degree $\lceil 4c_i(N-1)/\varepsilon \rceil$ satisfying

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

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

$$p_i(\mathbf{x}) - \hat{h}_i(\mathbf{x})q_i(\mathbf{x}) = q_i(\mathbf{x}) \left(\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \hat{h}_i(\mathbf{x}) \right) = q_i(\mathbf{x}) \left(\frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - h_i(\mathbf{x}) + \frac{\varepsilon}{2(N-1)} \right),$$

and

$$p_1(\mathbf{x}) + \left(\sum_{i=2}^N \hat{h}_i(\mathbf{x}) - (\rho - \varepsilon) \right) q_1(\mathbf{x}) = q_1(\mathbf{x}) \left(\frac{p_1(\mathbf{x})}{q_1(\mathbf{x})} + \sum_{i=2}^N h_i(\mathbf{x}) - \rho + \frac{\varepsilon}{2} \right).$$

Therefore, for any $\mathbf{x} \in \mathbf{K}$,

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

and

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

Note that

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

Hence, by Theorem 3.6, there exists a constant C_2 depending on g_j 's such that whenever

$$\begin{aligned} k &\geq C_2 \exp \left[\left(3^{D(N, \varepsilon)+1} \omega^{D(N, \varepsilon)} D(N, \varepsilon)^2 n^{D(N, \varepsilon)} \frac{(\rho^{\max} - \rho + \frac{3\varepsilon}{4}) q^{\max}}{4(N-1) q^{\min}} \right)^{C_2} \right] \\ &= C_2 \exp \left[\left(3(3\omega n)^{D(N, \varepsilon)} D(N, \varepsilon)^2 \frac{(4(\rho^{\max} - \rho) + 3\varepsilon)(N-1) q^{\max}}{\varepsilon q^{\min}} \right)^{C_2} \right] \\ &= O \left(\exp \left[\frac{1}{\varepsilon^{3C_2}} (3\omega n)^{\frac{4(N-1)C_1 C_2}{\varepsilon}} \right] \right), \end{aligned}$$

$(\rho - \varepsilon, \hat{h}_2, \dots, \hat{h}_N)$ 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 4.1. 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 $\mathbf{r}^\top \boldsymbol{\alpha} \equiv 0 \pmod{2}$ for all $\boldsymbol{\alpha} \in \mathcal{A}$. Moreover, we define the associated set $\overline{\mathcal{A}} := \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid \mathbf{r}^\top \boldsymbol{\alpha} \equiv 0 \pmod{2}, \forall \mathbf{r} \in \mathcal{R}(\mathcal{A})\}$.

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 [13, 29]. For any $\mathbf{r} \in \mathbb{Z}_2^n$, we define the map $\theta_{\mathbf{r}}: \mathbb{R}[\mathbf{x}] \rightarrow \mathbb{R}[\mathbf{x}]$ by $\theta_{\mathbf{r}}(f)(x_1, \dots, x_n) = f((-1)^{r_1}x_1, \dots, (-1)^{r_n}x_n)$. 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 $\text{supp}(f) \subseteq \{\boldsymbol{\alpha} \in \mathbb{N}^n \mid \mathbf{r}^\top \boldsymbol{\alpha} \equiv 0 \pmod{2}, \forall \mathbf{r} \in \mathcal{R}\}$.

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

$$(4.1) \quad \mathcal{A}_i := \text{supp}(p_i) \cup \text{supp}(q_i) \cup \bigcup_{j=1}^m \text{supp}(g_j),$$

and let

$$(4.2) \quad \mathcal{A}_1 := \text{supp}(p_1) \cup \text{supp}(q_1) \cup \bigcup_{i=2}^N \mathcal{A}_i.$$

For the remainder of this section, we denote the set of sign symmetries of \mathcal{A}_i by \mathcal{R}_i .

Given $\mathbf{r} \in \mathbb{Z}_2^n$, let $\mathbf{r}(\mathbf{a}) := ((-1)^{r_1}a_1, \dots, (-1)^{r_n}a_n)$ 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 4.2. 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):

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

Theorem 4.3. Under Assumption 1.1, it holds $\rho^s = \rho$.

Proof. Clearly, we have $\rho^s \leq \rho$ by Theorem 3.1. To show the converse, fix an arbitrary $\varepsilon > 0$ and let $(\rho - \varepsilon, h_2, \dots, h_N)$ be the feasible solution to (D) provided in the proof of Theorem 3.1 so that

$$(4.3) \quad \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \frac{\varepsilon}{N-1} \leq h_i(\mathbf{x}) \leq \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}, \quad \forall \mathbf{x} \in \mathbf{K}.$$

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 $\sum_{\mathbf{r} \in \mathcal{R}_i} \theta_{\mathbf{r}}(\mathbf{x}^\alpha) = 0$ for each $\alpha \in \mathbb{N}^n \setminus \overline{\mathcal{A}_i}$. In fact, as $\alpha \in \mathbb{N}^n \setminus \overline{\mathcal{A}_i}$, there exists $\tilde{\mathbf{r}} \in \mathcal{R}_i$ such that

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

Therefore, we have $\tilde{h}_i \in \mathbb{R}[\overline{\mathcal{A}_i}]$. Moreover, it follows from Lemma 4.2 and (4.3) that

$$(4.4) \quad \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} - \frac{\varepsilon}{N-1} \leq \tilde{h}_i(\mathbf{x}) \leq \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})}, \quad \forall \mathbf{x} \in \mathbf{K}.$$

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

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

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

Proof. It is straightforward to verify from the definitions. \blacksquare

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

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

We now point out that (SP) cannot be viewed as a consequence of (P) equipped with invariant measures. Indeed, the latter would involve additional constraints

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

since for any $\alpha \in \overline{\mathcal{A}_1} \setminus \bigcap_{i=2}^N \overline{\mathcal{A}_i}$, there exists $j \in [N] \setminus \{1\}$ such that $\alpha \notin \overline{\mathcal{A}_j}$ which implies $\int_{\mathbf{K}} \mathbf{x}^\alpha q_1 d\mu_1 = \int_{\mathbf{K}} \mathbf{x}^\alpha q_j d\mu_j = 0$ by Lemma 4.4. 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

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

It could be easily seen that $B_k^{\mathcal{A}}$ is block-diagonal up to appropriate row/column permutations [29]. We define the sparse quadratic module $Q_k(\mathbf{g}, \mathcal{A})$ associated with \mathcal{A} by

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

where $d_0 := 0$, for $s \in \mathbb{N}$, $[\mathbf{x}]_s := [1, x_1, x_2, \dots, x_n^s]$ 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

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

The following theorem is a sign symmetry adapted version of Theorem 2.2.

Theorem 4.5. [29, Theorem 6.11] *Let $f \in \mathbb{R}[\mathbf{x}]$ and $\mathcal{A} = \text{supp}(f) \cup \bigcup_{j=1}^m \text{supp}(g_j)$. 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 4.6. By a similar argument as for Theorem 4.5 (see [29]), one can actually show that if $f \in Q_k(\mathbf{g})$, then $f \in Q_k(\mathbf{g}, \mathcal{A})$ with $\mathcal{A} = \text{supp}(f) \cup \bigcup_{j=1}^m \text{supp}(g_j)$.

Theorem 4.7. *The following statements hold true:*

- (i) For $k \geq d_{\min}$, $\rho_k^s \leq \rho_k$. Moreover, if $\mathcal{R}_1 = \mathcal{R}_2 = \dots = \mathcal{R}_N$, then $\rho_k^s = \rho_k$.
- (ii) Under Assumption 1.1 and the Archimedean condition, $\rho_k^s \nearrow \rho$ as $k \rightarrow \infty$.

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 $\mathcal{R}_1 = \mathcal{R}_2 = \dots = \mathcal{R}_N$. To show $\rho_k^s \geq \rho_k$, 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 4.6, we deduce that

$$(4.5) \quad p_i(\mathbf{x}) - \tilde{h}_i(\mathbf{x}) q_i(\mathbf{x}) \in Q_k(\mathbf{g}, \mathcal{A}_1).$$

Moreover, from $p_1(\mathbf{x}) + \left(\sum_{i=2}^N h_i(\mathbf{x}) - c \right) q_1(\mathbf{x}) \in Q_k(\mathbf{g})$ and Remark 4.6 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 $(\text{SD}k)$, which implies $\rho_k^s \geq \rho_k$.

(ii). It is clear that $\rho_k^s \leq \rho_{k+1}^s$ 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 4.3. 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

$$(4.6) \quad p_i(\mathbf{x}) - h_i(\mathbf{x})q_i(\mathbf{x}) = p_i(\mathbf{x}) - \bar{h}_i(\mathbf{x})q_i(\mathbf{x}) + \frac{\bar{\varepsilon}}{N-1}q_i(\mathbf{x}) \geq \frac{\bar{\varepsilon}}{N-1}q_i(\mathbf{x}) > 0.$$

By Theorem 4.5, 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 $p_i - h_iq_i \in Q_{k_i}(\mathbf{g}, \mathcal{A}_i)$. 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 4.5 again, there exists $k_1 \in \mathbb{N}$ such that

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

Let $k_\varepsilon := \max_{1 \leq i \leq N} k_i$. Then $(\rho - \varepsilon, h_2, \dots, h_N)$ is feasible to $(\text{SD}k)$ with $k = k_\varepsilon$, which implies $\rho_{k_\varepsilon}^s \geq \rho - \varepsilon$. As ε is arbitrary, we prove the convergence of $\{\rho_k^s\}$ to ρ . \blacksquare

Note that the optimum ρ_k^s of $(\text{SD}k)$ may depend on which ratio being chosen as p_1/q_1 . We give an example to illustrate this phenomenon.

Example 4.8. 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 $\{(x, y, z) \in \mathbb{R}^3 \mid 1 - x^2 - y^2 - z^2 \geq 0\}$. We consider three cases: (1) $p_1 = x^2 + y^2 - yz$, $q_1 = 1 + 2x^2 + y^2 + z^2$; (2) $p_1 = y^2 + x^2z$, $q_1 = 1 + x^2 + 2y^2 + z^2$; (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 4.9. To guarantee $\rho_k^s = \rho_k$, one may simply letting $\mathcal{A}_1 = \dots = \mathcal{A}_N$ by Theorem 4.7.

Table 1
Computational results for Example 4.8.

| k | $\sup(\mathbf{D}k)$ | $\sup(\mathbf{SD}k)$ | | |
|-----|---------------------|----------------------|----------|----------|
| | | 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 |

Finally, the dual of $(\mathbf{SD}k)$ reads as

$$(\text{SP}k) \quad \begin{cases} \inf_{\mathbf{y}_i} \sum_{i=1}^N L_{\mathbf{y}_i}(p_i) \\ \text{s.t. } B_k^{\mathbf{A}_i} \circ \mathbf{M}_k(\mathbf{y}_i) \succeq 0, \quad i \in [N], \\ B_{k-d_j}^{\mathbf{A}_i} \circ \mathbf{M}_{k-d_j}(g_j \mathbf{y}_i) \succeq 0, \quad i \in [N], j \in [m], \\ L_{\mathbf{y}_1}(q_1) = 1, \\ L_{\mathbf{y}_i}(\mathbf{x}^\alpha q_i) = L_{\mathbf{y}_1}(\mathbf{x}^\alpha q_1), \quad \forall \alpha \in \mathbb{N}_{2k-\max\{\deg(q_1), \deg(q_i)\}}^n \cap \overline{\mathcal{A}_i}, i \in [N] \setminus \{1\}. \end{cases}$$

4.2. Sparse SDP relaxations by exploiting both correlative sparsity and sign symmetries. In this subsection, we present sparse SDP relaxations for (\mathbf{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 (\mathbf{SRFO}) proposed in [3]. We first describe the correlative sparsity pattern in (\mathbf{SRFO}) . For $I \subseteq [n]$, $\mathbf{x}(I)$ denotes the set of variables x_i with $i \in I$.

Assumption 4.10. 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 4.11. If \mathbf{K} is compact and one knows some $M > 0$ such that $M - x_1^2 - \dots - x_n^2 \geq 0$ for all $\mathbf{x} \in \mathbf{K}$, then we can always add redundant constraints $M - \sum_{j \in I_i} x_j^2 \geq 0, i \in [N]$ to the description of \mathbf{K} so that Assumption 4.10 (iii) is satisfied.

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

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

and $\pi_i: \mathcal{M}(\mathbf{K})_+ \rightarrow \mathcal{M}(\mathbf{K}_i)_+$ 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 $i \neq j$ and $I_i \cap I_j \neq \emptyset$, 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 $\pi_{ij} : \mathcal{M}(\mathbf{K}_i)_+ \rightarrow \mathcal{M}(\mathbf{K}_{ij})_+$ is defined in an obvious similar manner. Moreover, let us define the sets

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

and

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

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

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

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

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

Theorem 4.12. [3, Theorems 3.1 and 3.2] *Let Assumption 4.10 hold and assume that $q_i > 0$ on \mathbf{K}_i for $i \in [N]$. It hold $\rho^c = \rho$ and $\rho_k^c \nearrow \rho$ as $k \rightarrow \infty$.*

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

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

Proposition 4.13. *Under Assumption 4.10 (iii), the strong duality holds between (CP k) and (CD k) for all $k \geq d_{\min}$.*

Proof. Let $\{\mu_i\}$ be a feasible solution of (CP) (see the proof of [3, Theorem 3.1] for the existence). For each $i \in [N]$, let $\mathbf{y}_i = (y_{i\alpha})_{\alpha \in \mathbb{N}_{2k}^{I_i}}$ be such that $y_{i\alpha} = \int_{\mathbf{K}_i} \mathbf{x}^\alpha d\mu_i$ for $\alpha \in \mathbb{N}_{2k}^{I_i}$. Then, $\{\mathbf{y}_i\}$ is a feasible solution of (CP k). As Assumption 4.10 (iii) holds, according to the proof of [3, Theorem 2.2], the feasible set of (CP k) is compact, which implies that the optimal solution set of (CP k) is nonempty and bounded. Therefore, (CD k) is strictly feasible [28] and the strong duality holds [26, Theorem 4.1.3]. ■

Let

$$(4.7) \quad \mathcal{A} := \bigcup_{i=1}^N (\text{supp}(p_i) \cup \text{supp}(q_i)) \cup \bigcup_{j=1}^m \text{supp}(g_j),$$

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

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

Theorem 4.14. *Let Assumption 4.10 hold and assume that $q_i > 0$ on \mathbf{K}_i for $i \in [N]$. It holds $\rho^{\text{cs}} = \rho$.*

Proof. Since $\mathcal{M}(\mathbf{K}_i)_+^{\mathcal{R}}$ is a subset of $\mathcal{M}(\mathbf{K}_i)_+$ for each $i \in [N]$, we immediately have $\rho^{\text{cs}} \geq \rho$. For the converse, fix an arbitrary $\varepsilon > 0$ and suppose that $\{\mu_i\}$ is any feasible solution of (CP) with $\sum_{i=1}^N \int_{\mathbf{K}_i} p_i d\mu_i < \rho + \varepsilon$. 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 4.4, one can check that $\pi_{ij}(q_i d\tilde{\mu}_i) = \pi_{ji}(q_j d\tilde{\mu}_j)$. Moreover, by Lemma 4.4 again, $\sum_{i=1}^N \int_{\mathbf{K}_i} p_i d\tilde{\mu}_i = \sum_{i=1}^N \int_{\mathbf{K}_i} p_i d\mu_i < \rho + \varepsilon$. As $\varepsilon > 0$ is arbitrary, $\rho^{\text{cs}} \leq \rho$. ■

For each $i \in [N]$ and each $k \in \mathbb{N}$, define a binary matrix $B_{i,k}^A$ 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

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

The sign symmetry adapted version of (CP k) is given by

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

Theorem 4.15. *Let Assumption 4.10 hold and assume that $q_i > 0$ on \mathbf{K}_i for $i \in [N]$. It holds $\rho_k^{\text{CS}} = \rho_k^c$ for all $k \geq d_{\min}$. Consequently, $\rho_k^{\text{CS}} \nearrow \rho$ as $k \rightarrow \infty$.*

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

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

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

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

The dual of (CSP k) reads as

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

Proposition 4.16. *Under Assumption 4.10 (iii), the strong duality holds between (CSP k) and (CSD k) for all $k \geq d_{\min}$.*

Proof. Note that for any feasible solution $\{\mathbf{y}_i\}$ of (CSP k), only the entries $y_{i\alpha}$, $\alpha \in \overline{\mathcal{A}}$, are involved in (CSP k). By the proof of Theorem 4.15, the corresponding point $\{\mathbf{y}'_i\}$ is feasible to (CP k). Then, the proof of Proposition 4.13 indicates that the feasible set of (CSP k) is compact. Hence, the strong duality holds between (CSP k) and (CSD k) as proved in Proposition 4.13. \blacksquare

5. Numerical experiments. In this section, we conduct numerical experiments to test the performance of the sign symmetry adapted approaches against the approaches in [3]. We use the Julia package TSSOS¹ [14] to build the SDP relaxations and rely on MOSEK [2] to solve them. Throughout this section, k stands for the relaxation order. All numerical experiments were performed on a desktop computer with Intel(R) Core(TM) i9-10900 CPU@2.80GHz and 64G RAM.

5.1. Comparison of (SD k) with (D k). We begin by presenting three examples without correlative sparsity.

Example 5.1. Consider the problem

$$(5.1) \quad \min_{\mathbf{x} \in \mathbb{R}^3} \sum_{a=1/M}^{1-1/M} \frac{p_a(\mathbf{x})}{q_a(\mathbf{x})} \quad \text{s.t. } x_1^2 + x_2^2 + x_3^2 = 3,$$

where

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

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 ten zeros: $\{(1, \pm 1, \pm 1), (\pm a, 1, 0), (0, \pm a, 1), (1, 0, \pm a)\}$ [23]. Then, it is obvious that the optimum of (5.1) is $M - 1$. Moreover, we can rewrite $q_a(\mathbf{x})$ as

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

Then, by the arithmetic-geometric inequality, we see that $q_a(\mathbf{x}) > 0$ for all feasible points \mathbf{x} . Hence, (5.1) satisfies Assumption 1.1 for all positive integers M and d . For $M \in \{6, 8, 10\}$ and $d \in \{2, 3, 4\}$, we solve (5.1) using the minimum relaxation order $k = 3d$ and present the computational results in Table 2. From the table, we see that by exploiting sign symmetries, we gain speedup by $1 \sim 2$ magnitudes.

Example 5.2. Consider the following problem adapted from [3]:

$$(5.2) \quad \min_{\mathbf{x} \in \mathbb{R}^n} - \sum_{i=1}^N \frac{1}{\sum_{j=1}^n (x_j^2 - a_{ij})^2 + c_i} \quad \text{s.t. } 60 - \sum_{i=1}^n (x_i^2 - 5)^2 \geq 0.$$

The data $\{a_{ij}\}, \{c_i\}$ are given in [1, Table 16] with $N = 30, n \in \{5, 10\}$. For $n = 5$, the optimum of (5.2) is -10.4056 , and for $n = 10$, the optimum is -10.2088 . We present the

¹TSSOS is publically available at <https://github.com/wangjie212/TSSOS>.

Table 2
Computational results for (5.1).

| M | sup(D k)/Time | | | sup(SD k)/Time | | |
|-----|------------------|------------|------------|-------------------|------------|------------|
| | $d = 2$ | $d = 3$ | $d = 4$ | $d = 2$ | $d = 3$ | $d = 4$ |
| 6 | 5.00/0.31s | 5.00/3.85s | 5.00/36.1s | 5.00/0.06s | 5.00/0.38s | 5.00/1.17s |
| 8 | 7.00/0.43s | 7.00/4.25s | 7.00/52.4s | 7.00/0.09s | 7.00/0.54s | 7.00/1.84s |
| 10 | 9.00/0.54s | 9.00/5.34s | 9.00/126s | 9.00/0.12s | 9.00/0.71s | 9.00/2.43s |

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) by exploiting sign symmetries, we gain speedup by $1 \sim 3$ magnitudes.

Table 3
Computational results for (5.2). The symbol ‘-’ indicates that MOSEK runs out of memory.

| n | sup(D k)/Time | | | sup(SD k)/Time | | |
|-----|------------------|----------------|---------------|-------------------|----------------|----------------|
| | $k = 3$ | $k = 4$ | $k = 5$ | $k = 3$ | $k = 4$ | $k = 5$ |
| 5 | -49.4493/3.53s | -12.4737/27.8s | -10.4056/306s | -49.4493/0.19s | -12.4736/1.52s | -10.4056/5.66s |
| 10 | $k = 2$ | $k = 3$ | $k = 4$ | $k = 2$ | $k = 3$ | $k = 4$ |
| | -25.2408/4.68 | -19.2391/1158s | - | -25.2408/0.11 | -19.2391/3.12s | -10.2088/46.8s |

Example 5.3. Now we carry out randomly generated problems of the form

$$(5.3) \quad \min_{\mathbf{x} \in \mathbb{R}^n} - \sum_{i=1}^N \frac{1}{f_i^2 + 1} \quad \text{s.t.} \quad x_1^2 + \cdots + x_n^2 \leq 1,$$

which are constructed as follows: 1) randomly choose a subset of nonconstant monomials \mathcal{M} from $[\mathbf{x}]_d$ with prescribed probability ξ ; 2) randomly assign 3 elements in \mathcal{M} to each f_i coupled with random coefficients in $[0, 1]$. It is clear that the optimal value of (5.3) is $-N$. We denote the set of (5.3) generated in such a way by $\text{RANDSRFO}(N, n, d, \xi)$ and construct the following 12 instances²:

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

We solve those instances using the minimum relaxation order $k = d$ and present the computational results in Table 4. From the table, we see that by exploiting sign symmetries, we gain speedup by $1 \sim 2$ magnitudes.

²They are available at <https://wangjie212.github.io/jiewang/code.html>.

Table 4

Computational results for (5.3). The symbol ‘-’ indicates that MOSEK runs out of memory.

| | P_1 | P_2 | P_3 | P_4 | P_5 | P_6 |
|------------------------|--------------|--------------|--------------|--------------|--------------|--------------|
| sup(DK)/Time | -10.00/24.9s | -10.00/22.7s | -10.00/20.5s | -8.000/19.7s | -8.000/19.3s | -8.000/19.3s |
| sup(SDK)/Time | -10.00/1.97s | -10.00/0.92s | -10.00/0.89s | -8.000/1.57s | -8.000/1.41s | -8.000/1.61s |
| | P_7 | P_8 | P_9 | P_{10} | P_{11} | P_{12} |
| sup(DK)/Time | -6.000/311s | -6.000/331s | -6.000/370s | - | - | - |
| sup(SDK)/Time | -6.000/3.04s | -6.000/2.18s | -6.000/2.79s | -12.00/19.7s | -12.00/19.9s | -12.00/19.8s |

5.2. Comparison of (CDk) with (CSDk). We will proceed with two examples for which both correlative sparsity and sign symmetries are present.

Example 5.4. Consider

$$(5.4) \quad \min_{\mathbf{x} \in \mathbb{R}^{2N+2}} \sum_{i=1}^N \frac{(x_{2i-1}^2 + x_{2i}^2 + x_{2i+1}^2)x_{2i-1}^2 x_{2i}^2 x_{2i+1}^2 + x_{2i+2}^8}{x_{2i-1}^2 x_{2i}^2 x_{2i+1}^2 x_{2i+2}^2}$$

$$\text{s.t. } x_{2i-1}^2 + x_{2i}^2 + x_{2i+1}^2 + x_{2i+2}^2 = 4, \quad i \in [N].$$

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 $(\pm 1, \pm 1, \pm 1, \pm 1)$ [15]. Therefore, the optimum of (5.4) equals $4N$. It is clear that Assumption 4.10 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 (5.4) and present the computational results in Table 5. It can be seen from the table that by exploiting sign symmetries, we gain extra speedup by several times.

Table 5

Computational results for (5.4). The symbol ‘*’ indicates that MOSEK encounters numerical issues.

| N | sup(CDk)/Time | | sup(CSDk)/Time | |
|-----|------------------------|-------------|-------------------------|-------------|
| | $k = 4$ | $k = 5$ | $k = 4$ | $k = 5$ |
| 20 | 16.65*/1.38s | 80.00/10.6s | 19.06*/0.49s | 80.00/0.73s |
| 40 | 37.71*/3.25s | 160.0/15.8s | 37.71*/0.62s | 160.0/1.79s |
| 60 | 52.17*/6.97s | 240.0/28.9s | 57.75*/1.51s | 240.0/3.43s |

Example 5.5. Consider

$$(5.5) \quad \min_{\mathbf{x} \in \mathbb{R}^{2N+1}} \sum_{i=1}^N \frac{p_i(\mathbf{x})}{q_i(\mathbf{x})} \quad \text{s.t. } x_{2i-1}^2 + x_{2i}^2 + x_{2i+1}^2 = 3, \quad i \in [N],$$

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 5.1, it is easy to see that the optimum of (5.5) equals N . Moreover, Assumption 4.10 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 (5.5) using the minimum relaxation order $k = 3d$ and present the computational results in Table 6. Again, it can be seen from the table that by exploiting sign symmetries, we gain extra significant speedup.

Table 6

Computational results for (5.5).

| N | sup(CD k)/Time | | | sup(CSD k)/Time | | |
|-----|-------------------|-------------|-------------|--------------------|-------------|-------------|
| | $d = 2$ | $d = 3$ | $d = 4$ | $d = 2$ | $d = 3$ | $d = 4$ |
| 5 | 5.000/0.28s | 5.000/2.58s | 5.000/23.1s | 5.000/0.06s | 5.000/0.34s | 5.000/1.06s |
| 10 | 10.00/0.56s | 10.00/4.35s | 10.00/44.1s | 10.00/0.12s | 10.00/0.80s | 10.00/2.05s |
| 20 | 20.00/1.17s | 20.00/9.50s | 20.00/95.8s | 20.00/0.25s | 15.00/1.62s | 20.00/4.40s |

5.3. Comparison of (CD k) with (CSD k) and the epigraph approach. In this subsection, we further compare (CD k) with the epigraph approach which translates (SRFO) into a polynomial optimization problem:

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

Note that (SEA) inherits the correlative sparsity and sign symmetries of (SRFO) which can be exploited to obtain sparse SDP relaxations for (SEA) [30].

Example 5.6. Consider

$$(5.6) \quad \max_{\mathbf{x} \in \mathbb{R}^{N+1}} \sum_{i=1}^N \frac{1}{100(x_{i+1}^2 - x_i^2)^2 + (x_i^2 - 1)^2 + 1} \quad \text{s.t. } 16 - x_i^2 \geq 0, \quad i \in [N+1],$$

which is modified from the well-known Rosenbrock problem. Clearly, the optimum of (5.6) equals N and Assumption 4.10 holds with $I_i = J_i = \{i, i+1\}$, $i \in [N]$. We solve (5.6) for $N \in \{100, 200, 400\}$. It turns out that the SDP relaxations (CD k) and (CSD k) attain global optimality at $k = 2$ while the epigraph approach attains global optimality at $k = 4$. We present the computational results in Table 7, from which we can see that the sparse SDP relaxation (CSD k) is the most efficient approach.

Table 7

Computational results for (5.6). ‘EPI’ indicates the lower bounds obtained by the epigraph approach.

| N | EPI/Time | sup(CD k)/Time | sup(CSD k)/Time |
|-----|-------------|-------------------|--------------------|
| | $k = 4$ | $k = 2$ | $k = 2$ |
| 100 | 100.0/2.30s | 100.0/0.15s | 100.0/0.06s |
| 200 | 200.0/6.94s | 200.0/0.58s | 200.0/0.21s |
| 400 | 400.0/20.8s | 400.0/1.72s | 400.0/0.97s |

Example 5.7. Consider

$$(5.7) \quad \min_{\mathbf{x} \in \mathbb{R}^{N+s}} \sum_{i=1}^N \frac{\sum_{j=1}^s x_{i+j-1} x_{i+j}}{1 + \sum_{j=1}^{s+1} j x_{i+j-1}^2} \quad \text{s.t. } 1 - x_i^2 \geq 0, \quad i \in [N+s],$$

for some $s \geq 1$. It is clear that Assumption 4.10 holds with $I_i = J_i = \{i, i+1, \dots, i+s\}$, $i \in [N]$. We solve (5.7) 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 and moreover, the sparse SDP relaxation (CSD k) is the most efficient approach.

Table 8

Computational results for (5.7). ‘EPI’ indicates the lower bounds obtained by the epigraph approach.

| N | EPI/Time | sup(CD k)/Time | sup(CSD k)/Time |
|-----|---------------|-------------------|--------------------|
| | $k = 3$ | $k = 3$ | $k = 3$ |
| 20 | -4.5892/23.9s | -4.6173/61.6s | -4.6173/23.3s |
| 40 | -8.7429/62.0s | -8.8778/105s | -8.8778/52.8s |
| 60 | -12.901/100s | -13.138/148s | -13.138/70.8s |

Finally, we would like to mention that the epigraph approach behaves very badly on the examples presented in the previous subsections.

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 multi-user MIMO system, the base station can multiplex signals intended to 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. [21] 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

$$(6.1) \quad \max_{\mathbf{z} \in \mathbb{C}^n} \sum_{i=1}^N \frac{\mathbf{z}^H A_i \mathbf{z}}{\mathbf{z}^H B_i \mathbf{z}} \quad \text{s.t.} \quad \|\mathbf{z}\|^2 = 1,$$

where $A_i, B_i \in \mathbb{C}^{n \times n}$ are positive semidefinite Hermitian matrices, \mathbf{z}^H denotes the conjugate transpose of \mathbf{z} , n is the 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 (6.1) also appears in the sparse Fisher discriminant analysis in pattern recognition (see [4, 5, 32]). The single generalized Rayleigh quotient optimization problem corresponds to the classical eigenvalue problem and can be solved in polynomial time [20]. However, when $N \geq 2$, solving (6.1) is much more challenging and requires sophisticated techniques (see [16, 31, 34] for the real case of $N = 2$). In particular, Primolevo et al. [21] restricted \mathbf{z} in (6.1) to be columns of an identity matrix to give a suboptimal solution.

Here, we convert (6.1) into a real problem by specifying the real and imaginary parts of the complex variables and coefficients appearing in (6.1). 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 (6.1) becomes

$$(6.2) \quad \max_{\mathbf{x}, \mathbf{y} \in \mathbb{R}^n} \sum_{i=1}^N \frac{\mathbf{x}^\top A_{i,1} \mathbf{x} - 2\mathbf{x}^\top A_{i,2} \mathbf{y} + \mathbf{y}^\top A_{i,1} \mathbf{y}}{\mathbf{x}^\top B_{i,1} \mathbf{x} - 2\mathbf{x}^\top B_{i,2} \mathbf{y} + \mathbf{y}^\top B_{i,1} \mathbf{y}} \quad \text{s.t.} \quad \|\mathbf{x}\|^2 + \|\mathbf{y}\|^2 = 1.$$

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

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

We solve one instance of (6.2) with (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 by exploiting sign symmetries, we gain around twice speedup.

Table 9
Computational results for (6.2).

| (n, N) | sup(Dk)/Time | | sup(SDk)/Time | |
|----------|--------------|-------------|---------------|-------------|
| | $k = 2$ | $k = 3$ | $k = 2$ | $k = 3$ |
| (3, 20) | 406.7/0.61s | 406.7/12.7s | 406.7/0.38s | 406.7/5.91s |
| (4, 10) | 26.54/1.03s | 20.10/107s | 26.54/0.67s | 20.10/50.5s |
| (5, 5) | 20.12/2.43s | 19.11/531s | 20.12/1.57s | 19.11/339s |

REFERENCES

- [1] M. M. ALI, C. KHOMPATRAPORN, AND Z. B. ZABINSKY, *A numerical evaluation of several stochastic algorithms on selected continuous global optimization test problems*, Journal of global optimization, 31 (2005), pp. 635–672.
- [2] E. D. ANDERSEN AND K. D. ANDERSEN, *The Mosek Interior Point Optimizer for Linear Programming: An Implementation of the Homogeneous Algorithm*, in High Performance Optimization, H. Frenk, K. Roos, T. Terlaky, and S. Zhang, eds., vol. 33 of Applied Optimization, Springer US, 2000, pp. 197–232.
- [3] F. BUGARIN, D. HENRION, AND J. B. LASSERRE, *Minimizing the sum of many rational functions*, Mathematical Programming Computation, 8 (2016), pp. 83–111.
- [4] M. M. DUNDAR, G. FUNG, J. BI, S. SATHYAKAMA, AND B. RAO, *Sparse Fisher discriminant analysis for computer aided detection*, in Proceedings of the 2005 SIAM International Conference on Data Mining (SDM), H. Kargupta, C. Kamath, J. Srivastava, and A. Goodman, eds., Philadelphia, PA, 2005, Society for Industrial and Applied Mathematics, pp. 476–480.
- [5] E. FUNG AND M. NG, *On sparse Fisher discriminant method for microarray data analysis*, Bioinformatics, 2 (2007), pp. 230–234.
- [6] F. GUO, L. WANG, AND G. ZHOU, *Minimizing rational functions by exact jacobian SDP relaxation applicable to finite singularities*, Journal of Global Optimization, 58 (2014), pp. 261–284.
- [7] R. HARTLEY, F. KAHL, C. OLSSON, AND Y. SEO, *Verifying global minima for l_2 minimization problems in multiple view geometry*, International Journal of Computer Vision, 101 (2013), pp. 288–304.
- [8] D. JIBETEAN AND E. DE KLERK, *Global optimization of rational functions: a semidefinite programming approach*, Math. Program., 106 (2006), pp. 93–109.
- [9] F. KAHL, S. AGARWAL, M. K. CHANDRAKER, D. KRIEGMAN, AND S. BELONGIE, *Practical global optimization for multiview geometry*, International Journal of Computer Vision, 79 (2008), pp. 271–284.
- [10] M. KORDA AND D. HENRION, *Convergence rates of moment-sum-of-squares hierarchies for volume approximation of semialgebraic sets*, Optimization Letters, 12 (2018), pp. 435–442.
- [11] M. KORDA, D. HENRION, AND C. N. JONES, *Convergence rates of moment-sum-of-squares hierarchies for optimal control problems*, Systems & Control Letters, 100 (2017), pp. 1–5.
- [12] J. B. LASSERRE, V. MAGRON, S. MARX, AND O. ZAHM, *Minimizing rational functions: A hierarchy of approximations via pushforward measures*, SIAM Journal on Optimization, 31 (2021), pp. 2285–2306.
- [13] J. LÖFBERG, *Pre- and post-processing sum-of-squares programs in practice*, IEEE Transactions on Automatic Control, 54 (2009), pp. 1007–1011.
- [14] V. MAGRON AND J. WANG, *TSSOS: a Julia library to exploit sparsity for large-scale polynomial optimization*, in The 16th Effective Methods in Algebraic Geometry Conference, 2021.
- [15] T. MOTZKIN, *The arithmetic-geometric inequality*, in Inequalities, O. Shisha, ed., New York, 1967, Academic Press, pp. 205–224.
- [16] V.-B. NGUYEN, R.-L. SHEU, AND Y. XIA, *Maximizing the sum of a generalized Rayleigh quotient and another Rayleigh quotient on the unit sphere via semidefinite programming*, Journal of Global Optimization, 64 (2016), pp. 399–416.
- [17] J. NIE, *An exact jacobian sdp relaxation for polynomial optimization*, Mathematical Programming, 137 (2013), pp. 225–255.
- [18] J. NIE, J. DEMMEL, AND M. GU, *Global minimization of rational functions and the nearest GCDs*, Journal of Global Optimization, 40 (2008), pp. 697–718.
- [19] J. NIE AND M. SCHWEIGHOFER, *On the complexity of Putinar’s positivstellensatz*, Journal of Complexity, 23 (2007), pp. 135 – 150.
- [20] B. N. PARLETT, *The Symmetric Eigenvalue Problem*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1998.
- [21] G. PRIMOLEVO, O. SIMEONE, AND U. SPAGNOLINI, *Towards a joint optimization of scheduling and beamforming for MIMO downlink*, in 2006 IEEE Ninth International Symposium on Spread Spectrum Techniques and Applications, 2006, pp. 493–497.
- [22] M. PUTINAR, *Positive polynomials on compact semi-algebraic sets*, Indiana University Mathematics Journal, 42 (1993), pp. 969–984.
- [23] B. REZNICK, *Some concrete aspects of Hilbert’s 17th problem*, in Contemporary Mathematics, vol. 253, American Mathematical Society, 2000, pp. 251–272.
- [24] S. SCHAIBLE AND J. SHI, *Fractional programming: The sum-of-ratios case*, Optimization Methods and

- Software, 18 (2003), pp. 219–229.
- [25] A. SHAPIRO, *Semi-infinite programming, duality, discretization and optimality conditions*, Optimization, 58 (2009), pp. 133–161.
- [26] A. SHAPIRO AND K. SCHEINBER, *Duality, optimality conditions and perturbation analysis*, in Handbook of Semidefinite Programming - Theory, Algorithms, and Applications, H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., Kluwer Academic Publisher, Boston, 2000, pp. 67–110.
- [27] A. TIMAN, *Theory of Approximation of Functions of a Real Variable*, International Series of Monographs on Pure and Applied Mathematics, Pergamon, 1963.
- [28] M. TRNOVSKÁ, *Strong duality conditions in semidefinite programming*, Journal of Electrical Engineering, 56 (2005), pp. 1–5.
- [29] J. WANG, V. MAGRON, AND J.-B. LASSERRE, *TSSOS: A moment-SOS hierarchy that exploits term sparsity*, SIAM Journal on Optimization, 31 (2021), pp. 30–58.
- [30] J. WANG, V. MAGRON, J. B. LASSERRE, AND N. H. A. MAI, *Cs-tssos: Correlative and term sparsity for large-scale polynomial optimization*, ACM Transactions on Mathematical Software, 48 (2022), pp. 1–26.
- [31] X. WANG, L. WANG, AND Y. XIA, *An efficient global optimization algorithm for maximizing the sum of two generalized Rayleigh quotients*, Computational and Applied Mathematics, 37 (2018), pp. 4412–4422.
- [32] M. C. WU, L. ZHANG, Z. WANG, D. C. CHRISTIANI, AND X. LIN, *Sparse linear discriminant analysis for simultaneous testing for the significance of a gene set/pathway and gene selection*, Bioinformatics, 25 (2009), pp. 1145–1151.
- [33] W.-Y. WU, R.-L. SHEU, AND Ş. İ. BIRBIL, *Solving the sum-of-ratios problem by a stochastic search algorithm*, Journal of Global Optimization, 42 (2008), pp. 91–109.
- [34] L.-H. ZHANG, *On optimizing the sum of the rayleigh quotient and the generalized rayleigh quotient on the unit sphere*, Computational Optimization and Applications, 54 (2013), pp. 111–139.