J Glob Optim (2008) 40:697–718
DOI 10.1007/s10898-006-9119-8

ORIGINAL PAPER

Global minimization of rational functions and the nearest GCDs

Jiawang Nie · James Demmel · Ming Gu

Received: 17 October 2006 / Accepted: 25 October 2006 / Published online: 16 December 2006
© Springer Science+Business Media B.V. 2006

Abstract This paper discusses the global minimization of rational functions with or without constraints. We propose sum of squares relaxations to solve these problems, and study their properties. Some special features are discussed. First, we consider minimization of rational functions without constraints. Second, as an application, we show how to find the nearest common divisors of polynomials via unconstrained minimization of rational functions. Third, we discuss minimizing rational functions under some constraints which are described by polynomials.

Keywords Rational function · Polynomial · Global minimization · Sum of squares (SOS) · Greatest common divisor · Quadratic module

1 Introduction

Consider the problem of minimizing a rational function

$$r^* = \min_{x \in \mathbb{R}^n} \quad r(x) := \frac{f(x)}{g(x)} \tag{1.1}$$

$$s.t. \quad h_1(x) \geq 0, \ldots, h_m(x) \geq 0, \tag{1.2}$$

where $f(x)$, $g(x)$, $h_i(x) \in \mathbb{R}[X]$. Here $\mathbb{R}[X]$ is the ring of real polynomials in $X = (x_1, \ldots, x_n)$. Our goal is to find the global minimum r^* of the rational function $r(x)$,

J. Nie (✉)· M. Gu
Department of Mathematics, University of California, Berkeley, CA 94720, USA
e-mail: njw@math.berkeley.edu

J. Demmel
Department of Math & EECS, University of California, Berkeley, CA 94720, USA
e-mail: demmel@cs.berkeley.edu

M. Gu
e-mail: mgu@math.berkeley.edu

and if possible, one or more global minimizer(s) x^* such that $r(x^*) = r^*$. This contains a broad class of nonlinear global optimization problems. Without loss of generality, assume that $g(x)$ is not identically zero and is nonnegative on the feasible set, otherwise we can replace $\frac{f(x)}{g(x)}$ by $\frac{f(x)g(x)}{g^2(x)}$.

When $n = 1$ and there are no constraints, i.e., the case of one-dimensional unconstrained minimization, the problem is simple. As we can see, γ is a lower bound for $r(x)$ if and only if the univariate polynomial $f(x) - \gamma g(x)$ is nonnegative, i.e.,

$$f(x) - \gamma \cdot g(x) \geq 0 \quad \forall x \in \mathbb{R}.$$

As is well-known, a univariate polynomial is nonnegative if and only if it can be written as sum of squares (SOS) of polynomials [28]. This poses a convex condition [24,32] (actually it is a linear matrix inequality (LMI)) on γ for given $f(x), g(x)$. Thus the problem (1.1) can be reformulated as maximizing γ subject to a particular LMI. Therefore the problem (1.1) can be solved efficiently as a semidefinite program (SDP) [5,37].

However, when $n > 1$, the problem (1.1) can be very hard even if there are no constraints, which is due to the difficulty that a nonnegative multivariate polynomial might not be a sum of squares of polynomials [28]. Even in the special case that $\deg(f) = 4$ and $\deg(g) = 0$, that is, $r(x)$ becomes a multivariate polynomial of degree 4, to find its global minimum is NP-hard, as mentioned in Nesterov [19]. So we need some approximations of nonnegative polynomials to find an approximate minimum value (often a guaranteed lower bound) and extract approximate minimizer(s). One frequently used technique in polynomial optimization is to approximate nonnegative polynomials by sum of squares of polynomials, i.e., SOS relaxations. We refer to [15,23,24].

To test nonnegativity of a general polynomial of degree 4 or higher is NP-hard in n, the number of variables. For instance, for a given generic $n - by - n$ symmetric matrix $A = (a_{ij})_{n,n}$, to test whether the quartic homogenous polynomial

$$[x^2]^T A [x^2] := \sum_{i,j=1}^{n} a_{ij} x_i^2 x_j^2$$

is nonnegative is NP-hard [7]. The matrix A is said to be *co-positive* if $x^T Ax$ is nonnegative for every nonnegative vector $x \in \mathbb{R}^n$. However, to test whether a polynomial is a sum of squares of polynomials can be determined efficiently by solving a semidefinite program [23,24]. Recently, there has been much work on finding the global minimum of polynomial functions via sum of squares (SOS) relaxations (also called SDP or LMI relaxation). We refer to [11,15,21,23,24] and the references therein for work in this area. Our goal is to use SOS relaxations to solve the global minimization problems (1.1) and (1.2).

Recently, Jibetean and De Klerk [10] discussed SOS methods to minimize rational functions. They showed how to get lower bounds of rational functions by applying SOS relaxations in a natural way. To be more precise, for problems (1.1) and (1.2), they proposed to find a lower bound by computing maximum γ such that

$$f(x) - \gamma\, g(x) \text{ is a sum of squares about } x.$$

The above condition essentially poses a particular LMI constraint for γ. So such a maximum γ can be found by solving a particular SDP. However, [10] did not discuss on how to find the global minimizer(s), which is often more interesting than lower bounds. In this paper, we revisit the SOS methods for problems (1.1) and (1.2), and discuss the special features of SOS relaxations and their dual problems. Interestingly, from the optimal dual solutions, we find that the minimizers can be extracted. As an application, we show how to find nearest common divisors by globally minimizing a rational function.

Throughout this paper, we will use the following notations. $\mathbb{R}$ ($\mathbb{C}$) is the field of real (complex) numbers. $\mathbb{N}$ is the set of nonnegative integers. For any complex number z, $\bar{z}$ denotes its complex conjugate. For any integer vector $\alpha \in \mathbb{N}^n$, define $x^\alpha := x_1^{\alpha_1} \cdots x_n^{\alpha_n}$ and $|\alpha| := \alpha_1 + \cdots + \alpha_n$. $\sum \mathbb{R}[X]^2$ denotes the cone of sums of squares of polynomials in $\mathbb{R}[X]$. For any real matrix (or vector) A, A^T denotes its transpose. For a symmetric matrix W, $W \succeq 0$ ($\succ 0$) means that W is positive semidefinite (definite). For any two given matrices U and V of the same size, their inner product $U \bullet V$ is defined as $U \bullet V := \sum_{i,j} U_{ij} V_{ij}$. For any $x \in \mathbb{R}^n$, its two norm $\|x\|_2$ is defined as $\sqrt{x_1^2 + \cdots + x_n^2}$.

This paper is organized as follows. Section 2 discusses the method of SOS relaxation and the special features in minimizing rational functions without constraints. Section 3 shows the application of minimizing rational functions in finding nearest GCDs. Section 4 then discusses SOS relaxations and the special features in constrained case. In Sect. 5, we draw some conclusions and discuss some possible future work.

2 SOS relaxation

In this section, we discuss the global minimization of (1.1) without any constraints. The constrained case will be handled in Sect. 4.

Obviously, γ is a lower bound for r^* if and only if the polynomial $f(x) - \gamma g(x)$ is nonnegative. By approximating $f(x) - \gamma g(x)$ by SOS polynomials, we get the following SOS relaxation

$$\begin{aligned} r^*_{\text{sos}} := \sup_{\gamma} \quad & \gamma \\ s.t. \quad & f(x) - \gamma g(x) \in \sum \mathbb{R}[X]^2. \end{aligned}$$

For any feasible γ, we immediately have $r(x) \geq \gamma$ for every $x \in \mathbb{R}^n$. Thus every feasible γ (including r^*_{sos}) is a lower bound for $r(x)$, i.e., $r^*_{\text{sos}} \leq r^*$. Also see [10] for SOS relaxations for minimizing rational functions without constraints.

Let $2d = \max(\deg(f), \deg(g))$ (it must be even for $r(x)$ to have a finite minimum) and $m(x)$ be the column vector of monomials up to degree d

$$m(x)^T = \left[1, x_1, \ldots, x_n, x_1^2, x_1 x_2, \ldots, x_n^2, x_1^3, \ldots, x_n^d \right].$$

Notice that the dimension of vector $m(x)$ is $\binom{n+d}{d}$. Then $f(x) - \gamma g(x)$ is SOS if and only there exists a symmetric matrix $W \succeq 0$ of length $\binom{n+d}{d}$ such that [24,32] the identity holds:

$$f(x) - \gamma g(x) \equiv m(x)^T W m(x). \tag{2.1}$$

Now we write $f(x) = \sum_{\alpha \in F} f_\alpha x^\alpha$ and $g(x) = \sum_{\alpha \in F} g_\alpha x^\alpha$, where F is a finite subset of $\mathbb{N}^n$, i.e., F is the *support* of polynomials $f(x)$ and $g(x)$.

Throughout this paper, we index the rows and columns of matrix W by monomial powers up to degree d, i.e., the indices for the entries in W have the form (α, β) where $\alpha, \beta \in \mathbb{N}^n$. For fixed $\alpha \in F$, we define the monomial base matrix B_α as follows (see [15])

$$B_\alpha(\eta, \tau) = \begin{cases} 1, & \text{if } \eta + \tau = \alpha, \\ 0, & \text{otherwise.} \end{cases}$$

When $n = 1$, the B_α are Hankel matrices. Now we can see that (2.1) holds if and only if

$$f_\alpha - \gamma g_\alpha = B_\alpha \bullet W, \quad \forall \alpha \in F.$$

Therefore the SOS relaxation of problem (1.1) is essentially the following SDP:

$$r^*_{\text{sos}} := \sup_{\gamma, W} \quad \gamma \tag{2.2}$$
$$s.t. \quad f_\alpha - \gamma g_\alpha = B_\alpha \bullet W, \quad \forall \alpha \in F \tag{2.3}$$
$$W \succeq 0. \tag{2.4}$$

Notice that the decision variables are γ and W instead of x. We refer to [5,37] for the theory and applications of SDP.

Now let us derive the dual problem of SDP (2.2) and (2.3). Its Lagrange function is

$$\begin{aligned} \mathcal{L}(\gamma, W, y, S) &= \gamma + \sum_{\alpha \in F} (f_\alpha - \gamma g_\alpha - < B_\alpha, W >) y_\alpha + W \bullet S \\ &= \sum_{\alpha \in F} f_\alpha y_\alpha + \left(1 - \sum_{\alpha \in F} g_\alpha y_\alpha\right)\gamma + \left(S - \sum_{\alpha \in F} y_\alpha B_\alpha\right) \bullet W, \end{aligned}$$

where $y = (y_\alpha)$ and W are dual decision variables (Lagrange multipliers). Here $S \succeq 0$ corresponds to the constraint $W \succeq 0$. Obviously it holds

$$\sup_{\gamma, W} \ \mathcal{L}(\gamma, W, y, S) = \begin{cases} \sum_{\alpha \in F} f_\alpha y_\alpha, & \text{if } \sum_{\alpha \in F} g_\alpha y_\alpha = 1, \\ & \quad\ \sum_{\alpha \in F} y_\alpha B_\alpha = S, \\ +\infty, & \text{otherwise.} \end{cases}$$

Therefore, the dual of problems (2.2) and (2.4) is

$$r^*_{\text{mom}} := \inf_y \quad \sum_{\alpha \in F} f_\alpha y_\alpha, \tag{2.5}$$
$$s.t. \quad \sum_\alpha g_\alpha y_\alpha = 1, \tag{2.6}$$
$$M_d(y) \succeq 0 \tag{2.7}$$

where the matrix $M_d(y) := \sum_\alpha y_\alpha B_\alpha$ is called the *d-th moment matrix* of y. For an integer k, the kth moment matrix $M_k(y)$ of a monomial-indexed vector $y = (y_\alpha)$ is defined as

$$M_k(y) = (y_{\alpha+\beta})_{0 \le |\alpha|, |\beta| \le k}.$$

We refer to [15] for a more detailed description of moment matrices. (2.5)–(2.7) can also be considered as a generalization of moment approaches in [15], except the equality (2.6).

From the derivation of dual problems (2.5) and (2.7) we immediately have that $r^*_{\text{sos}} \le r^*_{\text{mom}}$, which is referred to as *weak duality* in optimization duality theory. Actually we have stronger properties for the SOS relaxation (2.2)–(2.4) and its dual (2.5)–(2.7) as summarized in the following theorem, which is similar to Theorem 3.2 in [15].

Theorem 2.1 *Assume that the SOS relaxation* (2.2)–(2.4) *has a feasible solution* (γ, W). *Then the following properties hold for the primal problems* (2.2)–(2.4) *and its dual* (2.5)–(2.7)*:*

(i) *Strong duality holds, i.e.,* $r^*_{\text{sos}} = r^*_{\text{mom}}$, *and* $f(x) - r^*_{\text{sos}} g(x)$ *is SOS.*
(ii) *The lower bound* r^*_{sos} *obtained from SOS relaxation* (2.2)–(2.4) *is exact, i.e.,* $r^*_{\text{sos}} = r^*$ *if and only if* $f(x) - r^* g(x)$ *is SOS.*
(iii) *When* $r^*_{\text{sos}} = r^*$ *and* $u^{(j)}$ $(j = 1, \ldots, t)$ *are global minimizers, then every monomial indexed vector* y *of the following form*

$$y \in \left\{ \sum_{j=1}^{t} \theta_j m_{2d}(u^{(j)}) : \theta_j \ge 0, \sum_{j=1}^{t} \theta_j = 1 \right\}$$

is an optimal solution to (2.5)–(2.7).

Proof (i) The result can be obtained from the standard duality theory of convex programming [29, Sect. 30], if we can show that there exists a vector $\hat{y}$ such that $\sum_\alpha g_\alpha \hat{y}_\alpha = 1$ and $M_d(\hat{y}) \succ 0$. Let μ be a measure on $\mathbb{R}^n$ with strictly positive density everywhere on $\mathbb{R}^n$ and finite moments, i.e., $|\int x^\alpha \mathrm{d}\mu| < \infty$ for all $\alpha \in \mathbb{N}^n$ (e.g., one density function can be chosen as $\exp(-\sum_{i=1}^n x_i^2)$). Define the vector $y = (y_\alpha)$ as follows:

$$y_\alpha = \int x^\alpha \mathrm{d}\mu < \infty.$$

Then we claim that

$$0 < \tau := \sum_\alpha g_\alpha y_\alpha = \int g(x) \mathrm{d}\mu < \infty.$$

The second inequality is obvious since all the moments of μ are finite. For the first inequality, for a contradiction, suppose $\tau \le 0$, that is,

$$\int g(x) \mathrm{d}\mu \le 0.$$

Since $g(x)$ is assumed to be nonnegative everywhere and μ has positive density everywhere, we must have that $g(x)$ should be identically zero, which is a contradiction. Now we prove that $M_d(y)$ is positive definite. For any monomial-indexed nonzero vector q with the same length as $M_d(y)$ (corresponding to a nonzero polynomial $q(x)$), it holds that

$$q^T M_d(y) q = \sum_{0 \le |\alpha|, |\beta| \le d} y_{\alpha+\beta} q_\alpha q_\beta = \int \left(\sum_{0 \le |\alpha|, |\beta| \le d} x^{\alpha+\beta} q_\alpha q_\beta \right) \mathrm{d}\mu = \int q(x)^2 d\mu > 0.$$

Now let $\hat{y} = y/\tau$, which obviously satisfies that $\sum g_\alpha \hat{y}_\alpha = 1$ and $M_d(\hat{y}) \succ 0$. In other words, the problem (2.5)–(2.7) has an interior point. Therefore, from the duality theory of convex optimization, we know that the strong duality holds, i.e., $r^*_{\text{sos}} = r^*$ and the optimal solution set of (2.2)–(2.4) is nonempty.

As already shown in (i), the optimal solution set of (2.2)–(2.4) is nonempty, which implies the conclusion in (ii) immediately.

(iii) When $r^*_{\text{sos}} = r^*$, the optimal value in (2.5)–(2.7) is also r^*, by strong duality as established in (i). Now choose an arbitrary monomial-indexed vector y of the form

$$y = \sum_{j=1}^{t} \theta_j m_{2d}(u^{(j)})$$

for any θ such that $\theta_j \geq 0$, $\sum_{j=1}^t \theta_j = 1$. Then we have

$$\sum_{\alpha \in F} f_\alpha y_\alpha = \sum_{j=1}^{t} \theta_j f(u^{(j)}) = \sum_{j=1}^{t} \theta_j r^* = r^*.$$

Obviously $M_d(y) = \sum_{j=1}^t \theta_j m_d(u^{(j)}) m_d(u^{(j)})^T \succeq 0$. So y is a feasible solution with optimal objective value. Thus y is a optimal solution to (2.5)–(2.7). □

The information about the minimizers of (1.1) can be found from the optimal solutions to the dual problem (2.5)–(2.7). Suppose $y^* = (y^*_\alpha)$ (where $y^*_{(0,\ldots,0)} \neq 0$) is one minimizer of (2.5)–(2.7) such that the moment matrix $M_d(y^*)$ has rank one. Then there is a vector w, with the same length as $M_d(y^*)$, such that

$$M_d(y^*)/y^*_{(0,\ldots,0)} = ww^T,$$

where the left-hand side is the called *normalized* moment matrix, with the $(1, 1)$ entry being 1. Set $x^* := [\, w(2), w(3), \ldots, w(n+1)\,]$. So for any monomial-index α, it holds that $w(\alpha) = (x^*)^\alpha$. Now plug the point x^* into the rational function $r(x)$, evaluate it, then we can see that

$$r(x^*) = \frac{f(x^*)}{g(x^*)} = \frac{\sum_\alpha f_\alpha (x^*)^\alpha}{\sum_\alpha g_\alpha (x^*)^\alpha} = \frac{\sum_\alpha f_\alpha y^*_\alpha}{\sum_\alpha g_\alpha y^*_\alpha} = r^*_{\text{mom}} = r^*_{\text{sos}}.$$

In other words, we get a point x^* at which the evaluation of objective $r(x)$ equals the lower bound r^*_{sos}. Therefore, x^* is a global minimizer and r^*_{sos} equals the global minimum r^*. When $M_d(y^*)$ (with $y^*_{(0,\ldots,0)} \neq 0$) has rank more than one, but it satisfies some *flat extension* condition ($\text{rank}(M_k(y^*)) = \text{rank}(M_{k+1}(y^*))$ for some integer $0 \leq k < d$), there is more than one global minimizer (the number equals the rank of the moment matrix), and they can be found numerically by solving an eigenvalue problem. We refer to [4,9] for more details about the flat extension condition and extracting minimizers. When it happens that $y^*_{(0,\ldots,0)} = 0$, we can not normalize the moment matrix $M_d(y^*)$ to represent some measure, which might be due to the case that the infimum of $r(x)$ is attained at infinity. For instance, consider the example $r(x) := 1/(1 + x^2)$ for $n = 1$. The optimal solution is $y^* = (0, 0, 1)$, which can not be normalized.

Remark 2.2 When $f(x)$ and $g(x)$ have real common zeros, the solution to the dual problems (2.5)–(2.7) is not unique. To see this, suppose w is such that $f(w) = g(w) = 0$, and y^* is an optimal solution to (2.5)–(2.7). Now let $\hat{y} = m_{2N}(w)$, which is not zero since $\hat{y}_{(0,\ldots,0)} = 1$. Then $\sum_\alpha f_\alpha \hat{y}_\alpha = \sum_\alpha g_\alpha \hat{y}_\alpha = 0$ and $M_N(\hat{y}) \succeq 0$. So we can see that $y^* + \hat{y}$ is another feasible solution with the same optimal value. In such situations, some extracted points from the moment matrix $M_N(y^* + \hat{y})$ may not be global minimizers, but they might be the common zeros of $f(x)$ and $g(x)$.

In the following we show some examples of minimizing rational functions via SOS relaxations. The problems (2.2)–(2.4) and its dual (2.5)–(2.7) are solved by *YALMIP* [17] which is based on *SeDuMi* [35]. They can also be solved by softwares like *SOSTOOLS* [26] and *GloptiPoly* [8].

Example 2.3 Consider the global minimization of the rational function

$$\frac{(x_1^2+1)^2 + (x_2^2+1)^2}{(x_1+x_2+1)^2}.$$

Solving (2.2)–(2.4) yields the lower bound $r^*_{\text{sos}} \approx 0.7639$. The solution y^* to (2.5)–(2.7) is

$$y^* \approx (0.2000,\ 0.1236,\ 0.1236,\ 0.0764,\ 0.0764,\ 0.0764,\ 0.0472,\ 0.0472,\\ 0.0472,\ 0.0472,\ 0.0292,\ 0.0292,\ 0.0292,\ 0.0292,\ 0.0292).$$

The rank of moment matrix $M_2(y^*)$ is one, and we can extract one point $x^* \approx (0.6180, 0.6180)$. The evaluation of $r(x)$ at x^* shows that $r(x^*) \approx 0.7639$. So x^* is a global minimizer and 0.7639 is the global minimum (approximately, ignoring rounding errors).

Example 2.4 Consider the global minimization of the rational polynomial

$$\frac{x_1^4 - 2x_1^2x_2x_3 + (x_2x_3+1)^2}{x_1^2}.$$

The lower bound given by (2.2)–(2.4) is $r^*_{\text{sos}} \approx 2.0000$. The solution y^* to (2.5)–(2.7) is

$$y^* \approx (1.0859,\ -0.0000,\ -0.0000,\ -0.0000,\ 1.0000,\ 0.0000,\ -0.0000,\ 0.8150,\ -0.0859,\\ 0.8150,\ -0.0000,\ -0.0000,\ -0.0000,\ -0.0000,\ 0.0000,\ -0.0000,\ -0.0000,\ -0.0000,\\ -0.0000,\ -0.0000,\ 1.0859,\ 0.0000,\ -0.0000,\ 0.8150,\ 0.0859,\ 0.8150,\ 0.0000,\ 0.0000,\\ -0.0000,\ -0.0000,\ 2.3208,\ -0.0000,\ 0.1719,\ 0.0000,\ 2.3208).$$

The moment matrix $M_2(y^*)$ does not satisfy the flat extension condition, and no minimizers can be extracted. Actually one can see that 2 is the global minimum by observing the identity

$$f(x) - 2g(x) = (x_1^2 - x_2x_3 - 1)^2.$$

The lower bound 2 is achieved at $(1, 0, 0)$ and hence is the global minimum. There are infinitely many global minimizers.

The relationship between the bounds is $r^*_{\text{mom}} = r^*_{\text{sos}} \le r^*$. But it may happen that $r^*_{\text{sos}} < r^*$, just as in SOS relaxations for minimizing polynomials. Let us see the following example.

Example 2.5 Consider the global minimization of the rational function

$$\frac{x_1^4x_2^2 + x_1^2x_2^4 + x_3^6}{x_1^2x_2^2x_3^2}.$$

The lower bound given by (2.2)–(2.4) is $r^*_{sos} = 0$, and the solution y^* to (2.5)–(2.7) is

$$y^*_{(2,2,2)} = 1, \;\; y^*_\alpha = 0, \quad \forall \alpha \neq (2,2,2).$$

The global minimum $r^* = 3$ because

$$x_1^4x_2^2 + x_1^2x_2^4 + x_3^6 - 3x_1^2x_2^2x_3^2 \geq 0 \quad \forall x \in \mathbb{R}^3$$

and $r(1,1,1) = 3$. So in this example, the SOS lower bound $r^*_{\text{sos}} < r^*$. Actually for any $0 < \gamma \leq 3$, the polynomial

$$x_1^4x_2^2 + x_1^2x_2^4 + x_3^6 - \gamma x_1^2x_2^2x_3^2$$

is nonnegative but not SOS. The proof is the same as to prove that the Motzkin polynomial

$$x_1^4x_2^2 + x_1^2x_2^4 + x_3^6 - 3x_1^2x_2^2x_3^2$$

is not SOS [28].

2.1 What if $r^*_{\text{sos}} < r^*$?

From Theorem 2.1, we know that $r^*_{\text{sos}} = r^*$ if and only if the polynomial $f(x) - r^*g(x)$ is SOS. But sometimes $f(x) - r^*g(x)$ might not be SOS, as we observed in Example 2.5. In this subsection, we discuss how to minimize a rational function $r(x)$ when $r^*_{\text{sos}} < r^*$. Here we generalize the big ball technique introduced in [15], but we must be very careful about the zeros of the denominator $g(x)$ in $r(x)$.

Suppose we know that at least one global minimizer of $r(x)$ belongs to the ball $B(c,\rho) := \{x \in \mathbb{R}^n : \rho^2 - \|x-c\|_2^2 \geq 0\}$ with center c and radius $\rho > 0$. Let $\pi(x) := \rho^2 - \|x-c\|_2^2$. Then we immediately have that $r^* = \min\limits_{x\in\mathbb{R}^n} r(x) = \min\limits_{x\in B(c,\rho)} r(x)$. In practice, we can often choose the center $c = 0$ and radius ρ big enough. So the original unconstrained minimization problem (1.1) becomes the constrained problem

$$\min_{x\in B(c,\rho)} \quad r(x).$$

One natural SOS relaxation of this constrained problem is

$$r_N^* := \sup_{\gamma} \quad \gamma \tag{2.8}$$

$$s.t. \quad f(x) - \gamma g(x) \equiv \sigma_0(x) + \sigma_1(x)\pi(x) \tag{2.9}$$

$$\deg(\sigma_1) \leq 2(N-1), \;\; \sigma_0(x), \sigma_1(x) \in \sum \mathbb{R}[X]^2. \tag{2.10}$$

Similar to the dual of (2.2)–(2.4), the dual problem of (2.8)–(2.10) can be found to be

$$\hat{r}_N^* := \inf_y \quad \sum_{\alpha \in F} f_\alpha y_\alpha \tag{2.11}$$

$$s.t. \quad \sum_\alpha g_\alpha y_\alpha = 1, \tag{2.12}$$

$$M_N(y) \succeq 0, \tag{2.13}$$

$$M_{N-1}(\pi * y) \succeq 0 \tag{2.14}$$

where π is the vector of the coefficients of polynomial $\pi(x)$. For a general polynomial $p(x) = \sum_\alpha p_\alpha x^\alpha$, the generalized moment matrix $M_k(p * y)$ is defined as

$$M_k(p * y)(\beta, \tau) := \sum_\alpha p_\alpha y_{\beta+\tau+\alpha}, \; 0 \leq |\beta|, |\tau| \leq k.$$

We have the following theorem for the SOS relaxation (2.8)–(2.10) and its dual (2.11)–(2.14), which is similar to Theorem 3.4 in [15].

Theorem 2.6 *Assume that $r^* > -\infty$ and at least one global minimizer of $r(x)$ lies in the ball $B(c, \rho)$. If the numerator $f(x)$ and denominator $g(x)$ of $r(x)$ have no common real zeros on $B(c, \rho)$, then the following holds:*

(i) *The lower bounds converge: $\lim_{N\to\infty} r_N^* = r^*$.*
(ii) *For N large enough, there is no duality gap between (2.8)–(2.10) and its dual (2.11)–(2.14), i.e., $r_N^* = \hat{r}_N^*$.*
(iii) *For N large enough, $r_N^* = r^*$ if and only if $f(x) - r^* g(x) = \sigma_0(x) + \sigma_1(x)\pi(x)$ for some SOS polynomials σ_0, σ_1 with $\deg(\sigma_1) \leq 2(N-1)$.*
(iv) *If $r_N^* = r^*$ for some integer N and $u^{(j)}$ $(j = 1, \ldots, t)$ are global minimizers on $B(c, \rho)$, then every vector y of the following form*

$$y \in \left\{ \sum_{j=1}^t \theta_j m_{2N}(u^{(j)}) : \theta_j \geq 0, \sum_{j=1}^t \theta_j = 1 \right\}$$

is an optimal solution to (2.11)–(2.14).

Proof (i) For any fixed $\gamma < r^*$, we can see that $f(x) - \gamma g(x) > 0$ on $B(c, \rho)$ if $g(x) \neq 0$ (we have assumed that $g(x)$ is nonnegative). When $g(x) = 0$, we must have $f(x) \geq 0$. Otherwise assume $f(u) < 0$ at some point u with $g(u) = 0$. Then in a neighborhood of u, the rational polynomial $r(x)$ has a singularity at u, and hence is unbounded from below, which contradicts the assumption that $r^* > -\infty$. Thus $g(x) = 0$ implies $f(x) \geq 0$ on $B(c, \rho)$. So we have that

$$f(x) - \gamma g(x) \geq 0, \quad \forall x \in B(c, \rho).$$

Since $\gamma < r^*$, $f(x) - \gamma g(x) = 0$ implies that $f(x) = g(x) = 0$, which is not possible. Therefore, the polynomal $f(x) - \gamma g(x)$ is positive on ball $B(c, \rho)$. Now by Putinar's Theorem [27], there exist SOS polynomials σ_0, σ_1 with degree high enough such that

$$f(x) - \gamma g(x) \equiv \sigma_0(x) + \sigma_1(x)\pi(x).$$

So in (2.8)–(2.10), γ can be chosen arbitrarily close to r^*. Therefore we have shown the convergence of lower bounds r_N^*.

(ii) Similar to the proof of Theorem 2.1, it suffices to show that the problems (2.11)–(2.14) has a strictly feasible solution. Let μ be a measure with uniform distribution on $B(c, \rho)$. Define vector $y = (y_\alpha)$ as follows:

$$y_\alpha := \int x^\alpha \mathrm{d}\mu.$$

Now we show that $M_N(y)$ and $M_{N-1}(\pi * y)$ are positive definite. $M_N(y) \succ 0$ can be shown in the same way as in the proof of (i) in Theorem 2.1. Now we show that $M_{N-1}(\pi * y) \succ 0$. For any nonzero monomial-indexed vector q of the same length as $M_{N-1}(\pi * y)$ (it corresponds to a nonzero polynomial $q(x)$ up to degree $N-1$), it holds that

$$q^T M_{N-1}(\pi * y) q = \int q(x)^2 \pi(x) d\mu = \frac{1}{\mathrm{Vol}(B(c, \rho))} \int_{B(c,\rho)} q(x)^2 \pi(x) \mathrm{d}x > 0$$

which implies that $M_{N-1}(\pi * y)$ is positive definite. In the above, $\mathrm{Vol}(B(c, \rho))$ denotes the volume of the ball $B(c, \rho)$. Since $g(x)$ is not identically zero and always nonnegative, $g(x)$ can not be always zero on $B(c, \rho)$ and hence

$$\sum_\alpha g_\alpha y_\alpha = \int g(x) \mathrm{d}\mu = \frac{1}{\mathrm{Vol}(B(c, \rho))} \int_{B(c,\rho)} g(x) \mathrm{d}x > 0.$$

Now set the vector $\hat{y} = y / \sum_\alpha g_\alpha y_\alpha$. Then can see that $\hat{y}$ is an interior point for the dual problems (2.11)–(2.14).

(iii) For any fixed $\hat{\gamma} < r^*$, from the previous arguments we know that the polynomial $f(x) - \gamma g(x)$ is positive on $B(c, \rho)$. Then by Putinar's Theorem, there exist SOS polynomials $s_0(x), s_1(x)$ with $\deg(\sigma_1)$ high enough such that

$$f(x) - \hat{\gamma} g(x) \equiv s_0(x) + s_1(x)\pi(x).$$

This means that the primal convex problem (2.8)–(2.10) has a feasible solution. From (ii) we know its dual problem (2.11)–(2.14) has a strict interior point. Now apply the duality theory of standard convex programming, then we know the solution set of (2.8)–(2.10) is nonempty. And notice that r^* is obviously an upper bound for all r_N^*.

When $r_N^* = r^*$, we know r_N^* is optimal. For N sufficiently large, by (ii), the primal problem (2.8)–(2.10) is guaranteed to have a solution. So there exist SOS polynomials $\sigma_0(x), \sigma_1(x)$ with $\deg(\sigma_1) \le 2(N-1)$ such that

$$f(x) - r^* g(x) \equiv \sigma_0(x) + \sigma_1(x)\, \pi(x).$$

The "if" direction is obvious.

The proof of (iv) is the same as (iii) of Theorem 2.1. □

Remark 2.7 In Theorem 2.6, we need the assumption that the numerator $f(x)$ and denominator $g(x)$ have no real common zeros on ball $B(c, \rho)$ to show convergence $\lim_{N\to\infty} r_N^* = r^*$. When they have common real zeros, for any $\gamma < r^*$, the polynomial $f(x) - \gamma g(x)$ is not strictly positive on $B(c, \rho)$ and hence Putinar's Theorem can not be applied. In such situations, the convergence is not guaranteed (see Remark 4.6). However, in case of two variables, i.e., $n = 2$, if $f(x)$ and $g(x)$ have at most finitely many real common zeros on $B(c, \rho)$, we still have $\lim_{N\to\infty} r_N^* = r^*$; furthermore, if the global minimizers of $r(x)$ are finite, then finite convergence holds, i.e., there exists $N \in \mathbb{N}$ such that $r_N^* = r^*$. Please see Theorem 4.8 in Sect. 4. Notice that the ball $B(c, \rho)$ satisfies both conditions (i) and (ii) there.

Remark 2.8 When $f(x)$ and $g(x)$ have common zeros on $B(c, \rho)$, the solution to the dual problems (2.11)–(2.14) is not unique. In such situations, some extracted points from the moment matrix $M_N(y^*)$ may not be global minimizers and they might be the common zeros of $f(x)$ and $g(x)$ (see Example 2.10).

Example 2.9 Consider the global minimization of the rational function (obtained by plugging $x_3 = 1$ into Example 2.5)

$$\frac{x_1^4 x_2^2 + x_1^2 x_2^4 + 1}{x_1^2 x_2^2}.$$

Choose $c = 0$ and $\rho = 2$. For $N = 3$, the lower bound given by (2.8)–(2.10) is $r_3^* = 3$, and the solution to (2.11)–(2.13) is

$$y^* = (1, 0, 0, 1, 0, 1, 0, 0, 0, 0, 1, 0, 1, 0, 1, 00, 0, 0, 0, 0, 1, 0, 1, 0, 1, 0, 1).$$

The moment matrix $M_3(y^*)$ has rank 4, and satisfies the flat extension condition. Four points are extracted: $(\pm 1, \pm 1)$. They are all global minimizers.

Example 2.10 Consider the global minimization of the rational function (obtained by plugging $x_2 = 1$ into Example 2.5)

$$\frac{x_1^4 + x_1^2 + x_3^6}{x_1^2 x_3^2}.$$

Choose $c = 0$ and $\rho = 2$. For $N = 4$, the lower bound given by (2.8)–(2.10) is $r_4^* = 3.0000$, and the solution to (2.11)–(2.13) is

$$y^* \approx (2.8377,\ 0,\ 0,\ 1,\ 0,\ 0,\ 1.0008,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,$$
$$1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1,\ 0,\ 1).$$

The moment matrix has rank 6 and satisfies the flat extension condition. Six points are extracted:

$$(\pm 1.0000, \pm 1.0000), \qquad (0.0000, \pm 0.0211).$$

The evaluation of $r(x)$ at these points shows that the first four points are global minimizers. The last two points are not global minimizers, but they are approximately common zeros of the numerator and denominator (see Remark 2.2).

3 Nearest greatest common divisor

This section discusses the application of minimizing rational functions to find the nearest common divisors of univariate polynomials.

Let $p(z)$ and $q(z)$ be two monic complex univariate polynomials of degree m such that

$$p(z) = z^m + p_{m-1} z^{m-1} + p_{m-2} z^{m-2} + \cdots + p_1 z + p_0, \tag{3.1}$$
$$q(z) = z^m + q_{m-1} z^{m-1} + q_{m-2} z^{m-2} + \cdots + q_1 z + q_0. \tag{3.2}$$

Their coefficients p_i, q_j are all complex numbers. When $p(z), q(z)$ have common divisors, their greatest common divisor (GCD) can be computed exactly by using Euclid's algorithm or other refined algorithms [2,3]. These algorithms need to assume that all the coefficients of $p(z)$ and $q(z)$ are error-free, and return the exact GCD. However, in practice, it is more interesting to compute the GCD of two polynomials whose coefficients may not be known exactly. In such situations, we often get the trivial common divisor (the constant polynomial 1) if we apply exact methods like Euclid's algorithm.

So instead, we will seek the smallest possible perturbations of the coefficients of $p(z)$ and $q(z)$ that cause their GCD to be nontrivial, say $z-c$ for some c. See [12,13,34, Sect. 6.4] for a discussion of this problem. Our contribution is to solve the associated global optimization problem by the methods we have introduced in the preceding section, instead of finding all the real critical points (zero gradient) as suggested in [12,13].

Throughout this paper, we equip the polynomials $p(z), q(z)$ with $\|\cdot\|_2$ norm of their coefficients, i.e., $\|p\|_2 = \sqrt{\sum_{k=0}^{m-1} |p_k|^2}$, $\|q\|_2 = \sqrt{\sum_{k=0}^{m-1} |q_k|^2}$. The perturbations made to $p(z), q(z)$ are measured similarly. The basic problem in this section is what is the minimum perturbation such that the perturbed polynomials have a common divisor? To be more specific, suppose the perturbed polynomials have the form

$$\hat{p}(z) = z^m + \hat{p}_{m-1}z^{m-1} + \hat{p}_{m-2}z^{m-2} + \cdots + \hat{p}_1 z + \hat{p}_0 \tag{3.3}$$

$$\hat{q}(z) = z^m + \hat{q}_{m-1}z^{m-1} + \hat{q}_{m-2}z^{m-2} + \cdots + \hat{q}_1 z + \hat{q}_0 \tag{3.4}$$

with common zero c, i.e., $\hat{p}(c) = \hat{q}(c) = 0$. The perturbation are measured as

$$\mathcal{N}(c, \hat{p}, \hat{q}) = \sum_{i=0}^{m-1} |p_i - \hat{p}_i|^2 + \sum_{j=0}^{m-1} |q_j - \hat{q}_j|^2.$$

The problem of finding nearest GCD can be formulated as to find $(c, \hat{p}, \hat{q})$ such that $\mathcal{N}(c, \hat{p}, \hat{q})$ is minimized subject to $\hat{p}(c) = \hat{q}(c) = 0$.

We can see that $\mathcal{N}(c, \hat{p}, \hat{q})$ is a convex quadratic function in $(\hat{p}, \hat{q})$. But the constraints $\hat{p}(c) = \hat{q}(c) = 0$ are nonconvex. However, if the common root c is fixed, the constraints $\hat{p}(c) = \hat{q}(c) = 0$ are linear with respect to $(\hat{p}, \hat{q})$, and the reduced quadratic program has a solution with closed form. $\mathcal{N}(c, \hat{p}, \hat{q})$ is a convex quadratic function about $(\hat{p}, \hat{q})$. It can be shown [13] that

$$\min_{(\hat{p},\hat{q}):\hat{p}(c)=\hat{q}(c)=0} \quad \mathcal{N}(c, \hat{p}, \hat{q}) = \frac{|p(c)|^2 + |q(c)|^2}{\sum_{i=0}^{m-1} |c^2|^i}.$$

Therefore the problem of finding nearest perturbation becomes the global minimization of a rational function

$$\min_{c \in \mathbb{C}} \quad \frac{|p(c)|^2 + |q(c)|^2}{\sum_{i=0}^{m-1} |c^2|^i} \tag{3.5}$$

over the complex plane. Karmarkar and Lakshman [13] proposed the following algorithm to find the nearest perturbation:

Algorithm 3.1 (Nearest GCD algorithm [13])

Input Monic polynomials $p(z), q(z)$.

Step 1 Determine the rational function $r(x_1, x_2)$

$$r(x_1, x_2) := \frac{|p(c)|^2 + |q(c)|^2}{\sum_{k=0}^{m-1} (x_1^2 + x_2^2)^k}, \quad c = x_1 + \sqrt{-1}x_2.$$

Step 2 Solve the polynomial system $\frac{r(x_1,x_2)}{\partial x_1} = \frac{r(x_1,x_2)}{\partial x_1} = 0$. Find all its real solutions inside the box: $-B \leq x_1, x_2 \leq B$ where $B := 5\max(\|p\|^2, \|q\|^2)$. Choose the one $(\hat{x}_1, \hat{x}_2)$ such that $r(\hat{x}_1, \hat{x}_2)$ is minimum. Let $c := \hat{x}_1 + \sqrt{-1}\hat{x}_2$.

Step 3 Compute the perturbation of coefficient

$$\lambda_j := \frac{\bar{c}^j p(c)}{\sum_{k=0}^{m-1} |c^2|^k}, \qquad \mu_j := \frac{\bar{c}^j q(c)}{\sum_{k=0}^{m-1} |c^2|^k}.$$

Output: The minimum perturbed polynomials with common divisors are returned as

$$\hat{p}(z) = z^m + \sum_{k=0}^{m-1} (p_k - \lambda_k) z^k, \ \hat{q}(z) = z^m + \sum_{k=0}^{m-1} (q_k - \mu_k) z^k.$$

The most expensive part in the algorithm above is Step 2. Karmarkar and Lakshman [13] proposed to use numerical methods like Arnon and McCallum [1] or Manocha and Demmel [18] to find all the real solutions of a polynomial system inside a box.

However, in practice, it is very expensive to find all the real solutions of a polynomial system inside a box. So in this section, we propose to solve (3.5) by SOS relaxations introduced in the previous section instead of finding all the real solutions of a polynomial system. The SOS relaxation of problem (3.5) is the following:

$$\begin{aligned} \sup \quad & \gamma \\ s.t. \quad & f(x_1, x_2) - \gamma\Big(\sum_{i=0}^{m-1} (x_1^2 + x_2^2)^i\Big) \text{ is SOS} \end{aligned}$$

where $f(x_1, x_2) = |p(x_1 + \sqrt{-1}x_2)|^2 + |q(x_1 + \sqrt{-1}x_2)|^2$.

In the following examples, we solve the global optimization problem via SOS relaxation (2.2)–(2.4) and its dual (2.5)–(2.7). In all the examples here, the global minimizers can be extracted and the big ball technique introduced in Sect. 2.1 is not required.

Example 3.2 (Example 2.1 [13]) Consider polynomials

$$\begin{aligned} p(z) &= z^2 - 6z + 5, \\ q(z) &= z^2 - 6.30z + 5.72. \end{aligned}$$

Solving SOS relaxation (2.2)–(2.4) and its dual (2.5)–(2.7), we find the global minimum and extract one minimizer:

$$r^* \approx 0.0121, \qquad c^* = x_1^* + \sqrt{-1}x_2^* \approx 5.0971,$$

which are the same as found in [13].

Example 3.3 Consider polynomials

$$\begin{aligned} p(z) &= z^3 - 6z^2 + 11z - 6, \\ q(z) &= z^3 - 6.24z^2 + 10.75z - 6.50. \end{aligned}$$

Solving SOS relaxation (2.2)–(2.4) and its dual (2.5)–(2.7), we get the lower bound and extract one point

$$r^*_{\text{sos}} \approx 0.0563, \qquad (x_1^*, x_2^*) \approx (3.5725, 0.0000).$$

Evaluation of $r(x)$ at x^* shows that $r(x^*) \approx r^*_{\text{sos}}$, which implies that $c^* \approx 3.5725$ is a global minimizer for problem (3.5).

Example 3.4 Consider polynomials

$$p(z) = z^3 + z^2 - 2,$$
$$q(z) = z^3 + 1.5z^2 + 1.5z - 1.25.$$

Solving SOS relaxation (2.2)–(2.4) and its dual (2.5)–(2.7), we find the lower bound $r^*_{\text{sos}} \approx 0.0643$ and extract two points

$$x^* \approx (-1.0032, 1.1011), \qquad x^{**} \approx (-1.0032, -1.1011).$$

The evaluations of $r(x)$ at x^* and x^{**} show that $r(x^*) = r(x^{**}) \approx r^*_{\text{sos}}$, which implies that x^* and x^{**} are both global minimizers. So $c^* = -1.0032 \pm \sqrt{-1} \cdot 1.1011$ are the global minimizers of problem (3.5).

4 Constrained minimization

In this section, we discuss the global minimization of a rational function subject to constraints described by polynomial inequalities. Consider the problem

$$r^* := \min_{x \in \mathbb{R}^n} \quad r(x) := \frac{f(x)}{g(x)} \tag{4.1}$$

$$s.t. \quad h_1(x) \geq 0, \ldots, h_m(x) \geq 0 \tag{4.2}$$

where $f(x), g(x), h_i(x)$ are all real multivariate polynomials in x. Without confusion, we still let r^* denote the minimum of the objective function subject to the constraints. If some h_i are rational functions, we can reformulate the constraints $h_i(x) \geq 0$ equivalently as some polynomial inequalities (one should be careful with the zeros of $h_i(x)$). Denote by S the feasible set. S is in the form of a *basic closed semialgebraic set*. Here we assume that $g(x)$ is not identically zero on S, and $g(x)$ is nonnegative on S (otherwise, replace $\frac{f(x)}{g(x)}$ by $\frac{f(x)g(x)}{g^2(x)}$).

When $g(x) \equiv 1$ (or a nonzero constant), problem (4.1)–(4.2) becomes a standard constrained polynomial optimization problem. Lasserre [15] proposed a general procedure to solve this kind of optimization problem by a sequence of SOS relaxations. To be more specific, for each fixed positive integer N, we seek γ as large as possible such that $f(x) - \gamma$ has the representation

$$f(x) - \gamma \equiv \sigma_0(x) + \sigma_1(x)h_1(x) + \cdots + \sigma_m(x)h_m(x)$$

with all $\sigma_i(x)$ being SOS and $\deg(\sigma_i h_i) \leq 2N$. Obviously each γ is a lower bound of $f(x)$ on S. Denote by f_N^* the maximum γ under these conditions, which is also a lower bound. These lower bounds f_N^* converge to the minimum of $f(x)$ on S under a certain constraint qualification condition (see Assumption 4.1 below). A convergence rate is given in [22]. We refer to [9,15,16,21,23,24] for more introductions on SOS methods for polynomial optimization.

When $g(x)$ is a nonconstant polynomial nonnegative on S, Lasserre's procedure can be generalized in a natural way. Also see [10] for SOS relaxations for minimizing rational functions subject to semialgebraic constraints. For each fixed positive integer N, consider the SOS relaxation

$$r_N^* := \sup \quad \gamma \tag{4.3}$$

$$s.t. \quad f(x) - \gamma g(x) \equiv \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) h_i(x), \tag{4.4}$$

$$\deg(g_i) \leq 2N - d_i, \;\; \sigma_i(x) \in \quad \sum \mathbb{R}[X]^2 \tag{4.5}$$

where $d_i = \lceil \deg(h_i)/2 \rceil$. For any feasible γ above, it is obvious that $f(x) - \gamma g(x) \geq 0$ on S and so $r(x) \geq \gamma$. Thus every such γ (including r_N^*) is a lower bound of $r(x)$ on S.

We denote by $M(S)$ the set of polynomials which can be represented as

$$\sigma_0(x) + \sigma_1(x)h_1(x) + \cdots + \sigma_m(x)h_m(x)$$

with all $\sigma_i(x)$ being SOS. $M(S)$ is called the *quadratic module* generated by polynomials $\{h_1, \ldots, h_m\}$. A subset M of polynomial ring $\mathbb{R}[X]$ is a quadratic module if $1 \in M$, $M + M \subset M$ and $p^2 M \subset M$ for all $p \in \mathbb{R}[X]$. Throughout this section, we make the following assumption for $M(S)$:

Assumption 4.1 (Constraint qualification condition) *There exist $R > 0$ and SOS polynomials $s_0(x), s_1(x), \ldots, s_m(x) \in \sum \mathbb{R}[X]^2$ such that*

$$R - \|x\|_2^2 = s_0(x) + s_1(x)h_1(x) + \cdots + s_m(x)h_m(x).$$

Remark 4.2 When the assumption above is satisfied, the quadratic module $M(S)$ is said to be *Archimedean*. Obviously, when this assumption holds, the semialgebraic set S is contained in the ball $B(0, \sqrt{R})$ and hence compact, but the converse might not be true. See Example 6.3.1 in [6] for a counterexample. Under this assumption, Putinar [27] showed that every polynomial $p(x)$ positive on S belongs to $M(S)$.

Remark 4.3 When Assumption (4.1) does not hold, we can add in S one redundant constraint like $R - \|x\|_2^2 \geq 0$ for R sufficiently large (e.g., a norm bound is known in advance for one global minimizer). Then the new quadratic module is always Archimedean. In practice, when the value of R is unknown, we can choose an initial guess R_1, say, $R_1 = 1$, and then compute the minimum value f_1^* of $f(x)$ on $S \cap B(0, \sqrt{R_1})$ (we may need solve a sequence of SOS relaxations (4.3)–(4.5) to get this minimum). After that, choose a larger $R_2 = 2R_1$, and compute the minimum value f_2^* of $f(x)$ on $S \cap B(0, \sqrt{R_2})$. Repeat this process. Then we can get a sequence of minima $\{f_N^*\}$. The sequence $\{f_N^*\}$ is decreasing because the minimum value of $f(x)$ on $S \cap B(0, \sqrt{R})$ is decreasing when R is increasing. We terminate the computation when the sequence $\{f_N^*\}$ converges. See Example 4.4 for how this process works.

Example 4.4 Consider the problem

$$\min_x \quad \frac{x_1^2 x_2 + x_1 x_2^2 + 8}{2x_1 x_2}$$
$$s.t. \quad x_1, x_2 \geq 0.$$

Applying arithmetic–geometric mean inequality, we can easily see that the global minimum $f^* = 3$. The feasible set S here is not compact. Now we illustrate how to compute the minimum f^* by adding redundant constraint $R - \|x\|_2^2 \geq 0$. We set $R_1 = 1$, $R_2 = 2$, $R_3 = 4$, $R_4 = 8$, $R_5 = 16$ and compute the minimum values f_k^* of $f(x)$ on $S \cap B(0, \sqrt{R_k})$ for $k = 1, 2, 3, 4, 5$. By SOS implementations, we get the following table

k	1	2	3	4	5
R_k	1	2	4	8	16
f_k^*	8.7071	5.0000	3.4142	3.0000	3.0000

Convergence is obtained at $k = 5$.

Similar to the derivation of (2.5)–(2.7), the dual problem of (4.3)–(4.4) can be found to be

$$\inf_y \quad \sum_{\alpha \in F} f_\alpha y_\alpha \tag{4.6}$$

$$s.t. \quad \sum_{\alpha} g_\alpha y_\alpha = 1, \tag{4.7}$$

$$M_N(y) \succeq 0, \tag{4.8}$$

$$M_{N-d_i}(h_i * y) \succeq 0, \; i = 1, \ldots, m. \tag{4.9}$$

The properties of SOS relaxations (4.3)–(4.5) and (4.6)–(4.9) are summarized as follows:

Theorem 4.5 *Assume that the minimum r^* of $r(x)$ on S is finite, and $f(x) = g(x) = 0$ has no solutions on S. Then the following holds:*

- (i) *Convergence of the lower bounds:* $\lim_{N\to\infty} r_N^* = r^*$. *If, furthermore, S has nonempty interior, then* (ii) *and* (iii) *below are true.*
- (ii) *For N large enough, there is no duality gap between* (4.3)–(4.5) *and its dual* (4.6)–(4.9).
- (iii) *For N large enough, $r_N^* = r^*$ if and only if $f(x) - r^* g(x) \equiv \sigma_0(x) + \sum_{i=1}^m \sigma_i h_i(x)$ for SOS polynomials $\sigma_i(x)$ with* $\deg(\sigma_i h_i) \leq 2N$.
- (iv) *If $r_N^* = r^*$ for some integer N and $u^{(j)}$ $(j = 1, \ldots, t)$ are global minimizers on S, then every vector y of the following form*

$$y \in \left\{ \sum_{j=1}^t \theta_j m_{2N}(u^{(j)}) : \theta_j \geq 0, \sum_{j=1}^t \theta_j = 1 \right\}$$

is an optimal solution to (4.6)–(4.9).

Proof (i) For any $\gamma < r^*$, we have that the polynomial

$$\vartheta_\gamma(x) := f(x) - \gamma g(x)$$

is nonnegative on S. When $\vartheta_\gamma(u) = 0$ for some point $u \in S$, we must have $f(u) = g(u) = 0$, since otherwise $g(u) > 0$ ($g(x)$ is assumed to be nonnegative on S) and

$r(u) = \gamma < r^*$, which is impossible. Therefore $\vartheta_\gamma(x)$ is positive on S. By Putinar's Theorem [27], there exist SOS polynomials $\sigma_i(x)$ of degree high enough such that

$$\vartheta_\gamma(x) \equiv \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) h_i(x).$$

Therefore the claim in (i) is true.

(ii)–(iv): The proof here is almost the same as for Theorem 2.6. In a similar way, show that (4.3)–(4.5) has a feasible solution, and (4.6)–(4.9) has an interior point. Then apply the duality theory of convex programming. In (iv), check every y with given form is feasible and achieves the optimal objective value. □

Remark 4.6 In Theorem 4.5, we made the assumption that $f(x)$ and $g(x)$ have no common zeros on S. But sometimes $f(x)$ and $g(x)$ may have common zeros, and it is also possible that the minimum r^* is attained at the common zero(s) (in this case, $f(x)$ and $g(x)$ are of the same magnitude order around the common zero(s)). In such situations, we can not apply Putinar's Theorem and might not have convergence. For a counterexample, consider problem

$$\begin{aligned} \min \quad & r(x) := \frac{1+x}{(1-x^2)^2} \\ s.t. \quad & (1-x^2)^3 \geq 0. \end{aligned}$$

The global minimum is $r^* = \frac{27}{32}$ and the minimizer is $x^* = -\frac{1}{3}$. However, for any $\gamma < \frac{27}{32}$, there do not exist SOS polynomials $\sigma_0(x), \sigma_1(x)$ such that

$$1 + x - \gamma(1-x^2)^2 \equiv \sigma_0(x) + \sigma_1(x)(1-x^2)^3.$$

Otherwise, for a contradiction, suppose they exist. Then the left hand side vanishes at $x = -1$ and so does the right-hand side. So $x = -1$ is a zero of $\sigma_0(x)$ with multiplicity greater than one, since σ_0 is SOS. Hence $x = -1$ is a multiple zero of the left-hand side, which is impossible since the derivative of $1 + x - \gamma(1-x^2)^2$ at $x = -1$ is 1. This counterexample is motivated by the one given by Stengle [33], which shows that the polynomial $1 - x^2$ does not belong to the quadratic module $M((1-x^2)^3)$ since $1 - x^2$ is not strictly positive on $\{x : (1-x^2)^3 \geq 0\}$. On the other hand, if we can know in advance that the global minimum is not attained where the denominator $g(x)$ vanishes, one way to overcome this difficulty is to add more constraints which keep the global minimizers but eliminate the zeros of $g(x)$.

Remark 4.7 When $f(x)$ and $g(x)$ have common zeros on S, the solution to the dual problems (4.6)–(4.9) is not unique. In such situations, some extracted points from the moment matrix $M_N(y^* + \hat{y})$ may not be global minimizers but they might be the common zeros of $f(x)$ and $g(x)$ (see Remark 2.2).

When $n = 2$, i.e., in case of two variables, the distinguished representations of nonnegative polynomials by Scheiderer [30] are very useful. Under some conditions on the geometry of the feasible set S, convergence or even finite convergence holds if $f(x)$ and $g(x)$ have finitely many common zeros on S. This yields our next theorem.

Theorem 4.8 *Suppose $n = 2$. Let $Z(f, g) = \{u \in S : f(u) = g(u) = 0\}$ and Θ be the set of global minimizer(s) of $r(x)$ on S. We have convergence $\lim_{N\to\infty} r_N^* = r^*$ if $\Omega = Z(f, g)$ is finite and satisfies at least one of the following two conditions:*

(i) *Each curve $\mathcal{C}_i = \{x \in \mathbb{C}^2 : h_i(x) = 0\}$ $(i = 1, \ldots, m)$ is reduced and no two of them share an irreducible component. No point in Ω is a singular point of the curve $\mathcal{C}_1 \cup \cdots \cup \mathcal{C}_m$.*
(ii) *Each point of Ω is an isolated real common zero of $f(x) - r^* g(x)$ in $\mathbb{R}^2$, but not an isolated point of the feasible set S.*

Furthermore, if $\Omega = Z(f, g) \cup \Theta$ is finite and satisfies at least one of (i) *and* (ii)*, then we have finite convergence, i.e., there exists an integer N such that $r_N^* = r^*$.*

Proof First, assume that $\Omega = Z(f, g)$ is finite and satisfies at least one of (i) and (ii). For any $\gamma < r^*$, we have that the polynomial

$$\vartheta_\gamma(x) := f(x) - \gamma g(x)$$

is nonnegative on S. When $\vartheta_\gamma(u) = 0$ for some point $u \in S$, we must have $f(u) = g(u) = 0$, since otherwise $g(u) > 0$ and $r(u) = \gamma < r^*$, which is impossible. By the assumption in the theorem, the nonnegative polynomial $\vartheta_\gamma(x)$ has at most finitely many zeros on S. Now applying Corollary 3.7(if (i) holds) or Corollary 3.10 (if (ii) holds) in [30], we know that there exist SOS polynomials $\sigma_i(x)$ of degree high enough such that

$$\vartheta_\gamma(x) \equiv \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) h_i(x).$$

Second, assume that $\Omega = Z(f, g) \cup \Theta$ is finite and satisfies at least one of (i) and (ii). Consider the polynomial $\vartheta_{r^*}(x) := f(x) - r^* g(x)$, which is nonnegative on S. When $\vartheta_{r^*}(u) = 0$ for some $u \in S$, we must have either $f(u) = g(u) = 0$ or $r(u) = r^*$. Thus polynomial $\vartheta_{r^*}(x)$ has at most finitely many zeros on S. Corollary 3.7(if (i) holds) or Corollary 3.10 (if (ii) holds) in [30] implies that there are SOS polynomials $\sigma_i(x)$ with $\deg(\sigma_i h_i) \le 2N$ (N is large enough) such that

$$\vartheta_{r^*}(x) \equiv \sigma_0(x) + \sum_{i=1}^{m} \sigma_i(x) h_i(x)$$

which completes the proof. □

Example 4.9 Consider the problem

$$\begin{aligned} \min_x \quad & \frac{x_1^4 x_2^2 + x_1^2 x_2^4 + 1}{x_1^2 x_2^2}, \\ s.t. \quad & x_1, x_2 \ge 0, \quad 1 - x_1^2 - x_2^2 \ge 0. \end{aligned}$$

SOS relaxations (4.3)–(4.5) with order $N = 3$ yields lower bound $r_3 \approx 5.000$, and we can extract one point $x^* \approx (0.7071, 0.7071)$ from the dual solution to (4.6)–(4.9). $r(x^*) \approx 5.0000$ shows that it is a global minimizer.

Example 4.10 Consider the problem

$$\min_x \quad \frac{x_1^4 + x_1^2 + x_3^6}{x_1^2 x_3^2},$$
$$s.t. \quad x_1, x_3 \geq 0, \quad 1 - x_1^2 - x_3^2 \geq 0.$$

SOS relaxation (4.3)–(4.5) with order $N = 3$ yields lower bound $r_3 \approx 3.2324$, and we can extract one point $x^* \approx (0.6276, 0.7785)$ from the dual solution to (4.6)–(4.9). $r(x^*) \approx 3.2324$ shows that it is a global minimizer.

Example 4.11 Consider the problem

$$\begin{aligned}
\min_x \quad & \frac{x_1^3 + x_2^3 + 3x_1x_2 + 1}{x_1^2(x_2+1) + x_2^2(1+x_1) + x_1 + x_2}, \\
s.t. \quad & 2x_1 - x_1^2 \geq 0, \quad 2x_2 - x_2^2 \geq 0, \\
& 4 - x_1x_2 \geq 0, \quad x_1^2 + x_2^2 - \frac{1}{2} \geq 0.
\end{aligned}$$

SOS relaxation (4.3)–(4.5) with order $N = 2$ yields lower bound $r_2^* = 1$ and we can extract three points

$$(0, 1), \quad (1, 0), \quad (1, 1)$$

from the dual solution to (4.6)–(4.9). The evaluations of $r(x)$ at these three points show that they are all global minimizers.

Example 4.12 Consider the problem

$$\begin{aligned}
\min_x \quad & \frac{x_1^4 + x_2^4 + x_3^4 + x_1^2 + x_2^2 + x_3^2 + 2x_1x_2x_3(x_1 + x_2 + x_3)}{x_1^3 + x_2^3 + x_3^3 + 2x_1x_2x_3}, \\
s.t. \quad & x_1^4 + x_2^4 + x_3^4 = 1 + x_1^2x_2^2 + x_2^2x_3^2 + x_3^2x_1^2, \\
& x_3 \geq x_2 \geq x_1 \geq 0.
\end{aligned}$$

SOS relaxation (4.3)–(4.5) with order $N = 3$ yields $r_3^* \approx 2.0000$ and we can extract two points

$$x^* \approx (0.0000, 0.0000, 1.0000), \qquad x^{**} \approx (-0.0032, 0.9977, 0.9974)$$

from the dual solution to (4.6)–(4.9). x^* is feasible and $r(x^*) \approx 2.0000$ implies that x^* is a global minimizer. Now x^{**} is not feasible, but if we round x^{**} to the nearest feasible point we get $(0, 1, 1)$, which is another global minimizer since $r(0, 1, 1) = 2$.

Example 4.13 Consider the problem

$$\begin{aligned}
\min_x \quad & \frac{x_1^2 + x_2^2 + x_3^2 + x_4^2 + 2(x_2 + x_3 + x_1x_3 + x_1x_4 + x_2x_4) + 1}{x_1 + x_4 + x_1x_2 + x_2x_3 + x_3x_4}, \\
s.t. \quad & x_1^2 + x_2^2 - 2x_3x_4 = 0, \\
& 4 - x_1^2 - x_2^2 - x_3^2 - x_4^2 \geq 0, \\
& x_1, x_2, x_3, x_4 \geq 0.
\end{aligned}$$

SOS relaxation (4.3)–(4.5) with order $N = 3$ yields $r_2^* \approx 2.0000$ and we can extract one point

$$x^* \approx (0.0002, 0.0000, 0.0000, 0.9998)$$

from the dual solution to (4.6)–(4.9). $r(x^*) \approx 2.0000$ implies that x^* is a global minimizer (approximately). Actually the exact global minimizer is $(0, 0, 0, 1)$.

5 Conclusions and future work

This paper studies the global minimization of rational functions with or without constraints. Sum of squares relaxations are proposed to solve these problems. We discuss unconstrained and constrained minimization in Sects. 2 and 4, respectively. One application in finding the nearest common divisors of univariate polynomials is shown in Sect. 3. For constrained minimization, under some conditions, the convergence of SOS relaxations can be shown when the numerator and denominator have no common zeros on the feasible set. When the numerator and denominator have common zeros, the convergence might not hold. A counterexample is given in Remark 4.6.

The implementations of SOS relaxations rely on SDP solvers. Regarding the question on what size problems (the number of variables and degree) can be solved in practice using this approach, let us recall the relation between SOS polynomials and the resulting SDPs. For a polynomial $p(x)$ of degree $2d$ in n variables, $p(x)$ being SOS yields an LMI whose size is $\binom{n+d}{d}$. This number can be huge for moderate n and d, say, $n = d = 10$. However, when d is fixed, the size of the resulting LMI is polynomial in n. So the SOS approach is still tractable. There are also many numerical experiments on the size of polynomial problems which can be solved by SOS relaxations. See [23,24]. On the other hand, if the polynomials are sparse, then the size of the resulting LMI can be reduced significantly. For practical problems, the sparsity must be exploited to have efficient computations. We refer to [14,25,36] for work in this area.

The global minimization of rational functions has wide applications. Besides finding the nearest common divisors of univariate polynomials, in linear system theory, filter design can be formulated as an optimization related to rational functions [20]. Also in approximation theory, people are often interested in using rational functions to approximate a given function. This problem can be formulated as the global minimization of rational functions. In future work we will seek more interesting applications of rational functions.

There are also some other possible generalizations of applying SOS relaxations to minimize rational functions. Obviously, polynomials whose exponents are integers (they can be negative) are rational functions. SOS relaxations can also be generalized to minimize polynomials having rational exponents. For example, the global minimization of function

$$x_1^2 + x_1x_2 + \sqrt{x_1} + \sqrt[3]{x_2} + x_2^2$$

can be equivalently transformed to the global minimization of polynomial

$$z_1^8 + z_1^4 z_2^3 + z_1^2 + z_2 + z_2^6$$

by coordinate transformation $x_1 = x_1^4, x_2 = z_2^3$. Here we must pay attention to the domain of variables. This might be an interesting future work.

There are also some other convex relaxation methods for polynomial optimization problems, e.g., reformulation-linearization technique (RLT) by Sherali and Tuncbilek [31]. The RLT method generates polynomial implied constraints to be included in the original polynomial optimization problem, and successively linearizes the resulting problem by linear programming (LP) relaxation. They use the obtained lower bounds in branch-and-bound algorithms. The convergence of RLT methods is proven when the feasible set is compact. It is also possible to generalize methods like RLT to solve the global optimization problem of rational functions. For example, problems (4.1)–(4.2) can be reformulated as the following polynomial optimization problem

$$\begin{aligned} \min_{x\in\mathbb{R}^n, t\in\mathbb{R}} \quad & t \\ s.t. \quad & h_1(x) \geq 0, \ldots, h_m(x) \geq 0, \\ & t \cdot g(x) - f(x) \geq 0, \quad \ell \leq t \leq u, \end{aligned}$$

where ℓ is chosen sufficiently small and u is chosen sufficiently large. Then RLT methods can be applied to solve this new polynomial optimization problem. It is future work to study the special features of RLT methods to minimize rational functions.

Acknowledgements The authors would like to thank Bernd Sturmfels and two referees for helpful comments.

References

1. Arnon, D., McCallum, S.: A polynomial time algorithm for the topological type of a real algebraic curve. J. Symbol. Comput. **5**, 213–236
2. Brown, W., Traub, J.: On Euclid's algorithm and the theory of subresultants. J. ACM **18**, 505–514
3. Collins, G.: Subresultants and reduced polynomial remainder sequences. J. ACM **14**, 128–142
4. Curto, R.E., Fialkow, L.: The truncated complex K-moment problem. Trans. Amer. Math. Soc. **352**, 2825–2855 (2000)
5. de Klerk, E.: Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications. Applied Optimization Series, vol. 65, 300 pp. Kluwer Academic Publishers, Dordrecht (2002)
6. Delzell, C., Prestel, A.: Positive Polynomials, Monographs in Mathematics. Springer, Berlin (2001)
7. Garey, M., Johnson, D.: Computers and Intractability: A guide to the theory of NP-completeness. W.H. Freeman and Company, New York (1979)
8. Henrion, D., Lasserre, J.: GloptiPoly: Global optimization over polynomials with Matlab and SeDuMi. ACM Trans. Math. Soft. **29**, 165–194 (2003)
9. Henrion, D., Lasserre, J.: Detecting global optimality and extracting solutions in GloptiPoly. In: Henrion, D., Garulli, A. (eds.) Positive Polynomials in Control, Lecture Notes on Control and Information Sciences. Springer Verlag, Berlin (2005)
10. Jibetean, D., de Klerk, E.: Global optimization of rational functions: a semidefinite programming approach. Math. Prog. **106**(1), 93–109 (2006)
11. Jibetean, D., Laurent M.: Semidefinite approximations for global unconstrained polynomial optimization. To appear in SIAM J. Optim.
12. Karmarkar, N., Lakshman, Y.: Approximate polynomial greatest common divisors and nearest singular polynomials. In: Proceedings ISSAC'96 (Zurich, Switzerland). ACM Press, New York
13. Karmarkar, N., Lakshman, Y.: On approximate GCDs of univariate polynomials. J. Symbol. Comput. **26**, 653–666 (1998)
14. Kojima, M., Kim, S., Waki, H.: Sparsity in sums of squares of polynomials.. Math. Program. **103**(1), 45–62 (2003)
15. Lasserre, J.: Global optimization with polynomials and the problem of moments. SIAM J. Optim. **11**(3), 796–817 (2001)
16. Laurent, M.: Semidefinite representations for finite varieties. To appear in Math. Prog.

17. Löfberg, J.: YALMIP: A MATLAB toolbox for rapid prototyping of optimization problems. Website: http://control.ee.ethz.ch/~joloef/yalmip.php
18. Manocha, D., Demmel, J.: Algorithms for interesting parametric and implicit curves I: simple intersections. ACM Trans. Graphics. **13**, 73–100
19. Nesterov, Y.: Squared functional systems and optimization problems. In: Frenk, H. et al. (eds.) High Performance Optimization, pp. 405–440. Kluwer Academic Publishers, Dordrecht (2000)
20. Nie, J., Demmel, J.: Shape optimization of transfer functions. In: Hager, W., Pardalos, P., Huang, S. et al. (eds.) Multiscale optimization methods and applications, pp. 313–326. Springer series on nonconvex optimization and its application series, Berlin (2005)
21. Nie, J., Demmel, J., Sturmfels, B.: Minimizing polynomials via sum of squares over the gradient ideal. Math. Prog. Ser. A **106**(3), 587–606 (2006)
22. Nie, J., Schweighofer, M.: On the complexity of Putinar's Positivstellensatz. To appear in J. Complexity.
23. Parrilo, P., Sturmfels, B.: Minimizing polynomial functions. In: Basu, S., Gonzalez-Vega, L. (eds.) Proceedings of the DIMACS Workshop on Algorithmic and Quantitative Aspects of Real Algebraic Geometry in Mathematics and Computer Science (March 2001), pp. 83–100. American Mathematical Society, Providence RI (2003)
24. Parrilo, P.: Semidefinite programming relaxations for semialgebraic problems. Math. Prog. Ser. B **96**(2), 293–320 (2003)
25. Parrilo, P.: Exploiting structure in sum of squares programs. In: Proceedings for the 42nd IEEE Conference on Decision and Control. Maui, Hawaii (2003)
26. Prajna, S., Papachristodoulou, A., Seiler, P., Parrilo, P.: SOSTOOLS User's Guide. http://www.mit.edu/~parrilo/SOSTOOLS/
27. Putinar, M.: Positive polynomials on compact semi-algebraic sets. Ind. Univ. Math. J. **42**, 203–206 (1993)
28. Reznick, B.: Some concrete aspects of Hilbert's 17th problem. Contemporary Mathematics, vol 253, pp. 251–272. American Mathematical Society, Providence, RI (2000)
29. Rockafellar, R.: Convex Analysis, 2nd edn. Princeton University Press, Princeton, UJ (1970)
30. Scheiderer, C.: Distinguished representations of non-negative polynomials. J. Algebra **289**, 558–573 (2005)
31. Sherali, H., Tuncbilek, C.: A global optimization algorithm for polynomial programming problems using a Reformulation-Linearization Technique. J. Glob. Optim. **2**, 101–112 (1992)
32. Shor, N.: Quadratic optimization problems. Soviet J. Comput. Systems Sci. **25**, 1–11 (1987)
33. Stengle, G.: Complexity estimates for the Schmüdgen positivstellensatz. J. Complexity **12**, 167–174 (1996)
34. Stetter, H.: Numerical Polynomial Algebra. Society of Industrial and Applied Mathematics (2004)
35. Sturm, J.F.: SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. Optim. Method Softw. **11 & 12**, 625–653 (1999)
36. Waki, H., Kim, S., Kojima, M., Muramatsu, M.: Sums of squares and semidefinite programming relaxations for polynomial optimization problems with structured sparsity. To appear in SIAM J. Optim.
37. Wolkowicz, H., Saigal, R., Vandenberghe, L. (eds.): Handbook of Semidefinite Programming. Kluwer Academic Publishers, Dordrecht (2000)