Constrained Optimization using Sum-of-Squares Techniques

Javad Lavaei, Somayeh Sojoudi, and Amir G. Aghdam
Department of Electrical and Computer Engineering, Concordia University
Montréal, QC Canada H3G 1M8
{j_lavaei, s_sojoudi, aghdam}@ece.concordia.ca

Abstract

Motivated by many control applications, this paper deals with the optimization of a rational function subject to a number of rational inequalities. First, the problem of finding the infimum of a given polynomial function is formulated as a sum-of-squares (SOS) problem, which can be handled efficiently by existing software tools. The results obtained are then extended to the rational functions. The problem of finding the infimum of a rational function subject to some inequalities in the form of some other rational functions is then investigated. To this end, the infimum of the rational objective function is computed to determine whether it is finite or not. If it is finite, a simple SOS formulation is presented in a way similar to the polynomial case. In the case when the infimum is not finite but some mild \textit{a priori} knowledge is available about either the constraints or the solution, the problem is formulated completely as SOS. The efficacy of the proposed methods are demonstrated in three numerical examples.

I. INTRODUCTION

Optimization often appears in many practical problems, and has attracted many researchers in the area of control systems. Optimization problems can be categorized as constrained and unconstrained, where the constraints can be in the forms of equalities and inequalities. An important class of optimization problems involves minimization of a rational function, and in some cases subject to certain rational inequalities. Problems of this kind arise in several practical applications, some of which are listed below:

- The high-performance decentralized control design problem, where a set of local controllers is desired to be obtained for the minimum achievable performance index, can be formulated as the computation of the global optimum of a polynomially constrained optimization problem [1].
- The problem of identifying the state-space model of a structural dynamic system satisfying some constraints can be translated to the minimization of a rational function subject to some rational constraints [2]. The main concern in this problem is how to find the global solution as opposed to a local one.
- In constrained model predictive control, where it is desired to predict the controlled variables over a future horizon, the minimization of a polynomial subject to some polynomial constraints is to be carried out in order to treat the problem [3], [4].
- Certain robust control problems such as parametric stability margin computation, can be formulated as checking the positiveness of a polynomial on a hyperrectangle, as pointed out in [5].
- The minimum norm problem which is investigated in the literature intensively, turns out to be equivalent to finding the global optimum of a polynomially constrained optimization problem [6].
- Minimization of a rational function is inevitably required in the problem of optimal model reduction [7].

The above practical applications point to the viable role of the aforementioned optimization problem in the real-world systems. This paper deals with the optimization of a rational function subject to a number of constraints by means of sum-of-squares (SOS) techniques [8], [9], [10]. The problem of finding the infimum of a polynomial is first considered, and SOS formulations are presented accordingly. The obtained SOS problems can be solved by using a number of softwares quite efficiently. The proposed approach is then extended to the case of finding the infimum of a rational function. Finally, the problem of obtaining the infimum of a rational function over a region defined by some other rational functions is investigated. As the first step, it is checked whether the objective function has a lower bound or not, and in the case of boundedness, a simple SOS formulation is presented. For the case when the objective function is unbounded from below, it is shown that if some \textit{a priori} knowledge is available, the problem can be solved efficiently. This \textit{a priori} knowledge can be the radius of a ball which contains the region defined by the rational functions, if exists. However, the knowledge on the lower bound of the infimum to be found suffices to solve the aforementioned problem, in general.

This paper is organized as follows. First, the existing methods to solve a rationally constrained optimization problem are pointed out in Section II, and their advantages and disadvantages are discussed thoroughly. The problem of finding the infimum of a rational function is then investigated in Section III, and is extended to the constrained case in Section IV. Some illustrative numerical examples are presented in Section V. Finally, some concluding remarks are given in Section VI.

This work has been supported by the Natural Sciences and Engineering Research Council of Canada under grant RGPIN-262127-03.
II. Preliminaries

Consider a polynomial \( f(x) \), where \( x = [x_1 \ x_2 \ \cdots \ x_n] \), and denote its infimum with \( \alpha_x \). In the recent years, the problem of finding \( \alpha_x \) has been investigated in the literature intensively from the viewpoint of semidefinite programming (SDP). These works will be surveyed below, and their drawbacks will be highlighted.

The work [11] states that \( \alpha_x \) is equal to the maximum value of \( \alpha \), such that the polynomial \( f(x) - \alpha \) is nonnegative. Then, in order to alleviate the complexity of the problem, it relaxes the condition of nonnegativity of \( f(x) - \alpha \) to being SOS. This relaxation is made based on the obvious fact that any SOS polynomial is nonnegative, however its converse is not necessarily true. Hence, the work [11] proposes the new problem of maximization of \( \alpha \) such that \( f(x) - \alpha \) is SOS, which can be handled by the relevant softwares. Note that the obtained solution is a lower bound for \( \alpha_x \). Although this approach works satisfactorily to some degree, it can be very conservative in general, due to the aforementioned relaxation. As an example, the infimum obtained for the polynomial:

\[
x_1^4x_1^2 + x_1^2x_2^4 + 1 - 3x_1^2x_2^2
\]

(1)

by utilizing this method is equal to \(-\infty\), while the exact infimum is 0. It is shown in [12] that for any integer \( d \geq 2 \), the ratio of the volume of the nonnegative non-SOS homogeneous polynomials of degree \( 2d \) and that of the SOS homogeneous polynomials with the same degree rapidly grows towards infinity, as \( n \) goes to infinity. This implies that the relaxation used in [11] and some other relevant papers is not always well-established.

As a remedy for the drawback associated with [11], the technique of using some \( \alpha \) priori knowledge of the minimizer of \( f(x) \) is exploited in [13]. Assume that \( x_\epsilon \) is known to be inside a ball of radius \( r \) centered at the origin. It is a direct consequence of Putinar’s theorem [14] that \( \alpha_x \) is equal to the maximum value of \( \alpha \), for which there exist two SOS polynomials \( \phi_1(x) \) and \( \phi_2(x) \) with the following property:

\[
f(x) - \alpha = (r^2 - xx^T) \phi_1(x) + \phi_2(x)
\]

(2)

Note that the coefficients of the polynomials \( \phi_1(x) \) and \( \phi_2(x) \) are in terms of \( \alpha \). The advantage of this formulation is that it is a SOS problem. Nevertheless, this method cannot address the following questions systematically:

- Does \( f(x) \) have an infimum?
- Can \( f(x) \) attain its infimum, if it exists (i.e. does there exist a finite point corresponding to that infimum)?
- If the infimum is attainable, how can the radius \( r \) be determined?

The open questions given above (specially the last one) make this approach ad-hoc in general. This technique is also utilized in [15].

The method proposed in [16] attempts to eliminate the gap between SOS polynomials and nonnegative polynomials, which is useful in resolving the deficiency of the work [11]. Consider a nonnegative polynomial \( p(x) \). It is shown in [16] that for any \( \epsilon > 0 \), there exists a number \( r(p, \epsilon) \) such that the polynomial:

\[
p(x) + \epsilon \sum_{i=1}^{r(p, \epsilon)} \sum_{j=1}^{n} \frac{x_j^{2i}}{i!}
\]

(3)

is SOS. Note that \( r(p, \epsilon) \) depends on \( p(x) \) and \( \epsilon \). This nice result incorporates the nonnegative polynomials into the SOS ones.

The work [17] considers the problem of minimizing a polynomial \( f(x) \). It perturbs \( f(x) \) by a penalty function as:

\[
f(x) + \epsilon \sum_{i=1}^{n} x_i^{2\sigma+2}
\]

(4)

where \( 2\sigma \) denotes the degree of \( f(x) \). The method proposed in [17] asserts the following advantages of the perturbed \( f(x) \) given by (4):

- The infimum of the perturbed \( f(x) \) approaches that of \( f(x) \), as \( \epsilon \) goes to zero.
- Although \( f(x) \) may not attain its infimum, the perturbed \( f(x) \) always attains the corresponding infimum.

An algorithm is then proposed in [17] to obtain the infimum of the perturbed \( f(x) \). However, as pointed out in [18], the required computational cost is huge, which restricts its applications to small-sized problems. Besides, it may have the problem of ill-conditioning like many other penalty-based approaches.

The results of [17] have further been developed in [18]. It is shown that the infimum of the perturbed \( f(x) \) given by (4) is inside a ball. The radius of this ball is also obtained in [18]. Next, the ball technique mentioned earlier is employed to find the infimum of the perturbed \( f(x) \). Nonetheless, there are some shortcomings with regard to this approach. First of all, the radius of that ball is proportional to \( \frac{1}{\epsilon^2} \), which is usually very large. Moreover, some of the values used in the formulation are in terms of \( \frac{1}{\epsilon^2} \). These result in an ill-conditioned optimization problem, for which \( \epsilon \) should be considered neither small (due to the mentioned difficulties) nor large (due to the required accuracy). However, unlike the other existing methods which seek a lower bound for \( \alpha_x \), the work in [18] presents an upper bound for it.
It is shown in [10] that if the infimum of \( f(x) \) is attainable, then it is equal to the maximum value of \( \alpha \) for which there exist a SOS polynomial \( \phi_0(x) \) and polynomials \( \phi_1(x), \ldots, \phi_n(x) \) such that:

\[
f(x) - \alpha = \phi_0(x) + \phi_1(x) \frac{\partial f(x)}{\partial x_1} + \cdots + \phi_n(x) \frac{\partial f(x)}{\partial x_n}
\]

(5)

It is stated in [10] that if a certain condition does not hold, the degrees of the polynomials \( \phi_0(x), \ldots, \phi_n(x) \) should ideally be assumed infinity. In other words, in that case the infimum will be obtained asymptotically (in infinite iterations).

The work [19] deals with the global optimization of a polynomial \( f(x) \). One of the requirements of the approach in [19] is that \( f(x) \) should be bounded from below. This has tried to improve the approach in [10] which is unable to deal with the polynomials whose infimums are unattainable. Indeed, it has introduced the notion of principal gradient tentacle, as opposed to the gradient variety used in [10]. For the polynomial \( f(x) \), its gradient tentacle is defined to be:

\[
S(\nabla f(x)) = \{ x : \| \nabla f(x) \| \leq 1 \}
\]

(6)

where

\[
\| \nabla f(x) \|^2 = \left( \frac{\partial f(x)}{\partial x_1} \right)^2 + \cdots + \left( \frac{\partial f(x)}{\partial x_n} \right)^2
\]

(7)

It is then stated that if \( f(x) \) has isolated singularities only at infinity, or alternatively if \( S(\nabla f(x)) \) is compact, then \( \alpha_* \) is equal to the maximum value of \( \alpha \) for which there exist two SOS polynomials \( \phi_1(x) \) and \( \phi_2(x) \) such that:

\[
f(x) - \alpha = \phi_1(x) + (1 - \| \nabla f(x) \|^2 \| x \|^2) \phi_2(x)
\]

(8)

However, the degrees of the polynomials \( \phi_1(x) \) and \( \phi_2(x) \) are sometimes infinity, i.e., \( \alpha_* \) will be obtained through an asymptotical convergence. Some other drawbacks are pointed out in [19]. First of all, if the infimum does not exist, this method will not detect it, and will lead to a wrong solution. Furthermore, in the case when the infimum is not attainable, this method can be very time-consuming. In comparison with other existing methods, one can easily infer that the term \( \| \nabla f(x) \|^2 \| x \|^2 \) will increase the numerical complexity of the problem noticeably (because of its degree).

Consider now the problem of computing the infimum of a given rational function \( \frac{f(x)}{h(x)} \) over the region \( \mathcal{D} \) defined by:

\[
\mathcal{D} = \{ x : g_1(x) \geq 0, \ldots, g_k(x) \geq 0 \}
\]

(9)

One of the most important results presented in the literature concerning this problem is Putinar's theorem. This theorem requires that the following qualification be satisfied:

**Qualification 1:** There exist SOS polynomials \( z_0(x), z_1(x), \ldots, z_k(x) \), such that the set of all vectors \( x \) satisfying the inequality:

\[
z_0(x) + z_1(x)g_1(x) + \cdots + z_k(x)g_k(x) \geq 0
\]

(10)

is compact.

Since Putinar's theorem will essentially be required in this paper, it is given below.

**Theorem 1:** [14] Assume that Qualification 1 holds for the polynomials \( g_1(x), \ldots, g_k(x) \). If a function \( p(x) \) is strictly positive over the region \( \mathcal{D} \), then there exist SOS polynomials \( \bar{z}_0(x), \bar{z}_1(x), \ldots, \bar{z}_k(x) \) with the following property:

\[
p(x) = \bar{z}_0(x) + \bar{z}_1(x)g_1(x) + \cdots + \bar{z}_k(x)g_k(x)
\]

(11)

The problem of finding the infimum of \( \frac{f(x)}{h(x)} \) over \( \mathcal{D} \) is tackled in [15], by assuming that:

i) \( \mathcal{D} \) is the closure of some compact open connected set.

ii) Qualification 1 holds.

iii) \( f(x) \) and \( h(x) \) have no common real roots in \( \mathcal{D} \).

It then exploits Putinar's theorem to conclude that the infimum of \( \frac{f(x)}{h(x)} \) over the region \( \mathcal{D} \) is equal to the supremum of \( \alpha \) for which there exist SOS polynomials \( \phi_0(x), \phi_1(x), \ldots, \phi_n(x) \) such that:

\[
f(x) - \alpha h(x) = \phi_0(x) + \phi_1(x)g_1(x) + \cdots + \phi_n(x)g_n(x)
\]

(12)

The above assumptions confine the application of this approach. In fact, they can be very restrictive, and their verification (especially the requirement (iii)) is not straightforward in the case of a rational function (as opposed to a polynomial). For the unconstrained optimization (i.e., when \( \mathcal{D} \) spans the whole space), this method utilizes the technique of the big ball. As pointed out earlier, this method is problematic, as it is unknown whether the infimum is attainable, or how to find its radius, if exists. Similar techniques are used in [20], but the problem is converted to a dual SDP in order to compute the minimizers, in addition to the infimum.

The work [21] considers the problem of minimizing a polynomial \( f(x) \) over the region \( \mathcal{D} \). While the works [15], [20] require the compactness of \( \mathcal{D} \), the method proposed in [21] eliminates this restrictive assumption. It is shown that if the minimum
occurs at a Karush-Kuhn-Tucker (KKT) point, then \( \alpha_* \) is equal to the maximum of \( \alpha \) for which there exist SOS polynomials \( \phi_1(x), \phi_2(x), \ldots, \phi_d(x) \) and polynomials \( \psi_1(x), \ldots, \psi_n(x), \chi_1(x), \ldots, \chi_k(x) \) such that:

\[
f(x) - \alpha = \sum_{i=1}^{d} \phi_i(x) g_i(x) + \sum_{i=1}^{n} \psi_i(x) \left( \frac{\partial f(x)}{\partial x_i} - \sum_{j=1}^{k} \lambda_j \frac{\partial g_j(x)}{\partial x_i} \right) + \sum_{i=1}^{k} \chi_i(x) \lambda_i g_i(x)
\]

Note that the term \((j_1 \ldots j_k) = (i-1)2\) in the above relation implies that \(j_1, \ldots, j_k\) are the digits of the number \(i-1\) in the base 2. Furthermore, \( \lambda_i \)'s play the role of Lagrangian multipliers. Aside from the complexity of the above formulation, as indicated in [21], the assumption that the minimum occurs at a KKT point is not trivial and cannot be relaxed, in general. For instance, this method is not able to find the minimum of \( x_1^2 \) subject to the constraint \( x_1^2 \geq 0 \).

It is evident that the above-mentioned methods require to make certain assumptions, which can be very restrictive, in general. Furthermore, they either present complicated formulas or lead to very conservative results. In this paper, novel approaches will be presented to find the infimum of a rational function, and also the infimum of a rational function over a region. The proposed approaches are SOS-based and far simpler than the existing methods.

III. THE GLOBAL SOLUTION OF AN UNCONSTRAINED OPTIMIZATION PROBLEM

In this section, it is desired first to present a simple methodology for finding the infimum of a polynomial, without making any assumption. Then, the procedure will be extended to the case of finding the infimum of a rational function.

A. The infimum of a polynomial

Consider a polynomial \( f(x) \). It is obvious that the infimum of this polynomial, denoted by \( \alpha_* \), can be obtained from the following relation:

\[
\alpha_* = \{ \sup(\alpha) : f(x) - \alpha \geq 0, \ \forall x \in \mathbb{R}^n \}
\]

When \( f(x) \) is of odd degree, \( \alpha_* \) is equal to \(-\infty\). Thus, assume that its degree is \( 2\sigma \), where \( \sigma \) is a positive integer. Let \( \mu \) represent a slack variable. Rewrite the function \( f \left( \frac{x}{\mu} \right) \) as \( \mu^{-2\sigma} f(x, \mu) \), where \( f(x, \mu) \) is a polynomial. A tight lower bound for \( \alpha_* \) will be given in the next theorem.

Theorem 2: Let \( \alpha_0 \) denote the maximum value of \( \alpha \) for which there exist a polynomial \( \phi_1(x, \mu) \) and a SOS polynomial \( \phi_2(x, \mu) \) such that:

\[
f(x, \mu) - \alpha \mu^{2\sigma} = (1 - xx^T - \mu^2) \phi_1(x, \mu) + \phi_2(x, \mu)
\]

Then, \( \alpha_0 \) is a lower bound for \( \alpha_* \).

Proof: It is straightforward to show that:

\[
\alpha_* = \{ \sup(\alpha) : f(x, \mu) - \alpha \mu^{2\sigma} \geq 0, \ \forall x \in \mathbb{R}^n, \ \mu \in \mathbb{R} \}
\]

On the other hand, one can easily conclude that \( f(x, \mu) - \alpha_0 \mu^{2\sigma} \) is a homogeneous polynomial of degree \( 2\sigma \). Therefore, the relation:

\[
\tilde{f}(\lambda x, \mu) - \alpha_0 (\lambda \mu)^{2\sigma} := \lambda^{2\sigma} (f(x, \mu) - \alpha_0 \mu^{2\sigma})
\]

holds for any real number \( \lambda \). It results from the equation (15) that \( \tilde{f}(x, \mu) - \alpha_0 \mu^{2\sigma} \) is nonnegative for any \( x \) and \( \mu \) satisfying the equality \( xx^T + \mu^2 = 1 \). Using the scaling property (17), one can deduce that \( \tilde{f}(x, \mu) - \alpha_0 \mu^{2\sigma} \) is nonnegative for \( x \) and \( \mu \), as long as the inequality \( xx^T + \mu^2 \neq 0 \) holds. As a consequence of this result and by noting that the homogeneous polynomial \( \tilde{f}(x, \mu) - \alpha_0 \mu^{2\sigma} \) is equal to zero at the origin, it can be concluded that \( \tilde{f}(x, \mu) - \alpha_0 \mu^{2\sigma} \) is always nonnegative. The proof follows now from the relation (16).

Theorem 2 presents a simple SOS formulation for calculating the infimum of a polynomial, and can be easily solved by using a number of software, e.g. YALMIP or SOSTOOLS [22], [23]. Note that the discrepancy between \( \alpha_0 \) and \( \alpha_* \) depends on the possibility of representing the polynomial \( \tilde{f}(x, \mu) - \alpha_0 \mu^{2\sigma} \) (which is nonnegative inside the closed unit ball) as the one given in (15). In fact, Putinar’s theorem states that if \( \tilde{f}(x, \mu) - \alpha_0 \mu^{2\sigma} \) is positive, such representation is possible. However, in the case when it is nonnegative, Putinar’s theorem cannot be used (see the counterexamples in [24]). It is worth mentioning that the proposed method has been applied to several problematic examples, and it could solve all of them accurately (i.e., \( \alpha_* = \alpha_0 \)), as shown in Section V.

Remark 1: One can follow the procedures proposed in [25] and [26] to find bounds on the degrees of the polynomials \( \phi_1(x, \mu) \) and \( \phi_2(x, \mu) \) given in Theorem 2 (for obtaining \( \alpha_0 \)). Note that if the degrees of \( \phi_1(x, \mu) \) and \( \phi_2(x, \mu) \) are not chosen sufficiently large, the solution of the SOS problem presented in Theorem 2 may be visibly different from the exact value of \( \alpha_0 \). However, any value obtained, no matter how far from the exact value is, can be considered as a lower bound for \( \alpha_0 \) and consequently, \( \alpha_* \).

The result of Theorem 2 presents a lower bound for the infimum, rather than an exact value for it. The following theorem presents an efficient method to find the infimum \( \alpha_* \) precisely.
Theorem 3: For any $\varepsilon > 0$, let $\alpha^\varepsilon_0$ denote the maximum value of $\alpha$ for which there exist two SOS polynomials $\phi_1^x(x, \mu)$ and $\phi_2^x(x, \mu)$ such that:

$$f(x, \mu) + \varepsilon - \alpha \left( \mu^{2\sigma} + \varepsilon^2 \right) = (1 - x\mu^T - \mu^2) \phi_1^x(x, \mu) + \phi_2^x(x, \mu)$$  \hspace{1cm} (18)

Then $\alpha^\varepsilon_0$ either equals $\alpha_*$ or converges to $\alpha_*$ as $\varepsilon \to 0$.

Proof: Denote the infimum of the rational function $\frac{f(x, \mu) + \varepsilon}{\mu^{2\sigma} + \varepsilon^2}$ in the unit ball with $\bar{\alpha}_0^\varepsilon$. It can be easily verified that $\bar{\alpha}_0^\varepsilon$ satisfies the following relation:

$$\bar{\alpha}_0^\varepsilon = \left\{ \sup(\alpha) : \hat{\alpha}(x, \mu) + \varepsilon - \alpha \left( \mu^{2\sigma} + \varepsilon^2 \right) \geq 0, \forall (x, \mu) \in B \right\}$$  \hspace{1cm} (19)

where $B$ denotes the unit ball. Since $\mu^{2\sigma} + \varepsilon^2$ is always positive, Lemma 1 in [15] along with the above equation yield that:

$$\bar{\alpha}_0^\varepsilon = \left\{ \sup(\alpha) : \hat{\alpha}(x, \mu) + \varepsilon - \alpha \left( \mu^{2\sigma} + \varepsilon^2 \right) > 0, \forall (x, \mu) \in B \right\}$$  \hspace{1cm} (20)

(note that the difference between (19) and (20) is the inclusion of zero in the inequality in (19)). Thus, it follows from Putinar’s theorem that $\bar{\alpha}_0^\varepsilon$ is the same as the maximum value of $\alpha$ for which there exist two SOS polynomials $\phi_1^x(x, \mu)$ and $\phi_2^x(x, \mu)$ such that:

$$\hat{f}(x, \mu) + \varepsilon - \alpha \left( \mu^{2\sigma} + \varepsilon^2 \right) = (1 - x\mu^T - \mu^2) \phi_1^x(x, \mu) + \phi_2^x(x, \mu)$$  \hspace{1cm} (21)

It can be concluded from (18) and (21) that $\bar{\alpha}_0^\varepsilon$ is equal to $\alpha_0^\varepsilon$, i.e., the infimum of the rational function $\frac{f(x, \mu) + \varepsilon}{\mu^{2\sigma} + \varepsilon^2}$ in the unit ball. On the other hand, it can be shown that $\alpha_*$ is the same as the infimum of the rational function $\frac{f(x, \mu)}{\mu^{2\sigma} + \varepsilon^2}$. Since the numerator and denominator of this rational function are homogeneous of the same degree, the infimum corresponds to infinitely many points which can lie anywhere in the $n + 1$ dimensional space, unless it is exactly the origin. Therefore, the infimum of the rational function $\frac{f(x, \mu)}{\mu^{2\sigma} + \varepsilon^2}$ in the unit ball is $\alpha_*$. Let $(\hat{x}_*, \hat{\mu}_*)$ denote one of the minimizers corresponding to the infimum of the rational function $\frac{f(x, \mu)}{\mu^{2\sigma} + \varepsilon^2}$ in the unit ball. One can write:

$$\alpha_0^\varepsilon = \frac{\hat{f}(\hat{x}_*, \hat{\mu}_*) + \varepsilon}{\mu^{2\sigma} + \varepsilon^2} \leq \frac{\hat{f}(0, 0) + \varepsilon}{0 + \varepsilon^2} = \frac{1}{\varepsilon}$$  \hspace{1cm} (22)

Therefore $\frac{\hat{f}(\hat{x}_*, \hat{\mu}_*)}{\mu^{2\sigma} + \varepsilon^2} \leq \frac{1}{\varepsilon^2}$. Consequently:

$$\alpha_0^\varepsilon = \frac{\hat{f}(\hat{x}_*, \hat{\mu}_*) + \varepsilon}{\mu^{2\sigma} + \varepsilon^2} \leq \frac{\hat{f}(\hat{x}_*, \hat{\mu}_*)}{\mu^{2\sigma}} \geq \alpha_*$$  \hspace{1cm} (23)

On the other hand, it can be easily shown that the value of $\frac{f(x, \mu)}{\mu^{2\sigma} + \varepsilon^2}$ at any arbitrary point can be attained asymptotically by the function $f(x, \mu) + \varepsilon$ (by virtue of $\varepsilon \to 0$). This completes the proof.

It is to be noted that part of the proof of Theorem 3 relies on the fact that the infimums of $\frac{f(x, \mu)}{\mu^{2\sigma} + \varepsilon^2}$ and $\frac{f(x, \mu) + \varepsilon}{\mu^{2\sigma} + \varepsilon^2}$ can become arbitrarily close to each other by choosing a sufficiently small $\varepsilon$. However, a question may arise why the function $\frac{f(x, \mu)}{\mu^{2\sigma} + \varepsilon^2}$ (which has a simpler form) was not considered instead of $\frac{f(x, \mu) + \varepsilon}{\mu^{2\sigma} + \varepsilon^2}$. It is interesting to note that the statement is not valid for the above-mentioned function. For instance, consider the rational function $\frac{(\mu^{2\sigma} + \varepsilon^2)^2}{\mu^{2\sigma} + \varepsilon^2}$. The infimum of this rational function is equal to 4. In contrast, the (attainable) infimum of $\frac{(x^2 + \varepsilon^2)^2}{(x^2 + \varepsilon^2)^2 + \varepsilon^2}$ is equal to 0, no matter how small $\varepsilon$ is.

Theorem 3 presents a SOS problem, which leads to finding $\alpha_*$. It is to be noted that there are bounds on the degrees of the polynomials $\phi_1^x(x, \mu)$ and $\phi_2^x(x, \mu)$ (see Remark 1). In addition, one can obtain some bounds on the relative error between $\alpha^\varepsilon_0$ and $\alpha_*$. For instance, it is straightforward to show that in the case when the infimum $\alpha_*$ is attainable, there exists a positive number $\varepsilon_0$ such that $\alpha^\varepsilon_0$ is always between $\alpha_*$ and $\alpha_* + \sqrt{\varepsilon}$, for any $\varepsilon \in (0, \varepsilon_0)$.

B. The infimum of a rational function

It is desired now to find the infimum of the rational function $\frac{f(x)}{h(x)}$. Without loss of generality, suppose that $f(x)$ and $h(x)$ are coprime, otherwise one can pursue the existing methods to eliminate their greatest common divisor (GCD). The following lemma is borrowed from [15].

Lemma 1: If the value of the function $h(x)$ is negative at one point and positive at another point, then the infimum of $\frac{f(x)}{h(x)}$ is $-\infty$. At this point, it is required to check whether or not $h(x)$ changes sign. This can sometimes be inferred from the nature of the polynomial $h(x)$. For example, when $h(x)$ is the square of another function, it is always nonnegative. However, the nonnegativeness or nonnegativeness of $h(x)$ can be verified, in general, by using the method proposed in the previous subsection. More precisely, Theorems 2 and 3 can be employed to find the infimum of $h(x)$, leading to one of the following possibilities:

- The infimum of $h(x)$ is nonnegative.
- The infimum of $h(x)$ is negative and finite. In this case, the infimum of $\frac{f(x)}{h(x)}$ is $-\infty$. 


The infimum of \( h(x) \) is \(-\infty\). Compute now the infimum of \(-h(x)\). If it is negative (finite or infinite), it means that \( h(x) \) takes both negative and positive values, which implies that the infimum of \( \frac{f(x)}{h(x)} \) is \(-\infty\). Otherwise, the infimum of \( \frac{f(x)}{h(x)} \) is finite. In this case, negate both the numerator and the denominator of \( \frac{f(x)}{h(x)} \) in order to make the infimum of its denominator nonnegative.

Without loss of generality, assume that \( h(x) \) is always nonnegative. It is evident that the infimum of the function \( \frac{f(x)}{h(x)} \), denoted by \( \alpha_* \), can alternatively be obtained from the following relation (e.g., see [20]):

\[
\alpha_* = \{ \sup(\alpha) : f(x) - \alpha h(x) \geq 0, \quad \forall x \in \mathbb{R}^n \} \tag{24}
\]

Now, pursuing approaches similar to the previous subsection, the following two theorems will be obtained, which represent extension of the results of Theorems 2 and 3 to the case of rational functions.

**Theorem 4:** Consider a slack variable \( \mu \). Rewrite the rational function \( \frac{f(x)}{h(x)} \) as \( \frac{f(x,\mu)}{h(x,\mu)} \) where \( f(x,\mu) \) and \( h(x,\mu) \) are two polynomials. Let \( \alpha_\varepsilon \) denote the maximum value of \( \alpha \) for which there exist a polynomial \( \phi_1(x,\mu) \) and a SOS polynomial \( \phi_2(x,\mu) \) such that:

\[
f(x,\mu) - \alpha h(x,\mu) = (1 - xx^T - \mu^2) \phi_1(x,\mu) + \phi_2(x,\mu) \tag{25}
\]

Then \( \alpha_\varepsilon \) is a lower bound for \( \alpha_* \).

**Theorem 5:** For any \( \varepsilon > 0 \), let \( \alpha_{\varepsilon}^\varepsilon \) denote the maximum value of \( \alpha \) for which there exist two SOS polynomials \( \phi_1^{\varepsilon}(x,\mu) \) and \( \phi_2^{\varepsilon}(x,\mu) \) such that:

\[
f(x,\mu) + \varepsilon - \alpha \left( h(x,\mu) + \varepsilon^2 \right) = (1 - xx^T - \mu^2) \phi_1^{\varepsilon}(x,\mu) + \phi_2^{\varepsilon}(x,\mu) \tag{26}
\]

Then \( \alpha_{\varepsilon}^\varepsilon \) equals \( \alpha_* \), or converges to \( \alpha_* \) as \( \varepsilon \to 0 \).

**Remark 2:** In the case when \( f(x) \) and \( h(x) \) are homogeneous of the same degree, one can consider the following equation, instead of the one given in (25):

\[
f(x) - \alpha h(x) = (1 - xx^T) \phi_1(x) + \phi_2(x) \tag{27}
\]

and the one given below, instead of (26):

\[
f(x) + \varepsilon - \alpha \left( h(x) + \varepsilon^2 \right) = (1 - xx^T) \phi_1^{\varepsilon}(x) + \phi_2^{\varepsilon}(x) \tag{28}
\]

In other words, introducing a redundant variable \( \mu \) is unnecessary in this case.

**IV. The global solution of a constrained optimization problem**

Consider a rational function \( \frac{f(x)}{h(x)} \) and assume that it is desired to find the infimum of this function over the region \( D \) described by a given set of rational functions \( \frac{g_1(x)}{u_1(x)}, \frac{g_2(x)}{u_2(x)}, \ldots, \frac{g_k(x)}{u_k(x)} \) as follows:

\[
D = \left\{ x : \frac{g_1(x)}{u_1(x)} \geq 0, \ldots, \frac{g_k(x)}{u_k(x)} \geq 0 \right\} \tag{29}
\]

This problem will be investigated next.

**A. An objective function bounded from below**

As the first step to find \( \alpha_* \), one should verify whether \( \frac{f(x)}{h(x)} \) is bounded from below or not. This can sometimes be inferred from the nature of \( \frac{f(x)}{h(x)} \). However, in general one should compute the infimum of \( \frac{f(x)}{h(x)} \) by exploiting the method proposed in the previous section. Assume that \( \frac{f(x)}{h(x)} \) is bounded from below, and denote a lower bound on it with \( L_1 \) (\( L_1 \) can be considered as the infimum of \( \frac{f(x)}{h(x)} \) obtained using the approach given in Section III). Note that it is not important how tight the lower bound \( L_1 \) is. Find an arbitrary point \( \tilde{x}_0 \) belonging to the region \( D \). Define \( L_2 \) as \( \frac{f(\tilde{x}_0)}{h(\tilde{x}_0)} \). Consider now the following objective function:

\[
\Phi_1(x,\mu) = \frac{f(x)}{h(x)} + L_2 - L_1 \frac{k}{4} \sum_{i=1}^{k} \left( \mu_i^2 - \frac{g_i(x)}{u_i(x)} \right)^2 \tag{30}
\]

where \( \mu = [\mu_1 \mu_2 \cdots \mu_k] \). The following theorem presents one of the important properties of \( \Phi_1(x,\mu) \).

**Theorem 6:** Assume that \( (x_o,\mu_o) \) is a local minimum point of the function \( \Phi_1(x,\mu) \) given in (30). If \( \Phi_1(x_o,\mu_o) \leq L_2 \), then \( x_o \) is a local minimizer of the function \( \frac{f(x)}{h(x)} \) over the region \( D \), and moreover \( \Phi_1(x_o,\mu_o) \) is equal to \( \frac{f(\tilde{x}_0)}{h(\tilde{x}_0)} \).

**Proof:** It is obvious that \( \frac{u_i(x)}{g_i(x)} \), \( i = 1, 2, \ldots, k \), is nonnegative, because otherwise:

\[
\Phi_1(x_o,\mu_o) \geq \frac{f(x_o)}{h(x_o)} + L_2 - L_1 \frac{k}{4} \left( \mu_o \frac{g_i(x_o)}{u_i(x_o)} - \frac{u_i(x_o)}{g_i(x_o)} \right)^2 > L_1 + \frac{L_2 - L_1}{4} \times (0 + 2)^2 = L_2 \tag{31}
\]
where $\mu_{i_0}$ represents the $i_0$th element of $\mu$). This contradicts the assumption of $\Phi_1(x_0, \mu_0) \leq L_2$. Note that the above inequality is attained based on the fact that the summation of any positive number and its inverse is at least equal to 2. Hence, $\frac{u_i(x_0)}{g_i(x_0)}$ is nonnegative for any $i \in \{1, 2, ..., k\}$. Assume for now that the first-order and the second-order necessary conditions hold at $x = x_o$. One can write:

$$0 = \frac{\partial \Phi_1(x, \mu)}{\partial \mu_1} \bigg|_{(x_0, \mu_0)} = \mu_{i_0}(L_2 - L_1) \left( \mu_{i_0}^2 - \frac{g_i(x_0)}{u_i(x_0)} - \frac{u_i(x_0)}{g_i(x_0)} \right)$$

(32)

The above relation is met for either $\mu_{i_0} = 0$ or $\mu_{i_0}^2 - \frac{g_i(x_0)}{u_i(x_0)} - \frac{u_i(x_0)}{g_i(x_0)} = 0$. Nevertheless, $\mu_{i_0} = 0$ is infeasible, because otherwise:

$$\frac{\partial^2 \Phi_1(x, \mu)}{\partial \mu_1^2} \bigg|_{(x_0, \mu_0)} = (L_2 - L_1) \left( \mu_{i_0}^2 - \frac{g_i(x_0)}{u_i(x_0)} - \frac{u_i(x_0)}{g_i(x_0)} \right) + 2 \mu_{i_0}^2 (L_2 - L_1) \leq 0$$

(33)

This contradicts the assumption that $(x_o, \mu_0)$ is a local minimizer of the function $\Phi(x, \mu)$. As a result:

$$\mu_{i_0}^2 - \frac{g_i(x_0)}{u_i(x_0)} - \frac{u_i(x_0)}{g_i(x_0)} = 0, \quad i \in \{1, 2, ..., k\}$$

(34)

On using the equation (34) and the first-order necessary condition $\nabla \Phi(x, \mu) \big|_{(x_0, \mu_0)} = 0$, it is straightforward to show that $\nabla f(x_0) \big|_{x_0} = 0$. Therefore, the first-order necessary condition is satisfied for the point $x_0$ to be a minimizer of the function $f(x)$ over the domain $D$. Regarding the second-order necessary condition, the Hessian of the function $\Phi_1(x, \mu)$ is required to be obtained. One can write:

$$\frac{\partial^2 \Phi_1(x, \mu)}{\partial x_i \partial x_j} \bigg|_{(x_0, \mu_0)} = \frac{\partial^2 f(x)}{\partial x_i \partial x_j} \bigg|_{x_0} + \frac{L_2 - L_1}{2} \frac{\partial^2 g_i(x)}{\partial x_i} + \frac{\partial^2 u_i(x)}{\partial x_i} \bigg|_{x_0}$$

(35)

for any $i, j \in \{1, 2, ..., n\}$. Moreover:

$$\frac{\partial^2 \Phi_1(x, \mu)}{\partial x_i \partial \mu_j} \bigg|_{(x_0, \mu_0)} = -\mu_{i_0}(L_2 - L_1) \left( \frac{\partial g_i(x)}{u_i(x)} + \frac{\partial u_i(x)}{g_i(x)} \right) \bigg|_{x_0}, \quad i \in \{1, 2, ..., n\}, \quad l \in \{1, 2, ..., k\}$$

(36)

and it can be concluded from the equation (34) that:

$$\frac{\partial^2 \Phi_1(x, \mu)}{\partial \mu_i^2} \bigg|_{(x_0, \mu_0)} = 2 \mu_{i_0}^2 (L_2 - L_1), \quad i \in \{1, 2, ..., k\}$$

(37)

and

$$\frac{\partial^2 \Phi_1(x, \mu)}{\partial \mu_i \partial \mu_j} \bigg|_{(x_0, \mu_0)} = 0, \quad i, j \in \{1, 2, ..., k\}, \quad i \neq j$$

(38)

Hence, using the equations (35), (36), (37) and (38), and noting that the second-order condition holds for the function $\Phi_1(x, \mu)$ at the point $(x_0, \mu_0)$, one can deduce that:

$$\nabla^2 \Phi_1(x_0, \mu_0) = \begin{bmatrix}
\nabla^2 f(x_0) \bigg|_{x_0} + \frac{L_2 - L_1}{2} T(x_0)T(x_0)^T - (L_2 - L_1) \tilde{T}(\mu_0)T(x_0)^T \\
\tilde{T}(\mu_0)^2 (L_2 - L_1)
\end{bmatrix} > 0$$

(39)

where $T(x)$ is a $n \times k$ matrix, whose $(i, l)$ element is equal to:

$$\frac{\partial g_i(x)}{u_i(x)} + \frac{\partial u_i(x)}{g_i(x)}$$

(40)

and $\tilde{T}(\mu)$ is a $k \times k$ block diagonal matrix with the $(i, i)$ entry equal to $\mu_{i_0}$. The Schur complement formula can be applied to the matrix given in (39) to arrive at the following inequality:

$$0 < \nabla^2 f(x_0) \bigg|_{x_0}$$

(41)

So far, the theorem is proved for the case when $(x_0, \mu_0)$ satisfies the optimality conditions for the function $\Phi_1(x, \mu)$. Using a similar approach, it can be proved in general.

The application of Theorem 6 is twofold. Firstly, the terms $(\mu_i^2 - \frac{g_i(x)}{u_i(x)} - \frac{u_i(x)}{g_i(x)})^2, \quad i = 1, 2, ..., k,$ act as barrier functions. In other words, as soon as $\frac{g_i(x)}{u_i(x)}$ becomes negative, this barrier term increases the value of the objective function $\Phi_1(x, \mu)$ at least by 4. Thus, one can start from the interior point $x = x_0$ and an arbitrary $\mu$, and employ a proper numerical algorithm for minimizing the function $\Phi_1(x, \mu)$, in order to find a local minimum of the function $\frac{h(x)}{f(x)}$ over the region $D$. This method
can be envisaged as an exact penalty approach, as it is not a sequential optimization [28], [29]. This penalty function approach can be superior to many of the existing methods, such as the improved versions of the inverse barrier method [28]. Secondly, one can determine the global solution of the problem using SOS. This will be explained below.

**Corollary 1:** The infimum of the function $\Phi_1(x, \mu)$ is identical to that of the rational function $\frac{f(x)}{h(x)}$ over the region defined by $\mathcal{D}$.

**Proof:** The proof follows from Theorem 6 and on noting that Theorem 6 can be similarly applied to the case when the infimum of $\frac{f(x)}{h(x)}$ is unattainable.

**Remark 3:** Using the SOS-based method proposed in the previous section for finding the infimum of an unconstrained rational function, one can determine the infimum of the function $\Phi_1(x, \mu)$, which is, in fact, the infimum of the rational function $\frac{f(x)}{h(x)}$ over $\mathcal{D}$ (according to Corollary 1).

**B. An unbounded objective function over a compact region $\mathcal{D}$**

$\frac{f(x)}{h(x)}$ is regarded here as a function unbounded from below, for which the method explained in the previous subsection is inapplicable. Nevertheless, a few assumptions are required to be made in order to develop the relevant results. First, the region $\mathcal{D}$ is required to be compact. Next, it is assumed that there exist two nonidentical balls of the known radii $r_1$ and $r_2$ such that $\frac{f(x)}{h(x)}$ is bounded from below inside these two balls, and also both of the balls contain the region $\mathcal{D}$. Without loss of generality, assume that the center of both balls is the origin, and that $r_1 > r_2$. At this point, a lower bound for the rational function $\frac{f(x)}{h(x)}$ inside the ball of radius $r_1$ should be provided. Denote this lower bound to be found with $\hat{L}_1$. Note that $\hat{L}_1$ can sometimes be inferred from the structure of $\frac{f(x)}{h(x)}$. However, one can find the infimum of $\frac{f(x)}{h(x)}$ in the ball, and consider it as $\hat{L}_1$. In order to compute infimum of $\frac{f(x)}{h(x)}$ in the ball (or a lower bound on it), one can combine Putinar’s theorem and the technique used in the proof of Theorem 5 to conclude that a lower bound on this infimum is equal to the maximum value of $\alpha$ for which there exist two SOS polynomials $\phi_1^*(x, \mu)$ and $\phi_2^*(x, \mu)$ such that:

$$f(x, \mu) + \varepsilon - \alpha (\hat{h}(x, \mu) + \varepsilon^2) = (r_1^2 - xx^T - \mu^2) \phi_1^*(x, \mu) + \phi_2^*(x, \mu)$$

(42)

where $\varepsilon$ is chosen sufficiently small (it is assumed that $f(x)$ and $\hat{h}(x)$ are coprime). Define now the following objective function:

$$\Phi_2(x, \mu) = \frac{f(x)}{h(x)} + \frac{L_2 - \hat{L}_1}{4} \sum_{i=1}^k \left( \frac{c^2 - \mu_i^2}{\mu_i} \right)^2 - \frac{g_i(x)}{u_i(x)} - \frac{u_i(x)}{g_i(x)}$$

(43)

where $c = \sqrt{r_1^2 - r_2^2}$, and $L_2$ was defined earlier.

**Theorem 7:** The infimum of the function $\Phi_2(x, \mu)$ inside the ball of radius $r_1$ (i.e., $xx^T + \mu^T \mu \leq r_1$) is the same as the infimum of the rational function $\frac{f(x)}{h(x)}$ over $\mathcal{D}$.

**Proof:** The proof for the case when the infimum of the rational function $\frac{f(x)}{h(x)}$ over $\mathcal{D}$ is attainable will be given here, and can be simply extended to the general case. Denote the global minimizer of the function $\Phi_2(x, \mu)$ inside the ball of radius $r_1$ with $(x_\ast, \mu_\ast)$. It suffices to show that $x_\ast$ is the infimum of the rational function $\frac{f(x)}{h(x)}$ over $\mathcal{D}$. This will be carried out by pursuing the proof of of Corollary 1 and taking the following facts into consideration:

1) The conditions given below are fulfilled:

$$\frac{g_i(x_\ast)}{u_i(x_\ast)} \geq 0, \quad \left( \frac{c^2 - \mu_i^2}{\mu_i} \right)^2 - \frac{g_i(x_\ast)}{u_i(x_\ast)} - \frac{u_i(x_\ast)}{g_i(x_\ast)} = 0, \quad i \in \{1, 2, ..., k\}$$

(44)

where $\mu_i$ represents the $i^{th}$ entry of $\mu$. Thus, $x_\ast$ lies inside the region $\mathcal{D}$.

2) Since the ball of radius $r_2$ contains the region $\mathcal{D}$, $x_\ast$ is located inside this ball, i.e., $xx_\ast^T < r_2^2$.

3) For any arbitrary $x$, if $\frac{g_i(x)}{u_i(x)}$ is nonnegative, the equation:

$$\left( \frac{c^2 - \mu_i^2}{\mu_i} \right)^2 - \frac{g_i(x)}{u_i(x)} - \frac{u_i(x)}{g_i(x)} = 0, \quad i \in \{1, 2, ..., k\}$$

(45)

has a solution between 0 and $c$ for $\mu_i$. This is a consequence of the fact that the function $\frac{c^2 - \mu_i^2}{\mu_i}$ can take any nonnegative value, when $\mu_i$ changes from 0 to $c$.

4) Since $x_\ast$ is inside a ball with the radius $r_2$, and all of the entries of $\mu_\ast$ are between 0 and $c$, thus $xx_\ast^T + \mu_\ast \mu_\ast^T$ is at most equal to $r_2^2$.

Write the function $\Phi_2(x, \mu)$ introduced in (43) as $p_1(x, \mu) + p_2(x, \mu)$, where $p_1(x, \mu)$ and $p_2(x, \mu)$ are two coprime polynomials, and also $p_2(x, \mu)$ is always nonnegative. Note that if such representation is not possible, it implies that the infimum is found to be $-\infty$, as pointed out earlier in Section III. The following theorems can be concluded from the results of Theorem 7 and Putinar’s theorem by employing an approach similar to the proofs of Theorems 4 and 5.
Theorem 8: Let the infimum of $f(x)$ over $\mathcal{D}$ be denoted by $\alpha_*$. A lower bound for $\alpha_*$ is the maximum value of $\alpha$ for which there exist a polynomial $\phi_1(x, \mu)$ and a SOS polynomial $\phi_2(x, \mu)$ such that:

$$p_1(x, \mu) - \alpha p_2(x, \mu) = (r^2 - xx^T - \mu \mu^T) \phi_1(x, \mu) + \phi_2(x, \mu)$$

(46)

Theorem 9: For any $\varepsilon > 0$, let $\alpha_0^\varepsilon$ denote the maximum value of $\alpha$ for which there exist two SOS polynomials $\phi_1^\varepsilon(x, \mu)$ and $\phi_2^\varepsilon(x, \mu)$ such that:

$$p_1(x, \mu) + \varepsilon - \alpha (p_2(x, \mu) + \varepsilon^2) = (r^2 - xx^T - \mu \mu^T) \phi_1^\varepsilon(x, \mu) + \phi_2^\varepsilon(x, \mu)$$

(47)

Then $\alpha_0^\varepsilon$ either equals $\alpha_*$ or converges to $\alpha_*$ as $\varepsilon \to 0$.

C. An unbounded objective function over a non-compact region $\mathcal{D}$

Assume that a lower bound on the infimum $\alpha_*$ is available. Denote this lower bound with $\bar{L}_1$. It is to be noted that various simple approaches can be used to find a value for $\bar{L}_1$. For instance, there are several algorithms which find $\alpha_*$ from below by infinite iterations. One can pursue these algorithms in order to obtain a lower bound on $\bar{L}_1$ using a finite number of iterations. As an alternative, $\bar{L}_1$ can manually be obtained by using some well-known inequalities in small-sized problems. Note that the value of $\bar{L}_1$ is not of great importance, as long as it is a lower bound for $\alpha_*$ (i.e., it may be much smaller than $\alpha_*$). Denote the infimum of $\left(\frac{f(x)}{h(x)} - \bar{L}_1\right)^2$ over the region $\mathcal{D}$ with $\beta$. It can be easily verified that $\alpha_* = \sqrt{\beta + \bar{L}_1}$ (note that $f(x)/h(x) - \bar{L}_1 \geq 0$ over $\mathcal{D}$). Define now the following objective function:

$$\Phi_3(x, \mu) = \left(\frac{f(x)}{h(x)} - \bar{L}_1\right)^2 + L_2 + \frac{1}{4} \sum_{i=1}^{k} \left(\mu_i^2 - g_i(x) - u_i(x)\right)^2$$

(48)

where $L_2 = \left(\frac{f(x)}{h(x)} - \bar{L}_1\right)^2$. It is to be noted that $\Phi_3(x, \mu)$ is constructed for the function $\left(\frac{f(x)}{h(x)} - \bar{L}_1\right)^2$ in the same way that $\Phi_1(x, \mu)$ is formed for $f(x)/h(x)$. Note also that the term $\frac{L_2+1}{4}$ in the above expression for $\Phi_3(x, \mu)$ is due to the fact that $\left(\frac{f(x)}{h(x)} - \bar{L}_1\right)^2$ is always greater than $-1$. Similar to the previous case, the infimum of $\Phi_3(x, \mu)$ is the same as $\beta$. Therefore, one can pursue the methods given in Section III (for finding the infimum of a rational function) to obtain the infimum of $\Phi_3(x, \mu)$, and accordingly, use the equation $\alpha_* = \sqrt{\beta + \bar{L}_1}$ to find $\alpha_*$. 

Remark 4: In this paper, some simple SOS formulations are presented to find the infimum of both constrained and unconstrained rational functions. However, they are only able to determine the infimum as opposed to the minimizer(s), like many other approaches. Nevertheless, the dual of any SOS problem given here can simply be obtained, by using the Lagrangian technique and following the existing methodologies such as the one introduced in [20], [16]. The dual problem, which also has a SOS formulation, is capable of computing the minimizers. The corresponding computations are systematic, and the formulas are skipped here.

Remark 5: The problem of minimization of a rational function subject to some rational inequalities is investigated in this section. However, the existing SOS-based works consider only polynomial inequalities. As a simple remedy for this restriction, one may consider the following set instead of $\mathcal{D}$ given in (29):

$$\mathcal{D} = \{x : g_1(x)u_1(x) \geq 0, \ldots, g_k(x)u_k(x) \geq 0\}$$

(49)

It is to be noted that no matter if $\mathcal{D}$ and $\mathcal{D}$ are equivalent or not, the expression for $\mathcal{D}$ is more suitable for the proposed formulations. More precisely, the degree of the rational function $\frac{g_1(x)}{u_1(x)} + \frac{g_2(x)}{u_2(x)}$ corresponding to $\mathcal{D}$ can be much smaller than that of $g_i(x)u_i(x)$ corresponding to $\mathcal{D}$.

V. Numerical examples:

Example 1: Consider the polynomial $x_1^8 + x_2^8 + x_3^8 + x_1^4x_2^4 + x_1^4x_3^4 + x_2^4x_3^4 - 3x_1^2x_2^2x_3^2$. The infimum of this polynomial is equal to 0, which cannot be attained in a finite number of iterations by using the method given in [10], as pointed out there. Now, it is desired to apply the method proposed in this paper to this example. According to Theorem 2, a lower bound for the infimum of this polynomial can be considered as the maximum value of $\alpha$ for which there exists a polynomial $\phi_1(x, \mu)$, where $\mathbf{x} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}$, such that the polynomial:

$$x_1^8 + x_2^8 + x_3^8 + x_1^4x_2^4 + x_1^4x_3^4 + x_2^4x_3^4 - 3x_1^2x_2^2x_3^2 - (1 - x_1^2 - x_2^2 - x_3^2 - \mu^2)\phi_1(x, \mu) - \alpha \mu^8$$

(50)

is SOS. Using a 9th-order polynomial $\phi_1(x, \mu)$, one can then solve the above SDP problem by using an appropriate software such as YALMIP or SOSTOOLS, to obtain $\alpha_0 = 0$. On the other hand, the exact value of $\alpha_0$, as pointed out before, is equal to 0 as well. Thus, the proposed relaxation let to the exact solution in a finite number of iterations (indeed, in a few seconds).
Example 2: It is desired to find the infimum of the following rational function:

$$x_1^4 x_2^2 + x_2^4 x_3 + x_3^2$$

$$x_1^2 x_2 x_3$$

(51)

It can be observed that the denominator of this rational function is nonnegative, and that its numerator and denominator are both homogeneous of the same degree. Therefore, it can be concluded from Theorem 5 and Remark 2 that a lower bound on the infimum of this rational function is equal to the maximum value of \( \alpha \) for which there exists a polynomial \( \phi_1(x) \), where \( x = [x_1 \ x_2 \ x_3] \), such that the following polynomial is SOS:

$$x_1^4 x_2^2 + x_2^4 x_3 + x_3^2 - \alpha x_1^2 x_2^2 x_3^2 - (1 - x_1^2 - x_2^2 - x_3^2) \phi_1(x)$$

(52)

The value \( \alpha_0 = 3 \) is obtained by using YALMIP with a 9\(^{th}\) order \( \phi_1(x) \). According to Motzkin polynomial [20], [27], the exact solution is also \( \alpha_s = 3 \). However, using the technique given in [11], the lower bound 0 will be obtained.

Example 3: Consider the problem of minimizing the polynomial \( x_1^2 + (x_1 x_2 - 1)^2 \) under the constraint \( x_1 x_2 \geq 2 \). It is straightforward to show that the infimum of this constrained optimization problem is equal to 1, which is unattainable because it occurs as \( x_1 \) approaches zero, with \( x_2 = \frac{2}{x_1} \). Since the infimum is not attainable, the method [10], which is a combination of the SOS and gradient techniques, cannot treat this optimization problem. For the same reason, most of the rudimentary optimization algorithms fail to solve this problem. On the other hand, since the region defined by the constraint \( x_1 x_2 \geq 2 \) is not compact, the approach given in [20], [15] is ineffective. Besides, the technique of presuming a large ball for the solution is not applicable, as there is no ball to include the minimum point. In addition, the penalty-based methods cannot efficiently handle this constrained optimization problem due to the aforementioned difficulties, and more importantly, due to the fact that the infimum occurs on the boundary of the feasible region (i.e., \( x_1 x_2 = 2 \)), which results in an ill-conditioned Hessian matrix [29].

Now, let the method proposed in the present paper be applied to this example. It is obvious that \( x_1^2 + (x_1 x_2 - 1)^2 \) is nonnegative. Hence, one can choose any negative value for \( L_1 \), say \( L_1 = -3 \). Moreover, it is known that the point \( (x_1, x_2) = (2, 2) \) satisfies the constraint and results in the value 13 for the objective function. Consider now the following objective function:

$$\Phi_1(x, \mu) = x_1^2 + (x_1 x_2 - 1)^2 + 4 \left( \mu^2 - (x_1 x_2 - 2) - \frac{1}{x_1 x_2 - 2} \right)^2$$

(53)

where \( x = [x_1 \ x_2] \). As pointed out earlier, the infimum of \( \Phi_1(x, \mu) \) in the whole space is the same as that of \( x_1^2 + (x_1 x_2 - 1)^2 \) over the noncompact region defined by the inequality \( x_1 x_2 \geq 2 \). Using the SOS method proposed in Theorem 3 with \( \varepsilon = 10^{-10} \), and employing YALMIP software, one will arrive at the infimum \( \alpha_s = 1 \) very rapidly.

VI. Conclusions

In this paper, the problem of computing the infimum of a rational function subject to some rational inequalities is investigated. This problem plays a key role in control design and performance analysis for many real-world systems. It is first shown that the infimum of a polynomial function can be obtained by solving a simple SOS problem. The result is then extended to the case of a rational function. Finally, the problem of finding the infimum of a rational function over a region defined by some other rational functions (in the form of inequalities) is considered, and in the case when the objective function is bounded from below, a simple SOS formulation is presented similar to the one obtained for polynomial functions. In the case of an objective function unbounded from below, the problem is treated by assuming that certain \textit{a priori} knowledge is available. Three illustrative examples are given to clarify the proposed approaches and demonstrate their effectiveness.

References


