MATHEMATICS OF OPERATIONS RESEARCH
Vol. 00, No. 0, Xxxxx 0000, pp. 000–000
ISSN 0364-765X | EISSN 1526-5471 | 00 | 0000 | 0001

INFORMS
DOI 10.1287/xxxx.0000.0000

On the S-Lemma for Univariate Polynomials

Simai HE
Department of Management Science, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong, simaihe@cityu.edu.hk

Man Hong WONG
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, Hong Kong, mhwong@se.cuhk.edu.hk

Shuzhong ZHANG
Department of Industrial and Systems Engineering, University of Minnesota, Minneapolis, MN 55455, USA, zhangs@umn.edu

The so-called S-lemma has played an important role in optimization, both in theory and in applications. The significance of S-lemma is especially pronounced in control theory, robust optimization, and non-convex quadratic optimization. Hitherto, S-lemma is however established only in the domain of quadratic functions. In this paper we shall extend the notion of S-lemma to the class of univariate polynomials of any degree. The applications of the new S-lemma will be discussed as well.

Key words: S-lemma, nonnegative polynomials, sum-of-squares

1. Introduction The fundamental principle of the S-lemma (or the S-procedure) is to support the equivalence of the following two statements:

Statement 1

$$f(x) \geq 0 \quad \forall g_i(x) \geq 0, i = 1, \cdots, k; \tag{1}$$

Statement 2 *There exist nonnegative constants $\lambda_1, \cdots, \lambda_k$, not all zero, such that*

$$f(x) - \sum_{i=0}^{k} \lambda_i g_i(x) \geq 0. \tag{2}$$

It is evident that Statement 1 is a consequence of Statement 2. Conversely, if Statement 1 implies Statement 2, then the procedure is called *lossless*, in which case we say that the S-lemma is applicable. An informative review of the S-lemma and its applications in various fields of mathematics (functional analysis, rank-constrained optimization and generalized convexities) can be found in [10] and the references therein. When $k = 1$ and the discussion is limited to the real quadratic functions (his results were actually more general as we shall discuss later), Yakubovich [15] first proved the equivalence under a regularity condition:

THEOREM 1. *Let $f, g : \mathbf{R}^n \mapsto \mathbf{R}$ be quadratic functions and suppose that there is an $\bar{x} \in \mathbf{R}^n$ such that $g(\bar{x}) < 0$. Then the following two statements are equivalent:*

1. *There is no $x \in \mathbf{R}^n$ such that*
$$\begin{cases} f(x) < 0 \\ g(x) \geq 0; \end{cases}$$
2. *There is a nonnegative number $y \geq 0$ such that*
$$f(x) + yg(x) \geq 0 \; \forall\, x \in \mathbf{R}^n.$$

The proof is based on a convexity result due to Dines [4]:

THEOREM 2. *([4]) If $f,g:\mathbf{R}^n \mapsto \mathbf{R}$ are homogeneous quadratic functions, then the set $\{(f(x),g(x)) : x \in \mathbf{R}\} \subset \mathbf{R}^2$ is convex.*

In the literature, alternative proofs can be found for the afore-mentioned S-lemma (Theorem 1). One such alternative proof can be found in Sturm and Zhang [13] (see also Ben-Tal and Nemirovski [2]). That approach relies on a technical lemma regarding the construction of a specific rank-1 decomposition for a positive semidefinite matrix. Another alternative proof can also be based on Lemma 2.3 in [16]:

THEOREM 3. *([16]; see also [1]) Let A_1 and A_2 be positive semidefinite matrices in $\mathbf{R}^{n\times n}$. Then the following are equivalent:*

1. $\max\{x^TA_1x, x^TA_2x\} \geq 0$ *for all* $x \in \mathbf{R}^n$ *(resp.* >0 *for all* $x \neq 0$*);*
2. *There exist* $\mu_1 \geq 0, \mu_2 \geq 0, \mu_1+\mu_2 = 1$ *such that* $\mu_1 A_1 + \mu_2 A_2$ *is positive semidefinite matrices in* $\mathbf{R}^{n\times n}$ *(resp. positive definite).*

Efforts have been made to explore the conditions under which Statement 1 and Statement 2 are equivalent under more general conditions. If the quadratic functions are complex-valued, then Yakubovich [15] actually established the equivalence for the case $k=2$ (see also a derivation in Huang and Zhang [6]). In the real case, Polyak [12] discussed the case $k=2$ under additional assumptions. Recently, Hu and Huang [5] obtained the same result under different assumptions, and also generalized S-lemma to even order tenors. Along a different direction, Tuy and Tuan [14] applied the topological minimax theorem to establish a generalized S-lemma (k can be any integer) where x is restricted to an affine manifold. The results shed some light on the relationship between the S-lemma and the saddle point theorem of the minimax formulation. On the other hand, although the S-lemma has been largely limited to quadratic functions with a single quadratic inequality $g(x)$, its footprints can be found in quite a variety of fields, e.g., stability analysis of a dynamic system using Lyapunov functions (cf. [7]), estimating sum of ellipsoids in error estimating (cf. [12]), and applications in robust optimizations (cf. [3]).

Given the broad interest and wide applicability of the S-lemma, it is natural to consider extending the equivalence between Statements 1 and 2 to other classes of functions. One natural candidate would be the class of univariate polynomials. Indeed, this paper aims to present a first result along this line. We shall illustrate the usefulness of the new S-lemma for the univariate polynomial functions by a few examples in Section 6.

Before proceeding, note, however, that for the univariate polynomial case, a straightforward analogy of the S-lemma as in the quadratic case cannot hold. In other words, there cannot be such nonnegative constant y as in Statement 2 in general. Consider for instance $f(x)=(x+2)(x+3)(x-8)(x-10)$ and $g(x)=x(x+5)(x-1)(x-5)$. It is easy to see that $f(x)\geq 0$ whenever $g(x)\geq 0$ in this case. However, if a positive y would exist in such a way that $f(x)-yg(x)$ is a nonnegative polynomial, then we should be able to find it by solving an optimization problem: $\max_{f(x)-yg(x)\geq 0\,\forall x} y$, which, by a result of Nesterov [9] is a semidefinite program. As we may readily verify, the optimal value of the above problem turns out to be $-\infty$. This means that in general, Statement 1 and Statement 2 are not equivalent. Or putting differently, the original form of the S-lemma fails if $f(x)$ and $g(x)$ are not quadratic functions.

Since requiring y to be a constant is too strong for the statements to be equivalent even for univariate polynomials, we introduce a more relaxed condition to replace Statement 2:

Statement 3 *There exist nonnegative polynomial functions $\lambda_i(x)$, $i=0,1,\cdots,k$, not all identically zero, such that*

$$\lambda_0(x)f(x) - \sum_{i=1}^{k}\lambda_i(x)g_i(x) \geq 0 \quad \forall\, x \in \mathbf{R}. \tag{3}$$

Note that when $\lambda_0(x) \equiv 1$ and all $\lambda_i(x)$ are constants, Statement 3 reduces to Statement 2. Indeed in this paper we shall establish the equivalence between Statements 1 and 3 when $k=1$, under a regularity condition. Note that this result actually provides us with a polynomial-time verifiable certificate for Statement 1 when $k=1$, since we will also show that finding $\lambda_0(x)$ and $\lambda_1(x)$ in Statement 3 can be cast in the form of a *Linear Matrix Inequality*.

2. Definition of several order relations

Let $f(x)$ and $g(x)$ be two polynomials.

Definition 1. Three ordering relations are defined as follows:
(a) $f|_{g\geq 0} \geq 0$ signifies the implication: $g(x)\geq 0 \Rightarrow f(x)\geq 0$;
(b) $f|_{g>0} > 0$ signifies the implication: $g(x)>0 \Rightarrow f(x)>0$;
(c) $f|_{g\geq 0} > 0$ signifies the implication: $g(x)\geq 0 \Rightarrow f(x)>0$.
When all the above orders hold, $f|_{g>^*0} >^* 0$ is used to collectively represent them.

Definition 2. The ordering relations $\succeq$ and $\succ$ are defined as follows:
1. $f(x)\succeq g(x)$ signifies the fact that

$$\exists\, h_1(x)\geq 0, h_2(x)\geq 0 \text{ and } h_3(x)\geq 0 \text{ such that } h_1(x)f(x)=h_2(x)g(x)+h_3(x) \quad \forall x\in \mathbf{R}$$

where $h_1(x), h_2(x), h_3(x)$ are not constantly zero.
2. $f(x)\succ g(x)$ signifies the fact that

$$\exists h_1(x)\geq 0, h_2(x)\geq 0 \text{ and } h_3(x)>0 \text{ such that } h_1(x)f(x)=h_2(x)g(x)+h_3(x) \quad \forall x\in \mathbf{R}.$$

As a matter of notation, $p_1(x)|p_2(x)$ is to indicate the fact that polynomial $p_2(x)$ is divisible by $p_1(x)$, meaning that all the factors of p_1 are factors of p_2 too. Naturally, $\frac{p_2(x)}{p_1(x)}$ is also a polynomial.

3. Technical preparations

In this section we proceed to establish some relationships among these orderings, which will be used to prove the eventual new S-lemma in the next section.

Lemma 1. ***(Decomposition)*** *$f|_{g>^*0} >^* 0 \Longrightarrow \exists f_i(x)$ with $\deg f(x)\leq 2$, $i=1,\cdots,I$, such that*
(a) $f(x)=f_0(x)\prod_{i=1}^{I} f_i(x)$ and
(b) $f_i|_{g>^*0} >^* 0$ for all $i=0,1,\ldots,I$ and
(c) $f_0(x)>0$ for all $x\in\mathbf{R}$.

Proof. We prove this lemma by inductively reducing the number of roots of f, based on the observation that for any two polynomial functions $\theta_1(x)|\theta_2(x)$, $\theta_1|_{g>^*0} >^* 0$ and $\theta_2|_{g>^*0} >^* 0$, we always have $\frac{\theta_2(x)}{\theta_1(x)}|_{g>^*0} >^* 0$. If we can find $\hat{f}(x)|f(x)$ with $\hat{f}|_{g>^*0} >^* 0$, $1\leq \deg \hat{f}(x)\leq 2$, then there is a decomposition for f if and only if there is a decomposition for $\frac{f}{\hat{f}}$.

If $f(x)$ has no real root, then either $f>0$ or $f<0$. In the first case, the decomposition is trivial. In the second case, g has to be negative, therefore $f=f_0f_1$ with $f_0=-f$ and $f_1=-1$ is a decomposition.

If x_0 is a real root of $f(x)$ with even degree of multiplicity, then $(x-x_0)^2>0\ \forall x\neq x_0$ and $g(x_0)<0$, therefore $(x-x_0)^2|_{g>^*0} >^* 0$ holds trivially. Applying the inductive process on $\frac{f(x)}{(x-x_0)^2}$, we can further decompose the polynomial until no multiple root exists.

Suppose that x_0 is a real root of $f(x)$. If it is a multiple root, then $(x-x_0)^2>0\ \forall x\neq x_0$ and $g(x_0)<0$, therefore $(x-x_0)^2|_{g>^*0} >^* 0$ holds trivially. Note that $\frac{f(x)}{(x-x_0)^2}|_{g>^*0} >^* 0$, we can further decompose the polynomial function $\frac{f(x)}{(x-x_0)^2}$ by induction. If x_0 is a single root, assume $\{x: g(x)\geq 0\}=\bigcup_{r=1}^{R}[a_r,b_r]$, where $a_r\leq b_r<a_{r+1}$ (a_1 can be $-\infty$ and b_R can be $+\infty$). Define $b_0:=-\infty$ if a_1 is bounded, and $a_{R+1}:=+\infty$ if b_R is bounded. If x_0 falls in the same interval $[b_r,a_{r+1}]$ $(0\leq r\leq R)$ with another root x_1 of $f(x)$, then $[(x-x_0)(x-x_1)]|_{g>^*0} >^* 0$ and $(x-x_0)(x-x_1)|f(x)$. Notice that there has to be an even number of roots for any bounded intervals. If there is no other root of $f(x)$ in the same interval, the interval has to be $[-\infty,a_1]$ or $[b_R,+\infty]$. For the first case, noticing that $x_0<a_1$ and $g(x)<0$ for all $x<a_1$, we have $(x-x_0)|_{g>^*0} >^* 0$ and $(x-x_0)|f(x)$. For the second case, similarly we have $(x_0-x)|_{g>^*0} >^* 0$ and $(x_0-x)|f(x)$. $\square$

The contra-positivity of Definition 1 is presented as a lemma here without proof.

LEMMA 2.

$$f|_{g\geq 0} > 0 \Longleftrightarrow -g|_{-f\geq 0} > 0$$

and

$$f|_{g\geq 0} \geq 0 \Longleftrightarrow -g|_{-f> 0} > 0.$$

Similarly, we also have

LEMMA 3.

$$f(x) \succeq g(x) \Longleftrightarrow -g(x) \succeq -f(x)$$

and

$$f(x) \succ g(x) \Longleftrightarrow -g(x) \succ -f(x).$$

LEMMA 4. *Let* $\deg f(x) \leq 2$ *or* $f > 0$ *and* $\deg g(x) \leq 2$ *or* $-g > 0$. *Then*
(I) $f|_{g\geq 0} > 0 \Longrightarrow f(x) \succ g(x)$;
(II) $f|_{g\geq 0} \geq 0 \Longrightarrow f(x) \succeq g(x)$;
(III) $f|_{g> 0} > 0 \Longrightarrow f(x) \succ g(x)$.

Proof. Since the arguments are almost identical, we will only prove the lemma for *(I)*.

If $f > 0$, then setting $h_1 = 1$, $h_2 = 0$ and $h_3 = f$ we have $h_1 f = h_2 g + h_3$. If $-g > 0$, then $h_1 = 0$, $h_2 = 1$ and $h_3 = -g$ leads to $h_1 f = h_2 g + h_3$.

When $\deg f(x) = 0$ or $\deg g(x) = 0$, it is trivial.

We will explore the combinations where $\deg f(x) = 1, 2$ with $\deg g(x) = 1, 2$.

(i) $\deg f(x) = 1$ and $\deg g(x) = 1$. Let $f(x) = a_1 x + b_1$ and $g(x) = a_2 x + b_2$ with $a_1, a_2 \neq 0$. Then $f(x) = \frac{a_1}{a_2} g(x) + b_1 - \frac{a_1 b_2}{a_2}$. Take $h_1(x) = 1, h_2(x) = \frac{a_1}{a_2}$ and $h_3(x) = b_1 - \frac{a_1 b_2}{a_2}$ $|_{g\geq 0} > 0$ implies $a_1/a_2 > 0$ and $f(-b_2/a_2) > 0$. Therefore $h_2(x) \geq 0$ and $h_3(x) > 0$ for all $x \in \mathbf{R}$.

(ii) $\deg f(x) = 2$ and $\deg g(x) = 1$.
Let $f(x) = a_1 x^2 + b_1 x + c_1$ and $g(x) = a_2 x + b_2$ with $a_1, a_2 \neq 0$. Note that $g(x) > 0$ either when x approaches $+\infty$ or $-\infty$. Therefore $f(x) > 0$ when x approaches $+\infty$ or $-\infty$, which means $a_1 > 0$. Because of the affine mapping $x \to -x$, we only need to consider the case $a_2 > 0$. Let $x_0 = -\frac{b_2}{a_2}$ and $x_1 = -\frac{b_1}{2a_1}$. Then $f(x_1) = c_1 - \frac{b_1^2}{4a_1}$ and $f(x) = a_1(x - x_1)^2 + f(x_1)$. If $x_1 \geq x_0$, then we have $g(x_1) \geq 0$ so $f(x_1) > 0$. Therefore $h_1 = 1$, $h_2 = 0$ and $h_3 = f > 0$. If $x_1 < x_0$, then $f(x_0) > 0$ and $f'(x_0) > 0$. Take $h_1 = 1$, $h_2 = f'(x_0)/a_2$, and $h_3(x) = f(x) - f'(x_0)(x - x_0)$. Obviously $h_1, h_2 > 0$. For $h_3(x)$, since $f(x) \geq f'(x_0)(x - x_0) + f(x_0)$, we have $h_3(x) \geq f(x_0) > 0$ for all $x \in \mathbf{R}$. Also, it is easy to see that $h_3(x)$ is a strongly convex function because $f(x)$ is strongly convex.

(iii) $\deg f(x) = 1$ and $\deg g(x) = 2$. By Lemma 2 and Lemma 3, it follows directly from (ii).

(iv) $\deg f(x) = 2$ and $\deg g(x) = 2$. We split further our discussion into three subcases: (iv.a) $f(x)$ is convex and $g(x)$ is concave; (iv.b) both $f(x)$ and $g(x)$ are convex (when g is convex, f has to be convex); (iv.c) both $f(x)$ and $g(x)$ are concave. By Lemma 2 and Lemma 3, (iv.c) is equivalent to (iv.b). Therefore we will only discuss the first two cases.

(a) The two closed intervals (maybe empty) $I_f = \{x : f(x) \leq 0\}$ and $I_g = \{x : g(x) \geq 0\}$ will have to be disjoint. Therefore there exists an x_0 such that x_0 separates the two intervals. Without losing generality, we assume I_f to be on the left side and I_g to be on the right side. Then $f|_{x-x_0\geq 0} > 0$ and $(x - x_0)|_{g\geq 0} > 0$. In the cases (ii) and (iii), we know $f(x) \succ x - x_0$ and $x - x_0 \succ g(x)$, and therefore $f(x) \succ g(x)$. It is easy to verify that the h_1, h_2 functions so constructed are constants.

(b) If $g(x) \geq 0$ or $f(x) \leq 0$ for all $x \in \mathbf{R}$, then it is trivial. Otherwise, there exists a x_0 such that $g(x_0) < 0$ and $f(x_0) > 0$. Define

$$\hat{f}(x) = (x - x_0)^2 f\left(\frac{1}{x - x_0}\right)$$

and

$$\hat{g}(x) = (x - x_0)^2 g\left(\frac{1}{x - x_0}\right).$$

Because $g(x_0) < 0$, we know $\hat{g}(x)$ is concave quadratic. Similarly, $\hat{f}(x)$ is convex. For all $x \neq x_0$, if $\hat{g}(x) \geq 0$ then $g\left(\frac{1}{x-x_0}\right) \geq 0$, therefore $f\left(\frac{1}{x-x_0}\right) > 0$, and consequently $\hat{f}(x) > 0$. Also, $\hat{g}(x_0) > 0$ and $\hat{f}(x_0) > 0$ because both f and g are convex quadratic functions. So we have $\hat{f}|_{\hat{g}\geq 0} > 0$. It follows from (iv.a) that there exist nonnegative constants $\hat{h}_1$, $\hat{h}_2$, and $\hat{h}_3(x) > 0$ for all $x \in \mathbf{R}$ such that

$$\hat{h}_1 \hat{f} = \hat{h}_2 \hat{g} + \hat{h}_3.$$

Let $h_1 = \hat{h}_1$, $h_2 = \hat{h}_2$, $h_3(x) = x^2 \hat{h}_3\left(\frac{1}{x} + x_0\right)$. For any $x \neq 0$, let $y = \frac{1}{x} + x_0$. Noticing that $x = \frac{1}{y - x_0}$, we have $f(x) = x^2 \hat{f}\left(\frac{1}{x} + x_0\right)$ and $g(x) = x^2 \hat{g}(\frac{1}{x} + x_0)$. Therefore

$$h_1 f(x) - h_2 g(x) - h_3(x) = x^2 \left[\hat{h}_1 \hat{f}(y) - \hat{h}_2 \hat{g}(y) - \hat{h}_3(y)\right] = 0.$$

Taking the limit $x \to 0$, we know that the above equality also holds for $x = 0$. Therefore $h_1 f(x) = h_2 g(x) + h_3(x)$ for all $x \in \mathbf{R}$. For any $x \neq 0$, because $\hat{h}_3(y) > 0$, it follows that $h_3(x) = x^2 \hat{h}_3(y) > 0$. For $x = 0$, $h_3(0) > 0$ follows from the fact that $\hat{h}_3$ is strongly convex (see case (ii)).
□

LEMMA 5. *$f_i(x) \succ g(x)$ for $i = 1, \cdots, I \Longrightarrow \prod_{i=1}^{I} f_i(x) \succ g(x)$. Similarly, $f_i(x) \succeq g(x)$ for $i = 1, \cdots, I \Longrightarrow \prod_{i=1}^{I} f_i(x) \succeq g(x)$.*

Proof. We shall prove the first part of the lemma, while the second part can be done similarly. Consider first the case $I = 2$, and $f_1(x) \succ g(x)$ and $f_2(x) \succ g(x)$, i.e. there exist nonnegative $\theta_j^i(x)$ for $i = 1, 2$ and $j = 1, 2, 3$ such that

$$\theta_1^1(x) f_1(x) = \theta_2^1(x) g(x) + \theta_3^1(x) \quad \text{and} \quad \theta_1^2(x) f_2(x) = \theta_2^2(x) g(x) + \theta_3^2(x).$$

Then

$$\theta_1^1(x)\theta_1^2(x) f_1(x) f_2(x) = \left(\theta_2^1(x)\theta_2^2(x)[g(x)]^2 + \theta_3^1(x)\theta_3^2(x)\right) + \left(\theta_2^1(x)\theta_3^2(x) + \theta_2^2(x)\theta_3^1(x)\right) g(x).$$

Taking $h_1(x) = \theta_1^1(x)\theta_1^2(x)$, $h_2(x) = \theta_2^1(x)\theta_3^2(x) + \theta_2^2(x)\theta_3^1(x)$, and $h_3(x) = \theta_2^1(x)\theta_2^2(x)[g(x)]^2 + \theta_3^1(x)\theta_3^2(x)$, the claimed lemma follows for the case $I = 2$. If $I = 3$ then we let $\widetilde{f_1}(x) = f_1(x) f_2(x)$ and $\widetilde{f_2}(x) = f_3(x)$ and then apply the lemma for $I = 2$. The process can be repeated recursively this way to yield the desired result for any value of I. □

4. Proof of the S-lemma

In this section, we shall prove our new variants of S-lemma. The main results are Theorems 4 and 5, the proofs of which are based on the technical lemmas introduced in Section 3. Figure 1 is a flowchart to illustrate the main steps of the proof.

THEOREM 4 **(Almost S-lemma)**. *$f|_{g\geq 0} > 0 \Longrightarrow f(x) \succ g(x)$.*

Proof. We shall prove the theorem in a few steps. First, notice that

$$\begin{aligned} & f|_{g\geq 0} > 0 \\ \overset{\text{Lemma 1}}{\Longrightarrow} & \exists f_i(x) \text{ with } \deg f_i(x) \leq 2, i = 1, \cdots, I, \text{and } f_0(x) > 0\ \forall x \in \mathbf{R} \text{ such that} \\ & f(x) = f_0(x) \textstyle\prod_{i=1}^{I} f_i(x) \text{ and } f_i|_{g\geq 0} > 0. \end{aligned}$$

This further implies that

$$
\begin{aligned}
& -g|_{-f_i\ge 0} > 0 \text{ for } i=1,\cdots,I, \\
\overset{\text{Lemma 2}}{\Longrightarrow} \; & \forall i=1,\cdots,I, \exists g_j^i(x) \text{ with } \deg g_j^i(x)\le 2, j=1,\cdots,J, \text{and } g_0(x)>0\,\forall x\in\mathbf{R} \\
& \text{such that } -g(x)=g_0(x)\textstyle\prod_{j=1}^J g_j^i(x) \text{ and } g_j^i|_{-f_i\ge 0}>0 \\
\overset{\text{Lemma 1}}{\Longrightarrow} \; & \forall i=1,\cdots,I, g_j^i(x)\succ -f_i(x), j=1,\cdots,J \\
\overset{\text{Lemma 4}}{\Longrightarrow} \; & \forall i=1,\cdots,I, -g(x)=g_0(x)\textstyle\prod_{j=1}^J g_j^i(x)\succ -f_i(x) \\
\overset{\text{Lemma 3}}{\Longleftrightarrow} \; & \forall i=1,\cdots,I, f_i(x)\succ g(x) \\
\overset{\text{Lemma 5}}{\Longrightarrow} \; & f(x)=f_0(x)\textstyle\prod_{i=1}^I f_i(x)\succ g(x). \quad \square
\end{aligned}
$$

With a regularity condition, we present the following S-lemma.

THEOREM 5 **(S-lemma)**. *Assume* $\gcd(g,g')=1$ *(regularity condition). Then* $f|_{g\ge 0}\ge 0 \Longleftrightarrow f(x)\succeq g(x)$.

Proof.

"$\Longrightarrow$" The arguments in the proof in Theorem 4 remains valid for $f|_{g\ge 0}\ge 0$.

"$\Longleftarrow$" We are going to show a contradiction for "$f(x)<0$ when $g(x)\ge 0$", given that $\gcd(g,g')=1$. Note that $\gcd(g,g')=1$ implies that if x^* is a root of g, then x^* must not be a local maximum. On the other hand, consider $h_1f=h_2g+h_3$, as defined in the ordering relation $\succeq$. Hence $g(x^*)\ge 0$ implies that $h_1(x^*)f(x^*)\ge 0$. If $f(x^*)<0$, then $h_1(x^*)=0$. By the continuity of f and h_1, there exists $\epsilon>0$ such that $\forall\,\bar{x}\in[x^*-\epsilon,x^*+\epsilon]\setminus x^*$, we have $f(\bar{x})<0$ and $h_1(\bar{x})>0$. Thus, $h_2(\bar{x})g(\bar{x})=f(\bar{x})h_1(\bar{x})-h_3(\bar{x})<0$, and so $g(\bar{x})<0$ for all $\bar{x}\in[x^*-\epsilon,x^*+\epsilon]\setminus x^*$. Together with the fact that $g(x^*)=0$, it follows that x^* is a local maximum of g, which contradicts the assumption that $\gcd(g,g')=1$.

□

The new S-lemma may not hold if the regularity condition is removed, as shown by the following example. Let $f(x)=-1$, $g(x)=-x^2$. Note that $\gcd(g,g')=x$. We can choose $h_1(x)=x^2$, $h_2(x)=1$ and $h_3(x)=0$. Then we have $f(x)\succeq g(x)$, but $f|_{g\ge 0}\not\ge 0$.

5. Degree of h_1 and h_2 and extensions of the S-lemma Following the constructions in Lemmas 4 and 3, we establish bounds on the degree of h_1 and h_2, as shown in the theorem below:

THEOREM 6. *Let* $f(x)=f_0(x)\prod_{i=1}^I f_i(x)$ *and* $g(x)=g_0(x)\prod_{j=1}^J g_j^i(x)$, *where* $\deg f_0:=2r<\deg f$ *and* $\deg g_0:=2s<\deg g$. *Furthermore, letting* $h_1(x)$ *and* $h_2(x)$ *be defined under the relation* $f(x)\succ g(x)$. *Then*

1. $\deg h_1(x)=2s+2\left\lceil\frac{\deg f(x)-2r}{2}\right\rceil\left(\left\lceil\frac{\deg g(x)-2s}{2}\right\rceil-1\right)$;
2. $\deg h_2(x)=2r+2\left\lceil\frac{\deg g(x)-2s}{2}\right\rceil\left(\left\lceil\frac{\deg f(x)-2r}{2}\right\rceil-1\right)$.

Proof. The given decomposition of f and g follow from the proof of Theorem 4, where $\deg f_i(x)\le 2$ and $\deg g_j^i(x)\le 2$. Let $\tilde{f}(x)=\prod_{i=1}^I f_i(x)$ and $-\tilde{g}(x)=\prod_{j=1}^J g_j^i(x)$. We first consider the nonnegative function $\tilde{h}_1(x),\tilde{h}_2(x)$ for $\tilde{f}(x)$ and $\tilde{g}(x)$. Note that $\deg\tilde{f}(x)=\deg f(x)-2r$ and $\deg\tilde{g}(x)=\deg g(x)-2s$. From the proof of Lemma 4, there exist constants $\zeta_1^{(i,j)}>0$, $\zeta_2^{(i,j)}>0$ and $\theta^{(i,j)}(x)\ge 0$ with $\deg\theta^{(i,j)}(x)\le 2$ such that

$$\zeta_1^{(i,j)}g_j^i(x)=\zeta_2^{(i,j)}(-f_i(x))+\theta^{(i,j)}(x)$$

leading to

$$\prod_{j=1}^J\zeta_1^{(i,j)}g_j^i(x)=\prod_{j=1}^J\left(\zeta_2^{(i,j)}(-f_i(x))+\theta^{(i,j)}(x)\right)$$

which implies that one can find nonnegative polynomials $\widetilde{\theta}_2^i(x)$ and $\widetilde{\theta}_3^i(x)$ such that

$$-\tilde{g}(x)\prod_{j=1}^{J}\zeta_1^{(i,j)} = -\widetilde{\theta}_2^i(x)f_i(x)+\widetilde{\theta}_3^i(x). \tag{4}$$

Rearranging, we have

$$\widetilde{\theta}_2^i(x)f_i(x)=\tilde{g}(x)\tilde{\zeta}^i+\widetilde{\theta}_3^i(x),$$

where $\tilde{\zeta}^i=\prod_{j=1}^{J}\zeta_1^{(i,j)}$. Therefore,

$$\prod_{i=1}^{I}\widetilde{\theta}_2^i(x)f_i(x)=\prod_{i=1}^{I}\left(\tilde{g}(x)\tilde{\zeta}^i+\widetilde{\theta}_3^i(x)\right).$$

Hence, we may again rearrange to obtain nonnegative polynomials $\tilde{h}_2(x)$ and $\tilde{h}_3(x)$ to satisfy the following equality

$$\tilde{h}_1(x)\tilde{f}(x)=\tilde{h}_2(x)\tilde{g}(x)+\tilde{h}_3(x). \tag{5}$$

In (4), one can verify that there exists i such that $\deg\widetilde{\theta}_2^i(x)=2\left(\left\lceil\frac{\deg g(x)-2s}{2}\right\rceil-1\right)$ and $\deg\widetilde{\theta}_3^i(x)=2\lceil\frac{\deg g(x)-2s}{2}\rceil$. Thus, in (5), $\tilde{h}_1(x)=\prod_{i=1}^{I}\widetilde{\theta}_2^i(x)$ and

$$\deg\tilde{h}_1(x)=\left\lceil\frac{\deg f(x)-2r}{2}\right\rceil\deg\widetilde{\theta}_2^i(x)=2\left\lceil\frac{\deg f(x)-2r}{2}\right\rceil\left(\left\lceil\frac{\deg g(x)-2s}{2}\right\rceil-1\right).$$

Similarly,

$$\deg\tilde{h}_2(x)=\left\lceil\frac{\deg g(x)-2s}{2}\right\rceil\deg\widetilde{\theta}_2^i(x)=2\left\lceil\frac{\deg g(x)-2s}{2}\right\rceil\left(\left\lceil\frac{\deg f(x)-2r}{2}\right\rceil-1\right).$$

Multiplying $f_0(x)g_0(x)$ on both sides of (5), we recover $f(x)$ and $g(x)$ with $h_1(x)=g_0(x)\tilde{h}_1(x)$ and $h_2(x)=f_0(x)\tilde{h}_2(x)$ with their respectively claimed degree. □

Let us remark that it is a trivial case when $\deg f_0(x)=\deg f(x)$ or $\deg g_0(x)=\deg g(x)$, i.e. either f or g has no real root. In particular, we have a practical result for $\deg g(x)\le 2$:

COROLLARY 1. *Assume* $\deg g(x)\le 2$. *Then* h_1 *is a constant and*

$$\deg h_2(x)=2\left(\left\lceil\frac{\deg f(x)}{2}\right\rceil-1\right).$$

This is equivalent to the case when $s=0$ *in Theorem 6.*

Proof. $\deg g(x)\le 2$ implies $s=0$. Then $\deg h_1=0$ and $\left\lceil\frac{\deg g(x)-2s}{2}\right\rceil=1$. Note that $\left\lceil\frac{\deg f(x)-2r}{2}\right\rceil=\left\lceil\frac{\deg f(x)}{2}\right\rceil-r$. Then according to Theorem 6,

$$\deg h_2(x)=2r+2\left(\left\lceil\frac{\deg f(x)}{2}\right\rceil-r-1\right)=2\left(\left\lceil\frac{\deg f(x)}{2}\right\rceil-1\right).$$

□

As an immediate application, we note that the following two classical theorems on nonnegative univariate polynomials follow directly from the new S-lemma.

Corollary 2. *(Polya and Szego [11]) If a polynomial $p(x)$ that satisfies $p(x) \geq 0\,\forall\, x \geq 0$, then there exists nonnegative polynomials $\theta_1(x)$ and $\theta_2(x)$ such that*

$$p(x) = \theta_1(x) + x\theta_2(x) \qquad \forall\, x \in \mathbf{R}.$$

Proof. Take $g(x) = x$. By Corollary 1, $h_1 \geq 0$ is a constant. Having verified that $\gcd(g, g') = 1$, we can apply S-lemma, i.e., there exists $h_2(x) \geq 0$ such that $p(x) - h_2(x)g(x) = p(x) - xh_2(x) \geq 0$. By letting $h_2(x) = \theta_2(x)$ and $h_3(x) = p(x) - xh_2(x)$, we obtained the desired result. □

Corollary 3. *(Fekete (1935); cf. Lasserre [8]) The polynomial $p(x)$ satisfies $p(x) \geq 0\,\forall\, x \in [0,1]$ if and only if there exists nonnegative polynomials $\theta_1(x)$ and $\theta_2(x)$ such that*

$$p(x) = \theta_1(x) + x(1-x)\theta_2(x) \qquad \forall\, x \in \mathbf{R}.$$

Proof. Take $g(x) = -x(x-1)$ and then follow similar steps as in the proof of Corollary 2. □

Related to the original S-lemma, there is a useful result from Yuan [16] regarding the non-negativity of the maximum of two quadratic functions. It turns out that, for polynomials, this nonnegativity associates with our version of S-lemma as well.

Corollary 4. *(Lemma 2.3 of Yuan [16], or Theorem 5.1 of Ai et al. [1].) Let $f(x)$ and $g(x)$ be two polynomials. Then,*

$$\max\{f(x), g(x)\} \geq 0\,\forall\, x \in \mathbf{R} \Longleftrightarrow f|_{-g>0} \geq 0.$$

Furthermore, if $\gcd(g, g') = 1$, then there exist $h_1(x) > 0$ and $h_2(x) \geq 0$ such that

$$h_1(x)f(x) + h_2(x)g(x) \geq 0 \quad \forall x \in \mathbf{R}. \tag{6}$$

Proof.

$$\begin{aligned} & \max\{f(x), g(x)\} \geq 0 \quad \forall\, x \in \mathbf{R} \\ \Longleftrightarrow\ & \begin{cases} g(x) < 0 \Rightarrow f(x) \geq 0 \\ f(x) < 0 \Rightarrow g(x) \geq 0 \end{cases} \\ \Longleftrightarrow\ & f|_{-g>0} \geq 0. \end{aligned}$$

Note that the two statements in the second step are contra-positive of each other, and therefore they are reduced to one. By Theorem 5, (6) holds. □

Remark 1. When the regularity condition $\gcd(g, g') = 1$ is imposed, $f|_{-g>0} \geq 0$ is equivalent to $f|_{-g\geq 0} \geq 0$. Otherwise, it is not. Consider $f(x) = x^2 - 1$ and $g(x) = -x^2$. In that case, $f|_{-g>0} \geq 0$ is true while $f|_{-g\geq 0} \geq 0$ is not.

6. Examples of applications Similar to the traditional S-lemma, the new S-lemma enables us to replace a *relational* robustness condition (Statement 1) by an *absolute* robustness condition (Statement 3). The latter condition in the case of univariate polynomials is known to have an LMI representation (cf. [9]). Therefore, convex optimization formulation becomes possible. To illustrate this point, we present a few examples below.

Example 1 ***(Cash Flow Management)***
In asset-liability management, we are often required to allocate cash flows in different future periods in order to meet a set of obligations. Suppose that we need to decide the cash flow for a series of obligation O_i at the end of period $i = 0, \cdots, n$. Meanwhile, a series of cash flow, $b_1, \cdots, b_n$ has already been arranged. In case some of cash flows are positive, they can be used to fund the obligations. In order to justify the whole portfolio economically (i.e. a nonnegative future value) under different interest rate environments, we can formulate an optimization problem to find the minimal amount of cash required for each period:

$$\begin{array}{ll} \min & \sum_{i=0}^{n} a_i \\ \text{s.t.} & \sum_{i=0}^{n}(a_i + b_i - O_i)(1+x)^{n-i} \geq 0 \quad \forall x \in [r_1, r_2] \cup [r_3, r_4], \end{array}$$

where a_i's are the decision variables, and $[r_1, r_2]$ and $[r_3, r_4]$ (with $r_1 < r_2 < r_3 < r_4$) reflect different economic environments, say boom and bust. Recall that $x \in [r_1, r_2] \cup [r_3, r_4]$ can be represented by a polynomial. The constraint is an example of Statement 1.

Example 2 ***("Best-fit" Polynomials)***
How functions can be best approximated by polynomials is of central importance in approximation theory. This is achieved by minimizing the quantified error(s) between the given functions and the approximating polynomial. This idea can be illustrated by polynomial regressions.

Given a set of sample points, we can regress it with a polynomial $\hat{f}_i(x)$. If trials are repeated M times, we obtain $N(\leq M)$ possibly different justified $\hat{f}_i(x)$'s. Applying concepts from approximation theory, we can formulate an optimization problem for a "best-fit" polynomial based on the $\hat{f}_i(x)$'s: $\min \sum_{i=1}^{N} |f(x) - \hat{f}_i(x)|$, where $\hat{f}_i(x)$'s are polynomials that are already computed and $f(x)$ is the new unifying approximate polynomial function to be found in this model. In general, the approximation is measured on a particular range $\mathcal{I}$ (that can be an interval or a union of such). Then the formulation is

$$\begin{array}{ll} \min & \sum_{i=1}^{N} t_i \\ \text{s.t.} & -t_i \leq f(x) - \hat{f}_i(x) \leq t_i \quad \forall x \in \mathcal{I}, i = 1, \cdots, N. \end{array}$$

It can be easily seen that the constraints represent another example of Statement 1.

Example 3 ***(Moment Bounds)***
When we estimate the extremal (say maximum) expectation of a certainty quantity $\psi(x)$ based on the moments information of its underlying randomness $x \in \Omega$, it is called a moment bound problem. The estimation is often through its dual formulation. Assuming $\Omega \subseteq \mathbf{R}$, the problem can be formulated as follows:

$$\begin{aligned} \sup \mathbb{E}[\psi(x)] = \inf_{z_0, \cdots, z_n} & \sum_{i=0}^{n} m_i z_i \\ \text{s.t.} & \sum_{i=0}^{n} z_i x^i \geq \psi(x) \quad \forall x \in \Omega, \end{aligned}$$

where m_i are the i-th moment of x. Given more information about the range of x, the "for-all" ($\forall$) condition can be represented by some nonnegative polynomial. Then the constraint set can be cast as an instance of Statement 1, provided that $\psi(x)$ is also a (piecewise) polynomial.

As we mentioned in the beginning of the section, in all these examples, the constraints are in the form of Statement 1. Thus, by means of the new S-lemma we are able to formulate the problems by SDP based on Statement 3, since a nonnegative univariate polynomial can be characterized by LMI. This lends to *convex* optimization formulations of the original problems.

7. A numerical illustration To better understand how the decompositions in Theorem 4 work, let us consider the following numerical example.

Let $f(x)=(x+1)(x-3)(x-5)$ and $g(x)=x(x-1)(x-6)(x-8)(x-9)$. We have $f|_{g\geq 0}\geq 0$.

Decompose $f(x)=f_1(x)f_2(x)$, where $f_1(x)=x+1$ and $f_2(x)=(x-3)(x-5)$. Note that $f_1|_{g\geq 0}\geq 0$ and $f_2|_{g\geq 0}\geq 0$. By Lemma 2, $-g|_{-f_1\geq 0}\geq 0$ and $-g|_{-f_2\geq 0}\geq 0$. Now, decompose $-g(x)=g_1(x)g_2(x)g_3(x)$, where $g_1(x)=x(x-1)$, $g_2(x)=(x-6)(x-8)$ and $g_3(x)=-(x-9)$. By Lemma 2 again, we have

$$\begin{array}{lll} g_1|_{-f_1\geq 0}\geq 0, & g_2|_{-f_1\geq 0}\geq 0, & g_3|_{-f_1\geq 0}\geq 0,\\ g_1|_{-f_2\geq 0}\geq 0, & g_2|_{-f_2\geq 0}\geq 0, & g_3|_{-f_2\geq 0}\geq 0. \end{array}$$

Then by Lemma 4, construct $\theta_3^{(i,j)}(x)$, for $i=1,2$ and j=1,2,3, such that

$$\begin{aligned}
\theta_3^{(1,1)}(x) &= g_1(x)-(-f_1(x)) = x^2+1,\\
\theta_3^{(1,2)}(x) &= \frac{1}{2}g_2(x)-(-f_1(x)) = \frac{1}{2}x^2-6x+25,\\
\theta_3^{(1,3)}(x) &= g_3(x)-(-f_1(x)) = 10,\\
\theta_3^{(2,1)}(x) &= g_1(x)-\frac{1}{2}(-f_2(x)) = \frac{3}{2}x^2-5x+\frac{15}{2},\\
\theta_3^{(2,2)}(x) &= \frac{1}{2}g_2(x)-\frac{1}{2}(-f_2(x)) = x^2-11x+\frac{63}{2},\\
\theta_3^{(2,3)}(x) &= g_3(x)-(-f_2(x)) = x^2-9x+24.
\end{aligned}$$

The first three expressions and last three imply, respectively,

$$g_1(x)\cdot\frac{1}{2}g_2(x)\cdot g_3(x) = \left[\theta_3^{(1,1)}(x)+(-f_1(x))\right]\left[\theta_3^{(1,2)}(x)+(-f_1(x))\right]\left[\theta_3^{(1,3)}(x)+(-f_1(x))\right]$$

and so

$$\frac{1}{2}[-g(x)] = -\widetilde{\theta}_2^1(x)f_1(x)+\widetilde{\theta}_3^1(x). \tag{7}$$

Moreover, writing it differently,

$$g_1(x)\cdot\frac{1}{2}g_2(x)\cdot g_3(x) = \left[\theta_3^{(2,1)}(x)+\frac{1}{2}(-f_2(x))\right]\left[\theta_3^{(2,2)}(x)+\frac{1}{2}(-f_2(x))\right]\left[\theta_3^{(2,3)}(x)+(-f_2(x))\right],$$

and so

$$\frac{1}{2}[-g(x)] = -\widetilde{\theta}_2^2(x)f_2(x)+\widetilde{\theta}_3^2(x). \tag{8}$$

In the above expressions, we used

$$\begin{aligned}
\widetilde{\theta}_2^1(x) &= \left[\theta_3^{(1,1)}(x)\theta_3^{(1,2)}(x)+(-f_1(x))^2\right]+\theta_3^{(1,3)}(x)\left[\theta_3^{(1,2)}(x)+\theta_3^{(1,1)}(x)\right]\\
&= \frac{x^4}{2}-6x^3+\frac{83x^2}{2}-64x+286,\\
\widetilde{\theta}_3^1(x) &= \left[\theta_3^{(1,1)}(x)\theta_3^{(1,2)}(x)+(-f_1(x))^2\right]\theta_3^{(1,3)}(x)+\left[\theta_3^{(1,2)}(x)+\theta_3^{(1,1)}(x)\right](-f_1(x))^2\\
&= \frac{13x^4}{2}-63x^3+\frac{561x^2}{2}+6x+286,\\
\widetilde{\theta}_2^2(x) &= \left[\theta_3^{(2,1)}(x)\theta_3^{(2,2)}(x)+\frac{1}{4}(-f_2(x))^2\right]+\theta_3^{(2,3)}(x)\left[\frac{1}{2}\theta_3^{(2,1)}(x)+\frac{1}{2}\theta_3^{(2,2)}(x)\right]\\
&= 3x^4-\frac{179x^3}{4}+\frac{1019x^2}{4}-\frac{1335x}{2}+\frac{1521}{2},\\
\widetilde{\theta}_3^2(x) &= \left[\theta_3^{(2,1)}(x)\theta_3^{(2,2)}(x)+\frac{1}{4}(-f_2(x))^2\right]\theta_3^{(2,3)}(x)+\left[\frac{1}{2}\theta_3^{(2,1)}(x)+\frac{1}{2}\theta_3^{(2,2)}(x)\right](-f_2(x))^2\\
&= 3x^6-\frac{277x^5}{4}+\frac{2679x^4}{4}-\frac{13901x^3}{4}+\frac{40899x^2}{4}-\frac{32625x}{2}+\frac{22815}{2}.
\end{aligned}$$

Then from (7) and (8)

$$\widetilde{\theta}_2^1(x) f_1(x) \cdot \widetilde{\theta}_2^2(x) f_2(x) = \left[\frac{1}{2} g(x) + \widetilde{\theta}_3^1(x)\right] \left[\frac{1}{2} g(x) + \widetilde{\theta}_3^2(x)\right]$$

and

$$h_1(x) f(x) = h_2(x) g(x) + h_3(x)$$

where

$$\begin{aligned} h_1(x) &= \widetilde{\theta}_2^1(x) \widetilde{\theta}_2^2(x) \\ &= \frac{3x^8}{2} - \frac{323x^7}{8} + \frac{4163x^6}{8} - \frac{31291x^5}{8} + \frac{149435x^4}{8} - \frac{245467x^3}{4} + \frac{588557x^2}{4} - 239577x + 217503, \end{aligned}$$

and

$$\begin{aligned} h_2(x) &= \frac{1}{2}\left[\widetilde{\theta}_3^1(x) + \widetilde{\theta}_3^2(x)\right] \\ &= \frac{3x^6}{2} - \frac{277x^5}{8} + \frac{2705x^4}{8} - \frac{14153x^3}{8} + \frac{42021x^2}{8} - \frac{32613x}{4} + \frac{23387}{4}. \end{aligned}$$

We can verify that

$$2\left\lceil \frac{\deg f(x)}{2} \right\rceil \left(\left\lceil \frac{\deg g(x)}{2} \right\rceil - 1 \right) = 2 \cdot \left\lceil \frac{3}{2} \right\rceil \cdot \left(\left\lceil \frac{5}{2} \right\rceil - 1 \right) = 8 = \deg h_1(x)$$

and

$$2\left\lceil \frac{\deg g(x)}{2} \right\rceil \left(\left\lceil \frac{\deg f(x)}{2} \right\rceil - 1 \right) = 2 \cdot \left\lceil \frac{5}{2} \right\rceil \cdot \left(\left\lceil \frac{3}{2} \right\rceil - 1 \right) = 6 = \deg h_2(x).$$

8. Conclusions The main steps of the proof for the univariate S-lemma and its associated technical lemmas are schematized in Figure 1. In view of the new theoretical development of the S-lemma, three directions may deserve further exploration. One is to study if there is a certain form of the S-lemma in the domain of bivariate quartic polynomials. This motivation comes from the parallel comparison between the existence of sos-polynomials-certificate for nonnegative polynomials and that of the S-lemma. Another direction is to study the existence of S-lemma for univariate complex polynomials. This may not even be a well-posed proposition as there is non-unique way to define nonnegative univariate complex polynomials, but there is certainly a general interest around that topic. Last but not least, it is interesting to verify if the S-lemma holds if the univariate polynomials can accommodate more $g_i(x)$'s. In other words, one may wish to figure out if the equivalence between Statement 1 and Statement 3 holds for $k \geq 2$. It is our belief that there is ample room for further extensions and new applications.

References

[1] Ai, W., Y. Huang, S. Zhang. 2011. New results on hermitian matrix rank-one decomposition. *Mathematical Programming, Series A* **128** 253–283.

[2] Ben-Tal, A., A. Nemirovski. 2001. *Lectures on Modern Convex Optimization: Analysis, Algoritms, and Engineering Applications, MPS/SIAM Series on Optimization 2*. SIAM, Philadephia.

[3] Chen, L., S. He, S. Zhang. 2011. Tight bounds for some risk measures, with applications to robust portfolio selection. *Operations Research* **59**(4) 847–865.

[4] Dines, L.L. 1941. On the mapping of quadratic forms. *Bulletin of the American Mathematical Society* **47**(6) 494–498.

[5] Hu, S.L., Z.H. Huang. 2012. Theorems of the alternative for inequality systems of real polynomials. *Journal of Optimization Theory and Applications* DOI 10.1007/s10957-012-9993-9.

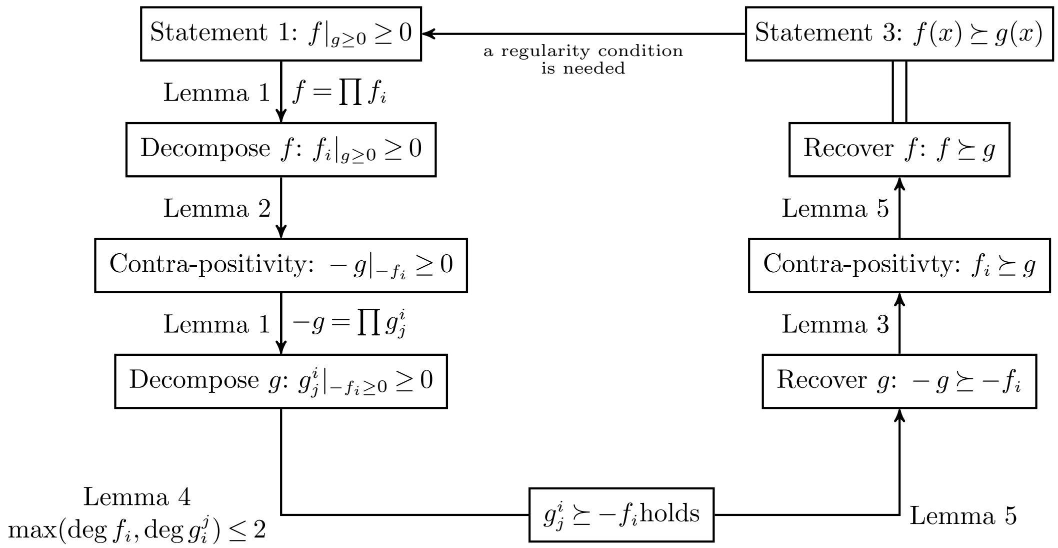

FIGURE 1. The flow of arguments in the proof of the S-lemma

[6] Huang, Y., S. Zhang. 2007. Complex matrix decomposition and quadratic programming. *Mathematics of Operations Research* **32**(3) 758–768.

[7] Jonsson, U.T. 2001. A lecture on the S-procedure.

[8] Lasserre, J.-B. 2010. *Moments, Positive Polynomials, and Their Applications, Imperial College Press Optimization Series Vol. 1*. Imperial College Press, London.

[9] Nesterov, Yu. 2000. Squared functional systems and optimization problems. H. Frenk, K. Roos, T. Terlaky, S. Zhang, eds., *High Performance Optimization Techniques*, chap. 17. Kluwer Academic Publishers.

[10] Pólik, I., T. Terlaky. 2007. A survey of the S-lemma. *SIAM Review* **49**(3) 371–418.

[11] Polya, G., G. Szego. 1976. *Problems and Theorems in Analysis II*. Springer-Verlag.

[12] Polyak, B.T. 1998. Convexity of quadratic transformations and its use in control and optimization. *Journal of optimization theory and applications* **99**(3) 553–583.

[13] Sturm, J.F., S. Zhang. 2003. On cones of nonnegative quadratic functions. *Mathematics of Operations Research* **28** 246–267.

[14] Tuy, H., H.D. Tuan. 2011. Generalized s-lemma and strong duality in nonconvex quadratic programming. http://www.math.ac.vn/library/download/e-print/11/pdf/htuy1101.pdf.

[15] Yakubovich, V.A. 1971. S-procedure in nonlinear control theory. *Vestnik Leningrad University* **1** 62–77.

[16] Yuan, Y.X. 1990. On a subproblem of trust region algorithms for constrained optimization. *Mathematical Programming* **47** 53–63.